Abstract

Effective assessment practice requires clearly defining and operationalizing terminology. We illustrate the importance of this practice by focusing on academic "undermatching"—when students enroll in colleges that are less academically selective than those for which they are academically prepared. Undermatching has been viewed as a potential obstacle in the United States' goal of increasing degree attainment but operationalizing undermatching is difficult. Using ELS: 2002, a national dataset from the U.S. Department of Education National Center for Education Statistics (NCES, 2014), we developed eight operationalizations of undermatching by altering three commonly used variables. We then compared the number and demographics of students who were identified as undermatched. Differences in operationalizations resulted in significant differences in undermatching by gender, race, parental education, and socioeconomic status. Results of this study illustrate the importance of the need to operationalize terminology used in assessment carefully and consistently.

AUTHORS

Ann M. Gansemer-Topf
Iowa State University

Jillian Downey
Truman State University

Ulrike Genschel
Iowa State University

Definitions Matter: Investigating and Comparing Different Operalionalizations of Academic Undermatching

Effective assessment practice requires clearly defining and operationalizing terminology, but assessment professionals often need to create their own definitions of student populations. For example, research investigating science, technology, engineering, and math (STEM) fields varies in who is included as a STEM major, with some studies including social science majors such as psychology; others limit the definition to hard sciences such as biology, chemistry, or engineering. First-generation students may be defined as those who have no college experience, those who have at least one parent without a college degree, or those who have no parents with a college degree. When these varying definitions are the subject of research studies the results may vary. Toutkoushian and Stollberg (2015) found that varying the definition of first generation altered the number of students who were identified as such—subsequently affecting policies and practices aimed at improving student success for this population of students.

Therefore, research that investigates how the operationalizations of variables may influence assessment results and implications is critical. This study focuses on a specific population—academically undermatched students—to highlight an often overlooked but essential assessment practice: clearly defining the terminology and methods. Academic "undermatching"—when students enroll in colleges that are less academically selective than those for which they are academically prepared—has been viewed as an impediment to degree attainment (Bowen, Chingos, & McPherson, 2009; Executive Office of the President, 2014). Researchers have operationalized undermatching in a variety of ways using a variety of datasets (e.g., Belasco & Trivette, 2015; Bowen et al., 2009; Heil, Reisel, & Attewell, 2014; Rodriquez, 2013; Smith, Pender, & Howell, 2013). Results of these studies have varied: Roderick, Coca, and Nagaoka (2011) found that approximately 62% of college-going students were likely to undermatch; Bowen et al. (2009) and Smith et al. (2013) concluded that 40% were likely to undermatch; and Belasco and Trivette calculated that about 28% were likely to undermatch. Each study was based on a different population of students.

CORRESPONDENCE

Email
anngt@iastate.edu

Rodriquez (2015) utilized one population of students and compared three approaches (acceptance rate, enrollment rate, and predicted rate) to undermatch and found that the percentages and characteristics of students defined as undermatched varied among the three approaches. Our study narrowed this variability further. We wanted to investigate if there were differences in undermatch when we used the same population and same approach but altered the variables within this approach. Our intent was to examine how small differences in operationalizations may change who is identified as undermatched. Given the importance for assessment professionals to clearly define their populations, our goal was to undertake research on a topic of national importance (i.e., academic undermatching) as a way to illustrate how variations in operationalizations might affect our assessment results and implications of these results. Using one national dataset, and operationalizing undermatching in eight different ways by altering similar variables, we sought to answer two research questions:

1. How consistent were different operationalizations in their ability to define students as undermatched?
2. In comparing different operationalizations of undermatching, were there differences in the demographic characteristics of students (gender, race/ethnicity, parental education, and income) for those classified as undermatched?

Assessment and Undermatch Research

Assessment is a valuable process that can guide institutional change, improvement, and strategic planning (Bresciani, Gardner, & Hickmott, 2012; Middaugh, 2011; Schuh, Biddix, Dean, & Kinzie, 2016) but this process requires developing goals that are clear, measurable, and meaningful (Banta, Jones, & Black, 2009; Bresciani et al., 2012; Suskie, 2010). Although developing clear and measurable goals is a consistent theme throughout the literature, less emphasis is placed on the importance of clarifying the terms and definitions within these outcomes or identifying the population that is being assessed. This lack of clarification in defining data and populations can lead to inconsistent data collection processes, measures, and interpretations (McLaughlin & Howard, 2004) that can undermine institutional efforts to effectively improve, change, or plan. Our study wanted to illustrate this point by focusing on academic undermatching.

In *Crossing the Finish Line*, Bowen et al.'s (2009) national study highlighted the negative relationship between undermatching and degree attainment. Holding all academic and demographic variables constant, students attending higher selective institutions were more likely to graduate than students at less selective institutions. Undermatched students are students who attend less selective schools; therefore, they are less likely to graduate.

