CollegeBoard ACCUPLACER®

Multiple Factors in College Placement Decisions

2019

Introduction

The purpose of this document is to provide guidance in using additional information when making placement decisions using scores from the next-generation ACCUPLACER® placement tests. This guidance is based on professionally responsible practice and upholds the standards of educational measurement. Additionally, when decisions are based on a combination of ACCUPLACER test scores and other information, sound measurement principles are closely observed.

In its *Code of Professional Responsibilities in Educational Measurement* (Section 6.7), the National Council on Measurement in Education (1995) states that

> Persons who interpret, use, and communicate assessment results have a professional responsibility to use multiple sources and types of relevant information about persons or programs whenever possible in making educational decisions.

As the publisher of ACCUPLACER, the College Board's commitment to our professional responsibility regarding the use of additional information is manifested in the *Guidelines on the Uses of College Board Test Scores and Related Data* (College Board, 2011, pp. 6-7). These *Guidelines* state that "test scores are useful as one means of predicting academic performance in college when considered with other relevant information." The College Board's commitment to our professional responsibility also allows us to uphold the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, Standard 12.10) which asserts that:

> In educational settings, a decision or characterization that will have major impact on a student should take into consideration not just scores from a single test but other relevant information.

Like with any assessment, ACCUPLACER test scores are not perfectly precise and should not be treated as though they are. They are useful as means of predicting performance in college courses when considered with other relevant information. That is to say that the College Board does not only support using additional relevant information when ACCUPLACER test scores are used for college placement, it also does not support the use of ACCUPLACER scores as the only basis for placement decisions.

Multiple Factors

When buying a car, one must not consider the price alone. Price may be used as a starting point, but one also must consider quality, total cost of ownership, reliability, and so forth. These are all relevant information about a car when considering whether the car suits you. These are factors to base our decision on when buying a car.

By the same token, multiple factors should be used when making high-stakes decisions such as college placement. These include high school records, life skills, passion for achieving goals, student's economic situation, etc. For the intent of this document, we will use the definitions discussed in the rest of this section.

Definitions

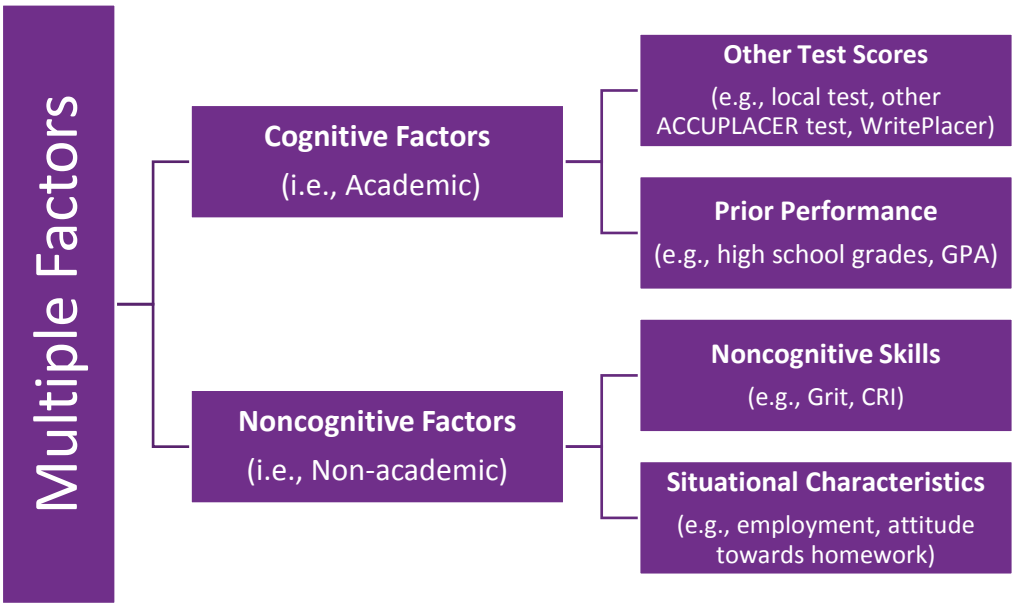
A *factor* is defined as one of the elements contributing to a particular result or situation (http://www.dictionary.com/browse/factor?s=t). In the context of college placement, relevant information that may be used in addition to ACCUPLACER test scores includes both cognitive and noncognitive factors. These factors are elements that are useful in predicting student performance in college courses, and thus should be considered in making placement decisions. *Cognitive factors* are academically related factors. These include current intellectual abilities as well as prior academic performance. Cognitive factors do not include behavioral, emotional, or psychological traits. *Noncognitive factors* include factors that are not related to academics, but they nevertheless affect students' performance in college courses. Some of these factors are measures.

