Report on Validity and Reliability of Calibrated Psychometric Employability Test

HireLabs

Tested Talent

Assessment Fund

Preface

The **Employability Test** is a psychometric assessment, which is being used in over 22 countries. The Employability Test was initially developed by HireLabs in 2010 with a long term objective to build a certification body that would serve both Employers (the demand-side) and the Job-seekers and their supporting networks (supply-side) of the labor market.

The Employability Test works in unison with UK-based **Tested Talent**, which provides online testing technologies. All statistical analysis and calculations present in this report are facilitated through the algorithms implemented within Tested Talent's technology.

The Employability Test is managed by the **HireLabs Test Creation Board** (HTCB), which has the following mandates:

1. To ensure that the test creation methodology is periodically reviewed and modified in order to keep current with labor market needs.

2. To maintain the integrity of the test creation process to meet validity and reliability criteria

The Tested Talent technology is a S.a.a.S (software As A Service), which can be licensed and white labeled by 3rd parties who like to launch their own assessment centers and testing tools.

Entities interested in launching their own white labeled assessment center can visit www.TestedTalent.com for more information.

The HTCB, which is funded by an independent body called the **Assessment Fund,** can be reached via email at st@hirelabs.sg

Abstract

This report presents the results and analysis of an extensive study conducted by HireLabs on the effective calibration of a **psychometric assessment called the Employability Test**. All testing was conducted online, and all data analysis and calculation were conducted using the technologies and algorithms of UK-based **Tested Talent**.

The objective is to study the effectiveness of customizing (or calibrating) a psychometric tool for the hiring needs of employers. The report examines multiple parameters of validity and reliability in situations where psychometric tests undergo a calibration process. The report also examines the method used to calibrate a psychometric test so that it is localized and is relevant for the ecosystem.

The controls and parameters put in place include the testing of *out-of-school youth* who have zero to two years (0 to 2 years) of work experience.

This report studies the correlation between a candidate's score on a calibrated psychometric test and the likelihood of a company to employ (and retain) that candidate for a duration that is higher than the company's current average.

The report also documents the methodology used to engage employers to participate in the study. It includes the process of calibrating the psychometric tests; the process of aggregating and scrubbing data; the statistical correlation and analysis on the data and how this relates to the validity and the reliability of a calibrated psychometric test.

The report concludes that an effective calibration satisfies multiple validity criteria, and also serves as a reliable screening tool.

Lastly, the report concludes that there is *moderate-to- strong* correlation between a high score (on a calibrated psychometric test) and a higher-than-average duration of employment.

HireLabs continues to advise employers that although the process of calibration does improve the validity and reliability of psychometric testing, employers will receive further clarity and insight through in-person interviews.

Introduction

In Q1 of 2010, HireLabs observed employers faced challenges when recruiting out-of-school youth who had 0 to 2 years of work experience. The challenges were magnified by the global financial crisis, which had led to unprecedented rates of youth unemployment. As a result, employers were experiencing a sharp increase in the number of job applications.

The supply of labor was far greater than the demand. As a result, the excessive supply of resumes was putting undue stress on the HR business units. The employers would be better served if they had a decision tool that would reduce the time for screening the large number of candidates. Especially if it can incorporate the employer's needs. Such tool can be a pre-employment testing tool.

In response to the employers' needs, HireLabs commissioned the creation of a new, calibrated psychometric test called the Employability Test. The objective of calibrated Employability Test was to evaluate the out-of-school-youth on a pre-defined set of competencies and assess their readiness for the job market. In order to introduce calibration, the **HireLabs Test Creation Board (HTCB)** made an active

decision to incorporate construct validity and content validity at the source.

For the calibration to be effective, the HTCB engaged the employers to select:

- The *competencies* that they would like to test.
- The *constructs* (for each competency) that they would like to measure in order to see if a candidate is employable.

By engaging the employers in calibrating the constructs, the HTCB intended to further enhance the correlation between construct validity and predictive validity of the Employability Test.

The HTCB incorporated guidance of the *Standards of Education and Psychological Testing (AERA/APA/NCME 1999)*, (p. 17), which states :

> "a sound validity argument integrates various strands of evidence into coherent account of the degree to which existing evidence and theory support the intended interpretation of the test scores".