Given the importance for assessment professionals to clearly define their populations, our goal was to undertake research on a topic of national importance (i.e., academic undermatching) as a way to illustrate how variations in operationalizations might affect our assessment results and implications of these results.

With the national push to raise completion rates for all students (ACE, 2013) and undermatching being viewed as an obstacle for degree completion (Executive Office of the President, 2014), a significant amount of attention has been focused on minimizing undermatch (Bastedo & Jaquette, 2011; Bowen et al., 2009; Hoxby & Avery, 2012; Roderick et al., 2008). Research on this topic has investigated if certain subpopulations are more likely to undermatch than others; results have been mixed. Rodriquez (2013) found that Latino students were more likely to undermatch than their White peers, and that low-income, first-generation students were also more likely to undermatch than students from middle- or high-income families with parents who had more than a high-school education. Bowen et al. (2009) found that African-American students were more likely to undermatch; Belasco and Trivette (2015) found that Latino and African-American students were less likely to undermatch. Belasco and Trivette also noted that females were more likely to undermatch than males, contradicting Smith et al.'s (2013) findings. This study attempted to narrow the variability among past studies by comparing how seemingly minor changes to the operationalizations alter who is classified as undermatched. We examined academic undermatching because, despite the significant amount of national attention focused on implications of undermatch for degree completion, the term undermatch remains difficult to define. Therefore, we determined that this topic would be an excellent example of how definitions matter. By engaging in this process, we hoped to reiterate the need for assessment professionals to engage in definitional rigor and clarity.

Methods

This quantitative study utilized the U.S. Department of Education's National Center for Education Statistics Education Longitudinal Study of 2002 (ELS: 2002; NCES, 2014) and the Barron's Selectivity Ratings (Barron's Educational Series, 2009). The ELS: 2002 captures students' demographics, high-school academic data, financial aid and college choice information (schools applied to and accepted at), the higher-education institution where the students enrolled, and the NCES selectivity classification of that institution. This dataset has been used in past studies of undermatching; (i.e., Belasco & Trivette, 2015; Rodriquez, 2013; Smith et al., 2013). Applying contrasting operationalizations to a dataset that had been used for undermatching provided us the opportunity to view if these differences changed who was defined as undermatched. We captured Barron's selectivity rating by merging that dataset with ELS.

Sample

The ELS data contained a nationally representative sample (N=11,840[1]). We used the panel sampling weights provided by NCES. Students with missing data were deleted, resulting in a sample of 8,020 students. Subsequent analysis demonstrated that this sample was not significantly different from the larger sample, and the sample size was sufficient enough to complete our analysis.

Undermatch Operationalization

In developing our operationalizations to answer Research Question 1, we reviewed previous literature that had statistically defined undermatching. After examining these multiple approaches, we chose to utilize the "eligibility frontiers" with the variables used by Bowen et al. (2009) and Belasco and Trivette (2015) to determine a student's access level. Both studies classified a student as undermatched if the selectivity level of school the student attended was less than the selectivity level of school for which the student had access. Bowen et al. (2009) and Belasco and Trivette (2015) created eligibility frontiers that utilized categorized information on high-school GPA and standardized test scores (i.e., SAT or ACT). To create an eligibility frontier, we considered only those students who applied to schools in the highest selectivity level. For each GPA and SAT score combination, a proportion was calculated indicating how likely it was for students that fall into that particular combination to be accepted to schools at the highest selectivity level. If this proportion was larger than a preselected threshold value, all students that fell into that particular category were deemed to have access to the highest-selectivity-level school. If the proportion was less than the threshold, the procedure was repeated for the next highest level of selectivity and continued until the highest level of access was determined for each GPA and SAT score combination. We chose to use the eligibility frontier approach to determine access because it allowed us to easily examine how changing minor pieces of the definition may influence undermatching. Additionally, this approach utilized GPA and SAT scores, two widely reported student characteristics.

Including more student-level variables may more accurately predict undermatch, but for our purpose we wanted to use a more parsimonious definition in order to examine how small changes in these few variables may change whether a student is defined as undermatched.

Other definitions of undermatching have used other high-school variables such as Advanced Placement credits, number of high-school credits (Rodriquez, 2013; Smith et al., 2013), or high school location (Hoxby & Avery, 2012). Including more student-level variables may more accurately predict undermatch, but for our purpose we wanted to use a more parsimonious definition in order to examine how small changes in these few variables may change whether a student is defined as undermatched.

Data Analysis

We converted all standardized test scores (i.e., ACT or SAT) into SAT scores. We then modified the following three factors (school selectivity classification, GPA and SAT categorization, and calculation of access probability) to examine if students were consistently identified as undermatched.