A *measure* is an instrument for measuring a construct. A thermometer, for example, is an instrument for measuring temperature. Academic measures are assessments that measure academic skills, such as standardized tests, grade point average or a writing sample with a scoring rubric. Noncognitive measures are instruments, such as student surveys or interviews, which show non-academic levels of interest and ability. Situational characteristics such as the amount of time a student dedicates for studying are not measures.

Taxonomy

Below is an organization of the multiple factors that should be considered for college placement using ACCUPLACER scores. This taxonomy of multiple factors is intended to help clarify the thinking in how each factor may be used.

Figure 1: Taxonomy of Multiple Factors Used in Making Placement Decision Using ACCUPLACER Tests



Multiple Approaches

There are two general approaches in using multiple factors for course placement using ACCUPLACER tests. The *additive approach* is when scores on some academic measures, or multiple of them, are added to the ACCUPLACER score, or a multiple of it. For example, placement decision on a credit-bearing English Composition course may be based on the average of the ACCUPLACER Reading and Writing scores. The *decision tree approach* is when you consider a factor one at a time based on one criterion at each juncture, until a final juncture when a placement decision is made. The decision tree approach is appropriate when using noncognitive factors. Further explanations and examples are provided below.

Additive Approach

The additive approach is appropriate when, in addition to an ACCUPLACER test score, other scores on some cognitive or academic measures are used in placement decisions. When this approach is used, the placement score is in the form of

$$S = a_1S_1 + a_2S_2 + a_3S_3 \ldots + a_nS_n + b$$

where S = placement score

S_i = score on measure i

a_i and b are scalars

for i = 1, 2, 3, …, n.

Example 1: Weighted Composite of Scores with the Same Range

In the example mentioned above, where the placement score for an English Composition course is the average of the ACCUPLACER Reading (R)and Writing (W) scores, the placement score is

$$S = \frac{R+W}{2} = \frac{1}{2}R + \frac{1}{2}W = 0.5R + 0.5W$$

As shown in this example, averaging the scores mean that the Reading and Writing scores are weighted equally in the placement score where the weight is 0.5 for each. In cases like this, the weights are assigned or decided upon by policy. Such policy is based on philosophical beliefs or theoretical basis. The weights applied to the different components must sum up to 1 because they are the proportions that the different components contribute to the composite score.

Note: In this example, each of the Reading and Writing scores have a range of [200, 300]. The next example discusses composite scores when the score ranges are different.

Example 2: Weighted Composite of Scores with Different Ranges

Suppose that the policy reflects a placement score composed of 20% of the Reading score and 80% WritePlacer® (i.e., Essay (E)) score. Note that the weights are 0.2 and 0.8, respectively. Before combining

the scores with the respective weights applied, the two scores must be on scales that have similar statistical characteristics. One way to accomplish this is to place the essay score *E*, which has scores of 1 through 8, on the Reading score scale that ranges from 200 to 300, with a mean of 250 and standard deviation of 20. Transforming the WritePlacer scores of 1—8 to the 200 to 300 scale may be accomplished by (1) mean-sigma, or (2) anchoring the end points of the scales.

