Codility_

Guide to Validation

Frequently asked questions about Codility's approach to validation.

Thank you for your interest in
Codility's approach to validation:
a complicated yet critical topic.

Neil Morelli, Ph.D.
Chief Industrial-Organizational
Psychologist

Codility_

Simply put, validation is the process of
verifying that an assessment means what it is
intended to mean. That makes validation vital
to how Codility provides value to your hiring
process; by helping you accurately identify the
right candidates for your roles.

The process of validation is collaborative.
Working together, we can verify that
assessments are accurate and fair, and
ultimately impactful towards your recruiting
goals.

With that said, my job is to oversee validation
at Codility and I look forward to partnering
with you to improve how you hire. When
you're ready to discuss more, let us know and
we'll be happy to help.

Contents:

Validation Overview

What is validation?

Validation is the process of gathering evidence that test results mean what they are intended to mean (Cizek, 2020).

For hiring, the intended meaning of scores is that they are job-related (Principles, 2018). For example, Codility is "valid" for recruiting if there is evidence that supports the following score interpretation: a high Codility score confirms that a candidate has enough skill in a required programming language that they are qualified for a technical job.

Because there is no 'one size fits all' approach to gathering validity evidence, several strategies exist for backing up the job-relatedness interpretation of tests. But according to the Uniform Guidelines on Employee Selection Procedures (1978), content and criterion validation are common.

Content validation demonstrates job-relatedness by rationally connecting a test's content (i.e., a test's instructions, questions, and scoring methods) to a job's content (i.e., the job's behaviors, activities, duties, and skills). Job experts typically review test content and rationally map or link it to job activities/duties. Criterion validation demonstrates job-relatedness by gathering performance data (e.g., supervisor ratings, production metrics, turnover) and correlating them with test scores to show a data-driven connection between test scores and the job.

How do we know our tests are valid?

Because validity is a property of the evidence about a test score's interpretation, and not a property of the test itself, validation is an ongoing process that needs to be updated as tests, jobs, and test purposes change.

However, "validity is the most important consideration in developing and evaluating selection procedures" (Principles, 2018, p. 4).

Therefore, Codility can provide historical validity evidence from several sources:

- How Codility tasks are designed

- Past studies verifying Codility's task content is job-related

- Process for conducting local validation studies Codility facilitates through professional services

Each of these evidence sources is explained below:

First, Codility demonstrates **validity evidence** during the task creation process. From the very beginning of the test-task development process, a specialized Codility Content Delivery Team identifies programming tasks that developers must perform on the job.

In other words, job experts who are knowledgeable about the work activities and duties for major engineering job families create test tasks to directly sample these critical work behaviors (e.g., writing algorithms, debugging, optimizing code).

Next, Codility has completed **several content validity studies** to document the link between Codility's task content and critical software engineering work behaviors. These studies were conducted by Industrial-Organizational Psychologists in collaboration with customers in the finance and healthcare industries.

A summary of these validity studies, as well as reliability and test fairness evidence, is cataloged in a comprehensive technical manual. Please ask your account executive or customer support manager for more information and access.

Finally, Codility has an **internal I-O Psychologist and customer support team** who can help manage local validation studies to collect and document validity evidence at your organization.

Localized studies, whether content or criterion-related, can provide strong defensibility that Codility assessments are valid (i.e., job-related and consistent with business necessity). Information about the scope and process of such a study (or studies) can be provided upon request.

Why is validation important?
Has Codility completed any validation studies in the past?

Validation is important because Codility scores are used to make important individual decisions about whom to hire.

An organization's success, and how its hiring fairness and legality are evaluated, are often the result of adding up these individual decisions.

Validation makes sure that the information provided by the test is the highest quality possible for making the best decisions.

In particular, validation provides these benefits:

Defensibility: Validation justifies and can legally defend the hiring decisions you make using test scores.

Fairness and Candidate Experience: Validation ensures the test is job-relevant and fair for the candidate, protecting their dignity and rights in the process.

Utility: Validation provides data that can estimate the financial impact of hiring decisions made using test scores.

Legal Compliance

Have your other clients ever had their use of Codility legally challenged? If so, how has this been addressed, or how would it be addressed?