[1]rounded to the nearest 10s by publication requirement of IES

Selectivity classification. This study included two measures of institutional selectivity: Barron's classification (Barron's Educational Series, 2009) and the selectivity variable found in NCES datasets. Barron's selectivity levels range from 0–6 and are based on high-school GPA, high-school rank, ACT/SAT scores, and acceptance rates. The NCES variable is used in national datasets. Ratings are 0–5 and based on admission policy (i.e., open or not), the number of applicants, number of students admitted, and the 25th and 75th percentiles of ACT/SAT scores.

High-school GPA categorization. One GPA categorization began at 2.0 with 0.3 point increases and was chosen because this was the categorization used by Belasco and Trivette (2015). The second GPA categorization began at 1.0 with 0.5 point increases and was chosen because these were the cutoffs used in other NCES datasets. We categorized SAT scores similar to Belasco and Trivette and Bowen et al. (2009) but did not want to significantly increase the number of operationalizations to compare; thus, we did not vary the SAT categorization cutoffs.

To answer Research Question 2, we limited our sample to only those classified as undermatched and then examined if gender, race/ethnicity, parental education, and socioeconomic status were affected similarly across operationalizations for those defined as undermatched.

Calculation of access probability. We calculated the access probabilities in two ways. The first calculation used all applications. For example, suppose we are considering the highest level of school selectivity. For a given GPA and SAT combination, the access probability was calculated by dividing the total number of acceptances by the total number of applications for all students in the GPA and SAT combination of interest. The second calculation of access probability aggregated over students (Belasco & Trivette, 2015): for students in a given GPA and SAT combination, the access probability was calculated by taking the total number of students that were accepted to at least one highest-selectivity-level school divided by the total number of students that applied to at least one highest-selectivity-level school. It is important to note that the first calculation of access probability used all applications but did not take into account the dependence of multiple observations from one student. In contrast, the second calculation of access probability aggregated all applications and acceptances over a student, thus not taking into account the total number of applications and acceptances for each student.

Regardless of the selectivity-level classification, high-school GPA categorization, or method used to calculate access probability, if the access probability for a given GPA and SAT combination and selectivity level was greater than or equal to 90% based on 10 or more observations, a student in that GPA and SAT combination was deemed to have access to that particular school selectivity level. If there were fewer than 10 observations, no conclusions were reached for the particular school selectivity level.

We obtained eight different operationalizations (O1, O2…O8) as a result of two levels for each of the three factors (see Table 1). For all eight operationalizations, an eligibility frontier was created that we used to categorize the level of school a student had access to.

Table 1

Description of Eight Operationalizations of Undermatching

Operationalization	Classification of School Selectivity	GPA Categorization	Access Probability Calculation
1	NCES	Start at 2.0, increase by 0.3	All Applications
2	NCES	Start at 2.0, increase by 0.3	Student Aggregate
3	Barron's	Start at 2.0, increase by 0.3	All Applications
4	Barron's	Start at 2.0, increase by 0.3	Student Aggregate
5	NCES	Start at 1.0, increase by 0.5	All Applications
6	NCES	Start at 1.0, increase by 0.5	Student Aggregate
7	Barron's	Start at 1.0, increase by 0.5	All Applications
8	Barron's	Start at 1.0, increase by 0.5	Student Aggregate

We then compared this level of access to the level of school the student first attended. If the level of school the student attended was less than the level of school to which they had access the student was classified as undermatched. To answer Research Question 1, we classified students as undermatched for all eight operationalizations and identified how often these different operationalizations agreed for each student.

Next, we examined gender, race/ethnicity, parental education, and socioeconomic status across all eight operationalizations by calculating the percentage identified as undermatch for all categories in each demographic variable. For example, if there were 4,230 females in the sample population, we examined what percentage of females were classified as undermatched using each operationalization. We then conducted a Pearson's chi-square test of independence (Agresti, 2012) to determine if there exists an association between each demographic variable and undermatching. A Pearson's chi-square test is used to establish if the outcomes of one variable are related to the outcomes of a second variable. For example, we conducted a Pearson's chi-square test for gender and operationalization 1, which told us if gender and being undermatched using operationalization 1 were associated. Comparing the outcomes of the chi-square tests across all eight operationalizations allowed us to determine if being undermatched was related to gender for all definitions or just a select few. To account for the multiple comparisons, we implemented a Bonferroni adjustment (Oehlert, 2000) at the individual variable level resulting in a level of significance of $\alpha/n = 0.05/8 = 0.00625$.

All operationalizations showed a statistically significant relationship between being undermatched and race/ethnicity as well as socio-economic status, meaning there was inconsistency across operationalizations.

To answer Research Question 2, we limited our sample to only those classified as undermatched and then examined if gender, race/ethnicity, parental education, and socioeconomic status were affected similarly across operationalizations for those defined as undermatched. We again calculated sample proportions to investigate if the demographic characteristics were similar across operationalizations.