Mean-Sigma

Mean refers to the average score and *sigma* refers to the standard deviation (sd). If the essay scores have an average of E_{mean} and standard deviation E_{sd}, transforming the essay score *E* to *E'* such that

$$E' = bE + a$$

where

$$b = \frac{20}{E_{sd}}$$

and

$$a = 250 - b * E_{mean}.$$

To illustrate, if E_{mean} = 4.43 and E_{sd} = 1.23, the b = 16.26 and a = 177.97. Thus, E' = 16.26**E* + 177.97. It is convenient to create a conversion table like the one in Table 1.

*Table 1: Transformed WritePlacer Score for Example 2**

WritePlacer Score (*E*)	Transformed WritePlacer Score (*E'*)	Rounded *E'*
1	194.23	200
2	210.49	210
3	226.75	227
4	243.01	243
5	259.27	259
6	275.53	275
7	291.79	292
8	308.05	300

* This table serves as an exemplar. Institutions should make their own score determinations.

Using the rounded values of E', the composite placement scores such that 20% is from the Reading score and 80% is from the WritePlacer score is

$$S = 0.2R + 0.8E'$$

Without using the transformation table (i.e., Table 1), the equation below may be used directly. Note that this equation is of the form presented on p. 3.

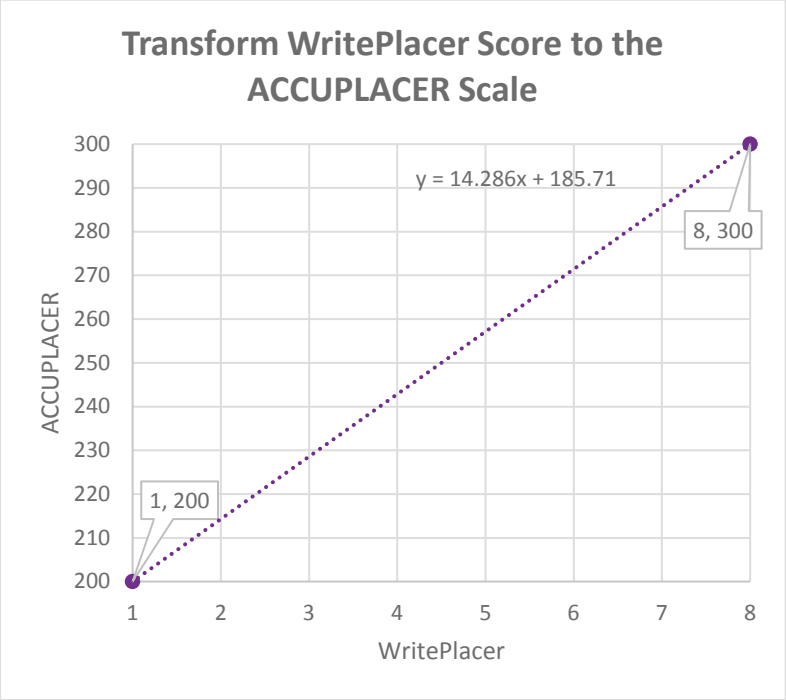
$$S = 0.2R + 0.8(16.26E + 177.97) = 0.2R + 13.00E + 14.23$$

The example described above is appropriate when the placement score is based on an ACCUPLACER test score combined with a locally developed test. The average and standard deviation must be known for each test for the population for which the policy is used.

Anchor End Points

If the mean and standard deviation of at least one of the tests is unknown, the transformation may be based on anchoring the end points of the two scales. Figure 2 illustrates a line determined by two points which are the result of mapping the lowest possible scores and the highest possible scores, respectively.