The HTCB also continued to act in accordance with the *Standards of Education and Psychological Testing (AERA/APA/NCME 1999)*, (p. 11), which states that the test maker is

> "responsible for furnishing relevant evidence and rationale in support of the intended test use".

In addition, the *Standards* (Standards 6.13, p. 70).states that

> "when substantial changes are made to a test, the tests documentation should be amended, supplemented, or revised to keep information for users current and to provide useful additional information of cautions".

Method

Employer Engagement

A list of 247 employers was identified, from the existing client base, in addition to employers on LinkedIn. These employers were located in 9 countries. The employer identification criteria included:

- The employer has an existing workforce of more than 50 employees.
- The employer has more than 5 active vacancies for entry level positions (that could be catered by the out-of-school youth with 0 to 2 years of experience).
- The employer has a work week of a minimum of 35 hours.
- The employer offers a salary that is market competitive.

The 247 employers were approached via email, which contained a link to the data collection survey. The employers contacted were incentivized to participate in the survey. The incentive was to receive complementary psychometric testing service for a limited period (30 days).

Out of the 247 employers contacted, 92 from 9 countries participated in the survey. **Figure 1** shows a breakdown of the various sectors and the companies who participated in the survey.

Figure 1: A sector-wise breakdown of 92 companies that participated in the survey.

Sector	*Number of companies surveyed*
Telecom	7
Healthcare	9
FMCG	9
Food & Beverage	17
Technology	21
Media & Entertainment	18
Financial	11

Seven Frequently Used Competencies by Employers

HireLabs has an open system for employers to select competencies that they would like to test in potential hires. For more information on HireLabs' open system, please review the elaborate list of competencies page made available to employers by following the link www.hirelabs.sg/ps/testcatalog.html

HTCB has observed that employers prefer different set of competencies at each level of the organizational chart. For the purpose of this study, the HTCB monitored the competencies that were selected by employers for hiring the out-of-school youth having 0 to 2 year of work experience. The HTCB observed that 7 competencies were most frequently selected by the employers. **Figure 2** lists and describes these 7 competencies used in the survey.

Figure 2: A list of 7 competencies prefered by employers (for testing candidates with 0 to 2 years of work experience).

1. *English Proficiency*
   Having knowledge of the structure and content of the English language including the meaning and spelling of words, rules of composition, and grammar.
2. *Computer Literacy*
   Using computers and computer systems (including hardware and software) to execute everyday office administrative tasks, such as entering data, processing information, and connecting and browsing the internet.
3. *Workplace Ethics*
   Having the understanding of what is considered as acceptable actions within the workplace.
4. *Critical Thinking*
   Using logic and reasoning to identify the strengths and weaknesses of solutions, conclusions, and work processes.
5. *Active Listening*
   Giving full attention to what other people are saying, taking time to understand the points being made, and asking questions as appropriate.
6. *Self-management*
   Knowing how to manage oneself in the workplace so that it increases one's productivity.
7. *Innovative Orientation*
   Being able to introduce elements of creativity into the existing work process so that it improves existing productivity level of that work.

The HTCB leveraged these 7 competencies as baseline for developing the Employability Test.

Construct identification

Once the HTCB established which competencies to leverage in the development of the Employability Test, the next step was to calibrate each competency to suit the needs of the employers.

A key driver in psychometric test development is constructs identification, or in simpler terms, knowing what to measure. The HTCB wanted to incorporate construct validity at the source, and for this to happen, the HTCB identified several common constructs for each of the 7 competencies and shared it with employers. The employers were then asked to select those constructs that they considered most relevant.

Figure 3 showcases a list of constructs that the employers selected (for each of the 7 competencies). The definitions of the construct have been adapted from US Department of Labor's ONet initiative.

Figure 3: A list of constructs (for each of the 7 competencies) selected by employers. The defination are adapted from ONet.

