QUALITY ASSURANCE STANDARDS

Framework and Outcomes Metrics

2018

APPROACH

With the growth of unaccredited educational providers, students are facing a barrage of unsubstantiated claims regarding the outcomes these different providers produce. Knowing where to invest their time and money in learning is becoming increasingly difficult. Similarly, employers and all levels of government face challenges determining from where to hire and which providers to allow to operate, respectively.

The U.S. Department of Education's launch of the Educational Quality through Innovative Partnerships (EQUIP) program also brought the federal government into the fray, as EQUIP allows nontraditional providers of education to partner with accredited colleges and universities in order to access federal aid for students.

These developments necessitate the emergence of a third-party quality assurance system for the health of the industry and its student, employer, and government stakeholders. The EQUIP program has jump-started the creation of such a system by introducing quality assurance entities (QAEs), which will monitor student outcomes at experimental sites to determine eligibility for federal loans and grants.

In order for a third-party quality assurance system to be successful, there must be a framework of quality assurance standards in place to allow programs to report their student outcomes in a transparent and trustworthy manner such that stakeholders can compare the outcomes from like programs.

This document outlines a framework for a new, outcomes-based system of quality assurance. Although no system can perfectly capture every dimension of an education or training program's quality, these standards allow institutions to report outcomes through the lens of the customers—the learners who pay to attend and the employers who hire those individuals afterward. To that end, we examine a range of near- and medium-term outcomes related to **learning, completion, placement, earnings, and satisfaction** that each program could claim to provide.

INTENDED USE OF THE STANDARDS

Institutions and auditors will use these standards to prepare and report program outcomes, provide transparent and trustworthy reports, and offer a useful comparison for similar programs. Institutions that accept and participate in these standards will be expected to use verified metrics in their promotional content and messaging.

It is important to note that these standards are intended only to establish what outcomes are measured and how. Compliance with these standards must be monitored and measured by independent auditors/QAEs, but neither these standards nor the review process are intended to make a judgement of quality (i.e., whether an outcome is good or bad). Rather, students, employers, funders/investors, accreditors, and other third-parties will make judgements for themselves based on accurate and comparable data.

RATIONALE FOR SELECTED OUTCOMES

Although institutions can choose which outcomes they report in line with the value they provide students and their programs' missions, a functional quality assurance system for postsecondary education should ultimately rely on multiple measures of results. For one, each additional measure acts as a check on possible ways to manipulate others. For example, a provider could offer short-term certificates that have a high completion rate but little labor market payoff. Looking only at completion would thus ignore a serious problem. Alternatively, a program could have very good earnings results but very few completers. In that case, the high rate of non-completion is a sign of concern that would be ignored in an earnings-only evaluation of results.

Conducting rigorous reviews of multiple measures provides benefits for education and training providers, as well as consumers. For educational providers, the process will generate a series of verified, accurate outcomes that could then be used in any marketing or promotional materials, or for benchmarking and continuous improvement efforts. In addition, because there are multiple options to report on student outcomes, providers can select the right measures for each of their particular programs. Similarly, students would have greater confidence that the numbers that are presented to sell them on a program are true and potentially comparable across the market, as providers would be held accountable for advertising outcomes inconsistent with their audited numbers.

These standards intentionally do not address two other key elements of an education or training offering—the financial stability and governance of the organization. Insolvent and mismanaged providers present a different set of quality concerns. Sudden closure could undermine a credential's value and create difficulties in verifying student attendance/completion. Any complete quality assurance system thus needs to combine investigation of a provider's learner results with a sharp eye to its financial health and governance. Given the different work and skills needed to assess finances/governance and student outcomes, it is entirely possible that these verification processes could be conducted by two separate entities. Having separate reviews makes sense as well because these outcomes standards operate at the programmatic level, while financial and governance questions look at the overall education institution or training provider.

APPLICABLE ENTITIES

These standards are intended to be applied broadly to higher education providers, including accredited and unaccredited institutions. As they are meant to measure and validate only the outcomes that a program claims, these standards are intentionally agnostic to institutional mission, structure, and other elements that might vary in a review of traditional institutions of higher learning. These standards accordingly do not restrict participation to education or training providers based on any type of qualifications related to facilities, faculty, or other structural elements. The goal of an agnostic approach to a provider's model or structure is for these standards to provide assurance about the value of enterprises with regards to their student outcomes.

Although these standards are agnostic to a provider's structure, they cannot operate at any level of learning more granular than a program. Individual courses of course can have value, but it is unrealistic to expect these standards to assess the value of each individual piece. These standards are accordingly designed to provide a quality assurance process at the program level. We chose program for several reasons. First, there is substantial evidence that outcomes within different program offerings at an institution of higher education may vary as much as they do across schools. Second, consideration of learning outcomes operates better at a programmatic level, where expected results are more specific and measurable. By contrast, institution-wide learning outcomes for a place with diverse offerings would have to be quite vague in order to capture learning that all students should accumulate. Third, some institutions may offer seemingly similar programs that are designed for different populations. For example, someone may offer a full-time, intensive web design program alongside a part-time one that targets a different population. It is likely that these similar programs would produce different results, if for no other reason than the timeframe for finding employment might be different.

