

# Technical Report

## Core Skills Entry-Level Assessment

---

February 1, 2017

Updated: June 12, 2017

# core score

a soft skills assessment

## Acknowledgments

Innovate+Educate adheres to a process focused on rigorous research methodology and statistical analysis in the development of their assessments. Our assessments meet the standards set forth by the U.S. Department of Labor's *Uniform Guidelines*, the American Educational Research Association (AERA), American Psychological Association (APA), the National Council on Measurement in Education (NCME)'s *Standards for Educational and Psychological Testing*, and the Society for Industrial Organizational Psychology's (SIOP) *Principles for the Validation and Use of Personnel Selection Procedures*.

Innovate+Educate wishes to thank the many partners that helped form our decision to bring an assessment portfolio to market that would impact economic advancement strategies for training, hiring, and advancement while remaining free to candidates seeking skill improvement. Our partners include the Innovate+Educate Board of Directors, Joyce Foundation, Workforce Services Greater Dallas, Workforce Connection of Central New Mexico, W. K. Kellogg Foundation, Walmart Foundation, White House National Economic Council, White House Office of Science Technology Policy (OSTP), and University Ventures.

Innovate+Educate wishes to thank Edan L. Jorgensen, Ph.D., and Eric T. Vincent, M.A., for their significant contributions to the development of the CSA-E assessment, as well as consultants and staff.

## Statement of Use

The Entry-Level Core Skills Assessment (CSA-E) can contribute to the full range of human resource areas, including recruiting, hiring, promotion, training, succession planning, mentoring, and performance evaluation. Individual employers ultimately determine the best means of using the assessment to achieve their business goals.

Under the *Uniform Guidelines on Employee Selection Procedures* (1978), employers are ultimately responsible for maintaining fair and equitable hiring practices that are appropriate, relevant, and justifiable. For employers who choose to use CSA-E as part of their selection process, Innovate+Educate recommends that it be used along with other criteria, such as education, academic credentials, prior work experience, reference checks, background checks, interviews, and internal or external recommendations. Using CSA-E as one of these multiple criteria augments the hiring process by adding new information on job candidates.

The use of the CSA-E for selection and/or succession planning purposes is only reliable if the results are current. Assessment scores for individuals can change over time due to the acquisition of knowledge, skills, and abilities obtained through training, education, and work experience. Therefore, the date the assessment was taken along with other relevant data (e.g., job performance, references, resume,

# core score

a soft skills assessment

traditional interview) should be considered for any high stakes decisions, such as selection and promotion of employees.

Another benefit of using CSA-E by employers is the identification of training and development opportunities for new and incumbent employees. Employers can use CSA-E to identify deficiencies in the competencies and address them with appropriate training interventions. This is especially helpful if there is a need to hire individuals who might not possess a skill set that is sufficient for their new jobs or expanded responsibilities.

Updates to this Technical Manual will be made as additional research becomes available.

## Table of Contents

Executive Summary .....	4
Assessment Themes and Competencies .....	5
Assessment Development: Background .....	9
Validity Considerations .....	10
Construct Validity .....	10
Content Validity .....	10
Criterion-Related Validity .....	11
A. Concurrent .....	11
B. Predictive .....	11
Validity Generalization .....	12
Understanding the Job – Job Analysis .....	12
Archival Research .....	12
Subject Matter Expert Groups .....	13
O*NET Analysis .....	14
Creating the Research Assessment: Item Development .....	16
Selecting a Sample .....	17
Performance Rating Scale (PRS) .....	17
Sample Size .....	18
Creating the CSA-E Assessment – Statistical Analysis .....	18
Sample Creation .....	19
Descriptive Data Analysis .....	19
Item Analysis .....	20
Regression Analysis .....	22
Total Score and Cutoff Score .....	22
Test-Retest Analysis .....	23
Conclusions .....	23

# core score

a soft skills assessment

Assessment Training and Administration .....	23
Ongoing Research .....	24
Individual Performance Metrics .....	24
Business Performance Metrics.....	24
Annual Assessment Review .....	24
References .....	25

## Executive Summary

The CSA-E was developed during 2016-2017 for a project with the retail services industries in a large metropolitan area of the United States. The project is focused on developing a competency-based hiring and training model for the region and participating employer partners. The CSA-E measures the five competencies (out of six competencies) from the National Retail Services Initiative (NRSI) Retail Competency Model that are applicable to entry-level jobs: 1) Communication; 2) Drive For Results; 3) Customer Service; 4) Adaptability; and 5) Critical Thinking (ACT Foundation, 2016). These competencies also align to the findings detailed in “Common Employability Skills: A Foundation for Success in the Workplace: The Skills All Employees Need, No Matter Where They Work,” produced by the National Network of Business and Industry Associations (March, 2015).

Initial research focused on analyzing relevant jobs from four sources:

- A review of the current literature;
- An analysis of 302 job descriptions from employers in the retail and service industries;
- Focus groups of job incumbents, managers, and executives in the retail/service sector industries; and
- An analysis of entry-level O\*NET occupations in the retail/customer service sector to identify the importance the knowledge, skills, and abilities (KSAs).

Based on information from the job analysis, a set of 135 items were developed related to the five competencies and aimed at discovering excellent from poor job performers. The job performance of 691 employees in the retail/service industry was rated by their supervisor or manager using a Performance Rating Scale to identify top performers and low or “contrast” performers. The 691 employees then took the set of 135 items during the Fall of 2016.

The researchers analyzed the item responses through an analysis of phi coefficients and concluded that the top and contrast groups truly differed in their response patterns to the 135 items. Further item analyses resulted in a final assessment version containing 125 items which validly measures the five competencies and distinguishes between top and low job performers.