Figure 2: Linear Transformation by Anchoring the End Points of the Score Scales



Using this linear transformation $E' = 14.29E + 185.71$, the composite score comprised of 20% Reading and 80% WritePlacer is

$$\begin{aligned} S &= .2R + .8E' \\ &= .2R + .8(14.29E + 185.71) \\ &= .2R + 11.43E + 148.57 \end{aligned}$$

Example 3: Composite Predictor

The College Board has always recommended that a predictive placement validity study be conducted on course placement policies that use ACCUPLACER scores. One year's worth of data from the institution is usually sufficient for such a study. Furthermore, the College Board offers to conduct this study free of charge for institutions through the Admitted Class Evaluation Service (ACES).

Suppose that after using both the Reading and Writing scores (such as in Example 1 above) for placement for a year a predictive placement validity study using multiple regression is performed. Included in the results of the study is a composite predictor (*CP*) score of the form

$$CP = Intercept + Slope_R * R + Slope_W * W$$

where the values of the intercept and slopes are computed from the data. Each value of *CP* is associated with an expected probability of success, and vice versa. The CP value for each expected probability of success by 0.05 is given in Table 2.

*Table 2: Probability of Success Associated with each Composite Predictor**

Expected Probability of Success	Composite Predictor
0.95	2.94
0.90	2.20
0.85	1.73
0.80	1.39
0.75	1.10
0.70	0.85
0.65	0.62
0.60	0.41
0.55	0.20
0.50	0.00
0.45	-0.20
0.40	-0.41
0.35	-0.62
0.30	-0.85
0.25	-1.10
0.20	-1.39
0.15	-1.73
0.10	-2.20
0.05	-2.94

* This table serves as an exemplar. Institutions should make their own score determinations.

The institution may choose the expected probability of success of students that they place in the course. If they pick 0.75, then they place students in the course if their *CP* scores are at or above 1.10.

Suppose the logistic regression procedure results in the following values:

Intercept = -4.345

$Slope_R$ = 0.004

$Slope_W$ = 0.019

A student who has a Reading score of 232 and a Writing score of 220 will have a *CP* value of 0.76. Thus, this student has an expected probability of success that is less than 0.70. Another student with a Reading score of 276 and a Writing score of 253 will have a *CP* value of 1.57. This student has an expected probability of success that is higher than 0.80.

The composite predictor may also be used with any number of academic measures indicating prior performance. These include previous grades or grade point average (GPA) previously collected.

Decision Tree Approach

A decision tree is a graphical depiction of decision-making that shows all the various alternatives and possible outcomes. Decision trees are usually represented graphically in a hierarchical structure that contains a starting node called the root and a group of branches or conditions that lead to other nodes until a final decision from each branch is reached. A decision tree can be used to visually and explicitly represent a decision-making process. It is a very beneficial tool for academic advisors working with students for proper course placement.

Decision Rules

Suppose that the ACCUPLACER Writing test is used for placement in an English Composition Class. The placement score being used, in the example is 236. This placement score might have been the result of a standard setting process or the institution might have been using information from the Skills Insights™ for ACCUPLACER Writing. If a student's score is below 236, it is recommended to use other information relevant to how the student would perform in the course, especially if the students score is lower, or close enough to the placement score. Factors relevant to succeeding in the specific course as well as those that are relevant in succeeding in college might be considered, such as number of hours of employment during the semester and attitude toward homework. Two questions come to mind:

1. What do we mean by "close enough" to the placement score?
2. What factors relevant to success should be considered?

Conditional Standard Error of Measurement (CSEM)

If a student takes a test multiple times, their observed scores will vary for many reasons. This variation may be due to how the student is feeling while taking the tests, the different test administration condition, or the time of day when the test was taken, as examples. The consistency of the student's observed scores is referred to as test score reliability. Standard error of measurement (SEM) can be considered a measure of inconsistency in test scores. A SEM value is an average across all observed scores while a conditional standard error of measurement (CSEM) is the estimated SEM for a particular (conditioned on) observed score.