1. English Proficiency	
Oral Comprehension	The ability to listen, understand information and ideas presented through spoken words and sentences.
Written Expression	The ability to communicate information and ideas in writing so others can understand.
Reading Comprehension	The ability to understand written sentences and paragraphs in work related documents.
2. Computer Literacy	
Troubleshooting	The ability to perform basic troubleshooting on computer systems (including hardware and software), enter data, and process information.
Word Processing Software	The ability to use collaborative editing software, Google Docs, Microsoft Word (or similar tools).
Presentation Software	The ability to use collaborative presentation software, Microsoft Power Point (or similar tools).
Email Software	The ability to use email software, IBM Lotus Notes, Gmail, Microsoft Outlook (or similar tools).
Internet Browsing Software	The ability to use internet search engines to surf the web and find information.
3. Workplace Ethics	
Compliance	The ability to use relevant information and personal judgment to determine whether events or processes comply with corporate policy, regulations, and standards.
Judging Qualities	The ability to assess the values, relevancy, and quality of things and people.
Self-Monitoring	The ability to monitor/assess performance of one's actions and behavior.
4. Active Listening	
Social Perceptiveness	The ability to be aware of others' reactions and understanding of why they react as they do.
Coordination	The ability to adjust one's actions in relation to others' actions.
Problem Sensitivity	The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing if there is a problem.

5. Critical Thinking	
Systems Evaluation	The ability to identify measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system.
Deductive Reasoning	The ability to apply general rules to specific problems to produce answers that make sense.
Inductive Reasoning	The ability to combine pieces of information to form general rules or conclusions (such as finding a relationship among seemingly unrelated events).
Information Ordering	The ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations).
Analyzing Data or Information	The ability to identify underlying principles, reasons, or facts of information by breaking down information or data into separate parts.
6. Innovative Orientation	
Fluency of Ideas	The ability to come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity).
Originality	The ability to come up with unusual or clever ideas on a given topic or situation, or to develop creative ways to solve a problem.
Service Orientation	The ability to actively find best solutions, or alternative solutions to help people.

7. Self-management	
Scheduling Work and Activities	The ability to schedule one's tasks, meetings, programs, and activities, as well as of other's.
Selective Attention	The ability to concentrate on a task over a period of time without being distracted.
Organizing, Planning, and Prioritizing Work	The ability to develop specific goals and plans to prioritize, organize, and accomplish one's work.
Multitasking	The ability to shift back and forth between two or more activities or sources of information.
Judgment and Decision Making	The ability to consider the relative costs and benefits of potential actions to choose the most appropriate option.

Construct calibration

Once the constructs (for each of the 7 competencies) were finalized in collaboration with the employers, the next step was to calibrate each construct. The construct calibration process is a very important step in customizing the test to the needs of the employer. The HTCB follows a rigorous calibration process, which in turn, further strengthens the constructs' validity.

The calibration process requires identifying a spectrum of possibilities or choices that scale up or down, similar to the *Likert scale*. The spectrum of possibilities includes real examples of what needs to be measured in the construct. Since the possibilities are listed on a spectrum (or range), it provides the employers an opportunity to indicate the minimum requirement expected from the new recruit. The calibration process goes into effect as the employers

select a possibility or choice on the spectrum that is aligned with their minimum recruitment requirement. With more employers participating in the survey, the calibration is further enhanced, and as a result the HTCB is able to get a better understanding of how to structure the questions in the psychometric Employability Test.

Since the HTCB's calibration process caters to the needs of the employers, it further reduces the validity gap, by focusing on specifics of what needs to be tested.

Figure 4 illustrates an example of a construct, and the spectrum of related possibilities and choices.

Figure 4: An example illustrating how employers participate in construct calibration to further strengthen construct valididty.

Competency English Proficiency	
Construct	*Reading comprehension*
Survey question for employers	What is the minimum level of reading and comprehension (in English) that you require from your employees having 0 to 2 years of experience?
Choice 1	At least read and understand internal day-to-day email communication
Choice 2	At least read and understand simple communication with customers
Choice 3	At least read and understand product operating instructions
Choice 4	At least read and understand trend reports and case studies
Choice 5	At least read and understand industry regulatory documents

Content validity

Once the constructs have been calibrated, the HTCB engaged its subject matter experts to create questions for each constructs. This leads to the creation of the Employability Test. The test creation process is not being included in this report to protect the integrity of the HTCB's test creation process.

The next step in content validity was to re-engage 10% of the employers (who participated in the construct calibration process) to review the Employability Test and provide their feedback on whether the content (and scenarios) used to create the questions were meeting their requirements.

Through following the previously described steps, the HTCB strengthened the final product. The employers needs were directly reflected in the final product by having them participated and contributed in construct identification, then in content calibration process and finally in content validation.