Results

Tables 2 and 3 show comparison of the eligibility frontiers based on O1–O4 and O5–O8. For readability purposes, we chose only to illustrate four operationalizations per table. One cell represents a given GPA and SAT categorization. Each cell is split into four quadrants. The numbers in each quadrant represent the operationalization (i.e., O1 is in the upper left quadrant in Table 2). The colors of each quadrant represent the school selectivity level a student had access to with darker colors corresponding to higher selectivity level schools. For example, students with a GPA between 2.3 and 2.6 and an SAT score between 1200 and 1290 using O1 had access to at most a Level 2 selectivity school. Using O2, students had access to a Level 3 school; O3 students had access to a Level 5 school and O4 students had access to a Level 4 school. One might assume that as SAT and GPA increase, the level of access also should increase. However, this lack of monotonicity (Johnson & Wichern, 2007) was present in all eight eligibility frontiers under consideration. For O1–O4, aside from Level 1 selectivity schools (indicated by blank cells), there is only one GPA and SAT combination (GPA from 3.2 to 3.5 and SAT between 1100 and 1190) for which all four operationalizations resulted in the same level of selectivity access. In comparing O5–O8, other than Level 1 selectivity schools, there are no GPA and SAT combinations that resulted in the same level of access (see Table 3).

We then examined how consistently students were identified as undermatched for the eight definitions (see Table 4). Of the 8,020 students, the proportions of classified students varied by definitions between 5.1% classified by one out of eight definitions, 8.6% by two out of eight, and 8.7% classified by all eight definitions. In the sample, 4,360 (54.3%) were classified as not undermatched by all eight definitions; 3,660 (45.7%) were classified as undermatched by at least one definition; 1,650 (20.6%) students were classified as undermatched by at least five definitions. Of the students classified as undermatched by at least one definition (n=3,660), 700 (19.1%) were consistently classified as undermatched using all eight definitions.

Likewise, we examined the sample proportions of undermatched students within categories of each demographic variable (e.g., male, females; see Table 5). Using O1, 15.2% of females would be considered undermatched compared to 39.4% if using O4. A similar pattern was found for men, with 15.1% of males classified as undermatched using O1 and 37.8% using O4. Operationalization 4—which used Barron's Selectivity Rating, calculated

Table 2

Comparison of Eligibility Frontiers for Operationalizations 1 through 4

SAT

GPA	<800	800-890	900-990	1000-1090	1100-1190	1200-1290	1300-1390	≥1400
<2.0	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
(2.0,2.3)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
(2.3,2.6)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
(2.6,2.9)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
(2.9,3.2)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
(3.2,3.5)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
(3.5,3.8)	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
≥3.8	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

Note: Number indicates operationalization. Colors indicate the level of selectivity to which a student has access. Darker colors indicate that the student has access to more highly selective school.

KEY

	Access to level 1
	Access to level 2
	Access to level 3
	Access to level 4
	Access to level 5

access probabilities by aggregating over students, and started with GPA at 2.0 and increased by .3—resulted in the highest sample proportion of students being classified as undermatched in all categories. Operationalization 5 had the lowest proportion of students defined as undermatched for females (14.9%), males (15.1%), African American (5.7%), Asian (14.2%), Biracial (15%), parents with some college (17.3%), parents with a college degree (12.8%), and socioeconomic status in the low- (16.3%) and middle-high income (10.7%). Operationalization 1 had the lowest proportion of students defined as undermatched for White (16.6%), Hispanic (16.6%), and parents with no college (16.8%).

A higher proportion of females were classified as undermatched as compared to males except when O5 was used. Results for race/ethnicity were mixed. Whites had the highest proportion of students defined as undermatched except for O1 and O5 when Hispanics had the highest proportion. Students identified as African American had the lowest proportion identified as undermatched and Pacific Islander the second lowest. When using O2, O5, and O6 a higher percentage of students whose parents had no college were classified as undermatched. For O1, O3, O4, O7, and O8 a higher percentage of students with parents who had some college were classified as undermatched. In comparing socioeconomic status, O1 and O5 had the highest proportion of low-income students whereas the other operationalizations had the highest proportion of middle-low income students. college were classified as undermatched. In comparing socioeconomic status, O1 and O5 had the highest proportion of low-income students whereas the other operationalizations had the highest proportion of middle-low income students.

For all definitions, between 43–50% of undermatched students had parents with a bachelor's degree or higher and less than 20% had parents with no college degree. Of those identified as undermatched over 80% were in the low-or middle-low income category, regardless of operationalization.

Table 3

Comparison of Eligibility Frontiers for Operationalizations 5 through 8

SAT

GPA	<800	800-890	900-990	1000-1090	1100-1190	1200-1290	1300-1390	≥1400
<1.0	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8
(1.0,1.5)	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8
(1.5,2.0)	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8
(2.0,2.5)	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8
(2.5,3.0)	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8
(3.0,3.5)	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8
≥3.5	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8	5 6 7 8

Note: Number indicates operationalization. Colors indicate the level of selectivity to which a student has access. Darker colors indicate that the student has access to more highly selective school.