Each ACCUPLACER score has an associated CSEM. The CSEM is a value computed to indicate the level of certainty about where a student's true score may lie given the score they obtained. For example, if a test taker receives a score of 250 on a next-generation placement test and the CSEM is 4.9, there is a 68% probability that the examinee's true score is within the 245.1 and 254.9 range. A smaller value of CSEM provides more precise measurement. In the context of decision rules, we can say that the

student's score is close enough to the placement score if the placement score is within one CSEM of the student's score.

Success-Relevant Factors

Both cognitive and noncognitive factors may be taken into consideration. For each factor variable, the attributes must be dichotomized to support a binary tree. An example of a factor that affects success in the course is the amount of time a student can dedicate to studying for the course. A student with fewer work hours might have more time to dedicate to studying for the course and doing homework. However, this might only be true for some students depending on their attitude towards homework. In this example, the first factor may be dichotomized as working less than half time vs. working at least half time. The second factor may simply be favorable vs. unfavorable.

Decision Tree

The decision process in this situation above is presented as a decision tree in Figure 3. The root is the ACCUPLACER score and GPA, and the three nodes are the CSEM, weekly work hours, and attitude towards homework. The attractiveness of a decision tree is that it can also be presented as a set of decision rules as presented in Table 3. Each node corresponds to a decision rule, where the last rule results in a terminal decision. Although there is no limit to how many nodes a decision tree can have, it is recommended to only include a few important factors. Note the root may be a composite score or a score that already has other factors based on the additive approach.