Post-testing results

Face validity

Since the launch of Employability Tests in Q4 of 2010, more than 300,000 candidates have been tested in over 22 countries. One of several key success factors for the HTCB was to ensure that its tests (which are in English) meet the face validation criteria.

Face validation refers to the extent to which a question (on the Employability test) is viewed or understood by the test taker. The HTCB included an option for the test takers to provide feedback while taking the test. Since testing is conducted online using a robust technology developed by UK-based Tested Talent, a collaborative effort was made to ensure face validity. Through this effort, Tested Talent included a feedback button on the live testing environment; this button was titled "I don't understand this question".

Tested Talent's reports indicate that for every 1,000 testing attempts, there are 16 clicks reported on the "I do not understand this question". This translates into 98.4% success in face validity.

Mean scores

The Tested Talent technology provides a wide range of capabilities for statistical analysis. The candidates were put into 3 categories, and a mean score was computed for each category. The 3 categories that were included for this study were:

- All the candidates tested (mean score = 59%)
- Candidates shortlisted (mean score = 72%)
- Candidates hired (mean score= 74%)

Figure 5 and **Figure 6** show a numerical and graphical comparison respectively, of the mean scores for each competency. This comparison was made possible through Tested Talent's technology.

As illustrated in **Figure 7**, the HTCB also noticed that there was a high variance of 1,500 basis points between the mean scores of all candidates tested and those who were hired. This high variance is indicative that the psychometric Employability Test can be used as a reliable pre-employment screening tool.

It is also important to highlight that in the *English proficiency* competency, the mean score of candidates shortlisted (74%) is higher than the mean score of the candidates hired (71%). When studying this HTCB variation in greater detail, the HTCB observed that since English is not the first language in the countries where the sample data was collected, such a variation is acceptable.

Figure 5: Numerical comparison of scores of the 3 categories for each competency. Comparison was made possible through Tested Talent's technology.

Competencies	All applicants (mean)	Shortlisted (mean)	Hired (mean)
English Proficiency	61%	74%	71%
Computer Literacy	54%	72%	74%
Workplace Ethics	69%	74%	74%
Critical Thinking	51%	69%	72%
Active Listening	62%	71%	75%
Self-management	64%	72%	76%
Innovative Orientation	55%	71%	76%
Overall Average	**59%**	**72%**	**74%**

Figure 6: Graphical comparison of scores of the 3 categories for each competency. Comparison was made possible through Tested Talent's technology.

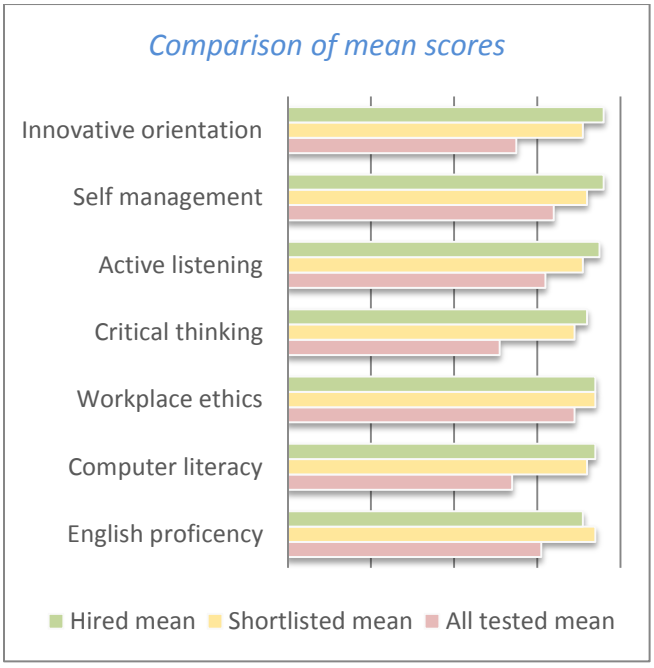

Figure 7: Score variance calculations of each competency by comparing the difference between the scores of all candidates tested against the scores of candidates hired. Variance calculations were made possible through Tested Talent's technology.