We recognize that any programmatic approach must still have some room for flexibility. For example, there may be instances where a provider operates different programs but they are closely related. In these cases, it is better to combine programs to reach sufficient sizes for evaluation than to simply not measure results. Ideally, program combination should strive to avoid lumping together programs of vastly different levels. For example, a program that is three months in length should not be combined with one that is 12 months.

These standards are agnostic to the level of programs considered. They check the outcomes one would want to see for a short-term certificate program the same as they would for a degree program. Despite this, we assume the initial application of these standards will likely be at career-oriented training programs rather than the traditional offerings that bedevil any discussion of outcome (e.g., liberal arts degrees).

SUPPORTING CLAIMS

Perfecting the indicators discussed in these standards takes a lot of work. For many, the underlying quality and accuracy will also be a function of what types of data a provider can access. Recognizing this, we refer programs and auditors to a supplementary document outlining the evidence that programs should use to develop their assertions. In addition to any evidence used in developing these claims, the document also clarifies that there should be some direct confirmation with students. In turn, third-party auditors will conduct full, annual reviews of programs in which they work to understand how the programs made their assertions and will, in turn, follow the guidance that the client followed to provide assurance around the claims.

Providers may wish to make claims about outcomes for subpopulations of the required standards. This is acceptable so long as they still report the complete set of outcomes and state clearly which subsegments are included in their marketing claims.

REPORTING LEVEL

Outcomes must be reported individually for each program and level within it. Online-only programs can be combined across campuses, but in-person offerings should be reported separately. The only exception would be multiple locations for a program in the same metropolitan area, which can be combined. Providers may additionally elect to aggregate programs across multiple regions, but location-specific data must be reported.

Aggregation
We recommend that programs that produce fewer than 30 learners in the denominator of a completion rate should adhere to the following guidelines for aggregation. These do not propose combining results for in-person programs across any areas except those within the same metro area.
- ▶ Combine the same program across up to five years until reaching a minimum size.
- ▶ If that does not work, combine closely related programs across the same year until reaching a minimum size.
- ▶ If that does not work, combine closely related programs across multiple years until reaching a minimum size.

One important caveat: Graduate-level programs should never be combined with undergraduate-level programs.

Subpopulations
Where sample sizes are large enough, programs should report results for students broken down into the following categories:
- ▶ Socioeconomic status
- ▶ Full- or part-time attendance
- ▶ Whether they already possess bachelor's degrees or higher
- ▶ Race
- ▶ Gender

Results do not need to be reported across categories (e.g., completion rates for white men) due to cell size concerns.

Subpopulations should follow the same aggregation process outlined above to generate adequate sample sizes. If none of these produce a sufficiently large number of learners, then the provider should at least report a provider-wide rate for a campus location.

COHORT DEFINITION
A provider should include all students who started at a program during a 12-month period. The provider would decide what dates define that period, and it should be consistent each year. For example, one provider could choose a measurement period that runs from September 1 of one year to August 30 of the next year, while another could choose a calendar year. The measurement window period should align across outcomes.

CHOICE OF REPORTED OUTCOMES
The broader the application of standards, the greater the chance that instances will arise where some of these measures may make less sense for a particular program or provider, relative to its mission and/or intended outcomes. For example, if a program's mission isn't aligned with boosting learners' earnings, that outcome metric may not be relevant. Recognizing this, programs may opt not to report certain standards measurements under the following conditions:
- ▶ They are not required to report one as part of any other requirement. This includes demands from states, the federal government, or licensing and certification bodies.
- ▶ They do not make any advertising, marketing, or promotional claims about the measure for which they do not make an audited claim.

LEARNING

Although students may consider getting a job, placing into graduate school, or attaining some other end goal to be the primary objective of a program, acquiring knowledge and skills gives them tools to achieve that goal. To that end, understanding student learning may be an important part of any consideration of quality.

Learning is a far more complex issue than any other suggested measure in these standards. Some elements that indicate learning will vary more by program type than a measure such as completion. This may be especially true when attempting to measure noncognitive, or "soft," skills relative to technical proficiency, which are generally considered to be more easily assessed. Given the potential for substantial variation in learning, these standards supply a process for measuring it. They offer general guiding principles, rather than focusing on what the exact measure should be.

CLEARLY ARTICULATED AND TRANSPARENT LEARNING OUTCOMES AND GOALS

Every program within a provider should first create a clear list of all learning outcomes it expects graduates to achieve. These could include both skills and specific content knowledge. Ideally, they should include statements of what students can do with the knowledge and skills. Providers should ensure that the learning outcomes are clear and not overly vague or generalized. For instance, instead of simply saying graduates should "demonstrate the ability to think critically," they should say something closer to "differentiate and evaluate theories and approaches to selected complex problems within the chosen field of study and at least one other field."