# core score

a soft skills assessment

A regression analysis was conducted to examine the extent to which the total assessment score is related to the Performance Rating Scale reporting and objective performance metrics. The validity coefficient for the CSA-E is .41 at a .001 confidence level. Thus, the assessment reliably differentiates high versus low success in job performance.

The final stage of the analysis used a quantitative strategy to evaluate the validity and reliability of the CSA-E. Highly successful employees scored significantly higher than less successful employees on the 125-item total score and across all five of the competencies.

Innovate+Educate has begun research on the alignment of new and existing training curricula to the CSA-E that will assist individuals with improving assessment scores and expanding career opportunities. More research is planned with additional industry sectors and populations, and to demonstrate additional reliability and validity evidence.

## Purpose

The purpose of this technical report is to describe the usage guidelines and research behind the development of the CSA-E in 2016. The general research methodology for the assessment is provided below along with specific details relating to the development of the CSA-E. The Assessment measures a unique set of competencies through theme scores.

## Assessment Themes and Competencies

Over the last seven years, Innovate+Educate has documented the competencies that employers from across the United States have consistently articulated are critical to entry, mid, and/or advanced role job success. These competencies have been cross-referenced with industry, labor, census, demand, and other relevant data sets to develop the targeted list of critical competencies to be assessed.

Translating research and philosophy into the business world can be accomplished through the development of assessments that measure the successful demonstration of competencies. With this approach, Innovate+Educate defines competency in any job as a configuration of an individual's consistent patterns of thoughts, feelings, and behaviors that characterize a person's ability to perform a job at acceptable or higher levels. People who perform at a high level in a job possess a pattern of thoughts, feelings, and behaviors that separate them from those who perform poorly. This total configuration of competencies can be broken down into conceptually distinct "themes" or "categories" which highlight key components of overall competency.

Innovate+Educate decided to initially develop the CSA-E for the retail services industries focused on developing a competency-based hiring and training model with participating employer partners in a large metropolitan region within the United States. It was decided to focus on the five out of six competencies

# core score

a soft skills assessment

from the National Retail Services Initiative (NRSI) Retail Competency Model that were applicable to entry-level jobs (ACT Foundation, 2016):

## 1. **Communication:** Listening; Non-verbal; Reading; Signaling; Speaking; Writing

### **Communication Theme:**

The Communication theme centers on the sharing of information between individuals, examining how information is both given and received. The CSA-E measures several methods of communication.

Verbal skills, including an individual's abilities to understand and take-in spoken communication as well as give information to others using spoken words are assessed. Writing skills, or an individual's skills communicating in written form, are also part of this theme.

In addition, indicators of an individual's abilities to understand information through reading is measured. Less formal forms of communication are also measured. For example, questions help assess an individual's skills using and understanding signals (such as hand gestures, sounds, and color coding tools) as well as non-verbal ways people communicate (such as body language).

## 2. **Drive for Results:** Appropriate Appearance; Brand Awareness; Compliance; Consistency; Dependability; Health & Safety; Industry Standards & Practices; Initiative; Organizational Awareness; Policies & Procedures; Proactive; Task Completion; Business Ethics; Conscientious; Honesty; Respectful of Diversity; Cultural Intelligence; Integrity; Personal Development; Accountable for One's Actions

### **Drive for Results theme:**

The Drives for Results theme covers several aspects of an individual's ability and skills used to accomplish their work. Some elements of this theme focus on getting the work done in ways that follow a standard process. For example, the following of safety procedures as well as the adherence to the rules and procedures of a place or work are measured.

There is also an aim to understand how the individual treats other people such as their respect for diversity and overall consideration for others in the workplace. Personal integrity, or honesty and truthfulness in the workplace is included here as well.

The Drives for Results theme also works to understand how the individual manages their own personal development within the work being accomplished. Finally, this theme looks to gather the individual's understanding of how their talents and abilities can best meet and assist with the goals of the organization where they work.

**3. Customer Service:** Anticipate Needs; Appreciate; Assist; Closing; Customer Needs; Customer Resolution; Problem-Solving Questions; Service Recovery; Sales/Selling; Customize the Customer Experience; Identify Root Causes; Basic Product Knowledge

**Customer Service theme:**

The Customer Service theme assesses the entry role individual's skills and abilities to assist those who are buying or using their employer's products or services.

Items in this theme aim to understand several aspects of how the entry role position can impact the customer experience. Areas such as being able to understand and see how they can meet the needs of customers to find solutions to problems is explored. This theme also captures how the individual might assist customers, the way they demonstrate appreciation, and ways they get to know the customer to make their experience as positive as possible. In addition, the individual's skills and abilities with guiding the customer through the sales process and their product/service knowledge is gathered.

**4. Adaptability:** Accepting of Change; Continuous Learning; Cooperating; Flexibility; Persisting; Supporting; Teamwork; Valuing Differences

**Adaptability theme:**

The Adaptability theme aims to understand the entry role individual's ability to adjust and be successful in a variety of work environments.

Here, both people skills and ability to handle situational aspects are assessed. For example, this theme aims to gain insight into an individual's tendencies towards aspects such as team work, cooperation with others, empathy, and helpfulness to peers. The ability to learn what to do when situations change, tendencies towards learning, and how stress is tolerated are additional types of elements evaluated in this theme.

**5. Critical Thinking:** Cause & Consequence; Locating Information; Mathematics; Observation; Problem Solving; Prioritization; Task/Service Balance

**Critical Thinking theme:**

The Critical Thinking theme looks at several aspects of rational thinking and reasoning skills.