Competencies	*Score variation (all applicants versus hired)*
English Proficiency	1,000 basis points
Computer Literacy	2,000 basis points
Workplace Ethics	500 basis points
Critical Thinking	2,100 basis points
Active Listening	1,300 basis points
Self-management	1,200 basis points
Innovative Orientation	2,100 basis points
Average overall variation	**1,500 basis points**

Predictive Validity and Correlation

In addition to assessing reliability, the HTCB also wanted to assess predictive validity. This measures the usefulness of test scores to predict future performance.

Data on employee performance is not always readily available. This is because employers are usually reluctant to volunteer accurate data on their employees' performance.

Since it was unlikely that employers would be willing to share the performance level of their employees (who were tested and hired), the HTCB decided to identify an alternative indicator to 'performance'. The HTCB substituted 'performance' with *employment duration* (in months) as a predictive indicator.

In order to reduce errors when calculating predictive validity, HTCB established the following conditions:

- Engage employers to understand average employment duration (in months) for an out-of-school youth with 0 to 2 years of work experience (this was accomplished in the initial survey).

- Re-engage the employers 3 months after the average employment duration, and verify if the employee (who was tested and hired) is still employed.

- Since employment durations vary across sectors, the HTCB opted to observe the predictive validity for each sector (and not as a whole).

Figure 8 showcases predictive validation data that was observed

Figure 8: A comparison of employment duration of 2 types of employees: those who were hired without being tested, and those who were hired after being tested. The data shows the percentage (%) of candidates who were tested and hired, and still employed 3 months longer than the average employment duration. Comparison was made possible through Tested Talent's technology.

Sector	*Average employment duration (non-tested candidates)*	*% candidates (tested) still employed 3+ months longer than average*
Telecom	15 months	75%
Healthcare	15 months	70%
FMCG	12 months	85%
Food & Beverage	12 months	80%
Technology	12 months	85%
Media & Entertainment	9 months	85%
Financial	12months	80%
Overall average	12 months	80%

The HTCB also wanted to understand the reliability of its psychometric Employability Test as a 'predictor' of a candidate who are viewed as more employable, see **Figure 9**. A good predictor of reliability can be measured by comparing selected variables, and are represented through the the *r value*. The HTCB compared the correlation between attaining a high score in a competency and its effect on employment longevity. As a generally accepted rule of thumb:

- *r value* > 0.7 indicates a strong relationship
- *r value* > 0.4 indicates a moderate relationship
- *r value* < 0.4 indicates a questionable relationship

Figure 9: The correlation coefficient (or *r value*) comparing a competency score as a predictor of employment duration. Comparison was made possible through Tested Talent's technology.

Competencies	Correlation coefficient as a predictor (r value)
English Proficiency	0.42
Computer Literacy	0.49
Workplace Ethics	0.48
Critical Thinking	0.69
Active Listening	0.51
Self-management	0.67
Innovative Orientation	0.68

The HTCB analyzed the correlation between each competency tested and employment duration (or longevity). Interestingly, there was *moderate correlation* observed in the following competencies:

- Active Listening (r = 0.49)
- Workplace Ethics (r = 0.48)
- English Proficiency (r = 0.42)

This means that although employers consider Active, Listening and English Proficiency to be important competencies to screen during the recruitment process, however, these three competencies are moderate predictors of whether the recruit would remain employed with the company for a period that the longer than average (by 3+ months).

Another observation made by the HTCB was that there is *moderate-to-strong correlation* in the following competencies:

- Innovation Orientation (r = 0.68)
- Critical Thinking (r = 0.69)
- Self-management (r = 0.67)

This means that in addition to being relevant competencies for pre-employment screening, these three competencies were helping the employers in other ways as well. The Innovation Orientation, Critical Thinking and Self-management competencies were also moderate-to-strong predictors of employment duration (or longevity).

Conclusion

This report discusses the validity and reliability of calibrated psychometric Employability Test.

The report provides details on the methodology used to ensure that several forms of validity are built into the process. The report also analyzes the results of all applicants and compares them against the results of all hired candidates.

The data shows the psychometric Employability Test score variation between all candidates who applied versus those who were hired is 1,500 basis points. This indicates that the Employability Test is an effective pre-employment screening tool.

Lastly the report attempts to analyze if a high test score is a predictor of employment longevity. The outcome displayed a mixed result of competencies that were moderately correlated, as well as competencies that had a moderate-to-strong correlation as a predictor for employment longevity.

For more information regarding this study, please email st@hirelabs.sg