CLEARLY DOCUMENTED LEARNING OUTCOMES

Articulating learning outcomes alone is insufficient. Providers must also have a process for concrete documentation that verifies the learning took place. This starts with a plan for how the provider will measure each student's ability to demonstrate these outcomes.

Although the exact process and documentation will vary, possible acceptable items include:
- Pre- and post-program assessments that demonstrate gains in learning and have identical administration conditions (e.g., a pre-program test cannot be given in half the time allotted for a final assessment)
- Post-program assessments that demonstrate mastery
- Portfolios of students' work
- Common assessments that are used by other providers (e.g., CLA+, NAEP, although the latter obviously would not be used here)
- Licensing exam results in fields where licenses are required (e.g., nursing)
- Industry-recognized certifications, where applicable

It is also useful to provide some guidance about what should not be considered acceptable demonstrations of learning:
- *Grades or GPA:* Too subjective, lack external validation, and provide little detail
- *Program completion:* Tells nothing about whether standards were high or lax
- *Student self-evaluations:* Lack quality external validation

INDEPENDENT, OBJECTIVE, AND EXTERNALLY VALIDATED ASSESSMENTS

Learning documentation should include external, independent processes that ensure that someone with no financial relationship to the provider reviewed the materials to evaluate whether they meet the claims of the provider. These standards do not presume the proper way to construct or administer assessments, but evidentiary standards are provided in a supplemental document.

ASSESSOR TESTS FOR RELIABILITY

Human graders will inevitably have inconsistent marks. Although good rubrics and training can partially solve this problem, assessment processes should also include structures that allow testing and controlling for inter-rater reliability.

ENSURE LEARNERS ARE WHO THEY SAY THEY ARE AND THE WORK IS ORIGINAL

Independent, well-documented learning will mean nothing if the people taking the assessments are not who they say they are. Protecting against this requires two things at minimum:
- Robust, upfront checks to verify a student's identity and protect privacy
- Periodic checks during and at the end of a program, course, or other unit of learning, to ensure that the person verified at the beginning is still the person completing the work, and that the work is original work

COMPLETION RATE

Completion rates provide a useful measure of how individual learners fare within a program, but completion rates alone should not be the sole outcome or measure of success for a program. It's even possible for individuals to attend a program and complete only part of it because, for example, a job offer midway through the program represents success for them. But it is also undeniable that program incompletion tends to strongly correlate with unwanted outcomes, such as defaulting on federal student loans. For these reasons, we include completion as one of several measures that provide information about program quality.

The proper formula for completion rates is a hotly debated topic in traditional higher education circles. The problem is that the conventional formula for tracking graduation rates fails to include students who attend part-time, does not give schools credit for individuals who transfer in, and dings their graduation rate for students who transfer out. The suggested definition below acknowledges these problems and tries to solve the question of part-time attendance and transfer students.

MEASURE DEFINITION
The percentage of learners who entered a given program in the same starting cycle and graduated from that program within the stated, anticipated time to completion.

COMPLETER DEFINITION
Institutions should only count students as completers if they have graduated from their programs. Individuals who have finished all required learning elements but have not been formally recognized as graduates should not be counted here. This is important for ensuring that providers that require the completion of a licensing or certification test in order to graduate only count students who pass that final hurdle.

These standards use different measurement windows for learners who attend full-time and those who attend part-time. To figure out the proper timeframe for tracking learners, a provider must generate an anticipated time to completion. How to do this based upon attendance intensity is discussed below.

Full-time students
This anticipated time to completion should be set by the school based upon the amount of time that it expects a student to finish a program with the full-time level of competencies, modules, credits, or other learning measurement units.

Part-time learners
Part-time completion is a more complex measure. Some providers may have a set schedule with a clear time to finish that is just extended beyond that of full-time attendees. Others may have setups where learners have flexibility to choose how much or how little of their coursework they want to attempt.

Because of these complexities, providers have some options for defining anticipated time to completion. First, they must decide if they want to generate individualized completion estimates. This option makes sense when a program allows students to pursue different levels of part-time attendance, such as three-quarter time, half-time, etc. It should not be used in instances where programs have set schedules and part-time learners cannot vary their attendance intensity.

If a program does not offer a set part-time schedule, it can choose to define anticipated time to completion in one of three ways:
1) **Use the historical mode**—under this option, providers would examine the mode time to completion for part-time learners during the past several years.
2) **Individually estimate based upon initial attendance intensity**—under this option, providers would examine learners' initial attendance intensity during the first learning period and project how long those individuals will take to finish if they continue at that pace uninterrupted. For example, if a program consists of 60 modules and a learner completes 10 modules in a three-month learning period, then the expected time to completion is 18 months (six cycles of 10, each lasting three months). This anticipated date of completion can also include expected breaks, such as if a provider does not offer necessary modules for a summer period.
3) **Use a multiple of the anticipated time to completion for full-time students**—under this option, a provider would simply take its full-time completion rate and adjust it for the attendance intensity of students. For example, if a full-time program takes 14 weeks to complete, the provider would estimate that a half-time student would take 28 weeks, a quarter-time student would take 56 weeks, etc.