Here, the ability to process information, make judgement calls, and create actions based on information rather than emotion are evaluated. Areas touched on in this theme time management, attention to detail, and prioritizing work. Additionally, this theme gains to understand the individual's skills when information needs to be gathered, understood, and put into action.

# core score

a soft skills assessment

These five competencies are aligned to another source document, Common Employability Skills: A Foundation for Success in the Workplace: The Skills All Employees Need, No Matter Where They Work, National Network of Business and Industry Associations, (March, 2015). Five of the skills listed in this document include:

**COMMUNICATION:** Maintaining open lines of communication with others

- Demonstrate sensitivity and empathy
- Listen to and consider others' viewpoints
- Recognize and interpret the verbal and nonverbal behavior of others
- Speak clearly, in precise language and in a logical, organized and coherent manner

**INITIATIVE:** Demonstrating a willingness to work and seek out new work challenges

- Take initiative in seeking out new responsibilities and work challenges, increasing the variety and scope of one's job
- Pursue work with energy, drive and effort to accomplish tasks
- Establish and maintain personally challenging, but realistic work goals
- Strive to exceed standards and expectations

**CUSTOMER FOCUS:** Actively look for ways to identify market demands and meet customer or client needs

- Understand and anticipate customer needs
- Provide personalized service with prompt and efficient responses to meet the requirements, requests and concern of customers or clients
- Be pleasant, courteous and professional when dealing with internal and external customers or clients
- Evaluate customer or client satisfaction

**ADAPTABILITY:** Displaying the capability to adapt to new, different, or changing requirements

- Be open to learning and considering new ways of doing things
- Actively seek out and carefully consider the merits of new approaches to work
- Embrace new approaches when appropriate and discard approaches that are no longer working
- Effectively change plans, goals, actions, or priorities to deal with changing situations

**CRITICAL THINKING:** Using logical thought processes to analyze and draw conclusions

- Identify inconsistent or missing information
- Critically review, analyze, synthesize, compare, and interpret information
- Draw conclusions from relevant and/or missing information
- Test possible hypotheses to ensure the problem is correctly diagnosed and the best solution is found



# core score

a soft skills assessment

Using a psychometrics-based approach to the study of competencies, Innovate+Educate measured competencies in order to enhance an organization's ability to recruit, select, retain, and train/develop top-performers.

## Assessment Development: Background

Innovate+Educate intentionally chose an approach to assessment development that addressed market needs for the assessment to identify both the best candidate for selection/promotion and the competency gaps for training and development purposes with multiple sources of validity evidence. In order to accomplish this, Innovate+Educate combined a norm- and criterion-referenced approaches to assessment design using various strategies to establish validity: content, criterion-referenced, and validity generalizability.

The *norm-referenced* approach that was used is based upon four decades of work pioneered by researchers at Ohio State University. These researchers discovered how the use of open-ended questions and recorded interviews in a person-centered approach provides rich insights into the inner self of individuals. At a time when almost all psychological research was focused on the abnormal or troubled population, Dr. Hall recognized that the majority of people were "normal" (Hall, 1946, p. 177). Therefore, he devoted his time and energy to the study of the best and brightest to understand and measure competencies and/or talents that make individuals perform at, or exceed, job-specific expectations.

The best way to study success in any field or job family is to study the recurrent patterns of thoughts, feelings, and behaviors of outstanding performers in that field. Such recurring patterns for outstanding performers would contrast with less successful individuals.

The *criterion-referenced* approach for assessment development was based upon a design that is directly linked to established critical criteria. This approach is important for establishing the content validity of the items used on an assessment (Shrock & Coscarelli, 2007). Innovate+Educate used the NRSI competencies (listed above) to determine the framework of competency constructs to measure, as they were established as critical to retail services and adjacent industries. Using the job analysis methodology (described below) that included the use of population data available through the O\*NET, the set of important knowledge, skills, abilities, and behaviors (KSABs) of successful employees were determined that would be measured by the assessment.

Assessments are developed to measure the competencies an individual possesses. Assessments are characterized by a defined set of questions presented in a standard format for all respondents. Typically, answers to the questions are quantitatively coded, or scored, to remove as much subjectivity as possible. Because Innovate+Educate employs a structured assessment technique in both administration and scoring, these assessments provide an equal playing field for all interviewees, or job seekers, and ensure that selection decisions are based on comparable data.

# core score

a soft skills assessment

For several reasons, assessments with multiple sources of validity evidence are superior for scientifically understanding current and future employees. The ability of these assessments to predict job performance clearly exceeds that of standard, unstructured job interviews (McDaniel, Whetzel, Schmidt & Maurer, 1994; Wiesner & Cronshaw, 1988). These assessments also demonstrate greater predictive validity than many other employee selection methods (Schmidt & Hunter, 1998). According to a study performed by Hermelin and Robertson (2001), structured assessments and cognitive ability tests had the strongest operational validity of the selection procedures tested. A meta-analysis conducted by Huffcutt and Arthur (1994) suggests that highly structured assessments, similar to those developed by Innovate+Educate, have validity levels comparable to cognitive ability tests.

## Validity Considerations

Assessments measure competencies for a particular job, or job family. Such assessments may be used to evaluate people who have previously been in the job or people who have never been in the job. The concept of validity refers to the general question, “How well does an instrument measure what it is intended to measure?” In relation to Innovate+Educate assessments, validity refers specifically to the question, “How well does the Innovate+Educate assessment measure competency for the job in question?” The U.S. Department of Labor *Uniform Guidelines* describes three major aspects of validity for selection instruments: construct, content, and criterion-related validity. In addition, validity generalization is a strategy related to applying validity to local, individual settings (Schmitt & Hunter, 1977).