KEY

(white)	Access to level 1
(light grey)	Access to level 2
(grey)	Access to level 3
(dark grey)	Access to level 4
(black)	Access to level 5

Table 4

Agreement of Eight Operationalizations (N=8,020)*

Students Classified As Undermatched	N*	%
Classified as undermatched by all 8	700	8.7
Classified as undermatched by 7/8	230	2.8
Classified as undermatched by 6/8	570	7.0
Classified as undermatched by 5/8	170	2.1
Classified as undermatched by 4/8	670	8.2
Classified as undermatched by 3/8	260	3.2
Classified as undermatched by 2/8	690	8.6
Classified as undermatched by 1/8	410	5.1
Not classified as undermatched by any	4360	54.3

* rounded to the nearest 10s by publication requirement of IES

NCES and Barron's classification systems produced significantly different results. When using NCES classifications, lower percentages of students were classified as undermatched compared to using Barron's. Operationalizations 4 and 8 were relatively similar suggesting that when using Barron's and calculating access probabilities by aggregating over student, only adjusting GPA, similar proportions were obtained.

Table 5

Proportion of Each Demographic Characteristics Defined as Undermatched Based on Eight Operationalizations (N=8,020)

		Operationalization							
	n	1	2	3	4	5	6	7	8
% Defined as undermatched		15.0	30.9	22.9	38.4	14.9	27.9	20.3	36.3
Gender									
Female	4,230	15.17	32.59	24.67	39.39	14.89	29.02	22.12	36.79
Male	3,510	15.13	29.32	20.97	37.78	15.10	27.18	18.43	36.21
Race/Ethnicity									
White	5,000	16.56	35.28	27.03	44.38	16.56	32.43	24.10	41.94
African American	850	6.15	14.32	8.28	16.45	5.68	11.48	6.86	14.91
Asian	780	15.38	24.87	18.59	31.92	14.23	22.69	17.18	30.26
Hispanic	290	16.61	31.83	19.03	32.53	17.65	27.68	16.96	31.49
Biracial	770	15.25	28.81	19.30	35.20	14.99	25.03	16.43	33.51
American Indian	30	12.9	22.58	16.13	29.03	12.9	19.35	16.13	25.81
Pacific Islander	20	8.70	30.43	8.70	34.78	8.70	26.09	8.70	34.78
Parental Education									
No College	1,310	16.77	35.15	23.51	37.83	17.61	30.63	20.75	34.53
Some College,	2,380	18.12	33.98	25.23	41.76	17.33	29.98	22.04	38.86
Bachelor's Degree	4,050	12.88	28.13	21.49	37.11	12.76	26.35	19.39	35.80
Socioeconomic									
Low (< $50,000)	3,230	16.59	32.53	23.22	37.83	16.34	28.93	20.47	34.82
Middle Low (50 – 100)	3,140	15.56	32.81	25.70	41.93	15.69	29.91	22.83	39.89
Middle High (100 – 200)	1,240	11.50	25.99	17.89	34.74	10.69	23.48	16.60	34.25
High (> $200,000)	400	9.45	18.16	14.18	28.36	9.45	17.41	10.95	27.86

Operationalizations 1 and 5 produced lower proportions of students identified as undermatched. These definitions both used NCES classification and calculated access probabilities using all applications and differed only by GPA.

Fewer than half of the students were consistently defined as undermatched for all eight operationalizations, thus illustrating the importance of clearly and formally defining variables.

We conducted chi-square tests for independence between each demographic variable and each operationalization (see Table 6). Statistically significant results were found for each demographic variable although the number of statistically significant results varied across demographic variables. For gender, three operationalizations (2, 3, 7) were statistically significant, suggesting a relationship between gender and being classified as undermatched when using these three operationalizations.

For parental education, O1–O6 were statistically significant while O7 and O8 were not. Thus, for operationalizations one through six parental education is associated with being undermatched, but this association is not present for definitions seven and eight. All operationalizations showed a statistically significant relationship between being undermatched and race/ethnicity as well as socio-economic status, meaning there was inconsistency across operationalizations.

We then limited our analysis to only those students who were classified as undermatched and examined the proportion of students in each demographic category (Table 7). Of those classified as undermatched a higher percentage of students were female for all definitions. The difference between females and males was greatest for O7 (57.5% vs. 39.7%) and least for O5 (52.9% vs. 44.5%). For race/ethnicity, White students were the highest proportion identified as undermatched. Approximately 7–10% of those undermatched identified as Asian or Biracial, 3–4% identified as African American or Hispanic and less than .5% were American Indian or Pacific Islander across all eight operationalizations. For all definitions, between 43–50% of undermatched students had parents with a bachelor's degree or higher and less than 20% had parents with no college degree. Of those identified as undermatched over 80% were in the low- or middle-low income category, regardless of operationalization.