MEASUREMENT PERIOD AND UPDATE CYCLE
These standards only expect providers to update their completion rates once per year. As noted, providers may optionally track, measure, and report cohorts in shorter intervals of time, but they must also report on an aggregate 12-month basis to comply with the standards. It is the expectation that a more frequent approach would cause too much difficulty for most providers because of staggered completion dates and the addition of multiple cohorts of completers each year.

A provider would have the option to choose the reset date for its completion cycle, but any provider with multiple programs would have to choose the same date for all programs, for the sake of simplicity.

Providers would be expected to update learner completion results once more after the first census in order to provide the percentage of learners who completed after 150 percent of anticipated time to completion.

To encourage number smoothing, these standards would combine up to three years' worth of cohorts of completion rates. The three years would be defined by the expected date of completion, not entry.

A further discussion of the measurement period and how often programs should refresh these results appears in the appendix.

EXCLUSIONS

In general, too many exclusions from a completion rate are worrisome and a potential sign of manipulation. That said, there may be some circumstances where including a student's results in that cohort is unfair. For example, learners called to active military service may be unable to finish their programs. They should be moved to later cohorts that afford enough time after ending their service to track their results.

Allowable exclusions:

EXCLUSION TYPE	RATIONALE
Death	Learners who have died cannot finish their programs.
Permanent disability	Learners who become disabled may not be able to finish their programs. Programs should substantiate this with appropriate documentation. See evidentiary guidelines for reference.
Employed before completion	It may be possible for some learners to achieve their desired outcomes of the program before completion. In these instances, completion is no longer a relevant measurement.

Circumstances that move students to later cohorts:

EXCLUSION TYPE	RATIONALE
Military deployment	Individuals who are deployed likely cannot finish their programs.
Pregnancy/parental leave	Pregnancy and/or parental leave are temporary circumstances. They do, however, mean that individuals may be unable to finish their programs for a period of time. Moving students to a later cohort is a fair way to approach this problem, as students are still counted, just at a later date.
Religious mission	This is accepted practice in traditional higher education. Students on religious missions are unlikely to be able to complete their coursework.
Disability	Learners who become disabled may need to move to a later cohort if they are able to continue the program. This is allowable and should be evaluated on a case-by-case basis. Programs should substantiate this with appropriate documentation—see evidentiary guidelines for reference.

Unacceptable exclusions:

EXCLUSION TYPE	RATIONALE
Not looking to finish	This is too vague a term and open to manipulation. These individuals are best sorted out by determinations on whether they are enrolled in a program. For example, someone who only tries an adult education course at a community college would not be in a program's cohort.

TRANSFER STUDENTS

Students who transfer into or out of a program represent one of the biggest challenges to accurate completion rates in traditional higher education.

It is likely that the question of transfer for programs that last only a few weeks or months is not relevant. The short enrollment period reduces the likelihood that students would be looking to move to other programs. And even if they were, they likely could not bring any credits or other measures of learning with them to other providers.

These standards recognize the potential pitfalls of transfers with suggestions for tackling this issue. This area will be updated as the standards are used in the field.

Transfers to other locations of the same program
Learners who stay in the same program but move across locations should be moved to the cohort of the new program. For example, a learner starts an intensive coding program in New York. Halfway through, the individual moves to San Francisco and enrolls in the same program there. That individual should be counted only in the San Francisco cohort and be removed entirely from the New York cohort.

Transfers into the program
Learners who move into a program after its beginning should be placed into a cohort that reflects the anticipated date of completion. This includes going into the corresponding full- or part-time cohort.

Transfers out of a program
Programs should have a substantial burden of proof to demonstrate that a transfer out of the program is counted as a success. For starters, the program has to demonstrate and document subsequent enrollment in a program elsewhere. But it must go further. Any successful transfer out must show that the learner is now enrolled in a program that accepted the majority of the learning units earned at the original program. Someone who starts at a new program with no carryover from the initial program should not be deemed a success. In addition, a program should not count as a success someone who transfers to another program and then never completes any additional learning units. Such a situation would suggest that the transfer did not stick and that perhaps the learner was not properly prepared for transferring. This condition also protects against the possibility of a program moving underperforming students out to improve numbers.

Verified transfers out of the program should be removed from the numerator and denominator. The exception is if a program has transfer as a stated part of its mission and/or views it as a success to transfer students. In these cases, a program could choose to report a transfer rate, or placement rate, alongside its completion rate (see below).

PLACEMENT RATE

Learners frequently cite obtaining jobs, advancing their careers, or continuing higher education as key reasons for pursuing any kind of postsecondary education or training.[1] This is especially true for career-oriented programs. For other programs, the desired outcome may be placement into another educational program (e.g., the purpose of a liberal arts program in history could be to place students into prestigious history Ph.D. programs). A placement rate thus provides a direct measurement of whether a program was successful in aiding learners with their overall goals.