### Construct Validity

The APA, in the *Standards for Educational and Psychological Testing* (2014, p. 11), defines a construct as “the concept or characteristic that a test is designed to measure.” Concepts and characteristics, being abstract or theoretical, are observed and measured indirectly. In the development of an Innovate+Educate assessment, competency for a particular type of job is the construct of interest.

Testing construct validity for an instrument is a statistical process of comparing that instrument to other instruments which have been previously established as valid measures of the construct of interest. The construct validity of an instrument may also be demonstrated when hypotheses concerning the nature of the construct are confirmed (e.g., high scorers are more likely than low scorers to exhibit successful on-the-job performance).

### Content Validity

Content validity is defined by the *Standards for Educational and Psychological Testing* (2014, p. 14) as “evidence from an analysis of the relationship between the content of a test and the construct it is intended to measure.” Content validity is demonstrated when people familiar with the construct of interest determine that the instrument contains items that appear to be reasonable measures of that construct.

# core score

a soft skills assessment

Content validity is not a statistical process. It is a logical, intuitive, inductive, and deductive process for creating a series of questions that seem like good measures of the construct of interest.

Evidence of content validity can be demonstrated through archival research, focus group sessions and an analysis of the important aspects of performance on the job. These procedures are preliminary steps in creating assessment items that are logically related to a job. It is important that the set of research items is large enough to ensure coverage of the job demands and of the characteristics and behaviors that define the individuals who meet those demands with excellence.

## **Criterion-Related Validity**

Criterion-related validity refers to the degree to which an instrument estimates some measure, generally of performance, that is external to the instrument itself. This type of validity is tested statistically, generally by examining correlations between the instrument and the criterion. The importance of having an accurate and reliable measure of the criterion cannot be overemphasized as a step in ensuring a meaningful assessment of criterion-related validity.

Criterion-related validity may be assessed in a concurrent or a predictive fashion, depending upon when the “predictor” and the “criterion” are measured.

### **A. Concurrent**

Establishing concurrent criterion-related validity entails assessing the strength of the relationship between an instrument and some performance criterion when both measures are obtained at the same time (i.e., concurrently). Innovate+Educate used a concurrent approach to assessment development. Individuals whose on-the-job performance (the criterion) has been assessed are measured using the research items from the Innovate+Educate Assessment at around the same time, and the strength of the relationship between the instrument (the Innovate+Educate Interview) and the criterion (on-the-job performance) is examined.

One challenge of using a concurrent validity model for instrument development is that there is a difference between the research sample (in this case, people who have met at least minimal performance standards on the job) and the target population (in this case, candidates for the job). This challenge is far outweighed, however, by the model’s temporal advantage. In a relatively short period of time, it provides a job-related selection instrument that can subsequently be cross-validated on the target population and refined if necessary.

### **B. Predictive**

In a predictive situation, the predictor (in this case, the Innovate+Educate interview) is assessed before the criterion (on-the-job performance) is measured within the population of interest (in this case, candidates for a job). An ideal study of predictive criterion-related validity would involve hiring a number of people using the methods currently available to an organization, assessing those people on the

# core score

a soft skills assessment

Innovate+Educate research assessment prior to their starting the job, and then assessing their performance on the job.

## Validity Generalization

Schmidt and Hunter (1977) established that variations in the validity for jobs and tests were largely due to errors. In some cases, validity can be established, and can be generalized to new settings without the need to conduct a local validation study. When a local validation study is deemed necessary, however, the information derived from a validity generalization study can be used to make substantive improvements to the design of the study and the conclusions drawn (Schmidt & Hunter).

Innovate+Educate used an analysis of data from the O\*NET database (see below) to establish further evidence of validity of the CSA-E and to support the generalizability of findings to similar jobs.

## Understanding the Job – Job Analysis

When creating an assessment for a specific job or job family, a job analysis is performed to obtain a deep understanding of the task requirements associated with the job of interest. In addition, the worker characteristics necessary to excel in the position are analyzed. Innovate+Educate engaged in three methods to establish the requirements of the jobs that are targeted by the CSA-E: archival research, subject matter expert groups, and an analysis of O\*NET data for relevant jobs.

### Archival Research

Innovate+Educate analysts studied existing research to enhance their approach to discovering the key characteristics of outstanding performers in a job. The entry-level competencies established in the NRSI Competency Model were developed and validated in 2016 to represent the knowledge, skills, abilities, and behaviors needed in retail and adjacent services industries. The five entry-level competencies are: 1) Communication; 2) Drive For Results; 3) Customer Service; 4) Adaptability; and 5) Critical Thinking (ACT Foundation, 2016). These competencies also align to the findings detailed in “Common Employability Skills: A Foundation for Success in the Workplace: The Skills All Employees Need, No Matter Where They Work,” produced by the National Network of Business and Industry Associations (March, 2015). In addition, Innovate+Educate reviewed existing assessments and peer-reviewed journal articles related to the job family of interest. This step encompassed a review of existing job descriptions for the target position(s). Three hundred and two job descriptions were obtained from employers and analyzed to formulate hypotheses regarding the types of abilities necessary to perform at a high level in entry-level/customer service-related roles.

## Subject Matter Expert Groups

To ensure the content validity of the CSA-E assessment being developed was relevant to those within the industries of interest, focus groups were conducted to explore the attributes of top performers. Content validity refers to the degree to which all the important aspects of the competencies to be measured are being addressed by the assessment. In this case, it refers to the inclusion of the measurement of all the worker attributes necessary to be successful in entry level/customer service jobs.