We're not aware of Codility ever being legally challenged, but several clients have worked with us in the past to gather validity evidence as prevention against legal challenges.

Codility developing test content that is job-related adheres to the standards enforced by the two main U.S. regulatory agencies for employment discrimination, the EEOC and OFCCP, and the standards codified in the Uniform Guidelines on Employee Selection Procedures.

However, the most legal defensibility is provided by gathering local evidence that Codility's use is job-related in a specific work context. As such, the best legal defensibility assurance is to conduct a local validation study that's suited to your needs regarding specific tests being used, their score interpretations, specific jobs families, intended employee and candidate audiences, etc.

Are you UGESP compliant? What about the EEOC or OFCCP?

Our content and its development generally adhere to the standards codified in the Uniform Guidelines on Selection Procedures. In addition, we have several anonymized validity studies that can be referenced, in addition to our more comprehensive technical manual.

These resources provide the information required by the UGESP and are commonly requested by regulatory agencies such as the EEOC and OFCCP in the US.

However, complying with the UGESP is an ongoing process that often requires gathering local validity evidence.

C_

How do you validate your assessments? What's involved in a "local validation study"? Are there any other options?

In most cases, we take a content validation strategy, where test content (e.g., instructions, examples, questions/prompts, and scoring methodology) is rationally and explicitly linked or mapped to job content (e.g., work activities and behaviors) by job experts.

Working collaboratively with our customers, we follow these general steps to capture and document this information. More information and details can be provided as part of our professional services offering.

Job analysis and job expert validation
Job documentation, O*NET records, and an initial group of job experts were reviewed to create a preliminary job duty list.

Job duty validation
This preliminary job duty list was validated by a separate, secondary group of job experts.

Work sample test item validation
Job experts provided ratings of the tasks and linkages to validated job duties.

Time limit and cutoff score reviews
(as necessary)
Job experts were asked if the allotted time for each task was enough for a minimally qualified candidate to complete. Job experts were also asked the number of points a minimally qualified candidate would receive out of a 0 to 100 score. Job experts were asked to complete the tasks and their average scores were reviewed.

We are also currently gathering data for a consortium criterion-related validation study. This study will correlate Codility scores with performance metrics for entry-level software engineering roles.

Once collected and analyzed, this data will be available for lightweight 'transportability studies' to generalize validity evidence from one job to other similar jobs.

Adverse Impact & Avoiding Bias

What evidence can you provide that your assessments do not have an adverse impact?

By default, Codility does not track personally identifiable information (PII) from applicants and is therefore unable to complete an adverse impact analysis.

However, a voluntary (optional) demographic questionnaire can be added to the candidate workflow in Codility so that we can monitor for fairness and adverse impact on an ongoing basis. With this data

available, Codility testing experts can perform an adverse impact analysis to determine the presence and severity of adverse impact, if there is any, based on age, ethnicity/race, and gender. The adverse impact analysis will be conducted in a manner to appropriately address Section 4D of the Uniform Guidelines.

How do Codility assessments avoid bias (e.g. gender-biased language, programming language bias)?

Codility mitigates bias from every client's hiring process by administering and scoring a sample of each candidate's coding capability exactly the same way, every time. In other words, standardized testing and scoring offered by Codility eliminate random errors and unconscious biases that crop up when using traditional, unstructured tests of candidate skills (e.g. human raters). This ensures that comparisons between candidates are reliable, fair, and based on true coding skills.

In addition, we take steps during and after the test creation process to identify and mitigate potential cultural and gender bias or adverse impact.

Here are a few ways we do this:

Third-Party Task Reviews
Codility utilizes a third-party professional copywriting/proofreading service when crafting and editing test statements and descriptions to mitigate biased or sensitive language.

Real-Life Tasks
Codility's Task Creation process is specifically designed to answer the question "How close is a task to 'real-life'." Codility tasks are based on real-life scenarios that resemble work that is common in everyday development.
Each test scenario is separately reviewed by professional software engineers who play the role of an applicant solving the target task for realism and fidelity to required skills.

Codility_

Where the best engineering teams are built.