

# Chapter 14

## Comparison of Published Assessments of Biological Experimentation as Mapped to the ACE-Bio Competence Areas



Anna J. Zelaya, Lawrence S. Blumer, and Christopher W. Beck

### 14.1 Introduction

Experimentation is fundamental to our work as life scientists. It is the core source of new knowledge in the life sciences and experimentation incorporates skills found in any list of undergraduate biology learning outcomes (American Association for the Advancement of Science, 2011; Clemmons et al., 2020). During the past two decades, increased focus on evidence-based learning and teaching has put increased emphasis on learning science by doing science, which means experimentation (American Association for the Advancement of Science, 2011; Boyer, 1998; National Research Council (NRC), 2003; Project Kaleidoscope (PKAL), 2002). Consequently, learning experimentation and assessing the effectiveness of teaching experimentation is essential for undergraduate life sciences education to gauge what students actually learn. Yet, the effectiveness of curricula in teaching experimentation is rarely assessed in courses, such as laboratory courses (Beck et al., 2014). Furthermore, even when experimentation is assessed, published assessment tools are not often used (Beck et al., 2014). Using published assessments improves our understanding of student learning of experimentation, as these assessments generally have been validated. In addition, when multiple studies use the same assessment, comparisons of approaches to teaching experimentation can be compared explicitly.

---

A. J. Zelaya · C. W. Beck

Department of Biology, Emory University, Atlanta, GA, USA  
e-mail: [anna.zelaya@emory.edu](mailto:anna.zelaya@emory.edu); [christopher.beck@emory.edu](mailto:christopher.beck@emory.edu)

L. S. Blumer (✉)

Department of Biology, Morehouse College, Atlanta, GA, USA  
e-mail: [lawrence.blumer@morehouse.edu](mailto:lawrence.blumer@morehouse.edu)

Once assessments are identified, they need to match learning outcomes. This provides evidence of what students know and can do as well as provides timely feedback to students during a course (Handelsman et al., 2007). However, identifying existing assessment tools that match the learning outcomes sought by an instructor might be the greatest barrier to genuine evidence-based teaching of biological experimentation. Current published assessments commonly used in biology courses represent a diverse array of tools ranging from descriptions of learning activities, self-reporting of student opinions, to multiple choice problem solving, and free response to prompts and assessment rubrics (Shortlidge & Brownell, 2016). This review is an attempt to align the commonly used assessments with defined competencies in biological experimentation.

The Basic Competencies of Biological Experimentation developed by the ACE-Bio Network (Pelaez et al., 2017; Chap. 1 in this volume) are a valuable starting point for biology educators to identify core competencies and assess the achievement of those outcomes in students. The network identified seven basic Competence Areas (*Identify, Question, Plan, Conduct, Analyze, Conclude, and Communicate*) that are components of experimentation. Each Competence Area contains two to ten Concepts and each concept contains one to nine Skill Statements. This framework of basic competencies in biological experimentation overlaps with some of the course-level learning outcomes of the BioSkills guide (Clemmons et al., 2020) that are based on the core competencies in the *Vision and Change* report (AAAS, 2011). However, the ACE-Bio framework is more detailed in elaborating Competence Areas, Concepts, and Skill Statements that describe biological experimentation. Here, we used this framework to categorize individual assessment items from assessments that address aspects of experimentation currently used in undergraduate biology courses. Mapping of assessments on this framework will allow instructors to better understand what is actually being assessed and education researchers to identify gaps in our arsenal of assessments related to experimentation.

## 14.2 Methods

We surveyed assessments of different aspects of experimentation currently used in undergraduate biology courses and categorized the assessment items using the framework of the Basic Competencies of Biological Experimentation (Pelaez et al., 2017; Chap. 1 in this volume). We limited our review to assessments that are freely available and documented in the biology education literature, starting with those suggested by Shortlidge and Brownell (2016) for the assessment of course-based undergraduate research experiences. We supplemented those references with additional assessments related to biological experimentation that have been published, including those collected by ACE-Bio participants in 2014. Our goal was not to include all possible assessments of biological experimentation, but to include a range of possible assessments that are used in biology courses. The complete list of assessments surveyed can be found in Table 14.1. In some cases, the references are

**Table 14.1** Assessments reviewed in this study categorized by type, student class level, and instrument availability

Assessment	Number of items surveyed	Type	Student class level	Availability	References
Instructional Practices Survey	24	Description of class activities – Survey	Non-majors; Introductory majors; Upper-level majors	Main article or supplementary materials	Beck and Blumer (2016)
LCAS	17	Description of class activities – Survey	Introductory majors; Upper-level majors	Main article or supplementary materials	Corwin et al. (2015)
BEDCI	14	Measurement of learning – Multiple choice	Introductory majors; upper-level majors	Download via link provided in main article	Deane et al. (2014)
TOSLS	28	Measurement of learning – Multiple choice	Majors and non-majors	Main article or supplementary materials	Gormally et al. (2012)
Modified Classroom Test of Scientific Reasoning	24	Measurement of learning – Multiple choice	Non-majors; Introductory majors	Main article or supplementary materials	Benford and Lawson (2001)
Molecular Biology Data Analysis Test	20	Measurement of learning – Multiple choice	Introductory level; Upper-level majors	Main article or supplementary materials	Rybarczyk et al. (2014)
SRBCI	12	Measurement of learning – Multiple choice	Introductory majors; Upper-level majors	Download via link provided in main article	Deane et al. (2016)
BioSQuaRE	29	Measurement of learning – Multiple choice	Introductory majors	Online through author	Stanhope et al. (2017)
RED	18	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Dasgupta et al. (2014)
RED Bird Assessment	3	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Dasgupta et al. (2014)
RED Drug Assessment	6	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Dasgupta et al. (2014)

(continued)