Innovate+Educate conducted subject matter expert (SME) groups composed of job incumbents, managers, and executives who are currently employed by various employers in the retail/service industry in a large metropolitan area. The participants were identified as successful employees with at least one year tenure in their current position. Managers and executives who participated did not necessarily need to have been in the target position at any point in time but had a good understanding of the job requirements and outcomes. Additionally, Innovate+Educate strived to include a sufficient number of women, men, and individuals of various ethnic and racial backgrounds who are representative of the larger population of employees in the retail/service industry.

**Table 1. SME Group Participant Descriptive Table**

Gender		Ethnicity		
Male	Female	White	Black	Hispanic/Latino
15	9	22	0	2
6	2	4	0	4
8	11	13	1	5
29	22	39	1	11

The facilitator of the SME group sessions asked standardized questions which were useful in obtaining a general understanding of the individuals and their jobs, as well as questions specific to the employer of each group. In addition to studying the tasks and behaviors inherent in the performance of that particular job, the relevant thoughts, feelings, and attitudes held by the incumbents and their managers were explored by the facilitator. The discussions were recorded and transcribed for further analysis. Innovate+Educate research analysts strived to obtain a general description of all incumbents and an understanding of the specific characteristics of high achievers in retail/customer service positions. Sample SME group questions included:

- Tell me about the qualities of an outstanding (target job position)?
- Think about someone that didn't make it as an employee. Why do you think

# core score

a soft skills assessment

they failed?

- If you had to put a percentage on the technical versus the managerial skills necessary for a (target position), what would it be?
- What is your background? Do you have a job-related background?
- Describe your company's culture. How does it differ from other similar companies?

Discussing characteristics necessary for successful job performance with incumbents allowed the researchers to identify conceptual themes. Themes are subgroups of items within an assessment which represent a particular concept. A theme is designed to represent related patterns of thoughts, feelings and behaviors. For example, it was discovered through the job and occupation analysis that customer service is an essential part of the role. Understanding this, items were created that represent this concept in the CSA-E assessment. Conceptual themes were then finalized after statistical analysis was completed on items completed by the population who took the research version of the assessment.

If SME groups were conducted at a location other than the client company, then separate on-site visits were arranged to understand the unique culture of that company. Although identical job titles might be utilized across different companies, a company's culture and expectations, and the details of a particular job, can create a large variance across companies in the expectations of those in a particular job title.

## **O\*NET Analysis**

The objective of the O\*NET analysis was to identify the knowledge, skills, and abilities (KSAs) with the greatest importance to entry-level jobs across the retail/customer service and related sectors from the O\*NET data. The O\*NET is an extensive online source for occupational information. (see "key attributes and characteristics of workers and occupations" at <http://online.onetcenter.org/>). The four job levels (entry, advanced, manager, and leader) are defined by a list of job titles, such as "Assistant Department Manager" (manager level) or Sales Associate (entry level). These job titles were mapped to O\*NET occupation codes using a proprietary text classification tool. This mapping was then manually reviewed to correct misclassifications. The O\*NET contains importance scores for 120 unique KSAs (35 skills, 33 knowledge factors, and 52 abilities), which have been calculated for 954 of the 1,110 occupations in the database. Importance "indicates the degree of importance a particular descriptor is to the occupation." It is measured on a scale from 1 (Not Important) to 5 (Extremely Important); however, to simplify the analysis, Innovate+Educate normalized these scores to a scale from 0 (Not Important) to 100 (Extremely Important).

An additional aspect of Innovate+Educate's methodology involved identifying work activities and work styles from the O\*NET and having them rated by the participants in the SME groups. The SMEs completed a series of ratings regarding the importance and frequency of work activities and work styles

# core score

a soft skills assessment

required for successful job performance. Ratings added quantitative data to the qualitative information gained from the focus group interviews. Items for the CSA-E assessment were then created to measure the important requirements of the job from a more objective standpoint. This information also helped to identify the degree to which a group of jobs was related.

KSAs from O\*NET used in the development of the CSA-E assessment for entry level workers include:

## 1. Communications:

- *Active Listening* — giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times.
- *Speaking* — Talking to others to convey information effectively.
- *Oral Comprehension* — the ability to listen to and understand information and ideas presented through spoken words and sentences.
- *Oral Expression* — the ability to communicate information and ideas in speaking so others will understand.
- *Speech Clarity* — the ability to speak clearly so others can understand you.

## 2. Drive for Results:

- *Social Perceptiveness* — Being aware of others' reactions and understanding why they react as they do.
- *Monitoring* — Monitoring/Assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action

## 3. Customer Service:

- *Customer and Personal Service* — Knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction
- *Service Orientation* — actively looking for ways to help people.

## 4. Adaptation:

- *Deductive Reasoning* — the ability to apply general rules to specific problems to produce answers that make sense.
- *Flexibility of Closure* — The ability to identify or detect a known pattern (a figure, object, word, or sound) that is hidden in other distracting material.

## 5. Critical Thinking:

- *Critical Thinking* — Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems.

# core score

a soft skills assessment

- *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).
- *Deductive Reasoning* — the ability to apply general rules to specific problems to produce answers that make sense.
- *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 there is a problem.

## Creating the Research Assessment: Item Development

Based on information from the job analysis, an initial set of research items aimed at discovering the themes that distinguish excellent from poor performers were drafted for the research version of the CSA-E assessment. This version included four to five times the number of items desired for the final version. In addition to creating original items, previously researched items that have demonstrated utility in identifying top performers were incorporated into the research version of the assessment. The wording of pre-existing items was scrutinized to ensure applicability to the current population. Previously researched items with poor face validity for the current population were excluded. Face validity is a subjective judgment of whether an item appears to be applicable to performance in the job for which the assessment is being designed. To be certain there were enough items representing each hypothesized attribute, the items were classified by the attribute they are intended to measure.