Figure 3: Decision Tree

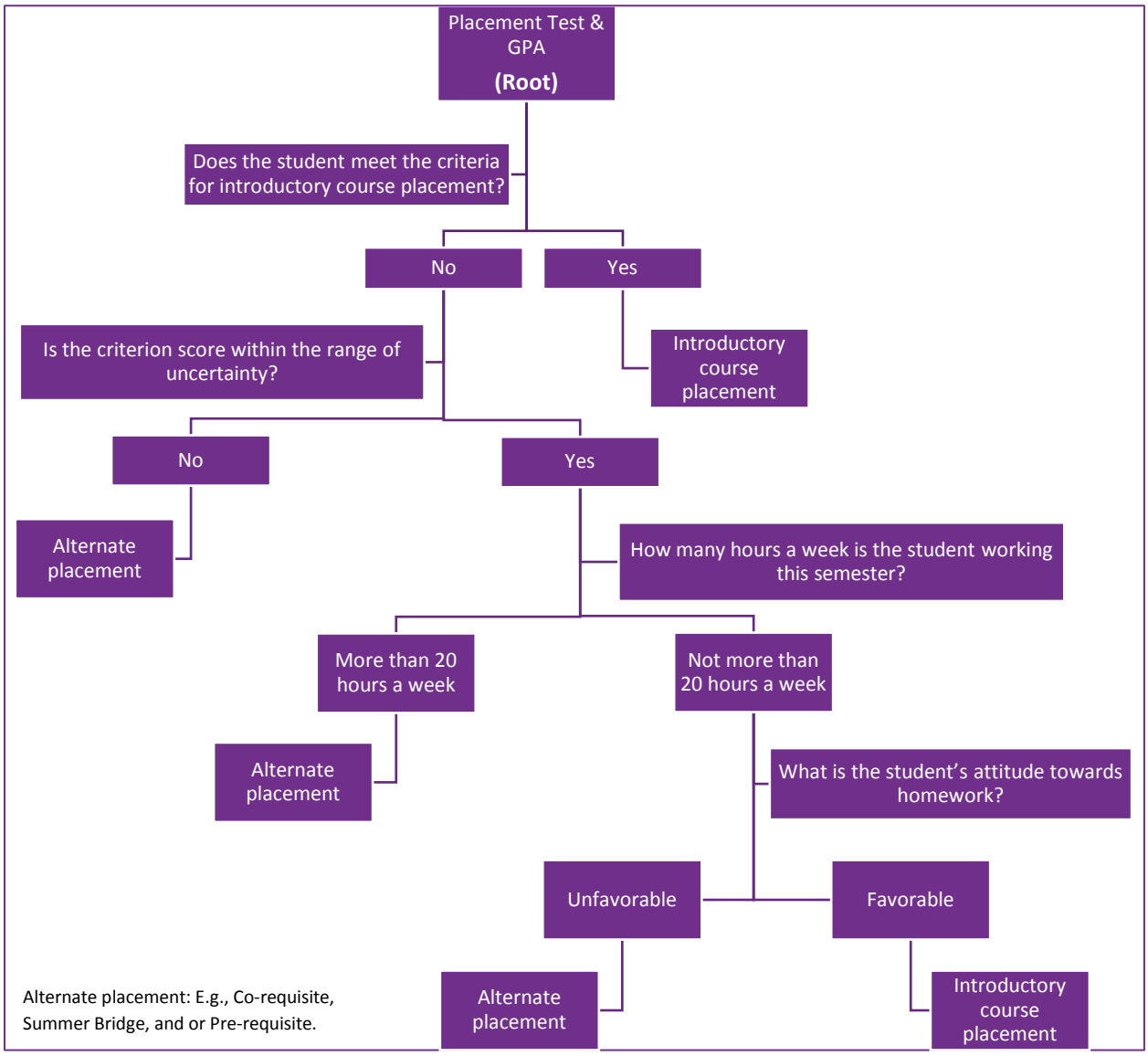


*Table 3: Decision Rules Set**

Rule	Variable	Possible Values	Decision Points*	Decisions	Terminal?
1	Score & GPA	[200, 300; 1.0-4.0]	<236; <2.67	Go to rule #2	No
			≥236; ≥2.67	Introductory Course Placement	Yes
2	CSEM	[0, 100]	<236 - Score	Go to rule #3	No
			≥236 - Score	Alternate Placement	Yes
3	Weekly Work Hours	[0, 168]	<20	Go to rule #4	No
			≥20	Alternate Placement	Yes
4	Attitude Towards Homework	Favorable, Unfavorable	Favorable	Introductory Course Placement	Yes
			Unfavorable	Alternate Placement	Yes

* This table serves as an exemplar. Institutions should make their own score determinations.

Data Mining

The decision tree illustrated above was constructed by considering factors that affect success for those students whose ACCUPLACER scores and GPA are close to meeting the criteria for the introductory course placement. Decision trees may also be generated using data mining tools when there is a considerably large data set that contains information on past decisions made based on ACCUPLACER tests scores and all the relevant factors to consider. This type of decision tree is considered empirically based.

Summary and Recommendations

Consistent with our professional responsibility as the publisher of ACCUPLACER, and upholding the *Standards* (AERA, APA, & NCME, 2014), the College Board recommends the use of multiple factors when using ACCUPLACER tests for placing students in credit-bearing courses. Factors that are relevant to students' college success are categorized into cognitive and noncognitive types. Cognitive factors are also called academic factors and are further categorized in to (1) performance on other test scores and (2) indicators of prior academic performance. Noncognitive factors are categorized into (1) noncognitive traits and measures and (2) situational characteristics.

Two approaches in using multiple factors are recommended here and are supported by the ACCUPLACER system. The additive approach is employed when the supplemental information being used is in the form of other tests or academic measures. The construction the weighted composites is based on principles and beliefs held by the institution responsible for making placement decisions. The construction of the composite predictor is based on empirical calculations from collected data.

The College Board unequivocally supports the use of noncognitive factors in placement decisions. The recommended use of noncognitive factors in placement decision is through decision tree approaches.

Decision rules are constructed with the ACCUPLACER score and CSEM considered in the first two nodes of the decision tree. The rest of the nodes are based on principles and beliefs held by the institution responsible for making placement decisions. When appropriate data are available, empirically based decision trees maybe constructed using data mining techniques.