---

1 The American Freshman: National Norms Fall 2015

MEASURE DEFINITION

The breakdown of all students eligible to receive a credential in the past 36 months and the percentage that within 180 days of exiting the program:

▶ Obtained new employment/placement in an occupation/academic program;
▶ Continued employment at a higher title or substantially higher salary;
▶ Obtained or continued self employment or contract work;
▶ Did not obtain or continue employment, contract work, or self-employment (including entrepreneurship); or
▶ Placed into another outcome relevant to the program mission or focus (as determined and reported by the program).

Exclusions for the reasons discussed later should be reported in a footnote to the measure.

This measure requires distinguishing between the first two bullets because they speak to different types of outcomes. For instance, if most of a program's successes come from individuals advancing in a current field, it may suggest that someone not currently employed in that field may want to consider a different program. Similarly, it would acknowledge the results with contract work or self-employment.

Having providers claim whether a placement is related to the training presents many subjective challenges. To address this, providers may report the top five most common titles, roles, or positions that students hold upon placement and the percentage of employed graduates that hold these titles.

MEASUREMENT PERIOD AND UPDATE CYCLE

These standards recommend placement rates within 90 and 180 days of completion as standard measurements. Some programs may wish to report and make claims about shorter (and/or longer) placement periods. This is acceptable if the institution can provide evidence to support these claims. Longer placement periods may be measured, provided that the 180-day standard is also reported. Importantly, these standards mandate that all promotional claims related to placement rate must be qualified with a timeframe. Programs should not make blind claims for placement rates (job or otherwise) without specifying the period of time measured.

Although 90 and 180 days are recommended as the standard timeframes for placement rate measurements, providers are only expected to report/update their placement rates once per year. A more frequent approach could cause too much difficulty because of staggered completion dates and the addition of multiple cohorts of completers each year. A provider has the option to choose the reset date for its placement rate, but any provider with multiple programs must choose the same date for all of its programs for the sake of simplicity.

To avoid substantial fluctuations and keep data up-to-date, placement rates should include cohort data from the 36 months preceding the reporting date. Cohort rates need not be measured/updated again after the first 90/180-day thresholds, but the reported placement rate should be an aggregate measure of placement rates from the past 36 months.

DENOMINATOR

The denominator should be comprised of all students who earned or were eligible to receive a credential in the past 36 months and were out of the program for at least 90 or 180 days. The criterion that students "were eligible to receive" ensures that students who complete enough credits or competencies are counted, so that programs do not have an incentive to delay awarding credentials. This requirement does not apply to students who may be eligible to receive credentials through a reverse transfer process.

NUMERATOR REQUIREMENTS

Placement rates should be measured within 90 and/or 180 days of program completion (or in line with the above guidance). Given the high number of students who attend postsecondary programs to achieve employment outcomes and the complexities around determining employment, the remainder of this section deals specifically with how to measure these outcomes. The validation criteria would extend to other forms of placement that programs may claim. Below are the conditions for determining whether a learner should be counted as a successful placement based upon: (1) new employment, (2) continued employment, or (3) contract/self-employment.

SITUATION	JOB REQUIREMENTS
New employment	▸ 32 hours a week or more ▸ Paid
Advancement within existing employment	▸ 32 hours a week or more ▸ Paid ▸ Advancement either through title change or pay raise beyond cost of living adjustment
Self-employment or contract work	▸ 32 hours a week or more (a learner who hits the 32-hour mark through a combination of part-time employment and self-employment is counted here, even if self-employment represents a minority of hours worked) ▸ Paid ▸ Duration of five weeks or longer

NONRESPONSE

It is likely that even aggressive, well-designed learner outreach will not yield a 100-percent response rate. These individuals should not be counted as successful placement. While some of them may have been successfully placed, it is unreasonable to count them as such without definitive proof. Providers also should not exclude nonresponsive learners from the placement rate denominator. Doing so creates unwanted incentives to intentionally avoid less successful learners so they do not show up in the data. If desired, nonresponse rate may be reported as a separate figure for context.

EXCLUSIONS

In general, too many exclusions from a placement rate are worrisome and a potential sign of manipulation.

Allowable exclusions:

EXCLUSION TYPE	RATIONALE
Death	Learners who have died cannot work.
Permanent disability	Learners who are disabled may not be able to work. Depending on the disability, this may be a permissible exclusion.

Circumstances that move students to later cohorts:

EXCLUSION TYPE	RATIONALE
Military deployment	Individuals who are deployed cannot also hold jobs. They may also be unable to be enrolled in further higher education, for example.
Pregnancy/parental leave	Pregnancy and/or parental leave are temporary circumstances. They do, however, mean that individuals may be unable to search for jobs or have other types of placements for a period of time. Moving students to a later cohort is a fair way to approach this problem, as students are still counted, just at a later date.
Religious mission	Students on religious missions will be temporarily out of the job market and unavailable for other forms of placement as well (e.g. further schooling). They should be moved to a new cohort.
Disability	Learners who become disabled may need to move to a later cohort if they are able to continue work/training at a later date. This is allowable and should be evaluated on a case-by-case basis. Programs should substantiate this with appropriate documentation—see evidentiary guidelines for reference.