Table 6

Chi-square Test Statistic and p-value for Testing Independence between Characteristic and Undermatch for Eight Operationalizations using a Bonferroni Adjustment Significance Level of α = 0.05/8 = 0.00625

	Op1	Op2	Op3	Op4	Op5	Op6	Op7	Op8
Gender (M & F Only) (n = 7740)	0.00 (0.985)	9.42 (0.002)*	14.63 (0.000)*	2.04 (0.154)	0.05 (0.818)	3.11 (0.078)	15.78 (0.000)*	0.25 (0.614)
Race/Ethnicity (White, African Amer, Asian, Hisp, & Birace) (n = 7680)	61.36 (0.000)*	167.63 (0.000)*	165.97 (0.000)*	268.13 (0.000)*	68.97 (0.000)*	176.33 (0.000)*	151.71 (0.000)*	252.83 (0.000)*
Parental Education (n = 7740)	35.29 (0.000)*	35.88 (0.000)*	12.10 (0.002)*	14.12 (0.001)*	33.11 (0.000)*	14.38 (0.001)*	6.53 (0.032)	8.73 (0.013)
Socio-economic (n = 8000)	28.67 (0.000)*	53.86 (0.000)*	48.94 (0.000)*	41.07 (0.000)*	33.56 (0.000)*	41.97 (0.000)*	44.65 (0.000)*	35.12 (0.000)*

*p < 0.00625

Table 7

Comparing Proportions of Students Defined as Undermatched in each Individual Operationalization with Students Who Were Identified as Undermatched in All Operationalizations

	Operationalization								
	1	2	3	4	5	6	7	8	All
n	1210	2480	1840	3080	1190	2230	1630	2910	700
Gender									
Female	53.28	55.72	56.89	54.09	52.85	54.97	57.49	53.45	57.68
Male	44.07	41.58	40.11	43.02	44.46	42.70	39.74	43.63	40.03
Race/Ethnicity									
White	68.63	71.19	73.57	71.93	69.38	72.52	73.96	71.92	71.74
African American	4.32	4.89	3.81	4.51	4.03	4.34	3.56	4.33	3.30
Asian	9.96	7.84	7.90	8.08	9.31	7.92	8.23	8.10	10.19
Hispanic	3.98	3.72	3.00	3.05	4.28	3.58	3.01	3.12	3.30
Biracial	9.71	8.93	8.07	8.76	9.65	8.59	7.74	8.82	8.46
Parental Education									
No College	18.17	18.55	16.73	16.03	19.30	17.91	16.65	15.48	18.51
Some College	35.77	32.65	32.70	32.22	34.56	31.92	32.19	31.72	33.72
Bachelor's Degree	43.32	46.06	47.47	48.80	43.37	47.81	48.28	49.81	45.34
Socioeconomic									
Low (< $50,000)	44.40	42.38	40.82	39.58	44.21	41.76	40.54	38.55	45.34
Middle Low (50 – 100)	40.50	41.58	43.92	42.67	41.28	41.99	43.98	42.95	41.18
Middle High (100 – 200)	11.78	12.97	12.04	13.92	11.07	12.98	12.59	14.52	10.90
High (> $200,000)	3.15	2.95	3.11	3.70	3.19	3.13	2.70	3.84	2.44

Discussion and Implications for Research and Practice

Fewer than half of the students were consistently defined as undermatched for all eight operationalizations, thus illustrating the importance of clearly and formally defining variables. In this section we will highlight our key findings and discuss how these findings can inform and improve assessment work.

Methods Influence Definitions

In past studies the percentage of students identified as undermatched varied from 28% to 62% and their demographic breakdowns differed (Belasco & Trivette, 2015; Bowen et al., 2009; Rodriquez, 2013; Smith et al., 2013). These variations are likely the result of studying different populations of students and applying different techniques. Our study illustrates that even using the same dataset and techniques but slightly changing variables can result in significant differences in the percentage and characteristics of students identified as undermatched. In other words, methods matter. This finding has implications for student learning assessment work. Competency in a certain discipline can be measured through various methods: standardized tests, comprehensive exams, or portfolios. The percentage of students who pass and the measure of student learning can vary based on assessment given. Therefore, when determining which student learning assessment to administer, it is important to consider the consequences of each of the methods.

Definitions Influence Subpopulations and Interpretations

Different operationalizations can tell different stories about subpopulations of students. For each category of students, the range of who is defined as undermatched varies significantly. Students whose parents have no college are defined as undermatched at the highest proportions for O2, O5, O6; students whose parents have some college have the highest proportions for O1, O3, O4, O7, and O8. The proportion of African Americans identified as undermatched ranges from 5.7% to 16.5%: three times as many African Americans were identified as undermatched using O4 compared to O1.

The study also illustrates the importance of recognizing limitations of the variables used in operationalizations and the consequences of these limitations on results and implications.