Multiple factors and recommended approaches are summarized in Table 4. Note that an additive approach, or any approach that will alter the final ACCUPLACER score based on noncognitive measures is not recommended.

Table 4: Multiple Factors and Their Recommended Uses

Factors		Examples	Recommended Approaches			
			Additive		Decision Tree	
			Weighted Composites	Composite Predictor	Decision Rules	Data Mining
Cognitive	Performance on Other Tests	• Other ACCUPLACER tests scores • ACCUPLACER retest scores • Local test scores • Other standardized test scores • Essays with scoring rubric	✓	✓	✓	✓
	Prior Performance Measures	• High school course grades • High school GPA		✓	✓	✓
Noncognitive	Noncognitive Traits and Measures	• Grit[1] • Social and emotional traits • Local surveys[i] • Conley Readiness Index (CRI)[2] • Other standardized noncognitive assessments			✓	✓
	Situational Characteristics	• Employment status • Number/amount of personal commitments • Attitudinal variables • Use of cognitive skills in day to day basis			✓	✓

[1] In this context, we use "grit" as a set of noncognitive traits that include perseverance, hardiness, ambition, need for achievement, and conscientiousness. The use of the Grit Scale is not recommended for use in college placement because of the high stakes nature of the decisions and students' ability to fake. This is consistent with the Grit Scale author's cautionary note when discussing the scale's limitations: "I also discourage the use of these scales in high-stakes settings where faking is a concern (e.g., admissions or hiring decisions)." (https://angeladuckworth.com/research/)

[2] https://www.pearson.com/us/higher-education/products-services-institutions/career-success-program/how-it-works/conley-readiness-index.html

Remarks on Validity and Reliability

The *Standards* (AERA, APA, & NCME, 2014) refer to validity as "the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests" (p.11). The rigor by which the ACCUPLACER tests are developed and the alignment studies performed addresses both the procedural and empirical aspect of content validity. In the context of placement, predictive validity is just as important if not more because it is the crux of how well the assessment performs its purpose. The College Board recommends that a predictive placement validity study be performed on each specific use of each ACCUPLACER test every year when sample size allows. In support of this recommendation, the College Board offers to conduct studies for institutions through Admitted Class Evaluation Service (ACES). Results of predictive placement validity studies from ACES provide information that will further improve upon the placement decision-making process based on ACCUPLACER scores and additional relevant factors.

One may consider many different criteria when selecting factors to include in the placement decision-making process. When considering the additive approaches the reliability of the measures being combined with the ACCUPLACER score is an important matter. It's always a concern when a less reliable measure is combined with the ACCUPLACER score due to the expected degradation of the reliability of the measure (i.e., the composite score) relative to the reliability of using the ACCUPLACER score alone. With reliability being a necessary condition for validity, there is a concern as to whether the validity of the placement decisions are thus degraded. Decision makers are faced with this dilemma when they feel that the additional measure is necessary to determine how student will perform in the course. This is exactly the situation when an essay score is added to a test score based on multiple-choice items. We thus conclude this document with the following words:

> The use of multiple measures itself does not necessarily improve the reliability and validity of the decisions. It is the logic by which the measures are combined that determines the accuracy and appropriateness of the decisions reached (Chester, 2003).

In memory of Mitchell Chester, 1952-2017

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*.

Chester, M. D. (2005). Making valid and consistent inferences about school effectiveness from multiple measures. *Educational Measurement: Issues and Practice, 24*(4), 40–52.

College Board (2011). *Guidelines on the Uses of College Board Test Scores and Related Data*. New York, NY: The College Board.

National Council on Measurement in Education. (1995). *Code of professional responsibilities in educational measurement*. Retrieved from http://www.ncme.org/ncme/NCME/NCME/Resource_Center/LibraryItem/Code_of_Professional_Responsibilitie.aspx