Unacceptable exclusions:

EXCLUSION TYPE	RATIONALE
Not looking for a job	Too vague and open to manipulation. However, learners' purposes for enrolling will be measured and reported for context.
Foreign students	Though these individuals may have a harder time finding employment, they presumably still sought the course for this purpose.
Full-time enrollment in a training program offered by a different provider	Graduate school placement or enrollment in additional training may be reported separately from job placement for applicable programs, but should not be excluded.
Nonresponsive students	This would encourage too much manipulation in the data.

See evidentiary guidelines for reference.

EARNINGS

Earnings are a crucial measure of value. They provide answers to key questions that students want answered, such as, "Will my investment pay off?" Earnings also serve as a check on completion data. If a program has high completion rates but poor earnings results, then there may not be sufficient value in the underlying program. Similarly, good earnings results may be cause for excusing seemingly subpar completion rates.

These standards acknowledge that several factors go into earnings. In addition to underlying program quality, earnings can be affected by geographic location, the types of occupations available to completers, and the program type. For these reasons, these standards recommend looking at earnings through four distinct measures:
- ▶ Absolute earnings
- ▶ Percent of students earning above a minimum threshold
- ▶ Change in earnings over time
- ▶ Earnings relative to net price paid

Having multiple measures helps protect against certain issues. For example, looking only at absolute earnings would otherwise disadvantage programs that may be tied to lower-wage occupations. But these programs may provide a good increase in wages or help secure a salary that is not low relative to what students paid upfront. From this data, outside rankings could construct return-on-investment (ROI) measures for multiple stakeholders (e.g., students, taxpayers) to better inform policy and enrollment decisions.

The ideal standard of evidence would be to obtain results from a government database, such as those held by the United States Social Security Administration or state unemployment insurance agencies.

MEASUREMENT DEFINITIONS
Absolute earnings: The median annual earnings of all individuals in the cohort and/or program graduates within the first full calendar year and after the fifth full calendar year post-graduation. Includes individuals with $0 earnings. These data should be annualized if reported on a quarterly basis.

Earnings threshold: The percentage of program graduates whose earnings within the first full calendar year and after the fifth full calendar year post-graduation are more than 200 percent of the federal poverty level for a one-person household in the continental United States. This measure assumes a single person household, regardless of actual circumstances. These measures use 200 percent because it more closely approximates the level of a high school graduate's earnings, which the U.S. Department of Education puts at roughly $25,000.

Earnings change over time: The median increase in earnings for program graduates within the first full calendar year and after the fifth full calendar year post-graduation, measured as the change in income from the last full calendar year prior to entering a program. The key part of this measure is obtaining prior data on earnings. This must be done through an information collection that occurs at the time of entry. Asking for past earnings several years after a learner graduated is unlikely to yield accurate data. Providers should obtain these data from either:
- ▶ Federal or state data systems capable of tracking this information; or
- ▶ Financial aid applications that require verified documentation of prior earnings through a tax form such as a W-2.

Earnings relative to net price paid: The typical annual earnings of a graduate within the first full calendar year and after the fifth full calendar year post-graduation expressed as a percentage of the typical amount paid for a program.

MEASUREMENT WINDOW
These standards recommend two measurement windows to judge earnings success. First is an early-term look that judges earnings results within the first full calendar year after completing a program. Second is a longer-term measure that looks at results in the fifth full calendar year after completing a program.

The longer window is considered to be better suited for accountability purposes because it affords time for learners to adjust to the workforce and also judges whether the program has ongoing usefulness. That said, we recognize that obtaining information from learners five years after leaving a program may be especially difficult and could be less relevant for short-term programs focused on more proximate outcomes.

MEASUREMENT PERIOD AND UPDATE CYCLE

These standards only expect providers to update their earnings results once per year. A provider would have the option to choose the reset date for its earnings data, but any provider with multiple programs would have to choose the same date for all of its programs for the sake of simplicity.

Based upon the date chosen, the provider would report completion results for all graduates within a cohort who have been out of school long enough to track within a window that includes at least one calendar year. In other words, if individuals graduate from a program on September 30, 2016, then the provider has until December 31, 2017, to report their earnings.

Providers would be expected to update earnings results at two census points: within at least one full calendar year after leaving a program and after the fifth full calendar year post-graduation. They would not be expected to collect intermediate data points for the purposes of these standards.

To encourage number smoothing, these standards would combine up to three years' worth of cohorts of earnings rates together. The three years would be defined by the end census date for earnings. In other words, if a program ends in September 2016, its census date would be on December 31, 2017, and so would not be counted in the data reported for 2017.