Our study also illustrates how the population of students can influence interpretations of results. In examining undermatch, the results and subsequent conclusions differ when comparing the demographics of undermatched students based on the total student population (Table 5) to demographics of undermatched students based on only those defined as undermatched (Table 7). For example, when examining O3 using the total student population (Table 5) similar proportions of students with parents with no college (23.5%) are as likely to be undermatched as students whose parents have some college (25.2%) and a Bachelor's degree (21.5%). Using the same operationalization but examining those students who are undermatched, almost half of the population (47.5%) of the students have parents with Bachelor's degrees versus 16.7% whose parents have no college (Table 7). The former results could be interpreted that parental education level is not related to undermatching whereas the latter may suggest that undermatching is more common for students whose parents have a college degree.

These variations in populations and subpopulations are similar to challenges faced in monitoring and reporting STEM results. Some definitions of STEM include majors such as psychology, which significantly increases the number of individuals in STEM, as well as the percentage of women and underrepresented students in STEM. Women are considered underrepresented in STEM but in some majors (e.g., biology) they may be equally or overrepresented. Additionally, whereas Asian Americans may be considered an underrepresented minority group within the college student population, they are not considered an underrepresented minority group population within STEM (NACME, n.d.). It is therefore critical that assessment professionals determine and delineate the populations for which they are reporting.

Recognize Limitations

The study also illustrates the importance of recognizing limitations of the variables used in operationalizations and the consequences of these limitations on results and implications. Because our study calculated undermatch based only on those students who had standardized test scores any student lacking this information was not included—potentially eliminating a significant number of undermatched students. For example, a student who is not considering college, or considering a college that does not require standardized test scores, may choose not to take a standardized exam. This student may be undermatched but because appropriate data to determine this undermatching was unavailable this student was not included in this study.

This too mirrors assessment practice. Institutions provide retention and graduation rates but these are often based on a cohort of full-time, direct-from-high-school students who begin in the fall. This restriction omits transfer students or students who begin part time. Most surveys provide students two choices for gender, overlooking students who identify as transgender. Low-income students may be defined as Pell-eligible while ignoring those students whose families may make less than $50,000 but did not receive a Pell Grant. Good assessment practice requires examining who may or may not be included and the implications of these decisions.

There is No "Perfect": Strive for Clarity and Consistency

Effective assessment practice requires clear and consistent definitions but many times assessment professionals examine student populations and outcomes that lack this clarity and consistency.

Changes in operationalizations produced varied results and also illustrate that no one definition is perfect. Our results make it difficult to identify the "best," "most valid," or "most reliable" operationalization. Nevertheless, the results provide insights into consequences of different decisions. Using all applications (versus student aggregate information) results in lower proportions of students classified as undermatched because the access probability was always smaller than when calculated aggregating over a student. A higher percentage of students were classified as undermatched when Barron's selectivity classification was used. Using NCES classification of data found within NCES-sponsored restricted datasets may be easier to use but because there are fewer selectivity categories it may also decrease the proportion of students identified as undermatched.

For researchers interested in a broad definition of undermatch, using Barron's classification and calculating the student aggregate provides the greatest likelihood of being defined as undermatched. Researchers wanting to be most consistent may include only those students who were defined as undermatched for each operationalization, recognizing that this approach also minimizes the sample size. Statistically speaking, the "most valid" may be O5 because it has the highest degree of monotonicity (i.e., as GPA and SAT scores increased so did the likelihood of being admitted into a more highly selective institution). There is not one approach but many. Decisions on which operationalization to use must be made within the context of the research study, its purpose, research questions, and potential implications.

Choosing definitions and providing rationale for these decisions is needed in assessment practice. Institutions differently define categories of students "at-risk" or "under represented" and then assess their success. The definitions of the student population (i.e., at-risk) and the definition of success (e.g., retention, GPA, graduation) can lead to different results, so it is necessary that the definitions be clearly articulated and used consistently.

With so many potential choices and approaches it is important to heed the advice from Schuh and Upcraft (2001) who remind us that no assessment is perfect but one can strive for "good enough." There may not exist a universal definition for many of the topics we want to assess and we may not achieve complete accuracy (Suskie, 2009). However, we can work toward a good enough definition—one for which there is a strong rationale, one that can most effectively address the assessment questions, and most critically, one that can assist us in achieving our higher-education missions and goals.

Conclusion

Effective assessment practice requires clear and consistent definitions but many times assessment professionals examine student populations and outcomes that lack this clarity and consistency. Assessment professionals create definitions for the concepts they are examining but the consequences of these definitions may often be overlooked or not understood. Using academic undermatching as an example, we created eight unique operationalizations of undermatch that subsequently led to different results and conclusions. This study contributes to effective assessment practice by reinforcing the importance and implications of clearly defining student populations, terms, and variables.

References

Agresti, A. (2012). *Categorical data analysis* (3rd. ed.). Hoboken, NJ: Wiley.