Three types of items were used on the research version of the assessment. The first type are closed-ended, Likert-type items. These items were written to measure the degree to which participants agreed with a statement. Participants' level of agreement is indicated by their rating on a seven-point agreement scale ranging from *strongly agree* to *strongly disagree*.

The second type are closed-ended that are accompanied by a "conceptual match," which is defined prior to the administration of the research interview. The conceptual match defines the class of responses that is hypothesized to be characteristic of the most highly successful performers in a particular role. These conceptual matches are defined by examination of previous research as well as through analysis of the responses to questions discussed in the focus groups. Responses consistent with the conceptual match for an item are scored with a "1"; responses not consistent with the conceptual match for an item are scored with a "0."

Professional item-writers developed the third type of item from the blueprint that was created with the results of the job analysis described above. The analysis used population data available through the O\*NET to determine the set of important knowledge, skills, abilities, and behaviors (KSABs) of successful employees that would be measured by the assessment. Four-option multiple choice items were designed to reflect the content of real jobs and to distinguish performance that is poor or less than acceptable from performance that is associated with success. To do this, each item was mapped to a specific KSAB in the blueprint that is categorized under the competencies, and a justification and reference to source materials



# core score

a soft skills assessment

was provided for the content of each item as well as all the options presented in the item. In addition, the items were written to tap varying levels of cognitive complexity, ranging from basic recall of terms, procedures, and policies, up to the most complex level of problem-solving. To do this, items at the more cognitively complex level used real-work scenarios, such as dilemmas, and the kinds of documents and forms regularly encountered in the work place.

Two additional types of items were included in the research version of the assessment:

- Job behaviors and conditions of the test taker's current employer;
- Demographic items, which enable a description of the sample, tracking of the assessment process, and a means to test adverse impact are also asked. Items included gender, ethnicity, and age.

In total, there were 135 items in the research version of the assessment.

## Selecting a Sample

As described above, the research methodology was based on the premise that top performers possess distinctly different competency levels than unsuccessful performers. Studying the differences in these competencies allows for the selection of top performers over unsuccessful or "contrast" performers. When creating an assessment, one of the key components is categorizing participants by performance. Comparing the responses between the groups of performers on research items leads to the ability to select top over contrast performers with the assessment. Therefore, these groups (by performance) must be established before developing the final assessment.

Innovate+Educate worked closely with client-partners to identify their top performers (people who exemplified the ideal in that role) and contrast performers (people who were unsuccessful in that role). The importance of selecting the right people into the sample and properly categorizing participants as top or contrast performers was a key step in the research design. Innovate+Educate worked with the employers using performance data, such as direct supervisor ratings on the Performance Rating Scale (PRS), to identify the research sample.

## Performance Rating Scale (PRS)

The Performance Rating Scale (PRS) is an evaluation of on-the-job performance from the perspective of an employee's direct manager or supervisor. It includes questions to which a manager responds on a scale of 1 through 10 (with 10 being high) and from "strongly disagree" to "strongly agree."

The employee is also rated based on productivity and job performance compared to other employees who perform the same job. Such ratings provided research analysts with an assessment of a manager's level of confidence in their ability to evaluate that employee's performance and whether the manager would hire more people like that employee. Adhering to strict criteria for sample selection helped make certain that the two samples truly represented the identified groups (top performers versus contrast performers).

In order for top performers to meet the criteria for inclusion in the study, they had to be identified by their manager as having an average rating greater than four on a one-to-five scale, receive an overall performance rating of at least eight on a one-to-ten scale, and be identified by their manager as someone who would be hired again. Contrast performers had to meet the criteria of an average rating below three on a five-point scale, receive a four or lower rating of overall performance on a 10-point scale from their manager, and be described as someone who would *not* be hired again. Additional information from the managers, such as tenure and additional comments were also considered in selecting the sample. Performance data was collected at the same time as the PRS forms and were statistically compared to the ratings of managers to determine accuracy.

Table 2. Performance by Top Verses Contrast

	Frequency	Percent
Contrast Performer	107	16%
Top Performer	236	50%
Total	343	66%
Non-Top/Contrast (Excluded)	354	

## Sample Size

A minimum of 450 people with corresponding performance metrics was required for the research phase of the assessment development. Within the sample, the ability to evaluate performance and slice the PRS data into top and contrast groups must be distinct so as not to blur the line between excellence and mediocrity in performance. Due to the strict guidelines for sample selection, employers nominated several more candidates than actually completed the research version of the assessment.

The researchers communicated to the various employers which employees were selected for inclusion in the study. The researchers also assisted the employer in communicating the next step in the process to the selected sample, including administering the research assessment within a designated time frame.

## Creating the CSA-E Assessment – Statistical Analysis

Scores were recorded and tracked across multiple employers. Once completed, assessment results were merged into one database, proofed for accuracy, and analyzed to examine the relationships among items and their utility for predicting on-the-job performance. This step resulted in the most refined understanding

of the characteristics which define outstanding performers and which separate them from less successful performers. Based on the results of this step, items were selected for inclusion in the final assessment.