A further discussion of the measurement period and how often providers should refresh these results appears in the appendix.

COHORT DEFINITION

Ideally, earnings data should include all learners who were eligible to complete a program during a 12-month window (i.e., following the same cohort definition outlined earlier). It may of course be unrealistic for some programs to capture this data from non-completers. Therefore, these standards recommend that institutions include data for all program graduates and encourage them to provide data for all students. The level of data and evidence provided should be qualified during the audit process, and institutions must not make false or misleading claims about earnings data or the population from which they collect and report data.

Earnings cohorts should include graduates, even if they are making $0 and do not meet one of the exclusion categories mentioned below.

EXCLUSIONS

In general, too many exclusions from earnings data are worrisome and a potential sign of manipulation.
That said, there may be some circumstances where including a student's results in that cohort is unfair. For example, learners called to active military service may see lower wages than they would have otherwise. They should be moved to later cohorts that afford enough time after ending their service to track their results.

Allowable exclusions:

EXCLUSION TYPE	RATIONALE
Death	Learners who have died cannot earn salaries/wages.
Permanent disability	Learners who become disabled may not earn salaries/wages.

Circumstances that move students to later cohorts:

EXCLUSION TYPE	RATIONALE
Military deployment	Individuals who are deployed may not have earnings related to their programs.
Religious mission	Individuals on religious missions likely will not have earnings related to their programs.

Unacceptable exclusions:

EXCLUSION TYPE	RATIONALE
Not looking for employment	Most individuals initially using these programs are looking for jobs. This claim is too vague and open to manipulation.
Foreign students	Although some learners may be studying on an education visa, they should still be employable where they hold working rights.
Nonresponsive students	This would encourage too much manipulation in the data.

See evidentiary guidelines for reference.

STAKEHOLDER SATISFACTION AND CONFIRMATION OF PURPOSE

Feedback from students and employers can provide vital insight into attitudes and opinions that could not otherwise be captured by the numbers. In particular, it can help to see if the reasons students are attending a given program match up with the mission of that program. It also represents a mechanism for bringing employer validation into the quality assurance system.

BENCHMARKING

To establish a benchmark for satisfaction, we must first understand students' purpose for attending, expectations, and desired outcomes. Upon admission to a program, but prior to beginning, students should complete an intake survey to capture why they are enrolling and what they hope to accomplish.

EXCLUSION TYPE	RATIONALE	RATIONALE	RATIONALE
Students	What is your purpose for attending? What are your expectations/ desired outcomes of the program?	After admission and before the program begins. Option to repoll students in later surveys in order to confirm or reassess their purposes and motivations.	Whether students' purposes and expectations match with the institution's mission will be a useful measure with regards to placement and satisfaction.

Upon completion of a program, this framework recommends two types of satisfaction measurements; one for students and another for employers who have hired program graduates.

POST-SEPARATION SURVEY LEVELS

Surveys should follow a similar approach to the one used for a net promoter score, which relies on a single, clear question. Alternative follow-up questions with more detail may also have some utility.

In this case, the survey would look something like the below:

WHO TO SURVEY	QUESTION	SCORE RANGE
Students	Knowing what you know now, how likely would you be to repeat this experience?	1 to 10
Employers	Knowing what you know now, how likely are you to hire graduates from this program again?	1 to 10

Optional additional student questions:

QUESTION	RESPONSE TYPE
Did this program fulfill your purpose for attending?	1 to 10
Did this program meet your expectations/desired outcomes?	1 to 10
If you secured a new job after completing this program, do you believe it was a direct result of having completed the program?	Yes, No, Don't Know
Was this program required for you to obtain this job?	Yes, No, Don't Know
How well did the program prepare you for this job?	1 to 10

Optional additional employer questions:

QUESTION	RESPONSE TYPE
Was program completion the deciding factor in the hiring or promoting of this individual?	Yes, No, Don't Know
Was the program relevant to placement in this job?	Yes, No, Don't Know
How well prepared was the graduate for this job?	1 to 10

MEASUREMENT PERIOD AND UPDATE CYCLE

These standards only expect providers to update their satisfaction results once per year. A provider would have the option to choose the reset date for its earnings data, but any provider with multiple programs would have to choose the same date for all of its programs for the sake of simplicity.

To encourage results smoothing, these standards would combine up to three years' worth of satisfaction surveys. The three years would be defined by the end date for conducting a survey.

A further discussion of the measurement period and how often providers should refresh these results appears in the appendix.

APPENDIX

DISCUSSION OF MEASUREMENT PERIOD AND UPDATE CYCLE: COMPLETION

Based upon the date chosen, the provider would report completion results for all learners within a cohort who have been enrolled long enough to meet the expected date of completion. In other words, if a program is one year and someone enrolls on September 1, 2016, then the provider should not consider that person a full-time enrollee in a program cohort until after September 1, 2017.