**Table 14.1** (continued)

Assessment	Number of items surveyed	Type	Student class level	Availability	References
RED Shrimp Assessment	4	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Dasgupta et al. (2014)
EDAT	10	Measurement of learning – Prompt with rubric	Non-majors	Main article or supplementary materials	Sirum and Humburg (2011)
E-EDAT	10	Measurement of learning – Prompt with rubric	Introductory majors; Upper-level majors	Main article or supplementary materials	Brownell et al. (2013)
TIED	20	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Killpack and Fulmer (2018)
CRBS	20	Measurement of learning – Research assignment with rubric	Non-majors; Introductory majors; Upper-level majors	Main article or supplementary materials	Kishbaugh et al. (2012)
The Rubric for Science Writing	15	Measurement of learning – Research assignment with rubric	Introductory majors; upper-level majors	Main article or supplementary materials	Timmerman et al. (2011)
Quantitative Literacy Rubric	9	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Speth et al. (2010)
BioTAP	15	Measurement of learning – Research assignment with rubric	Upper-level majors	Online through author	Reynolds et al. (2009)
Graph Rubric	12	Measurement of learning – Research assignment with rubric	Non-majors; Introductory majors; Upper-level majors	Main article or supplementary materials	Angra and Gardner (2018)
SPFA	18	Measurement of learning – Prompt with rubric	Not explicitly stated, but included major and non-majors	Main article or supplementary materials	Wilson and Rigakos (2016)

(continued)

**Table 14.1** (continued)

Assessment	Number of items surveyed	Type	Student class level	Availability	References
EDR	4	Measurement of learning – Prompt with rubric	Introductory majors	Main article or supplementary materials	Dirks and Cunningham (2006)
Graphing Quiz	12	Measurement of learning – Short response and multiple choice	Introductory majors	Main article or supplementary materials	Dirks and Cunningham (2006)
TIPS	34	Measurement of learning – Multiple choice	Introductory majors	Main article or supplementary materials	Dirks and Cunningham (2006)
<sup>a</sup> URSSA	45	Self-report of learning – Survey	Not explicitly stated	Download via link provided in main article	Weston and Laursen (2015)
BioVEDA	12	Measurement of learning – Multiple choice	Introductory majors	Main article or supplementary materials	Hicks et al. (2020)
<sup>a</sup> CURE-Survey	46	Self-report of learning – Survey	Not explicitly stated	Online through author	Lopatto (2008)
Self-Efficacy, Science Identity, Science Community	15	Self-report of learning – Survey	Non-majors; Introductory majors; Upper-level majors	Main article or supplementary materials	Estrada et al. (2011)
Experimental Control Exercises	7	Short answer and multiple-choice quiz, no rubric	Upper-level majors	Main article or supplementary materials	Shi et al. (2011)
Pre-Post Test for Analytical Skills	12	Short answer and multiple-choice quiz, no rubric	Non-majors; Introductory majors; Upper-level majors	Main article or supplementary materials	Picone et al. (2007)

The number of individual items surveyed in each assessment instrument is shown. The number of items surveyed in an assessment varied from 3 to 52

<sup>a</sup>Denotes assessment that contained multiple components, some of which were deemed not applicable e.g., related to affect or specific programmatic or demographic information. These sections were excluded, therefore the number of items surveyed may be different from the total number of items in the assessment

for the assessment instruments themselves while in others, the assessments are supplementary and used for a study that examined student competence area in biological experimentation as an outcome measure.

Each assessment instrument was reviewed first to determine whether its items related to the Basic Competencies of Biological Experimentation. Instruments that

do not assess biological experimentation (e.g., assessments of student affect with no items explicitly related to biological experimentation (Chemers et al., 2011; Glynn et al., 2011; Hanauer & Dolan, 2014; Hanauer & Hatfull, 2015; Semsar et al., 2011), student views of the nature of science (Halloun & Hestenes, 1998; Lederman et al., 2002)) were excluded since they do not measure students understanding, skills, or knowledge related to biological experimentation. For assessment instruments that were retained, we categorized each item in one (or more) of the seven Basic Competence Areas or “None of the above”. Furthermore, we identified the Concepts and Skill Statements that are being assessed, when possible. To deal with the fact that assessment items might not map to specific Concepts and Skill Statements, we added an “Other” category within each of the seven Basic Competence Areas to represent additional Concepts and also within each of the subsidiary Concepts to represent additional Skill Statements.

To align our codings of assessments using the Basic Competencies of Biological Experimentation framework, all three authors coded items from three assessments (Corwin et al., 2015; Gormally et al., 2012; Sirum & Humburg, 2011) that included the range of types of assessments in our dataset (see Table 14.1). Based on discussion of preliminary coding, we agreed to code in a hierarchical fashion such that we first determined whether an assessment item fit one or more Basic Competence Areas, then whether it fit one or more Concepts within those Competence Areas, and finally whether it fit one or more Skill Statements within those Concepts. The remainder of the instruments were coded by two of the three authors, with each author coding approximately two-thirds of the instruments. When coders disagreed in their coding of a particular item at the level of the Basic Competencies, the coders discussed the item to determine a consensus coding. We included differences among coders at the level of Concepts and Skill Statements as they were reflective of the ambiguity in coding many of the items in the assessments at these levels.

## 14.3 Results and Discussion

### 14.3.1 *Instruments for Assessing Competence Areas in Biological Experimentation*

The majority of assessments included in our study aimed to measure learning via multiple choice assignments or short answer writing prompts (with or without a rubric), while three were survey type assessments that measured affect or self-reported learning gains with some items explicitly related to biological experimentation (Table 14.1). The LCAS (Corwin et al., 2015) and the instructional practices survey (Beck & Blumer, 2016) explore student perceptions on the types of activities they performed in class. Many of the assessments have been used with students in both introductory and upper-level courses for biology majors, suggesting that they can be used to assess aspects of experimentation in a wide range of students.