American Council on Education (ACE). (2013). *National Commission on Higher Education Attainment*. Retrieved from http://www.acenet.edu/news-room/Pages/National-Commission-on-Higher-Education-Attainment.aspx

Banta, T.W, Jones, E.A, & Black, K. (2009) *Designing effective assessment: Principles and profiles of good practice*. San Francisco, CA: Jossey-Bass.

Barron's Educational Series. (2009). *Barron's Profiles of American Colleges of 2009*. Hauppauge, NY: Barron's Educational Series.

Bastedo, M.M., & Flaster, A. (2014). Conceptual and methodological problems in research on college undermatch. *Educational Researcher, 43*(2), 93–99. doi:0013189X1452303

Bastedo, M. N., & Jaquette, O. (2011). Running in place: Low-income students and the dynamics of higher education stratification. *Educational Evaluation and Policy Analysis, 33*(3), 318–339.

Belasco, A. S., & Trivette, M. J. (2015). Aiming low: Estimating the scope and predictors of postsecondary undermatch. *The Journal of Higher Education, 86*(2), 233–263.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton, NJ: Princeton University Press.

Bresciani, M. J., Gardner, M. M., & Hickmott, J. (2012). *Demonstrating student success: A practical guide to outcomes-based assessment of learning and development in student affairs*. Sterling, VA: Stylus Publishing.

Executive Office of the President. (2014). *Increasing college opportunity for low-income students: Promising models and a call to action*. Retrieved from https://www.google.com/#q=increasing+opportunity+for+low+income+students+US

Fosnacht, K. (2014). *Selectivity and the college experience: How undermatching shape the college experience among high-achieving students*. Paper presented at the American Educational Research Association, Philadelphia, PA.

Fosnacht, K. (2015). *Undermatching and the first-year experience: Examining effect heterogeneity*. Paper presented at the Association for the Student of Higher Education, Denver, CO.

Heil, S., Reisel, L., & Attewell, P. (2014). College selectivity and degree completion. *American Educational Research Journal, 51*(4) 1–23. doi: 10.3102/0002831214544298

Hoxby, C. M., & Avery, C. (2012). *The missing" one-offs": The hidden supply of high-achieving, low income students* (No. w18586). National Bureau of Economic Research.

Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Englewood Cliffs, NJ: Prentice Hall.

McLauglin, G. W., & Howard, R. D. (2004). *People, processes, and managing data* (2nd ed.). Tallahassee, FL: Association of Institutional Research.

Middaugh, M. F. (2011). *Planning and assessment in higher education: Demonstrating institutional effectiveness*. San Francisco, CA: Jossey-Bass.

National Action Council for Minorities in Engineering (n.d). *Underrepresented minorities in STEM*. Retrieved from http://www.nacme.org/underrepresented-minorities

National Science Foundation (2017). *Women, minorities, and persons with disabilities in science and engineering*. Retrieved from https://www.nsf.gov/statistics/2017/nsf17310/data.cfm

Oehlert, G. W. (2000). *A first course in design and analysis of experiments*. New York, NY: W.H. Freeman.

Roderick, M., Nagaoka, J., Coca, V., Moeller, E., Roddie, K., Gilliam, J., & Patton, D. (2008). *From high school to the future: Potholes on the road to College*. Retrieved from http://ccsr.uchicago.edu/publications/high-school-future-potholes-road-college

Roderick, M., Coca, V., & Nagaoka, J. (2011). Potholes on the road to college high school effects in shaping urban students' participation in college application, four-year college enrollment, and college match. *Sociology of Education, 84*(3), 178–211.

Rodriguez, A. (2013). *Unpacking the black box: Estimating the high school-level effects of undermatching among under-represented students* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (No. AAI3594848)

Rodriguez, A. (2015). Tradeoffs and limitations: Understanding the estimation of college undermatch. *Research in Higher Education, 56*(6), 566–594.

Schuh, J. H., Biddix, J. P., Dean, L. A., & Kinzie, J. (2016). *Assessment in student affairs*. San Francisco, CA: Jossey-Bass.

Schuh, J. H., & Upcraft, M.L. (2001). *Assessment practice in student affairs: An applications manual*. San Francisco, CA: Jossey-Bass.

Smith, J.I., Pender, M., & Howell, J. (2013). The full extent of academic undermatch. *Economics of Education Review, 32*, 247–261.

Suskie, L. (2010). *Assessing student learning: A common sense guide*. San Francisco, CA: Jossey-Bass.

Toutkoushian, R., & Stollberg, R. (2015). *Talking 'bout my generation: Defining 'First-Generation Students" in higher education research*. Paper presented at Association of Institutional Research, Denver, CO.

U.S. Department of Education, National Center for Education Statistics (NCES). (2014). *Education Longitudinal Study of 2002* (ELS: 2002). (Data file). Restricted data license.