To the extent practicable, providers should strive to have reporting dates that are not close to the anticipated completion date for cohorts. Where that is unavoidable—as it will be for certain programs with more flexible start-dates, for example— providers have the option to decide in what reporting year to include programs that have anticipated completion dates close to the reporting date. This is done to ensure there is enough time to properly tabulate results. For example, a cohort with an anticipated completion date of December 31, 2016, may not have collected all the data necessary for reporting results on January 1, 2017. In this case, that program can report its results on January 1, 2018.

An example illustrates how this works. A program resets its completion rates on January 1 of each year. It has been in operation since 2013 and has had entering cohorts of learners in July and September of each year. The full-time program lasts six months, while the part-time program uses an anticipated time to completion of one year.

Under that framework, here's how the cohorts would work for the completion data reported on January 1, 2017.

Full-time students

COHORT START DATE	ANTICIPATED COMPLETION DATE	COUNTED IN THE MEASURE?
May 1, 2013	October 31, 2013	No—cohort too old
September 1, 2013	February 28, 2014	Yes
May 1, 2014	October 31, 2014	Yes
September 1, 2014	February 28, 2015	Yes
May 1, 2015	October 31, 2015	Yes
September 1, 2015	February 29, 2016	Yes
May 1, 2016	October 31, 2016	Yes
September 1, 2016	February 28, 2017	No—cohort too new

Part-time learners:

COHORT START DATE	ANTICIPATED COMPLETION DATE	COUNTED IN THE MEASURE?
May 1, 2013	May 31, 2014	Yes
September 1, 2013	September 30, 2014	Yes
May 1, 2014	May 31, 2015	Yes
September 1, 2014	September 30, 2015	Yes
May 1, 2015	May 31, 2016	Yes
September 1, 2015	September 30, 2016	Yes
May 1, 2016	May 31, 2017	No—cohort too new
September 1, 2016	September 30, 2017	No—cohort too new

As the example shows, every cohort included in the placement rate is finished prior to the reporting date with enough time to conduct a measurement. Cohorts also do not receive future follow up. Data reported at the 12-month evaluation are included in the rate until they age out.

MEASUREMENT PERIOD AND UPDATE CYCLE: EARNINGS
To the extent practicable, providers should strive to have reporting dates that are not close to the final measurement dates for cohorts. If that is not avoidable, a provider has the option to decide in what reporting year to include programs that have anticipated completion dates close to the reporting date. This is done to ensure there is enough time to properly tabulate results. For example, a cohort with earnings tracked until December 31, 2016, may not have collected all the data necessary for reporting results on January 1, 2017. In this case, that program can report its results on January 1, 2018.

An example illustrates how this works. A program resets its earnings rates on January 1 of each year. It has been in operation since 2013 and had graduating cohorts of learners in May and September of each year. Under that framework, here's how the cohorts would work for the earnings data reported on January 1, 2017.

COHORT COMPLETION DATE	1 YEAR WINDOW ENDS	COUNTED IN THE MEASURE?
May 1, 2013	December 31, 2014	Yes
September 1, 2013	December 31, 2014	Yes
May 1, 2014	December 31, 2015	Yes
September 1, 2014	December 31, 2015	Yes
May 1, 2015	December 31, 2016	Yes
September 1, 2015	December 31, 2016	Yes
May 1, 2016	December 31, 2017	No—cohort too new
September 1, 2016	December 31, 2017	No—cohort too new

MEASUREMENT PERIOD AND UPDATE CYCLE: SATISFACTION

Based upon the date chosen, providers conduct retrospective satisfaction surveys of all students who left a program and the employers who hired them no sooner than one year after separation and no later than two years after separation. In other words, if someone leaves a program on September 30, 2016, then the provider would not survey that student until after September 30, 2017, but before September 30, 2018. The rationale for this waiting period is that in the aftermath and euphoria of completing a program, before students have moved into the next phase of life, it is easy to overvalue or overrate their satisfaction with the program on a simple net promoter score. This allows students to not only evaluate the program experience itself, but also its value to their lives after leaving the program.

To the extent practicable, providers should strive to have reporting dates that are not close to the final measurement dates for cohorts. If that is not avoidable, a provider has the option to decide in what reporting year to include programs that have anticipated completion dates close to the reporting date. This is done to ensure there is enough time to properly tabulate results. For example, students and employers whose satisfaction scores can be tracked until December 31, 2016, may not have collected all the data necessary for reporting results on January 1, 2017. In this case, that program can report its results on January 1, 2018.

An example illustrates how this works. A program resets its satisfaction data on January 1 of each year. It has been enrolling students since 2013. Under that framework, here's which students would be counted in the reported data as of January 1, 2018.

YEAR OF STUDENT SEPARATION	MEASUREMENT WINDOW BEGINS	MEASUREMENT WINDOW ENDS	COUNTED IN THE MEASURE?
2013	2014	2015	Yes
2014	2015	2016	Yes
2015	2016	2017	Yes
2016	2017	2018	No—cohort too new