### 14.3.2 *Mapping Assessments to Competence Areas*

The assessments that we mapped varied considerably in the number of Competence Areas covered by the assessment, ranging from one to all seven Competence Areas (Table 14.2). The URSSA (Weston & Laursen, 2015), CURE-Survey (Lopatto, 2008), CRBS (Kishbaugh et al., 2012), and Rubric for Science Writing (Timmerman et al., 2011) assess all 7 competencies. The URSSA (Weston & Laursen, 2015) and the CURE-Survey (Lopatto, 2008) are student self-reports and are designed for programmatic assessment by considering a large number of areas. In contrast, CRBS (Kishbaugh et al., 2012) and Rubric for Science Writing (Timmerman et al., 2011) are rubric banks or rubrics that instructors can use for assessing a broad range of competencies in student products, such as paper, posters, and presentations. At the other end of the spectrum, the assessments that covered the least number of competencies were the Shrimp Assessment of the RED (Dasgupta et al., 2014), which only covered 1 Competence Area, followed by the Modified CTSR (Benford & Lawson, 2001), EDAT (Sirum & Humburg, 2011), SRBCI (Deane et al., 2016), E-EDAT (Brownell et al., 2013), Experimental Control (Shi et al., 2011), TIED (Killpack & Fulmer, 2018), and Graph Rubric (Angra & Gardner, 2018), all of which only covered 2 of the 7 Competence Areas (Table 14.2). In general, these assessments focus on Plan and Conclude, except for the Graphic Rubric (Angra & Gardner, 2018), which focuses on Analyze (Fig. 14.1). It is possible that some assessments, like the CURE-survey (Lopatto, 2008), covered a high percentage of the Competence Areas, because the items tended to be phrased in broad or generic terms (e.g., “Write a research proposal”), which subsequently was coded as having the potential to cover many skills within the framework. Others that had a lower total percent coverage of the competency framework (e.g., E-EDAT (Brownell et al., 2013)), had more narrowly phrased questions that encompassed a specific skill (e.g., “Develop a hypothesis about what causes changes in poppy growth rate”) and subsequently was only categorized into one or fewer of the seven categories.

In most cases, when an assessment was scored as measuring a Competence Area, it considered multiple Concepts within each Competence Area (Table 14.2). Not surprisingly, however, we note a trade-off between the number of Competence Areas covered by an assessment and the proportion of items associated with a particular Competence Area. Assessments that covered a large number of Competence Areas tend to have fewer items associated with a particular Competence Area (Fig. 14.1). In contrast, assessments that focus on one or two Competence Areas had a high proportion of items concentrated in those Competence Areas.

From the perspective of individual Competence Areas, Plan and Conclude are covered by the most assessments (27 and 24 out of 30 instruments, respectively), indicating an emphasis on experimental design skills and drawing inferences from data in current assessments. Identify and Conduct are the least assessed of the Competence Areas (8 and 7 of 30 assessments, respectively) (Fig. 14.1). The nature of the Concepts and Skill Statements within Identify and Conduct might make them particularly difficult to assess. For example, many of the Skill Statements in Conduct

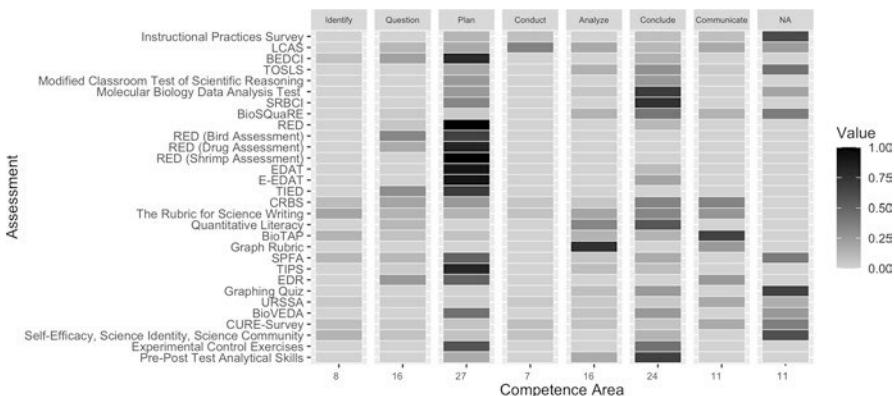
**Table 14.2** Assessment coverage of Competence Areas, Concepts, and Skill Statements

Assessment	Competence Areas	Concepts	Skill Statements	Competence Areas with multiple Concepts	Concepts with multiple Skill Statements
Instructional Practice Survey	4 (57%)	6 (17%)	8 (6%)	2 (29%)	2 (7%)
LCAS	6 (86%)	14 (39%)	20 (14%)	5 (71%)	5 (17%)
BEDCI	4 (57%)	13 (36%)	27 (20%)	3 (43%)	7 (24%)
TOSLS	4 (57%)	13 (36%)	32 (23%)	4 (57%)	8 (28%)
Modified Classroom Test of Scientific Reasoning	2 (29%)	6 (17%)	10 (7%)	2 (29%)	4 (14%)
Molecular Biology Data Analysis Test	3 (43%)	9 (25%)	19 (14%)	2 (29%)	7 (24%)
SRBCI	2 (29%)	5 (14%)	15 (11%)	2 (29%)	4 (14%)
BioSQuaRE	5 (71%)	7 (19%)	12 (9%)	2 (29%)	2 (7%)
RED	3 (43%)	10 (28%)	23 (17%)	2 (29%)	7 (24%)
RED (Bird Assessment)	2 (29%)	6 (17%)	9 (6%)	2 (29%)	3 (10%)
RED (Drug Assessment)	2 (29%)	6 (17%)	14 (10%)	1 (14%)	4 (14%)
RED (Shrimp Assessment)	1 (14%)	5 (14%)	8 (6%)	1 (14%)	2 (7%)
EDAT	2 (29%)	10 (28%)	21 (15%)	1 (14%)	6 (21%)
E-EDAT	2 (29%)	8 (22%)	20 (14%)	1 (14%)	6 (21%)
TIED	2 (29%)	8 (22%)	15 (11%)	2 (29%)	5 (17%)
CRBS	7 (100%)	26 (72%)	45 (33%)	7 (100%)	12 (41%)
The Rubric for Science Writing	7 (100%)	19 (53%)	43 (31%)	7 (100%)	10 (34%)
Quantitative Literacy Rubric	3 (43%)	6 (17%)	17 (12%)	2 (29%)	5 (17%)
BioTAP	6 (86%)	15 (42%)	35 (25%)	6 (86%)	8 (28%)
Graph Rubric	2 (29%)	5 (14%)	9 (7%)	2 (29%)	2 (7%)
SPFA	5 (71%)	22 (61%)	30 (22%)	5 (71%)	6 (21%)
TIPS	3 (43%)	6 (17%)	11 (8%)	1 (14%)	4 (14%)
EDR	2 (29%)	8 (22%)	12 (9%)	2 (29%)	2 (7%)
Graphing Quiz	2 (29%)	3 (8%)	4 (3%)	0 (0%)	2 (7%)
URSSA	7 (100%)	16 (44%)	23 (17%)	5 (71%)	5 (17%)
BioVEDA	3 (43%)	10 (28%)	21 (15%)	2 (29%)	7 (24%)
CURE-Survey	7 (100%)	30 (83%)	56 (41%)	7 (100%)	12 (41%)
Self-Efficacy, Science Identity, Science Community	5 (71%)	6 (17%)	11 (8%)	2 (29%)	3 (10%)

(continued)

**Table 14.2** (continued)

Assessment	Competence Areas	Concepts	Skill Statements	Competence Areas with multiple Concepts	Concepts with multiple Skill Statements
Experimental Control Exercises	2 (29%)	4 (11%)	8 (6%)	2 (29%)	2 (7%)
Pre-Post Test for Analytical Skills	3 (43%)	6 (17%)	12 (9%)	3 (43%)	4 (14%)



**Fig. 14.1 Heatmap of the coverage of Competence Areas by each assessment.** The values are the proportion of items in each assessment that address a given Competence Area. “NA” was assigned to items in an assessment if they could not be categorized in any of the Competence Area. The values at the bottom of each column are the total number of assessment instruments that addressed a given Competence Area

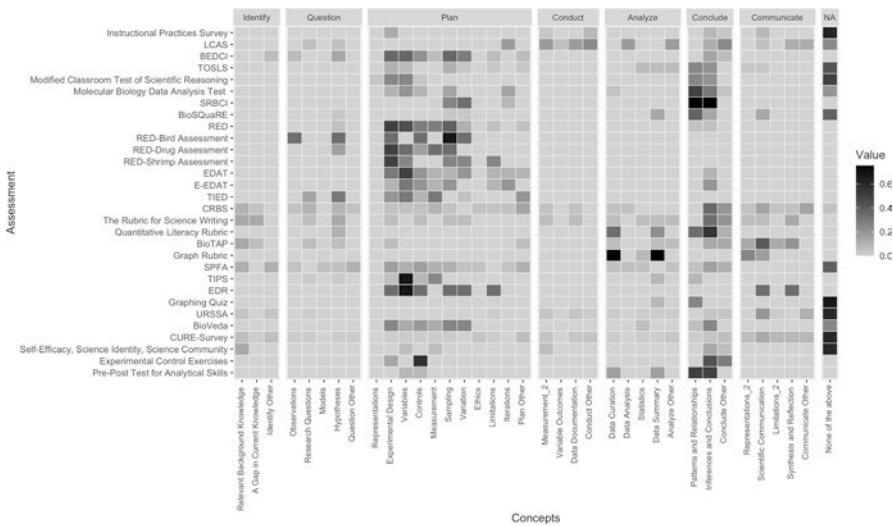
could only be observed by an instructor in a laboratory course or mentored research context.

Some assessments show a high proportion of items that do not fit into the ACE-Bio framework (Fig. 14.1). In some cases, items are related to student affect, student metacognition, faculty assessment practices, computational quantitative literacy, or are too general (Beck & Blumer, 2016; Estrada et al., 2011; Gormally et al., 2012; Lopatto, 2008; Stanhope et al., 2017; Wilson & Rigakos, 2016). Other assessments include items that are not currently considered in the ACE-Bio framework, but perhaps should be included (see below), such as collaboration skills and aspects of statistical literacy (Corwin et al., 2015; Gormally et al., 2012).

### 14.3.3 Mapping Assessments to Concepts

Similar to our mapping of assessments to Competence Areas, assessments were quite variable in the number of Concepts that are considered (Table 14.2). Some assessments focused on very few Concepts (4 or 5 out of 22 for the Experimental Control Exercise (Shi et al., 2011) and the Graph Rubric (Angra & Gardner, 2018), respectively). In contrast, the assessments that cover a broad range of Competence Areas also incorporate a high percentage of Concepts (e.g., CURE-Survey (Lopatto, 2008) and CRBS (Kishbaugh et al., 2012)). For most assessments, only a single Skill Statement was assessed for a particular Concept rather than multiple Skill Statements (Table 14.2). As with the Competence Areas, there is a trade-off between the breadth of an assessment and the proportion of items associated with a particular Concept.

Certain Concepts are well-represented in the assessments we surveyed. Within the *Plan* Competence Area, Concepts of Experimental Design, Variables, Controls, and Sampling have a high frequency of items (Fig. 14.2). The same is true for the Concepts of Data Curation and Data Summary within the *Analyze* Competence Area, and Patterns and Relationships and Inferences and Conclusions within the *Conclude* Competence Area (Fig. 14.2). However, some Concepts are conspicuously absent, even in Competence Areas that are often included. For example, the Concepts Representations and Ethics within the *Plan* Competence Area are



**Fig. 14.2 Heatmap of the coverage of Concepts within each Competence Area by each assessment.** The values are the proportion of items in each assessment instrument that addressed a given Concept. “NA” was assigned to items in an assessment if they could not be categorized in any of the Competence Areas. Within each Competence Area, “Other” tabulates items in an assessment that were categorized in a given Competence Area but did not address any of the specified Concepts in that Competence Area

infrequent even though *Plan* is commonly assessed (Figs. 14.1 and 14.2). Likewise, *Models* is uncommon within *Question* (Figs. 14.1 and 14.2).

Because items clearly fit within a particular Competence Area, but not necessarily within a Concept in that Competence Area, we created an “Other” category for each Competence Area. The frequency of items coded in these “Other” categories, especially within *Conclude* (Fig. 14.2), suggests the potential for expanding the ACE-Bio framework (see below).

#### 14.3.3.1 Gaps in Existing Assessments of Biological Experimentation

None of the assessments we reviewed were developed with the ACE-Bio framework as a guide. Consequently, the match between assessment items and the Competence Areas, Concepts, and Skill Statements is not perfect and is subject to interpretation. We have therefore limited most of our reporting of gaps to the level of Competence Areas, to the most general level of categorization. Among the seven basic Competence Areas, two are not well addressed by the assessments we surveyed, *Identify* and *Conduct* (Fig. 14.1). Fewer than one-half of the assessments include items that were categorized in the *Identify* or *Conduct* Competence Areas, and among the assessments that include items in these Competence Areas, the proportion of assessment items in either Competence Area is small. Similarly, among those assessments that include six or all seven Competence Areas (Table 14.2), few items assess Skill Statements in *Identify* or *Conduct* (Fig. 14.1). One exception is the LCAS (Corwin et al., 2015) for *Conduct*, but this assessment is limited to descriptions of class activities rather than students’ skills and knowledge. Another exception is the Rubric for Science Writing (Timmerman et al., 2011) for *Identify*. In both cases, the percentage of items addressing the Competence Area is 30–40%, but only for one of these two Competence Areas in each case. It is worth noting that the Concepts and Skill Statements in the *Identify* Competence Area (Pelaez et al., 2017; Table 1.3 in Chap. 1 in this volume) are relatively high-order (the ability to identify gaps and limitations in current knowledge) that require experiences uncommon among undergraduates and are infrequently expected learning outcomes in undergraduate courses (Cole & Beck’s Chap. 3 in this volume). The *Conduct* Competence Area may be assessed more readily by using in-class methods than by using assessments reviewed here, such as a laboratory practical, mid-experiment discussions with students, direct observation of students while conducting an experiment, or checking extemporaneous documentation in laboratory notebooks (Moore & Lynn, 2020). Within the other Competence Areas, even those that are well covered by assessment items, there are noticeable gaps. Within the *Plan* Competence Area, the Concepts of Representations, Ethics, and Limitations are not well addressed by any assessments (Fig. 14.2), even though these Concepts are considered important in undergraduate teaching of experimentation (Clemmons et al., 2020; Cole & Beck’s Chap. 3 in this volume; Diaz-Martinez et al., 2019). In *Analyze*, the Concept of Statistics (choosing and conducting the appropriate statistical test

and others) also is not well addressed by assessments, as assessments of statistical literacy like SRBCI (Deane et al., 2016) focus on *Conclude* (Fig. 14.2).

#### 14.3.4 *Gaps in ACE-Bio Framework of Competence Areas*

One of the most striking findings in our analysis is the frequency of assessment items that do not fit nicely in one of the ACE-Bio Competence Areas, and the number of assessment items that we categorized in a given Competence Area, but could not assign to a specific Concept or Skill Statement (Fig. 14.2). Some of this apparent mismatch is a result of assessments items that focus on quantitative literacy, but not experimentation. Similarly, many assessment items that we could not categorize in the framework address student affect (e.g., self-efficacy) in domains not directly related to biological experimentation. We did not code assessments that focused exclusively on the nature of science (e.g., Lederman et al., 2002), because they do not address experimentation. Although both quantitative literacy (Clemons et al., 2020) and student affect (Trujillo & Tanner, 2014) are important student outcomes, they do not necessarily fit within the framework of biological experimentation. None-the-less, we found five aspects of experimentation that are not an explicit part of the ACE-Bio framework that appear in assessments and represent potential gaps in the existing framework. We do not present these as criticisms of the framework but note that the framework should be viewed as a document that requires interpretation and therefore thoughtful clarification and modification. The Concept of creativity is a pre-cursor to or facilitator of the *Question* and *Plan* Competence Areas at the very least and plays an underlying role in *Conclude* and *Communicate*. Creativity could be addressed as an aspect of experimentation (Beno & Tucker's Chap. 20 in this volume). Similarly, modern biological research often requires or is greatly enhanced by collaboration. In addition, collaboration is a core competency in the *Vision and Change* report (AAAS, 2011) and program-level learning outcome in the BioSkills guide (Clemons et al., 2020). Collaboration is assessed in the LCAS in the context of course-based undergraduate research experiences (CURES) (Corwin et al., 2015). Yet, collaboration is not explicitly in the framework, but could be incorporated in a number of the Competence Areas, much like creativity. This gap and the possibility of incorporating collaboration in the existing Competence Areas is explored in more detail later in this volume (Chaps. 20 and 22).

The other potential gaps in the framework are more specific to individual Competence Areas. The articulation of hypotheses is well described in *Question* but making falsifiable predictions for each hypothesis is not. This important feature of experimentation appears in some assessments but is not part of the framework. Lastly, the Concept of Statistics is part of the *Analyze* Competence Area. However, interpretation of statistical tests is missing from the framework. Statistical interpretation is addressed in some assessments and could be more explicitly incorporated in the *Analyze* or *Conclude* Competence Areas.

## 14.4 Recommendations

### 14.4.1 *Recommendations for Instructors*

Choosing an assessment on experimentation that will be used in a course or program requires that an instructor first decide on the learning outcomes to be assessed. That is not a trivial issue since no one assessment will address every aspect of experimentation and the format of an assessment may limit its usefulness. The assessments that are available can be categorized in two groups – those that are narrowly focused and those that address the breadth of the experimentation Competence Areas. Narrowly focused assessments are best used as formative assessments or assessments for education research rather than for assigning grades in a course. Measuring learning with a prompt or narrowly focused assignment and a rubric (Angra & Gardner, 2018; Brownell et al., 2013; Speth et al., 2010) will permit instructors to assess specific aspects of experimentation, mainly in the *Plan* and *Conclude* Competence Areas. Objective tests of learning, such as multiple-choice tests (Bedford & Lawson, 2001; Deane et al., 2014, 2016; Dirks & Cunningham, 2006; Gormally et al., 2012; Picone et al., 2007; Rybarczyk et al., 2014; Shi et al., 2011; Stanhope et al., 2017), also may be used as measures of very specific learning outcomes related to experimentation. It is very tempting to use a rapidly scored test as means of assigning grades, but we recommend against that because tests are not authentic assessments of experimentation, scientific research is not assessed in this manner. Matching the assessment used to the learning outcome set for students is essential. If the learning outcome is student achievement in the ability to perform experimentation, then having them perform the activities that comprise the process of biological experimentation is the most authentic assessment (papers, posters, proposals, research seminars scored with a rubric (Kishbaugh et al., 2012; Reynolds et al., 2009; Timmerman et al., 2011). Measuring learning with a research assignment and a rubric will permit instructors address the broadest range of experimentation Competence Areas (Table 14.1) and also could be used as a means of assigning grades.

Instructors should ensure that any assessment that they use was designed for the level of their students. Those assessments that were developed for introductory students could be used with upper-level students (e.g., EDAT (Sirum & Humburg, 2011) with several caveats. First, instructors should be sure to administer the assessment at the beginning of the semester to determine whether there is a likelihood of a ceiling effect. Second, instructors should consider differences in the expectations of Competence Areas in experimentation for introductory and upper-level students (Cole & Beck's Chap. 3 in this volume). These differences in expectations also make assessments designed for upper-level biology majors unlikely to be useful for assessing experimentation in introductory courses. Finally, instructors need to remember that these assessments were validated with introductory students.

The timing of the use of specific assessments also matters, both within a course and within an undergraduate curriculum. Instructors might reasonably begin a

course with very narrow learning outcomes and focus on specific skills and build to more comprehensive learning outcomes (and more authentic assessments such as papers, posters, proposals, research seminars) as the course developed during a semester. In this case, starting with less authentic assessments may be completely appropriate if they were used to create the scaffolding for more authentic assignments in that course. However, more advanced undergraduate courses should focus on the most authentic assessments (assessments that are closest to the activities performed by working scientists) and score them with rubrics to cover a broad range of experimentation competencies. A summary of these recommendations is given in bulleted form below the discussion.

#### **14.4.2 Recommendations for Education Researchers**

Our analysis of current assessments for biological experimentation leads to several recommendations for education researchers (summarized as a bulleted list below). The gaps in assessments that address the basic Competencies of Experimentation provide an opportunity to develop new assessment tools or modify existing tools. The Competence Areas of Identify and Conduct are essential aspects of the experimentation process, but we need the tools to assess them. Authors of other chapters in this volume provided examples of work to address this deficiency that we have identified, as described in the Preface to this book. Similarly, there are opportunities to develop assessment tools to address the Concepts of Representations, Ethics, and Limitations within the Plan Competence Area and the Concept of Statistics (choosing and conducting the appropriate statistical test and others) within the Analyze Competence Area. Using the ACE-Bio framework can be an important starting point for developing general or more discipline-specific assessments in these areas (Dasgupta et al., 2016). In addition, by using the framework as a basis for assessment, the aspects of biological experimentation that are being assessed will be clearer.

Aligning expectations of student competencies in experimentation for students at different levels with assessments designed for students at those levels is essential for rigorous studies of student learning on experimentation. While some assessments are applicable to students across multiple levels, others are specific to students at either the introductory or upper-level (Table 14.1). Therefore, education researchers can develop new assessments, or validate existing assessments for students at different levels, that align with the expectations for students at those levels (Cole & Beck's Chap. 3 in this volume). For example, the EDAT (Sirum & Humburg, 2011) was designed for non-majors introductory biology. However, faculty do not necessarily expect students to have much first-hand experience with the Competence Area Plan (Cole & Beck's Chap. 3 in this volume), which the EDAT covers extensively (Fig. 14.1). Even rubrics for student assignments could be refined to better articulate the expectations for students at different levels. The CRBS (Kishbaugh et al., 2012) is an example of where this has been done effectively.

Finally, how student learning of one competence area in biological experimentation relates to their learning of other competence areas is unclear. Linkages and correlations between learning of different experimentation competencies would be informative for both teaching and assessing experimentation. From the perspective of assessment, high correlations between learning of different competence areas would allow researchers and instructors to assess fewer competence areas while at the same time getting a complete picture of student understanding of experimentation.

In summary, consider the following recommendations for instructors and education researchers:

**Recommendations for Instructors:**

- Choose assessment instruments that best match the learning outcome expectations for a course.
- Use narrowly focused assignments as formative assessments but not for grading.
- Use broad based authentic assessments of learning, research assignment with a rubric for grading.
- Scaffold learning outcomes and assessments within course and within curriculum.

**Recommendations for Education Researchers:**

- Develop new assessments to fill current gaps in the *Identify* and *Conduct* Competence Areas.
- Develop new assessments to fill current gaps in the Concepts of Representations, Ethics, and Limitations within the *Plan* Competence Area.
- Develop new assessments to fill current gaps in the Concept of Statistics (choosing and conducting the appropriate statistical test and others) within the *Analyze* Competence Area.
- Develop new assessments, or validate existing assessments for students at different levels, so that expectations of students and assessments align.
- Explore linkages and correlations between learning of different experimentation competencies.

## 14.5 Conclusions

By mapping current assessments in biological experimentation on the ACE-Bio competence areas, we have provided a tool for instructors to select the best available assessments to examine student learning of experimentation in their classes and identified avenues of future research related to the development of new assessments on experimentation. Through appropriate application of current assessments and development of new assessments, we hope to advance our understanding of how students become competent at experimentation.

Finally, how student learning of one competence area in biological experimentation relates to their learning of other competence areas is unclear. Linkages and correlations between learning of different experimentation competencies would be

informative for both teaching and assessing experimentation. From the perspective of assessment, high correlations between learning of different competence areas would allow researchers and instructors to assess fewer competence areas while at the same time getting a complete picture of student understanding of experimentation.

**Acknowledgments** This work was partially funded by a stipend from the National Science Foundation (NSF) grant #1346567 for the Advancing Competencies in Experimentation–Biology (ACE–Bio) Network. We thank Nancy Pelaez, Trevor Anderson and Stephanie Gardner for facilitating the work of the ACE–Bio Network. Preliminary feedback from Trevor Anderson, Annwesa Dasgupta, and others in the ACE–Bio Network helped improve this chapter. For this, we are grateful. This study and publication were funded by NSF grants DUE-1821184, HRD-1818458 and HRD-2010676 to Morehouse College and DUE-1821533 to Emory University. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the NSF.

## References

American Association for the Advancement of Science (AAAS). (2011). *Vision and change in undergraduate biology education: A call to action*. Retrieved from Washington, DC <https://live-visionandchange.pantheonsite.io/wp-content/uploads/2013/11/aaas-VISchange-web1113.pdf>

Angra, A., & Gardner, S. M. (2018). The graph rubric: Development of a teaching, learning, and research tool. *CBE Life Sciences Education*, 17(4), ar65. <https://doi.org/10.1187/cbe.18-01-0007>

Beck, C. W., & Blumer, L. S. (2016). Alternative realities: Faculty and student perceptions of instructional practices in laboratory courses. *CBE Life Sciences Education*, 15(4), ar52. <https://doi.org/10.1187/cbe.16-03-0139>

Beck, C. W., Butler, A., & Burke da Silva, K. (2014). Promoting inquiry-based teaching in laboratory courses: Are we meeting the grade? *CBE Life Sciences Education*, 13, 444–452.

Benford, & Lawson, A. E. (2001). *Relationships between effective inquiry use and the development of scientific reasoning skills in college biology labs*. Retrieved from <https://files.eric.ed.gov/fulltext/ED456157.pdf>. Accessed 3 Jan 2021.

Boyer, E. L. (1998). *The Boyer Commission on educating undergraduates in the research university, reinventing undergraduate education: A blueprint for America's research universities*. Stony Brook.

Brownell, S. E., Wenderoth, M. P., Theobald, R., Okoroafor, N., Koval, M., Freeman, S., ... Crowe, A. J. (2013). How students think about experimental design: Novel conceptions revealed by in-class activities. *Bioscience*, 64(2), 125–137. <https://doi.org/10.1093/biosci/bit016>

Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *Journal of Social Issues*, 67(3), 469–491. <https://doi.org/10.1111/j.1540-4560.2011.01710.x>

Clemmons, A., Timbrook, J., Herron, J., & Crowe, A. (2020). BioSkills guide: Development and national validation of a tool for interpreting the Vision and Change Core Competencies. *bioRxiv*, 2020. <https://doi.org/10.1101/2020.01.11.902882>

Corwin, L. A., Runyon, C., Robinson, A., & Dolan, E. L. (2015). The laboratory course assessment survey: A tool to measure three dimensions of research-course design. *CBE Life Sciences Education*, 14(4). <https://doi.org/10.1187/cbe.15-03-0073>

Dasgupta, A. P., Anderson, T. R., & Pelaez, N. J. (2014). Development and validation of a rubric for diagnosing students' experimental design knowledge and difficulties. *CBE Life Sciences Education*, 13(2), 265–284.

Dasgupta, A. P., Anderson, T. R., & Pelaez, N. J. (2016). Development of the Neuron Assessment for measuring biology students' use of experimental design concepts and representations. *CBE Life Sciences Education*, 15(2), ar10. <https://doi.org/10.1187/cbe.15-03-0077>

Deane, T., Nomme, K., Jeffery, E., Pollock, C., & Birol, G. (2014). Development of the Biological Experimental Design Concept Inventory (BEDCI). *CBE Life Sciences Education*, 13(3), 540–551. <https://doi.org/10.1187/cbe.13-11-0218>

Deane, T., Nomme, K., Jeffery, E., Pollock, C., & Birol, G. (2016). Development of the Statistical Reasoning in Biology Concept Inventory (SRBCI). *CBE Life Sciences Education*, 15(1), ar5–ar5. <https://doi.org/10.1187/cbe.15-06-0131>

Diaz-Martinez, L. A., Fisher, G. R., Esparza, D., Bhatt, J. M., D'Arcy, C. E., Apodaca, J., ... Olimpo, J. T. (2019). Recommendations for effective integration of ethics and responsible conduct of research (E/RCR) education into course-based undergraduate research experiences: A meeting report. *CBE Life Sciences Education*, 18(2), mr2. <https://doi.org/10.1187/cbe.18-10-0203>

Dirks, C., & Cunningham, M. (2006). Enhancing diversity in science: Is teaching science process skills the answer? *CBE Life Sciences Education*, 5(3), 218–226. <https://doi.org/10.1187/cbe.05-10-0121>

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *Journal of Educational Psychology*, 103(1), 206–222. <https://doi.org/10.1037/a0020743>

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, 48(10), 1159–1176. <https://doi.org/10.1002/tea.20442>

Gormally, C., Brickman, P., & Lutz, A. (2012). Developing a Test of Scientific Literacy Skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *CBE Life Sciences Education*, 11(4), 364–377. <https://doi.org/10.1187/cbe.12-03-0026>

Haloun, I., & Hestenes, D. (1998). Interpreting VASS dimensions and profiles for physics students. *Science Education*, 7, 553–577.

Hanauer, D. I., & Dolan, E. L. (2014). The project ownership survey: Measuring differences in scientific inquiry experiences. *CBE Life Sciences Education*, 13(1), 149–158. <https://doi.org/10.1187/cbe.13-06-0123>

Hanauer, D. I., & Hatfull, G. (2015). Measuring networking as an outcome variable in undergraduate research experiences. *CBE Life Sciences Education*, 14(4). <https://doi.org/10.1187/cbe.15-03-0061>

Handelsman, J., Miller, S., & Pfund, C. (2007). *Scientific teaching*. WH Freeman and Co.

Hicks, J., Dewey, J., Brandvain, Y., & Schuchardt, A. (2020). Development of the biological variation in experimental design and analysis (BioVEDA) assessment. *PLoS One*, 15(7), e0236098. <https://doi.org/10.1371/journal.pone.0236098>

Killpack, T. L., & Fulmer, S. M. (2018). Development of a tool to assess interrelated experimental design in introductory biology. *Journal of Microbiology & Biology Education*, 19(3). <https://doi.org/10.1128/jmbe.v19i3.1627>

Kishbaugh, T. L. S., Cessna, S., Jeanne Horst, S., Leaman, L., Flanagan, T., Gruber Neufeld, D., & Siderhurst, M. (2012). Measuring beyond content: A rubric bank for assessing skills in authentic research assignments in the sciences. *Chemistry Education Research and Practice*, 13(3), 268–276. <https://doi.org/10.1039/C2RP00023G>

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497–521. <https://doi.org/10.1002/tea.10034>

Lopatto, D. (2008). Exploring the benefits of undergraduate research: The SURE survey. In R. Taraban & R. L. Blanton (Eds.), *Creating effective undergraduate research programs in science* (pp. 112–132). Teacher's College Press.

Moore, J. E., & Lynn, T. W. (2020). Lab notebooks: Efficient grading and effective learning. Article 45. In K. McMahon (Ed.), *Advances in biology laboratory education*. Volume 41. Publication of the 41st Conference of the Association for Biology Laboratory Education (ABLE). <https://doi.org/10.37590/able.v41.art45>

National Research Council (NRC). (2003). *BIO 2010. Transforming undergraduate education for future research biologists*. National Academy Press.

Pelaez, N., Anderson, T. R., Gardner, S. M., Yin, Y., Abraham, J. K., Bartlett, E., Gormally, C., Hill, J. P., Hoover, M., Hurney, C., Long, T., Newman, D. L., Sirum, K., & Stevens, M. (2017). The basic competencies of biological experimentation: Concept-skill statements. *PIBERG Instructional Innovation Materials. Paper 4*. Retrieved October 18, 2021 from <https://docs.lib.psu.edu/pibergium/4/>

Picone, C., Rhode, J., Hyatt, L., & Parshall, T. (2007). Assessing gains in undergraduate students' abilities to analyze graphical data. *Teaching Issues and Experiments in Ecology*, 5, Research #1. <http://tiee.esa.org/vol/v5/research/picone/abstract.html>

Project Kaleidoscope (PKAL). (2002). *Recommendations for action in support of undergraduate science, technology, engineering, and mathematics: Report on reports*. Washington, DC.

Reynolds, J., Smith, R., Moskovitz, C., & Sayle, A. (2009). BioTAP: A systematic approach to teaching scientific writing and evaluating undergraduate theses. *BioScience*, 59(10), 896–903, 898. Retrieved from <https://doi.org/10.1525/bio.2009.59.10.11>

Rybczynski, B. J., Walton, K. L., & Grillo, W. H. (2014). The development and implementation of an instrument to assess students' data analysis skills in molecular biology. *Journal of Microbiology and Biology Education*, 15, 259–267.

Semsar, K., Knight, J. K., Birol, G., & Smith, M. K. (2011). The Colorado Learning Attitudes about Science Survey (CLASS) for use in biology. *CBE Life Sciences Education*, 10(3), 268–278. <https://doi.org/10.1187/cbe.10-10-0133>

Shi, J., Power, J. M., & Klymkowsky, M. W. (2011). Revealing student thinking about experimental design and the roles of control experiments. *International Journal for the Scholarship of Teaching and Learning*, 5, 1–16.

Shortridge, E. E., & Brownell, S. E. (2016). How to assess your CURE: A practical guide for instructors of course-based undergraduate research experiences. *Journal of Microbiology & Biology Education*, 17(3). <https://doi.org/10.1128/jmbe.v17i3.1103>

Sirum, K., & Humberg, J. (2011). The experimental design ability test (EDAT). *Bioscene*, 37, 8–16.

Speth, E. B., Momsen, J. L., Moyerbrailean, G. A., Ebert-May, D., Long, T. M., Wyse, S., & Linton, D. (2010). 1, 2, 3, 4: Infusing quantitative literacy into introductory biology. *CBE Life Sciences Education*, 9(3), 323–332. <https://doi.org/10.1187/cbe.10-03-0033>

Stanhope, L., Ziegler, L., Haque, T., Le, L., Vinces, M., Davis, G. K., ... Overvoorde, P. J. (2017). Development of a Biological Science Quantitative Reasoning Exam (BioSQuaRE). *CBE Life Sciences Education*, 16(4), ar66. <https://doi.org/10.1187/cbe.16-10-0301>

Timmerman, B. E. C., Strickland, D. C., Johnson, R. L., & Payne, J. R. (2011). Development of a 'universal' rubric for assessing undergraduates' scientific reasoning skills using scientific writing. *Assessment & Evaluation in Higher Education*, 36(5), 509–547. <https://doi.org/10.1080/02602930903540991>

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE Life Sciences Education*, 13(1), 6–15. <https://doi.org/10.1187/cbe.13-12-0241>

Weston, T. J., & Laursen, S. L. (2015). The Undergraduate Research Student Self-Assessment (URSSA): Validation for use in program evaluation. *CBE Life Sciences Education*, 14(3). <https://doi.org/10.1187/cbe.14-11-0206>

Wilson, K. J., & Rigakos, B. (2016). Scientific Process Flowchart Assessment (SPFA): A method for evaluating changes in understanding and visualization of the scientific process in a multidisciplinary student population. *CBE Life Sciences Education*, 15(4), ar63. <https://doi.org/10.1187/cbe.15-10-0212>

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

