

**Teacher Assessment Conceptions and Practices  
at an Independent School**



VANDERBILT  
UNIVERSITY®

John Comforto

A capstone project submitted in partial fulfillment of the requirements  
for the degree of Doctor of Education in Leadership and Learning in Organizations

August 1, 2022

## Contents

<b>Executive Summary</b>	<b>3</b>
<b>Introduction</b>	<b>5</b>
<b>Organizational Context</b>	<b>6</b>
<b>Area of Inquiry</b>	<b>7</b>
<b>Literature Review</b>	<b>9</b>
Beliefs and Practices	9
Assessment Conceptions	10
Assessment Practices	11
<b>Conceptual Framework</b>	<b>15</b>
<b>Project Questions</b>	<b>17</b>
<b>Data Collection</b>	<b>18</b>
Section 1 - Demographics	18
Section 2 - Assessment Conceptions	18
Sections 3 and 4 - Assessment Practices: Types of Assessment and Grading	19
<b>Data Analysis</b>	<b>20</b>
<b>Findings</b>	<b>24</b>
Finding 1	24
Finding 2	27
Finding 3	40
Summary	43
<b>Recommendations</b>	<b>47</b>
Recommendation 1	47
Recommendation 2	48
Recommendation 3	49
<b>Conclusion</b>	<b>50</b>
<b>References</b>	<b>51</b>
<b>Appendix A</b>	<b>56</b>
<b>Appendix B</b>	<b>62</b>

## Executive Summary

The Jade Coast Academy (pseudonym) is a K-12 independent school. In recent years, leaders at The Jade Coast Academy have noted the highly variable nature of classroom assessment practices at the school. The purpose of this study is to better understand and analyze that variability with attention to the relationship between the conceptions that teachers have about assessment (the meaning that teachers make about the nature and purpose of assessment) and the assessment practices that they employ in the classroom (the various ways that they measure and report student learning).

To examine assessment conceptions and practices, a closed-ended survey was administered to teachers at the school in the service of the following project questions:

- Project Question 1** In what ways do assessment conceptions vary across the institution?
- Project Question 2** In what ways do assessment practices vary across the institution?
- Project Question 3** Does a relationship exist between assessment conceptions and practices at the institution?

Survey responses were analyzed for significant differences in conceptions and practices by division, department, and years of teaching experience. Correlational analysis was performed to identify relationships between conceptions and practices at the school.

A summary of the findings, organized by project question, is presented below:

**Finding 1** | Analysis of the survey responses revealed that the most prominent conception of assessment at the school is one that situates assessment as a tool for the improvement of teacher instruction and student learning. A significant difference in conceptions was observed between the upper and middle school, where teachers in the upper school report a stronger orientation towards the improvement conception than teachers in the middle school.

**Finding 2** | Eleven assessment practices showed significant differences by division, department, and/or years of teacher experience. The most pronounced difference relates to the way that teachers calculate their semester grades, where teachers with fewer than four years of experience rely far less heavily on academic performance (as opposed to classroom behavior) than teachers with more than four years of experience. Practices that show the highest degree of school-wide variability, as measured by their standard deviation, include the usage of major examinations in the classroom and the assignment of zeros for missing work.

**Finding 3** | Twelve practices showed moderate-to-strong relationships with assessment conceptions, suggesting that a connection does exist between conceptions and practices at the

school. The strongest positive relationship exists between the improvement conception and the usage of major examinations.

Taken together, the findings support the anecdotal observations of leaders at the school regarding variability in assessment practices. To create greater coherence in assessment practices, the following recommendations are offered:

**Recommendation 1** | Develop and articulate an assessment policy that a) outlines the school's philosophy on and expectations around assessment, b) acknowledges the intentional differences in pedagogy and curriculum that exist between divisions and even departments, and c) considers a balance between professional autonomy and institutional alignment.

**Recommendation 2** | Engage teachers in reflection and community activity around assessment that a) is sustained over time, b) aims to make assessment conceptions explicit, and c) creates opportunities for teachers to participate in conversation about their assessment practices.

**Recommendation 3** | Examine and harness relationships between assessment conceptions and practices, starting with a review of the relationships that do not align with expectation, such as that between the improvement conception and the usage of major examinations in the classroom.

## Introduction

Assessment is the process of collecting, analyzing, and evaluating evidence of student learning (Remesal, 2011). Classroom assessment—a term that denotes the range of ways teachers measure learning in the classroom (Popham, 2020)—is the most common form of assessment, consuming approximately one-third of a teacher’s time (Brookhart et al., 2006; Stiggins & Conklin, 1992). This study examines classroom assessment from the perspective of teachers at The Jade Coast Academy, an independent school in [REDACTED].

Classroom assessment has been described as a final, closely guarded frontier of teacher autonomy in a profession increasingly subject to external directives (Feldman, 2019). