Educational Research and Development Journal
Fall 2023, Vol. 26, No. 2, pp. 62–84


# Student Perceptions of Assessment in Taiwan and the United States

Chuang Wang
Dawson Hancock
*University of North Carolina at Charlotte, Charlotte, North Carolina*

Jin-Jy Shieh
*University of Macau, Macau, China*

Jeremy Hachen
*Concord High School, Cabarrus County Schools, North Carolina*

## Abstract

The purpose of this study is to investigate the use of both formative and summative assessment in Taiwan and the United States. The focus is on the comparisons between undergraduate and graduate students and between U.S. and Taiwanese students in their attitudes toward the use of assessment in higher education. Responses from 349 undergraduate and graduate students from Taiwan and 97 undergraduate and graduate students from a large city in the United States were used for chi-square test, multivariate analysis of variance, and independent samples t-test. The two groups of students were comparable with respect to the distribution of gender and grade level. Results suggest no statistically significant difference with respect to the forms of assessment: self-assessment, peer assessment, and teacher assessment. As for the tools of assessment, attendance, class participation, homework, reading report, quizzes and exams were used more often in the classrooms in Taiwan whereas learning diary, essays, display, and projects were used more often in the United States. Students in Taiwan also agreed more to use a combination of all forms of assessments in comparison to their American counterparts. As for the contents of assessment, students in Taiwan endorsed more on the perception that assessments should include aspects of cognition, skill, and attitude. Results from this study provided information about the use of formative and summative assessment in Taiwan and the United States and student perspectives of its use. Findings from this study might be helpful for the U.S. and Taiwanese instructors to learn from each other and to meet their students' expectations in the use of assessment.

## Introduction

Although assessment may be designed as formative or summative, it is an important component in teaching and learning (Dunn & Mulvenon, 2019). According to McMillan (2013), assessment influences student learning, engagement, and motivation and provides information for the enhancement of instruction. Since students may not be able to fully understand the practical application of established standards, objectives, and stated learning goals, assessment helps them appreciate what is operationally expected from them (Yorke, 2003). Rowntree (1987) echoed this idea when he wrote that "the spirit and style of assessment defines the *de facto* curriculum" (p. 1). Students will allocate their efforts and attention on learning what they believe will be assessed (Gibbs, 2006; Stödberg, 2012). This indicates that assessment can be used to influence student learning. To improve student learning, it is pivotal for instructors to understand students' attitude toward the use of assessment and their perceptions of how assessment should be used in the classroom (Furnham et al., 2011). Student perceptions of assessment are influenced by the social

and cultural contexts in which the assessment is used (Huang & Asghar, 2016). The purpose of this study, therefore, is to understand student perceptions of assessment and how assessment is implemented in Taiwan and the United States so that instructors can enhance student learning through assessment. Research questions that guided this study are:

1. What are commonly used forms of assessment in higher education in Taiwan and the United States?
2. What are commonly used assessment activities employed by Taiwanese and U.S. instructors in higher education classrooms?
3. Is there a difference between undergraduate and graduate students in their perceptions of assessment?
4. Is there a difference between Taiwanese and U.S. students in their perceptions of assessment?

Answers to these questions will contribute to the understanding of assessment in socially and culturally different contexts. Findings of this study will help instructors in Taiwan and the United States to implement assessment properly so that student learning is enhanced. Policy makers may also benefit from this study by making assessment-related policies that are culturally appropriate in higher education.

## Theoretical Framework and Literature Review

This study is based on the theoretical framework of assessment and related studies of the relationship between assessment and learning outcomes. The context of this study, Taiwan and the United States, were introduced with respect to the use of assessment in higher education.

### Definition and Use of Assessment

Assessment, a term often used interchangeably with testing, measurement, and evaluation, was defined as "any act of interpreting information about student performance, collected through any of a multitude of a means or practices" (Brown, 2004; p. 304). In this study, we adopt Ghaicha's (2016) definition of assessment, which is closest to the classroom context used in this study:

> Assessment is operationally defined as a part of the educational process where [faculty] instructors appraise students achievements by collecting, measuring, analyzing, synthesizing and interpreting relevant information about a particular object of interest in their performance under controlled conditions in relation to curricula objectives set for their levels, and according to the procedures that are systematic and substantively grounded. (p. 213)

Assessment plays a pivotal role in education and is often set in a top-town model: international, national, regional, institutional, and classroom-level to hold governments, educational agencies (i.e., schools), programs, teachers, and students accountable to the public (Ghaicha, 2016). The focus of this study is classroom assessment, which is conducted by instructors: a) to make judgement of their students' academic performance; b) to diagnose their students' strengths and weakness; and c) to make adjustments to their curriculum and/or pedagogies.

Classroom assessment includes both formative and summative assessment. Summative assessment, also known as assessment of learning, provides evidence of achievement to the public (Stiggins, 2002). Formative assessment; however, is assessment for learning because the purpose of formative assessment is to gather information so that instructors can set learning goals, design classroom assessment that measures student progress, and to build student confidence in learning.

In this study, we use the term "assessment as learning" because we agree with Stiggins (2002) that "there are no good arguments against balancing our assessment of and for learning" (p. 2). Stiggins (2001) argued that students should be involved in formative assessment that grants students the confidence and ability to learn effectively. The purpose of this study, therefore, is to understand student perspectives of the use of classroom assessment so that educators can improve the teaching and learning process.

## Assessment and Learning Outcomes

Black and William (1998) reviewed 250 journal articles and books on the relationship between assessment and academic achievement and declared that assessment was an essential component of classroom instruction and did improve student learning. They suggested that the mean effect size of assessment could be as high as 0.4 to 0.7. Although there is long standing and modern research that shows assessment has a sizeable impact on student learning (Black & William, 1998; Li, 2016; Ramsden, 1997; Vanderlelie & Alexander, 2016; Xiao et al., 2022), the work of Black and William has come under critical review in more recent years. Bennett (2011) argues that the lack of a universally agreed upon definition of assessment prevents the topic from being accurately studied. Bennett's claim that assessment lacks codified practices leads him to doubt the impressive effect sizes suggested by Black and William (1998). A meta-analysis of assessment studies produced a weighted mean effect size of 0.2 (Kingston & Nash, 2011). This finding is significantly lower than previous estimates but still indicates that assessment can be a useful tool for improving student learning. For all the questions surrounding empirical studies on assessment, the general consensus remains that it can be an effective tool for advancing student learning (Bennett, 2011; Dunn & Mulvenon, 2019; Xiao et al., 2022).

## Assessment Forms

A multitude of assessment forms were identified in the literature, such as summative assessment (e.g., large-scale and norm-referenced public examinations, final examinations in a course), formative assessment (e.g., quizzes, essays, research projects, oral presentations, portfolios), educative assessment, student assessment, and classroom assessment (Ghaicha, 2016). The practice of formative peer review has been gaining more attention and practice in higher education (Hanrahan & Isaacs, 2001; van den Berg et al., 2006). This form of shared learning can take many forms and involves students providing feedback on each other's work (Vickerman, 2009). Topping's (1996) review of the literature on peer tutoring in higher education found that peer feedback and tutoring can be as effective as that given by faculty. Studies on peer assessment in higher education have also shown improvement in writing skills (van den Berg et al., 2006). Even with the increasing commonality of this practice within higher education, there are few published studies on how students view the practice (Hanrahan & Isaacs, 2001). Several studies on students perceptions of peer assessment indicate that even though some students struggle with assessing the work of others, most find the practice to be a positive experience and were motivated by formative peer assessment (Lindblom-Ylänne et al., 2006; Vickerman, 2009). To our knowledge, no studies have compared the forms of assessment commonly used in Taiwan and the United States. Therefore, this study aims to have an understanding of the forms of assessment used in Taiwan and the United States (Research Question 1).

## Student Perceptions of Assessment

More and more researchers have paid attention to student involvement in the assessment process (e.g., Peterson & Einarson, 2001; Holzinger et al., 2020). Students do have opinions, preferences, and their own perceptions as to how they are assessed in schools. In one study, students expressed a desire to have time in class to prepare for assessments and to have performance checklists (e.g., rubrics) explained to them in advance (Rieg, 2007). Students have also regularly reported multiple-choice assessments to be a favored method (Furnham et al., 2011; Holzinger et al., 2020; Struyven et al., 2005). According to another study, students expressed a belief that a variety of cognitive processes should be assessed including reproduction, critical thinking, problem solving, and comprehension (van de Watering et al., 2008). Some studies of student perceptions of assessment indicate that oral examination, group work, and evaluating the work of others are undesirable forms of assessment (Furnham et al., 2011; Holzinger et al., 2020; van de Watering et al., 2008). These studies were conducted in European countries such as Great Britain and Netherlands. Therefore, there is a need to conduct a study in Asia (e.g., Taiwan) and America (e.g., the United States) to understand these students' perceptions of the use of assessment and forms of assessment (Research Question 2). In addition, this study also aims to examine if there is a difference between Taiwanese and U.S. students with respect to their perspectives of the use of assessment (Research Question 4).

**Difference between Undergraduate and Graduate Students**

The assessment preferences of undergraduate students do not deviate from the research of students at-large. Anderson (1987) found that undergraduate students preferred true/false and multiple choice assessments to higher order thinking methods such as short answer and essay. This echoes the findings of other studies on student assessment preferences (Furnham et al., 2011; Struyven et al., 2005). Even though students stated a preference for easier surface methods of assessment, they rated the higher order thinking methods as a better gauge of learning (Anderson, 1987). A study of business students found that undergraduates prefer discrete facts and quick and easy problems (Phillips, 1999). Birenbaum (2007) found that the most favored form of assessment among undergraduate participants in Israel was examples of what would appear on future exams and guidance on how to prepare for summative assessments. The least preferred method in Birenbaum's (2007) study was higher-order thinking tasks that involved creativity, explanation, and integration of learned material. In an earlier study, Birenbaum (1997) found that undergraduate assessment preferences were not related to different subjects of study but were more closely aligned with individual learning styles.

A 2016 study found that undergraduate students did perceive formative tests as useful (Grosas et al., 2016). The undergraduates in the study also provided their beliefs about the use of assessment. The responses focused on learning for the final summative exam and attaining a better understanding of the exam process (Grosas et al., 2016). The participants did not indicate that informing the instructor of their current knowledge was an important aspect of the assessments (Grosas et al., 2016).

Although there is research on how undergraduate students perceive formative assessment, there is a dearth of literature on the attitudes of graduate students. Ferguson (2011) conducted a study of student perceptions on effective feedback that included both undergraduate and graduate students. He found a surprising amount of consistency across undergraduate and graduate students' attitudes towards effective feedback practices. He concluded that students have a high degree of agreement about what qualifies as effective assessment feedback and process. Both sets of students identified "brief written comments throughout" and "written summary/overview" as the most

useful assessment feedback methods and "group verbal feedback" as the least useful (Ferguson, 2011). A more recent systematic review suggested that students benefit from quizzes and feedback from peers and teachers (Morris et al., 2021). No previous studies have compared undergraduate and graduate students' perspectives of assessment, so this study is to fill in this gap by a comparison between these two groups of students (Research Question 3). To help readers better understand the context of this study, we present the context of assessment in Taiwan and the United States in the following sections.

## Assessment in Taiwan

Taiwan falls under the umbrella of Confucian Heritage Culture and has a history with deep roots in testing culture (Berry, 2011; Huang & Asghar, 2016). For centuries, pivotal assessment such as the Joint Entrance Examination of Universities and Colleges has determined future prospects for Taiwanese students (Berry, 2011; Huang & Asghar, 2016). Education is seen by many in Taiwan to be their only option for upward economic mobility (Huang & Asghar, 2016; Law, 2004). In order to prepare students for these highly consequential assessments, traditional teacher-centered practices revolved around rote learning methods such as knowledge reproduction and teacher-oriented classrooms (Lu et al., 2010; Tsai & Kuo, 2008). Formative assessment, or assessment for learning, is not a widely practiced or a clearly understood concept in Taiwan (Huang & Asghar, 2018).

Schools in Taiwan centered their teaching on the memorization of facts for high stakes assessments (Berry, 2011; Chou, 2019). Taiwan has more recently been encouraging the use of Assessment for Learning in their curriculum, but they are facing many of the same challenges when it comes to achieving full implementation (Berry, 2011).

Recent efforts in Taiwan have attempted to shift instructional practice away from rote didacticism and to encourage a more student-centered approach (Huang & Asghar, 2018). Educational authorities in Taiwan are looking to western education models and are borrowing several practices, including the use of formative assessment (Huang & Asghar, 2018). Schools are being encouraged to implement formative assessment practices as part of their ongoing efforts to replace exam-oriented education with quality oriented education (Huang & Asghar, 2018). Although assessment reform is a focal point of Taiwan's widespread efforts, many barriers to full implementation exist on both a macro and micro scale (Berry, 2011; Pham & Renshaw, 2015).

One issue stunting the development of formative assessment in Taiwan is a lack of definitive conceptualization of what this practice should look like in the classroom (Huang & Asghar, 2018). Policy doctrines that are passed down to the schools are often vague and do not offer clear explanation on how to implement formative assessment practices (Chou & Ching, 2012). The speed and breadth of top-down education reform has made it difficult for schools and teachers in Taiwan to develop appropriate structures of implementation and practice (Yang et al., 2002). The scarcity of professional development opportunities is a major barrier to conceptualization and application of formative assessment in Taiwan, as teachers are not prepared to implement these strategies into teaching and learning (Berry, 2011). Inadequate funding and a lack of quality research into formative assessment practices in Taiwan compound the challenges of effectively adopting this practice (Huang & Asghar, 2018).

Interviews of Chinese teachers by Tan and Chua (2015) revealed that the exam-driven culture inhibits their ability to adopt student-centered learning and formative assessment practices. Even though curriculum reform efforts have promoted the use of alternative assessment strategies, students and parents still perceive success on the high-stakes summative assessments as the most

important product of schooling (Yang et al., 2002). Culture at the school level has not changed much and a high-stakes test-oriented culture dominates teaching and learning in Tiawan (Huang & Asghar, 2018). Therefore, teachers in Taiwan are attempting reconcile the idea that assessment facilitates learning and personal development with their responsibilities to exam-oriented preparation (Huang & Asghar, 2018). It seems as if formative assessment will take a higher position in Taiwan only if it can be shown to raise test scores at a higher rate than drill and rote practice (Yang et al., 2002).

There are also large scale cultural obstacles to employing formative assessment in Confucian Heritage Societies. The hierarchy of power and the values of collectivism create unique conditions for formative assessment in classrooms (Brown & Gao, 2015; Tweed & Lehman, 2002). In these classrooms, the teacher is seen as the content expert, knowledge transmitter, and moral judge and the primary educational goal is academic achievement and knowledge reproduction (Tan & Chua, 2015). Children raised in a Confucian Heritage society are taught to be modest and to keep one's head down. This mentality can make some forms of formative assessment (e.g., self and peer assessment) difficult because students often find it culturally inappropriate to evaluate their own work or that of others (Pham & Renshaw, 2015). Furthermore, teachers often avoid formative assessment and opportunities for constructive feedback to ensure students are not singled out (Yin & Buck, 2015).

## Assessment in the United States

While Taiwan has only recently begun decentralizing many aspects of education (Huang & Asghar, 2016), education policy-making in the United States has always been a state and local issue (Flaitz, 2011). The United States is without a national curriculum or national assessment and the federal government has largely stuck to a role of providing supplemental funding (Isaacs, 2001). The various states further decentralize broad educational authority to local school districts, of which there are over 14,000 nationwide (National Center for Education Statistics, 2021). It is due to this comparatively extreme degree of decentralization and local autonomy that a uniform use of formative assessment strategies is not able to be identified (Isaacs, 2013).

The publication of *A Nation at Risk* initiated a sizeable intervention by the federal government in educational testing and assessment protocol. The United States passed the *No Child Left Behind Act* (NCLB) in 2002 in an attempt to counteract many of the fears about American education laid out in *A Nation at Risk*. Under the landmark legislation, states were required to adopt or develop standardized tests in literacy and mathematics (No Child Left Behind Act, 2002). The legislation also mandated that only objective knowledge be assessed. Assessment in the United States changed under NCLB by becoming more high-stakes for schools, basing assessment on standards, and increasing school accountability for public funding (Flaitz, 2011).

Assessment in the United States has been increasingly used as a way of holding local educational authorities accountable for the public funding they receive by making them show evidences of student achievement (Flaitz, 2011). The financial rewards and punishments of NCLB work against the expansion of formative assessment practices by directing schools' attention to student performance on summative assessment (Flaitz, 2011). Because state-mandated assessments are high-stakes for schools, many districts have implemented rigid pacing guides, benchmarks, and content standards that make it difficult for classroom teachers to use formative assessment to address student learning needs (Black & William, 2005). While other nations are seeking to expand the use of formative assessment as a tool to enhance student learning, NCLB has been moving the United States in the opposite direction (Flaitz, 2011).

The use of assessment in the United States varies widely and a minority of teachers are effectively using them to improve instruction (Sharkey & Murnane, 2006). Even though a growing number of districts are including assessment in their strategies for enhancing student learning (Sharkey & Murnane, 2006), it is still infrequently used in classrooms (Marsh, 2007). Many teachers simply are not trained in how to interpret student data to improve their instruction (Sharkey & Murnane, 2003). These challenges of implementation and the impacts of NCLB have pushed education in the United States to focus more on assessment of learning and away from formative assessment practices.

Formative assessment includes a wide range of tasks, activities, and observations that are used to gauge student learning and direct instructional practice (Black & William, 1998). Summative assessment and high-stakes testing can even be considered formative if the results are used to direct student learning (Yorke, 2003). The benefits of formative assessment to student learning are well established (Bennett, 2011; Black & William, 1998; Dunn & Mulvenon, 2019; Xiao et al., 2022), yet there are few studies that examine student perceptions and preferences of implementation (Birenbaum, 2007; Phillips, 1999). Additionally, there is not enough literature to construct a comparison of assessment practices across different cultures (Gilles et al., 2011), especially in the context of higher education (Xiao et al., 2022).

### Summary

The review of literature above suggests that understanding the student perceptions of assessment is valuable for both cultures as Taiwan attempts to encourage development and usage of formative assessment whereas the United States continues its path towards a high-stakes summative assessment. Our study seeks to better understand how students in Taiwan and the United States perceive assessment in education so that it may be implemented more effectively by the respective school systems. As the literature on the graduate students' perspectives of assessment is limited, we aim to fill in this gap by comparing the graduate students' perspectives with undergraduate students' perspectives of assessment in both Taiwan and the United States.

## Method

### Participants

Since the purpose of our study is to understand undergraduate and graduate students' perspectives of assessment in Taiwan and the United States, we chose higher education students in Taiwan and the United States as our sampling frame. Proportional stratified random sampling was used to recruit participants. The sampling of students in Taiwan was based on the fifteen public universities with education colleges in Taiwan. As the population of Taiwan is mainly distributed in the northern, central, southern, and eastern regions, these regions were used as strata for stratified random sampling. Ten of these universities were selected, including three traditional normal universities, three comprehensive universities but with education colleges, three education universities, and one comprehensive university restructured from a university of education. These 10 universities were chosen because the other five were small in size and that students in these 10 universities are quite representative of education-major students in Taiwan. A total of 491 education majored undergraduate and graduate students, which represent 3% percent of the target population, were selected to participate in this study by completing a questionnaire online. However, 349 student questionnaires were returned, with a response rate of 71%. Of these 349 students, 231 (66%) were undergraduate students and 114 (33%) were graduate students whereas

272 (78%) were female and 73 (21%) were male. Four students (1%) did not report their information about gender and grade level.

The sampling of students in the United States was from a single comprehensive university with a college of education in the southeastern region. Valid responses from 97 education majored undergraduate and graduate students were used with a response rate of 68%. Of these 97 students, 82 (86%) were undergraduate students and 13 (14%) were graduate students whereas 77 (79%) were female and 18 (19%) were male. Two students (2%) did not report their information about gender and grade level.

## Instrument

Guided by the theoretical framework that considers assessment as part of the learning process and Ghaicha's (2016) operational definition of assessment, we developed teacher and student questionnaires for this study. The development of the teacher questionnaire was conducted by interviewing 13 university professors from Taiwan, Macau, Chinese Mainland, Singapore, and the United States because the questionnaire was intended to be administered in these countries and regions. The contents of the interviews were sorted out and analyzed, and then the first draft of the questionnaire was drawn up with reference to theoretical frameworks in assessment, findings from previous research studies, and questionnaires used by other scholars (e.g., Holzinger et al., 2020). Afterwards, 29 teachers from Taiwan participated in a pilot study and provided comments on revisions to the questionnaire, including clarification of the terms used. After the teacher questionnaire was completed, the narrative was fine-tuned from the student perspective to develop a first draft of the student questionnaire. After the first draft was completed, three students from an institute of education in Taiwan were invited to review the statement of the questions. Another pilot study was conducted with 81 students from an institute of education in Taiwan. Finally, the questionnaire was revised according to the results of the pilot study to form a final version of the student questionnaire (see Appendix).

The questionnaire starts with demographic information which asks for the background information of the students. Items such as grade level (undergraduate versus graduate) and location (Taiwan versus the United States) helped us put the participants into separate groups to answer Research Question 3 (Is there a difference between undergraduate and graduate students in their perceptions of assessment?) and Research Question 4 (Is there a difference between Taiwanese and U.S. students in their perceptions of assessment?). In addition to demographic information, the questionnaire consists of two sections. Section One includes two questions. The first question asks for the methods their instructors used in this semester, and students were asked to choose from four choices: teacher-assessment, student peer assessment, student self-assessment, and combined teacher-student assessment. The second question gives 11 options for students to select the types of assessment their instructor used to assess learning outcomes. Students were allowed to choose more than one options for each question in this section. Responses to questions in this section help us answer Research Question 1 (What are commonly used forms of assessment in higher education in Taiwan and the United States?). Section Two consisted of 20 items about student attitude toward purposes of assessment (6 items), forms of assessment (5 items), use of assessment in grading (4 items), and the impact of assessment (5 items). Students were asked to rate on a four-point Likert scale from 0 (strongly disagree) to 3 (strongly agree). Responses to items in this section help us answer Research Questions 2 (What are commonly used assessment activities employed by Taiwanese and U.S. instructors in higher education classrooms?).

## Data Analysis

Descriptive statistics were used to report commonly used assessment methods and classroom assessment activities in Taiwan and the United States. Chi-square tests were employed to examine differences in the frequencies of the assessment methods and classroom assessment activities reported. These statistical procedures were used to answer Research Question One and Research Question Two because we wanted to compare the frequencies in Taiwan and the United States.

Inferential statistics, i.e., multivariate analysis of variance (MANOVA), were used to answer Research Questions 3 and 4. Specifically, MANOVA was used to examine mean differences in comparisons in students' responses between Taiwanese and U.S. students and between undergraduate and graduate students. Independent samples t-tests were used to examine differences at the item-level with Bonferroni's correction to reduce the probability of incorrectly rejecting the null hypotheses (Type I error). The familywise alpha level was set at .05.

## Results

**Research Question 1:** What are commonly used forms of assessment in higher education in Taiwan and the United States?

Frequencies of commonly used assessment methods used in the classroom were reported in Table 1.

**Table 1.**

*Frequencies of Forms of Commonly Used Assessment in Taiwan and the United States*

| | Taiwan | United States |
|---|---|---|
| Teacher Assessment | 314 (50.5%) | 71 (51.4%) |
| Peer Assessment | 116 (18.6%) | 21 (15.2%) |
| Self Assessment | 87 (14.0%) | 25 (18.1%) |
| Joint Assessment of Teacher and Students | 105 (16.9%) | 21 (15.2%) |

Data from both undergraduate and graduate students were combined in Table 1. Results from chi-square test suggested no statistically significant differences between students in Taiwan and the United States in their report of the form of assessment method used in their classrooms, $\chi^2$ ($df$ = 3, $n$ = 760) = 2.25, $p$ = .52. Results show that teacher assessment is the dominant form of assessment in both Taiwan and the United States. The other forms of assessment (i.e., peer assessment, self-assessment, and joint assessment of teacher and students) are evenly distributed with each accounting nearly 15% of the assessment.

**Research Question 2:** What are commonly used assessment activities employed by Taiwanese and U.S. instructors in higher education classrooms?

Frequencies of commonly used assessment activities used in the classroom were reported in Table 2.

**Table 2.**

*Frequencies of Commonly Used Assessment Tools in Taiwan and the United States*

| Activities | Taiwan | United States |
|---|---|---|
| Attendance* | 203 (11.5%) | 30 (6.4%) |
| Classroom Participation* | 170 (9.6%) | 31 (6.5%) |
| Homework* | 258 (14.6%) | 47 (10.0%) |
| Learning Diary* | 28 (1.6%) | 26 (5.6%)) |
| Portfolios | 108 (6.1%) | 32 (6.8%) |
| Essays* | 42 (2.4%) | 39 (8.3%) |
| Display* | 173 (9.8%) | 64 (13.7%) |
| Projects* | 65 (3.7%) | 55 (11.8%) |
| Oral Presentation | 250 (14.1%) | 67 (14.4%) |
| Reading Report* | 135 (7.6%) | 26 (5.6%) |
| Quizzes and Exams* | 337 (19.1%) | 51 (10.9%) |
| All Activities | 1769 (100%) | 468 (100%) |

*Note.* Proportions of assessment activities marked with asterisks were statistically and significantly different between classrooms in Taiwan and the United States.

Results from chi-square test suggested statistically significant differences between students in Taiwan and their counterparts in the United States in their report of the assessment activities used in their classrooms, $\chi^2$ ($df$ = 10, $n$ = 2237) = 141.73, $p$ < .001. It seems that attendance, classroom participation, homework, reading reports, and quizzes and exams were used more often in Taiwan whereas learning diary, essays, display, and projects were used more often in the United States to assess student learning.

**Research Question 3:** Is there a difference between undergraduate and graduate students in their perceptions of assessment of learning?

Descriptive statistics for student perceptions of assessment of learning were presented in Table 3 for both Taiwanese and U.S. undergraduate and graduate students' perspectives of assessment.

**Table 3.**

*Means and Standard Deviations of Student Perspectives of Assessment*

| | | Purpose | | Form | | Use | | Impact | |
|---|---|---|---|---|---|---|---|---|---|
| *Location* | *Level* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Taiwan | Undergraduate | 2.07 | 0.44 | 2.29 | 0.40 | 1.95 | 0.43 | 2.11 | 0.43 |
| | Graduate | 2.14 | 0.44 | 2.27 | 0.40 | 1.99 | 0.38 | 2.10 | 0.44 |
| U.S. | Undergraduate | 1.89 | 0.28 | 2.02 | 0.16 | 1.90 | 0.25 | 2.05 | 0.30 |
| | Graduate | 2.11 | 0.38 | 2.04 | 0.39 | 1.90 | 0.36 | 2.22 | 0.37 |

Results of MANOVA showed no statistically significant interaction effect between location and level, $F(4, 433) = 0.88$, $p = .48$, $\eta^2 = .01$ (small effect size). This non-significant interaction effect suggests that the difference between undergraduate and graduate students do not vary across location (i.e., Taiwan and the United States). The main effect of level was not statistically significant either, $F(4, 433) = 1.61$, $p = .17$, $\eta^2 = .02$ (small effect size). This non-significant main effect suggests that the difference between undergraduate and graduate students was not statistically significantly different from zero. That is to say, both Taiwanese and U.S. undergraduate students' perspectives of assessment do not differ from those of graduate students. The main effect of location; however, was statistically significant, $F(4, 433) = 6.24$, $p < .001$, $\eta^2 = .06$ (medium effect size). This significant main effect suggests that Taiwanese students (both undergraduate and graduate) endorsed more of the linear combination of the purpose, form, use, and impact of assessment in comparison to U.S. students. Tests of between-subjects effects showed that Taiwanese students endorsed more the forms of assessment than U.S. students, $F(1, 436) = 15.41$, $p < .001$, $\eta^2 = .03$ (small effect size). The difference in other aspects of assessment (i.e., purpose, use, and impact) are not statistically significant, which means that Taiwanese and U.S. students differ only in their perspectives of the form of classroom assessment.

To gain a more detailed understanding of specific similarities and differences between the students classified by location and level, Table 4 presents the results of the independent samples t-tests of the comparisons between undergraduate and graduate students at the item level. Since MANOVA did not show statistically significant interaction effects between location and level, we combined the Taiwanese and U.S. students in this comparison.

**Table 4.**

*Comparison between Undergraduate and Graduate Students' Perception of Assessment*

| | Undergraduate | | Graduate | | *Comparison* |
|---|---|---|---|---|---|
| *The purpose of assessment is to…* | *M* | *SD* | *M* | *SD* | *t* |
| equip students with the lifelong learning knowledge, skills and attitude. | 2.26 | 0.56 | 2.41 | 0.56 | -2.76** |
| help students improve their quality of learning. | 2.28 | 0.58 | 2.44 | 0.55 | -2.90** |
| differentiate students 'achievement for admissions to advanced learning/employment. | 1.88 | 0.80 | 1.96 | 0.76 | -1.00 |
| provide evidences for the society to determine whether the curriculum is appropriate or not. | 1.80 | 0.77 | 1.76 | 0.74 | 0.53 |
| determine whether students can pass the course. | 1.79 | 0.76 | 1.82 | 0.72 | -0.39 |
| provide feedback so students can know their learning progress and further learning direction | 2.34 | 0.56 | 2.37 | 0.53 | -0.58 |
| *Assessment Should Include…* | *M* | *SD* | *M* | *SD* | *t* |
| the aspects of cognitive, skill and attitude. | 2.31 | 0.57 | 2.18 | 0.60 | 2.23 |
| Students' self-assessment, peer assessment and teacher assessment. | 2.24 | 0.60 | 2.01 | 0.73 | 3.71*** |
| oral and written assessment. | 2.19 | 0.56 | 2.09 | 0.60 | 1.85 |
| both qualitative and quantitative assessment. | 2.28 | 0.53 | 2.22 | 0.59 | 1.20 |
| Clear and objective standards should be established for formative assessment | 2.38 | 0.52 | 2.41 | 0.54 | -0.67 |

| Grading | *M* | *SD* | *M* | *SD* | *t* |
|---|---|---|---|---|---|
| The grading weight of course work should be higher than final examination. | 2.37 | 0.59 | 2.49 | 0.57 | -2.27* |
| Students' grades should be given according to a specific distribution system (e.g. 20% of A). | 2.05 | 0.68 | 2.15 | 0.64 | -1.60 |
| Once students achieve certain expected achievement, they should be given correspondingly high scores. | 1.45 | 0.87 | 1.21 | 0.91 | 2.77** |
| Apart from marks or ranking, students should be given written feedback. | 1.93 | 0.70 | 1.95 | 0.67 | -0.43 |
| *Attitudes toward Assessment* | *M* | *SD* | *M* | *SD* | *t* |
| Assessment can promote teachers' teaching. | 2.10 | 0.62 | 2.23 | 0.64 | -2.14* |
| Assessment can reflect teaching effectiveness. | 1.96 | 0.65 | 2.11 | 0.63 | -2.41* |
| The ways of Assessment influence students' investment on their study time. | 2.15 | 0.63 | 2.08 | 0.60 | 1.07 |
| Student self-assessment helps students monitor their own learning progress. | 2.13 | 0.56 | 2.19 | 0.52 | -1.24 |
| Student peer-assessment can help students develop skills of evaluating and giving feedback. | 2.20 | 0.55 | 2.14 | 0.63 | 1.02 |

*Note.* * $p < .05$; ** $p < .01$; *** $p < .001$.

Results of independent samples t-test showed that graduate students endorsed more on the following items:

a) The purpose of assessment is to equip students with the lifelong learning knowledge, skills and attitude.

b) The purpose of assessment is to help students improve their quality of learning.

c) The grading weight of course work should be higher than final examination.
d) Assessment can promote teachers' teaching.
e) Assessment can reflect teaching effectiveness.

Results of independent samples t-test showed that undergraduate students endorsed more on the following items:

a) Assessment should include students' self-assessment, peer assessment and teacher assessment.

b) Once students achieve certain expected achievement, they should be given correspondingly high scores.

**Research Question 4:** Is there a difference between Taiwanese and U.S. students in their perceptions of assessment?

Table 5 presents the results of the independent samples t-tests of the comparisons between Taiwanese and U.S. students at the item level. Since MANOVA did not show statistically significant interaction effects between location and level, we combined the undergraduate and graduate students in this comparison.

**Table 5.**

*Comparison of Students' Perception of Assessment between Taiwan and United States*

| | **United States** | | **Taiwan** | | ***Comparison*** |
|---|---|---|---|---|---|
| *The purpose of assessment is to …* | *M* | *SD* | *M* | *SD* | *t* |
| equip students with the lifelong learning knowledge, skills and attitude. | 2.39 | 0.57 | 2.30 | 0.56 | 1.36 |
| help students improve their quality of learning. | 2.33 | 0.60 | 2.35 | 0.56 | -0.28 |
| differentiate students 'achievement for admissions to advanced learning/employment. | 1.88 | 0.67 | 1.92 | 0.81 | -0.43 |
| provide evidences for the society to determine whether the curriculum is appropriate or not. | 1.74 | 0.69 | 1.79 | 0.77 | -0.58 |
| determine whether students can pass the course. | 1.76 | 0.65 | 1.82 | 0.77 | -0.63 |
| provide feedback so students can know their learning progress and further learning direction. | 2.32 | 0.59 | 2.26 | 0.53 | -0.68 |
| *Assessment Should Include…* | *M* | *SD* | *M* | *SD* | *t* |
| the aspects of cognitive, skill and attitude. | 2.04 | 0.65 | 2.30 | 0.56 | -3.84*** |
| Students' self-assessment, peer assessment and teacher assessment. | 1.67 | 0.71 | 2.25 | 0.62 | -7.84*** |
| oral and written assessment. | 1.88 | 0.60 | 2.21 | 0.56 | -4.93*** |
| both qualitative and quantitative assessment. | 2.06 | 0.60 | 2.30 | 0.54 | -3.73*** |
| Apart from marks or ranking, students should be given written feedback. | 1.89 | 0.63 | 1.95 | 0.70 | -0.72 |
| Grading | *M* | *SD* | *M* | *SD* | *t* |
| Clear and objective standards should be established for assessment | 2.46 | 0.58 | 2.37 | 0.52 | 1.38 |
| The grading weight of course work should be higher than final examination. | 2.55 | 0.54 | 2.39 | 0.59 | 2.36* |
| Students' grades should be given according to a specific distribution system (e.g. 20% of A). | 2.22 | 0.59 | 2.05 | 0.68 | 2.23* |
| Once students achieve certain expected achievement, they should be given correspondingly high scores. | 0.91 | 0.88 | 1.46 | 0.86 | -5.37*** |
| *Attitudes toward Assessment* | *M* | *SD* | *M* | *SD* | *t* |
| Assessment can promote teachers' teaching. | 2.43 | 0.56 | 2.09 | 0.64 | 4.73*** |
| Assessment can reflect teaching effectiveness. | 2.24 | 0.58 | 1.97 | 0.65 | 3.79*** |
| The ways of assessment influence students' investment on their study time. | 2.00 | 0.55 | 2.14 | 0.63 | -2.03* |
| Student self-assessment helps students monitor their own learning progress. | 2.22 | 0.53 | 2.13 | 0.54 | 1.33 |
| Student peer-assessment can help students develop skills of evaluating and giving feedback. | 2.05 | 0.65 | 2.19 | 0.57 | -2.03* |

*Note.* * $p < .05$; ** $p < .01$; *** $p < .001$.

Results of independent samples t-test showed that Taiwanese students endorsed more on the following items:

a) Assessment should include the aspects of cognitive, skill and attitude.

b) Assessment should include students' self-assessment, peer assessment and teacher assessment.

c) Assessment should include oral and written assessment.

d) Assessment should include both qualitative and quantitative assessment.

e) Once students achieve certain expected achievement, they should be given correspondingly high scores.

f) The ways of assessment influence students' investment on their study time.

g) Student peer-assessment can help students develop skills of evaluating and giving feedback.

Results of independent samples t-test showed that U.S. students endorsed more on the following items:

a) The grading weight of course work should be higher than final examination.

b) Students' grades should be given according to a specific distribution system (e.g. 20% of A).

c) Assessment can promote teachers' teaching.

d) Assessment can reflect teaching effectiveness.

## Discussion

This study is limited due to the representativeness of the U.S. sample and the comparableness of the U.S. sample with the Taiwanese sample. We have tried to draw a representative sample from Taiwan because Taiwan is relatively small and we had access to the educational institutions in Taiwan. United States, however, is so large that it is impossible for us to collect a representative sample. Moreover, although results of statistical tests showed no significant differences between the Taiwanese and U.S. samples with regard to grade level and gender, these two samples are qualitatively different because of cultural background and previous educational experience. Therefore, readers should be cautious when interpreting the results from this study. The U.S. sample only represents education-majors in a southeastern university. Another limitation is that the data are self-reports and we are aware of the limitations of using self-report data. Still another limitation is the use of quantitative data analysis method only. Qualitative approaches such as interviews and classroom observations would be promising ways to triangulate the findings and understand more deeply of student perspectives.

### Results for Research Question 1

No statistically significant differences between students in Taiwan and the United States were found in their report of the form of assessment method used in their classrooms, which means that the commonly used assessment forms (i.e., teacher assessment, student peer assessment, student self-assessment, and joint assessment of teacher and students) are the same in Taiwan and the United States. This result suggests that the concept and methods of assessment are not unique in either context and that educators have communicated very well in this field. This finding challenged the claim from a previous study that formative assessment was not widely practiced or clearly understood in Taiwan (Berry, 2011). We believe that this is due to the recent innovations and exposure to U.S. education in Taiwan (Poole, 2016).

Results also showed that teacher assessment was the dominant form of assessment in both Taiwan and the United States, which echoed previous studies (e.g., Huang & Asghar 2018; Isaacs,

2013). As there were limited studies in the comparison between Taiwan and the United States in the forms of classroom assessment, future researchers should continue this line of research and do more comparison studies.

### Results for Research Question 2

Significant differences were noticed in the classroom assessment activities in Taiwan and the United States. In Taiwan, attendance, classroom participation, homework, reading reports, and quizzes and exams were used more often. In the United States, learning diary, essays, display, and projects were used more often to assess student learning. This finding indicates that classrooms in Taiwan are still more reliant on traditional assessment measures (quizzes and exams) and repetition (homework and attendance). This could be evidence that the reform efforts in Taiwan to encourage a more student-centered approach (Poole, 2016) are not having a significant impact on what students are experiencing in the classroom. There is a lack of full understanding among teachers of what formative assessment should look like in the classroom (Gu, 2014), and the results of this study show that the assessments being implemented in Taiwan still rely on more traditional and well-known methods.

These differences also reflect the contexts in each region. Taiwan has a history with deep roots in testing culture, and schools in Taiwan centered their teaching on the memorization of facts for high-stakes assessments (Berry, 2011). High-stakes testing plays a more significant role in Taiwan for college application than that in the United States, and the culture in Taiwan has thousands of years in history of using tests to select government officials (Berry, 2011; Brown & Gao, 2015). Education in Taiwan is the only option for upward economic mobility for the majority (Zhao & Qiu, 2012). No wonder students in Taiwan place more value on the use of quizzes and exams.

### Results for Research Question 3

The comparison between undergraduate and graduate students in both Taiwan and the United States suggested some trend in the differences: undergraduate students cared more about their grades and graduate students cared more about the quality of learning. What is common between them is the quality of assessment, which is consistent with the literature in that both undergraduate and graduate students had a high degree of agreement about what qualifies as effective assessment feedback and process (Ferguson, 2011).

### Results for Research Question 4

There were no significant differences between Taiwanese and U.S. students in their perspectives of the purpose of assessment even at the item level, which means that all students agree with the purpose of assessment. For the use of assessment in grading, students in Taiwan endorsed more on the item “Once students achieve certain expected achievement, they should be given correspondingly high scores.” This suggests that students in Taiwan value grades/scores more, which is an indication for surface learning. U.S. students, however, endorsed this item less, which is an indication for deep learning.

U.S. students endorsed more on the items that assessments can promote teaching and reflect teaching effectiveness, which is not surprising because a growing number of districts were including formative assessment in their strategies for enhancing student learning more than a decade ago (Lee et al., 2020; Sharkey & Murnane, 2006). Student attitudes in Taiwan about formative assessment were more focused on attaining a grade as a result of achieving success on

an assessment than their American counterparts. American students showed more of an attitude on how learning can be improved through the assessment process.

The use of assessment in the United States varies greatly, and only a small portion of teachers were effectively using them to improve student instruction (Marsh, 2007; Sharkey & Murnane, 2006). Many teachers were not trained in how to interpret student data to improve their instruction (Sharkey & Murnane, 2003). These challenges of implementation and the impacts of NCLB have pushed education in the United States to focus more on assessment of learning and away from formative assessment practices. This study suggests that U.S. students believe that assessment can promote the teachers' instruction and that assessment can reflect teaching effectiveness. Therefore, results from this study inform stake holders and policy makers about students' perceptions of the purpose, content, and use of assessment with the hope that all students learn effectively.

As for the contents of assessment, students in Taiwan endorsed more on the perception that assessments should include aspects of cognition, skill, and attitude. This result is consistent with results from a previous study in that students believed that a variety of cognitive processes should be assessed including reproduction, critical thinking, problem solving, and comprehension (van de Watering et al., 2008). Students in Taiwan also agreed more that assessments should include self-assessment, peer assessment, and teacher assessment, oral and written assessment, and both quantitative and qualitative assessments. The results differ from a former study in that students have identified oral examination, group work, and evaluating the work of others as undesirable forms of assessment (Furnham et al., 2011; van de Watering et al., 2008).

Assessment is usually used to increase student engagement because of the notion that assessment is for learning (Holmes, 2015). We argue that assessment is itself a learning process (i.e., assessment as learning). Formative assessment provides data to the instructor so that the instructor adjusts his/her teaching based on the results of formative assessment. Instructors can also tailor instruction to meet the needs of individual students. Therefore, students will be more likely engaged in learning and make progress in academic studies.

## Conclusions and Implications

The significance of this study lies in the implications for policymakers and teaching practices in both Taiwan and the United States. Taiwan has more recently been encouraging the use of assessment for earning in the curriculum, but educators in Taiwan are facing many of the same challenges when it comes to achieving full implementation (Berry, 2011). Results from this study provided information about student perceptions of formative and summative assessment, which might be helpful for the instructors in Taiwan to meet the challenges and use both formative and summative assessment in their classrooms efficiently.

Results of this study also showed that undergraduate students value their grades very much, which suggests that policymakers should put less weight on the use of grades in graduate school admission or job market. Policies should be made to guide all students (undergraduate and graduate) to pay more attention to the learning experience as well as the lifelong learning habit.

# References


Anderson, P. S. (1987). *Comparison of student attitudes about seven formats of educational testing with emphasis on the MDT multi-digit testing technique.* Paper presented at the Annual Meeting of the Mid-Western Educational Research Association, Chicago, IL.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5–25. https://doi.org/10.1080/03075071003731135

Berry, R. (2011). Educational assessment in mainland China, Hong Kong and Taiwan. In R. Berry & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (49–62). Springer.

Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher Education*, *33*(1), 71–84.

Birenbaum, M. (2007). Assessment and instruction preferences and their relationship with test anxiety and learning strategies. *Higher Education*, *53*, 749–768. https://doi.org/10.1007/s10734-005-4843-4

Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Black, P. & William, D. (2005). Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal, 16*9, 249–261. https://doi.org/10.1080/09585170500136218

Brown, H. D. (2004). *Language assessment: Principals and classroom practices.* Longman.

Brown, G. T., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, *2*(1), 1-19. doi: 10.1080/2331168X.2014.993836

Chou, M.-H. (2019). The impact of the English listening test in the high-stakes national entrance examination on junior high school students and teachers. *International Journal of Listening, 35,* 53–71. https://doi.org/10.1080/10904018.2019.1659142

Chou, Z., & Ching, G. S. (2012). *Taiwan education at the crossroad: When globalization meets localization.* Palgrave McMillan. http://dx.doi.org/10.1080/0046760x.2014.913698

Dunn, K. E., & Mulvenon, S. W. (2019). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, *14*(7), 1–11. https://doi.org/10.7275/jg4h-rb87

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education, 36*(1), 51–62. https://doi.org/10.1080/02602930903197883.

Flaitz, J. (2011). Assessment for learning: U.S. perspectives. In R. Berry & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 33–48). Springer.

Furnham, A., Batey, M., & Martin, N. (2011). How would you like to be evaluated? The correlates of students' preferences for assessment methods. *Personality and Individual Differences*, *50,* 259–263.

Ghaicha, A. (2016). Theoretical framework for educational assessment: A synoptic review. *Journal of Education and Practice*, *7*(24), 212–231.

Gibbs, G. (2006). How assessment frames student learning. In C. Bryan & K. Clegg (Eds.) *Innovative assessment in higher Education* (pp. 23–36). Routledge.

Gilles, J. L., Detroz, P., & Blais, J. G. (2011). An international online survey of the practices and perceptions of higher education professors with respect to the assessment of learning in the classroom. *Assessment & Evaluation in Higher Education, 36,* 719–733. https://doi.org/10.1080/02602938.2010.484880

Grosas, A. B., Raju, S. R., Schuett, B. S., Chuck, J. A., & Millar, T. J. (2016). Determining if active learning through a formative assessment process translates to better performance in summative assessment. *Studies in Higher Education, 41,* 1595–1611. https://doi.org/10.1080/03075079.2014.988704

Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a foreign language in a Chinese secondary school? *Assessment in Education: Principles, Policy & Practice, 21,* 286–305. https://doi.org/10.1080/0969594X.2013.836076

Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development, 20*(1), 53–70. https://doi.org/10.1080/07924360120043658

Holmes, N. (2015). Student perceptions of their learning and engagement in response to the use of a continuous e-assessment in an undergraduate module. *Assessment & Evaluation in Higher Education, 40*(1), 1–14. https://doi.org/10.1080/02602938.2014.881978

Holzinger, A., Lettner, S., Steiner-Hofbauer, V., & Melser, M. C. (2020). How to assess? Perceptions and preferences of undergraduate medical students concerning traditional assessment methods. *BMC Medical Education, 20*(312), 1–7.

Huang, Y.-S., Asghar, A. (2016). Science education reform in Confucian learning cultures: policymakers' perspectives on policy and practice in Taiwan. *Asia-Pacific Science Education, 2*(3), 1–22. https://doi.org/10.1186/s41029-016-0010-8

Huang, Y.-S., Asghar, A. (2018). Science education reform in Confucian learning cultures: Teachers' perspectives on policy and practice in Taiwan. *Cultural Studies of Science Education, 13*(1), 101–131.

Isaacs, T. (2001). Entry to university in the United States: The role of SATs and advanced placement in a competitive sector. *Assessment in Education: Principles, Policy & Practice, 8,* 391–406. https://doi.org/10.1080/09695940120089161

Isaacs, T. (2013). 150 years of statewide assessment in New York: Are the regents examinations still fit for purpose? *Assessment in Education: Principles, Policy, & Practice, 21,* 344–357. https://doi.org/10.1080/0969594X.2013.841641

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and practice, 30*(4), 28–37. https://doi.org/10.1111/j.1745-3992.2011.00220.x

Law, W. W. (2004). Translating globalization and democratization into local policy: Educational reform in Hong Kong and Taiwan. *International Review of Education, 50*(5–6), 497–524.

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education, 33*(2), 124-140. https://doi.org/10.1080/08957347.2020.1732383

Li, H. (2016). How is formative assessment related to students' reading achievement? Findings from PISA 2009. *Assessment in Education: Principles, Policy & Practice, 23*, 473–494. https://doi.org/10.1080/0969594X.2016.1139543

Lindblom-Ylänne, S., Pihlajamäki, H., & Kotkas, T. (2006). Self-, peer-and teacher-assessment of student essays. *Active Learning in Higher Education, 7*(1), 51–62. https://doi.org/10.1177/1469787406061148

Lu, T., Cowie, B., & Jones, A. (2010). Senior high school student biology learning in interactive teaching. *Research in Science Education, 40*, 267–289. https://doi.org/10.1007/s11165-008-9107-8

Marsh, C. J. (2007). A critical analysis of the use of formative assessment in schools. *Educational Research for Policy and Practice, 6*(1), 25–29. https://doi.org/10.1007/s10671-007-9024-z

McMillan, J. H. (2013). *Classroom assessment: Principles and practice for effective standards-based instruction*. Pearson.

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education, 9.* https://doi.org/10.1002/rev3.3292

National Center for Education Statistics (2021). *Digest of Education Statistics: 2021*. Institute of Education Sciences.

No Child Left Behind Act (2002). P.L. 107–110, 20 U.S.C. § 6319.

Peterson, M. W., & Einarson, M. K. (2001). What are colleges doing about student assessment: Does it make a difference? *The Journal of Higher Education, 72*(6), 629–669.

Phillips, F. (1999). Business students' learning preferences and associated task performance. *Journal of Education for Business*, *75*(1), 27–32. https://doi.org/10.1080/0883239909598986

Pham, T. H. T., & Renshaw, P. (2015). Formative assessment in Confucian heritage culture classrooms: Activity theory analysis of tensions, contradictions and hybrid practices. *Assessment & Evaluation in Higher Education, 40*(1), 45-59. https://doi.org/10.1080/02602938.2014.886325

Poole, A. (2016). 'Complex teaching realities' and 'deep rooted cultural traditions': Barriers to the implementation and internalization of formative assessment in China. *Cogent Education, 3*(1), 1-14. doi:10.1080/2331168X.2016.1156242

Ramsden, P. (1997). The context of learning in academic departments. In F. Marton, D. Hounsell, & N.J. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education* (198–217). Scottish American Press.

Rieg, S. A. (2007). Classroom assessment strategies: what do students at-risk and teachers perceive as effective and useful? *Journal of instructional Psychology*, *34*, 214–226.

Rowntree, D. (1987). *Assessing students: How shall we know them?* Routledge.

Sharkey, N. S. & Murnane, R. J. (2003). Learning from student assessment results. *Educational Leadership, 61*(3), 77–81.

Sharkey, N. S. & Murnane, R. J. (2006). Tough choices in designing a formative assessment system. *American Journal of Education, 112*, 572–588. https://doi.org/10.1086/505060

Stiggins, R. J. (2001). *Student-involved classroom assessment.* Prentice-Hall.

Stiggins, R. J. (2002). *Leadership for excellence in assessment-2003: A planning guide to the perfect assessment system.* Portland, OR: Assessment Training Institute.

Stödberg, U. (2012). A research review of e-assessment. *Assessment & Evaluation in Higher Education, 37*(5), 591–604. Taylor & Francis.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, *30*, 325–341. https://doi.org/10.1080/02602930500099102

Tan, C., & Chua, C. S. (2015). Education policy borrowing in China: Has the West wind overpowered the East wind? *Compare: A Journal of Comparative and International Education*, *45*, 686-704. doi:10.1080/03057925.2013.871397

Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education*, *32*, 321–345. https://doi.org/10.1007/BF00138870

Tsai, C.-C., & Kuo, P.-C. (2008). Cram school students' conceptions of learning and learning science in Taiwan. *International Journal of Science Education, 30*(3), 351–373.

Tweed, R. G., & Lehman, D. R. (2002). Learning considered within a cultural context: Confucian and Socratic approaches. *American Psychologist*, *57*(2), 1–107. https://doi.org/10.1037/0003-066X.57.2.89

van den Berg, I., Admiraal, W., & Pilot, A. (2006). Design principles and outcomes of peer assessment in higher education. *Studies in Higher Education*, *31*, 341–356. https://doi.org/10.1080/03075070600680836

van de Watering, G., Gijbels, D., Dochy, F., & Van der Rijt, J. (2008). Students' assessment preferences, perceptions of assessment and their relationships to study results. *Higher Education, 56*, 645–658. https://doi.org/10.1007/s10734-008-9116-6

Vanderlelie, J. J., & Alexander, H. G. (2016). Learning-oriented assessment increases performance and written skills in a second year metabolic biochemistry course. *Biochemistry and Molecular Biology Education, 44*, 412–420.

Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment & Evaluation in Higher Education*, *34*, 221–230. https://doi.org/10.1080/02602930801955986

Xiao, Y., Cai, Y., Ge, Q., Yang, Y. (2022). The potential of using formative assessment to enhance academic achievement in the Confucian-heritage culture: A comparison between Hong Kong and Shanghai. *The Asia-Pacific Education Researcher,* https://doi.org/10.1007/s40299-022-00702-0

Yang, J. H., Huang, I. T. C., & Aldridge, J. M. (2002). Investigating factors that prevent science teachers from creating positive learning environments in Taiwan. In S. W. Goh & M. Khine (Eds.), *Studies in educational learning environments: An international perspective* (pp. 217–224). World Scientific International Publishers. doi:10.1142/97898127771330010

Yin, X., & Buck, G. A. (2015). There is another choice: An exploration of integrating formative assessment in a Chinese high school chemistry classroom through collaborative action research. *Cultural Studies of Science Education, 10*, 719–752. https://doi.org/10.1007/s11422-014-9572-5

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477–501. https://doi.org/10.1023/A:102396702

Zhao, Y., & Qiu, W. (2012). Policy changes and educational reforms in China: decentralization and marketization. *On the Horizon, 20*, 313-323. doi:10.1108/10748121211272452

# Appendix

## Questionnaire of University Students' Perspectives on Classroom Assessment

This questionnaire consists of 3 sections as follows:

1. Basic information (3 questions);
2. Your experience of assessment at your university (2 questions);
3. Your perspectives on classroom assessment (20 questions);

1. Basic Information

1.1. Where is your university/institution?

1) Taiwan 2) United States

1.2. You are:

1) Freshman 2) Sophomore 3) Junior 4) Senior 5) Master's student
6) PGCE (postgraduate certificate in education) student Other：

1.3. Gender：

1) Male 2) Female

2. Your experiences of assessment at your university

2.1. What are the methods your instructor applies to the assessment of learning this semester? (Select all that apply)

1) Teacher-assessment 2) Student peer assessment
3) Student self-assessment 4) Combined teacher-student assessment

2.2. What are the items your instructor includes in the assessment of learning this semester? (Select all that apply)

1) Attendance
2) Class participation
3) Homework
4) Learning diary
5) Learning portfolios / e-portfolios
6) Essay
7) Display
8) Project
9) Oral presentation
10) Reading report / summary
11) Pencil-paper examination (e.g. quizzes, tests, mid-term and final examination)
12) Other:

3. Your perspectives on classroom assessment (20 questions)
   Do you agree with the following statements?
   Please rate the following statements with:
   0) Strongly Disagree; 1) Disagree; 2) Agree; 3) Strongly Agree

**Attitude toward the Purpose of Assessment**

1) The aims of assessment are to equip students with the lifelong learning knowledge, skills and attitude.
2) The aims of assessment are to help students improve their quality of learning.
3) The aims of assessment are to differentiate students 'achievement for admissions to advanced learning or employment.
4) The aims of assessment are to provide evidences for the society to determine whether the curriculum is appropriate or not.
5) The aims of assessment are to determine whether students can finally pass the course.
6) The aims of assessment are to provide feedback to students so that they can know their learning progress and further learning direction.

**Attitudes toward the Forms of Assessment**

1) Classroom assessment should include the aspects of cognitive, skill and attitude.
2) Classroom assessment should include all the following elements (students' self-assessment, peer assessment and teacher assessment).
3) Classroom assessment should include oral and written assessment.
4) Classroom assessment should include both qualitative and quantitative assessment.
5) Apart from marks or ranking, students should be given written feedback.

**Attitude toward the Use of Assessment in Grading**

1) Clear and objective standards should be established when conducting assessment.
2) The grading weight of course work (exclusive of final score) should be higher than final examination.
3) Students' grades should be given according to a specific distribution system (e.g. 20% of A；50% of B，5% of Failed).
4) Once students achieve certain expected achievement, they should be given correspondingly high scores.

**Attitude toward the Impact of Assessment**

1) Assessment can promote teachers' teaching.
2) Assessment can reflect teaching effectiveness.
3) The ways of assessment influence students' investment on their study time.
4) Student self-assessment helps students monitor their own learning progress.
5) Student peer-assessment can help students develop skills of evaluating and giving feedback.