## Sample Creation

The first phase of this process involved creating samples from the entire test population. One group represented one-third of the sample (called the holdout sample), and another group included the other two-thirds of the group (called the primary sample). Each of the two groups had fairly equal numbers of top and contrast performers. For this reason, the top group was randomly split into two groups (one-third and two-thirds) and the same was completed for the contrast group. The top holdout sample was then merged with the contrast holdout sample; the identical process was followed for the primary sample. This allowed for similar distribution of top and contrast participants and distribution of demographic diversity. The primary sample was used to determine which research items discriminated between top performers and contrast performers.

## Descriptive Data Analysis

Frequencies for demographic data were calculated to define the composition of the sample. Means, standard deviations, and correlations (uncorrected) are calculated as well for scores as a first step in the data analysis.

Table 3. Frequency of Research Sample by Gender

	Female	Male	N/A	Total
Employer A	46	22	4	72
Employer B	97	101	10	208
Employer C	16	7	2	25
Employer D	85	154	26	265
Employer E	6	5	0	11
Employer F	26	19	1	46
Total	276	308	43	627

# core score

a soft skills assessment

## Item Analysis

The researchers analyzed the item responses through an analysis of phi coefficients. Phi coefficients assess the degree to which items discriminate top and contrast group members at a level greater than what may be expected to have occurred by chance. If the likelihood of getting the obtained differences by chance alone is small, it is concluded that the top and contrast groups truly differ in their responses to the items. The results of these analyses are examined for differences in responses between top and contrast groups and statistical significance. In order to be selected for inclusion in the final version of the assessment, items were required to meet two criteria:

1. Top and contrast group responses must be statistically different from one another
2. Items hypothesized to be in the same theme must be correlated with one another (based on statistical analysis of inter-item correlations, item-total correlations, and Cronbach's alpha reliability values) so that the final assessment items create statistically significant differences between top performers and contrast performers on both the final assessment total score and on the theme scores.

## Themes

As described above, competencies represent the necessary characteristics for success in the job. The theoretical understanding of these concepts begins during the job analysis. The final composition of each competency, or theme, was determined through a series of statistical processes. Seeking empirical validation for hypothesized themes, researchers statistically evaluated each item. Each theme represents a concept that is more specific than the entire assessment. Themes are created through factor analysis and internal-consistency analysis (Cronbach's alpha). Items within a particular theme should group together on these analyses indicating that they measure the same concept, or competency. Factor analysis identifies the degree to which the themes are independent of one another. If they are too closely related, the themes may represent the same competency. Correlations between each competency to all other competencies and the total assessment were analyzed. In order for themes to be considered independent of one another, each theme had to exhibit a greater correlation with the entire assessment than with any other individual competency. Specifically, competency-to-competency correlations did not exceed a .40 level.

Themes are also examined for face validity, meaning whether the items within a theme seemed related conceptually and representative of the established concept or competency. Face validity is an important consideration during development, but it is examined again during the final assessment creation. It is desirable that themes included in the assessment each contain a substantive number of items to ensure adequate measurement and interpretation of results. While themes are designed for developmental purposes and to help clients address fit issues within their organizational culture, they were not, however, individually validated as predictors of performance.

The themes for the CSA-E are represented graphically for presentation to the examinee and/or hiring manager. Two lines are displayed across each theme on the *Competency Intensity Index* for comparison purposes. One line represents the average score on each theme for top participants, and the other line represents the average of the contrast participants. This allows for the comparison of candidates against the top and contrast groups. Once an individual completes the CSA-E, a line representing their score for each theme is added. It should be understood that candidates who do not meet the cutoff score are likely to be deficient in competencies needed to be successful on the job.

Table 4. Top and Contrast Mean Tests for the 5 Entry-level Assessment Competencies

		N	Mean	St. Dev.	t-value	p-value	d-value
Communication	Top	207	107.14	14.74	9.30	.00	.78
	Contrast	96	97.84	8.13			
Drive for Results	Top	186	78.62	20.72	4.86	.03	.32
	Contrast	88	73.76	6.28			
Customer Service	Top	137	116.14	23.41	22.17	.00	1.31
	Contrast	70	93.97	4.06			
Adaptation	Top	184	103.03	30.55	17.68	.00	.80
	Contrast	88	85.35	6.71			
Critical Thinking	Top	202	81.39	21.32	5.02	.02	.31
	Contrast	94	76.37	6.17			

Table 5. Internal Consistency: Alpha Coefficients for the Five Assessment Competencies

Competencies	Alpha
Communication	.71
Drive for Results	.83
Customer Service	.83

Adaptation	.79
Critical Thinking	.80

### Regression Analysis

Regression analysis was conducted to examine the extent to which the total assessment score is related to Performance Rating Scale (PRS) ratings and objective performance metrics. Assessment score should be a significant predictor, in the correct direction, of PRS ratings and provide objective performance data. The validity coefficient for the Innovate+Educate CSA-E assessment is .41 ( $p \leq .001$ ).

### Total Score and Cutoff Score

The CSA-E assessment is designed to assist clients in predicting performance for a particular job, or job family. These predictions are based upon overall assessment scores. The assessment also has a number of themes, and these themes also have scores. It is important, however, to keep in mind that the CSA-E was designed to use performance on these competencies together, in combination, to evaluate job candidates for selection purposes. They are not separately validated as individual predictors of performance and should not be the sole justification for a selection decision. Individual scores on competencies, however, can be used to identify gaps in competencies that can be remedied by a training and development plan.

Table 6. Top and Contrast Mean Tests for the Total Entry-Level Assessment Score

		N	Mean	St. Dev.	t-value	p-value	d-value	Alpha
Entry-Level	Top	137	515.01	103.06	88.17	.00	1.39	.81
	Contrast	70	426.84	15.67				

The cutoff score for passing the assessment is established by determining the total score that results in the best reclassification of the top and contrast participants as either passing or not passing. A reclassification rate of 85% (or a sum score of 524 [out of 615]) percent or greater is acceptable.

# core score

a soft skills assessment

In cases where a candidate's overall score meets a cutoff for recommendation but additional information from individual themes or items may cause some concern, the interpretation must be carefully conveyed to employers. It must be remembered that it is the overall assessment score that is validated for selection purposes. It would be appropriate to inform the employer that the candidate's overall score on the assessment indicates they have sufficient competency for the job but that additional information should be considered.

## Test-Retest Analysis

Test-retest reliability is an important step to ensure that assessment results remain consistent across administrations and time. Retest assessment participants are selected to represent an equal number of top and contrast individuals. Research is currently underway to retest a sample of the original population which took the research version of the assessment. It is expected that the two test results will be highly related to indicate that the assessment is reliable over time. Additionally, the data obtained from these retest interviews are often used during analyst training (see "Assessment Training and Administration" section below).

## Conclusions

The purpose of developing the CSA-E Entry-level Assessment was to create a tool that would have utility for predicting highly successful performers in entry-level, customer service oriented jobs. The 135-item research assessment was administered to over 691 job incumbents that had their performance evaluated by their manager. All research items and their follow-up items were based upon the key traits and behaviors of highly successful employees. Based on statistical analyses, the set of 135 items was refined and adjusted to include a final set of 125 items reflecting 5 competencies, or themes, which reliably differentiated highly successful versus less successful performance. The final stage of the analysis quantitatively evaluated the validity and reliability of the CSA-E entry-level assessment. Highly successful employees scored significantly higher than less successful employees on the 125-item total score and across all five competencies.

## Assessment Training and Administration

Training for assessment administration and interpretation contributes to the reliability of the scoring of selection interviews (Campion, Palmer, & Campion, 1997; Conway, Jako, & Goodman, 1995). Innovate+Educate requires such training for all employers and other entities who want to use the CSA-E assessment. All scoring is done by Innovate+Educate.

# core score

a soft skills assessment

## Ongoing Research

Research on the predictive validity of the CSA-E assessment is still an important step in the ongoing validation process. Because it is used as a tool for selection decisions, it is important to cross-validate the instrument for the target population (job candidates) by tracking the relationships between assessment scores and both individual and business performance metrics.

Innovate+Educate is available to help employers see a return on their investment. We encourage them to collect data on both individual performance and business performance indicators

### Individual Performance Metrics

In ongoing research on the validity of the CSA-E assessment, we would expect to see positive correlations between assessment results and performance in training and on the job. However, it is important to note that the use of the Innovate+Educate assessment as a part of the selection process will create a *restricted range* on assessment performance (that is, most people selected for the job in question will have performed with excellence on the assessment), and this restricted range on assessment results makes statistical assessments of the relationship between assessment performance and on-the-job performance more difficult.

### Business Performance Metrics

As the overall level of competencies increases in an organization, business performance metrics should trend positively and labor turnover and related costs should, in the long term, decrease. To conduct a complete examination of the relationship between assessment results and turnover, it is necessary to look not only at turnover rates but also at reasons for terminations, as indicated by a structured exit assessment and/or interview.

With a combination of both individual and business performance metrics, Innovate+Educate can assist employers with a return-on-investment study.

## Annual Assessment Review

In addition to continuous study of metrics provided by client partners, Innovate+Educate takes a formal approach to the review of assessment. The CSA-E assessment will be evaluated on a yearly basis to assure that it continues to be used as a valid and reliable instrument for employers and other entities working to assist individuals in preparing for the workforce.



## References

ACT Foundation (2016). *National retail services initiative: Competency model*. Retrieved April 20, 2017, from <http://actfdn.org/wp-content/uploads/2016/06/NRSI-Competency-Model-2016.pdf>

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

Campion, M. A., Palmer, D. K., and Campion, J. E. (1997). A review of structure in the employment interview. *Personnel Psychology, 50*, 655-702.

Conway J. M., Jako, R. A., & Goodman, D. F. (1995) A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology, 80*, 565-79.

Equal Employment Opportunity Commission (1985). *Uniform employee selection guidelines: Interpretation and clarification*. Washington, DC: Commerce Clearing House Inc.

Hall, William. E., (1946). Accentuate the positive. *Educational Leadership, Dec*, 175-179.

Hermelin, E., & Robertson, I. T. (2001). A critique and standardization of meta-analytic validity coefficients in personnel selection. *Journal of Occupational and Organizational Psychology, 74*, 253-277.

Huffcutt, A. I., & Arthur, W., Jr. (1994). Hunter and Hunter (1984) revisited: Assessment validity for entry-level jobs. *Journal of Applied Psychology, 79*, 184-190.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79*, 599-616.

Murphy, K. (2003). *Validity generalization: A critical review*. Lawrence Erlbaum, Mahwah, NJ.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62*, 529-540.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

Shrock, S. A., & Coscarelli, W. C. (2007). *Criterion-referenced test development: Technical and legal guidelines for corporate training and certification* (3rd ed.). Silver Spring, MD: International Society for Performance Improvement.

# core score

a soft skills assessment

Section 60-3, Uniform Guidelines on Employee Selection Procedure (1978); 43 FR 38295 (August 25, 1978).

Society for Industrial Organizational Psychology (2003). *Principles for the Validation and Use of Personnel Selection Procedures* (4<sup>th</sup> ed.).

Wiesner, W. H., & Cronshaw, S. F. (1988). *A meta-analytic investigation of the impact of Assessment format and degree of structure on the validity of the employment.*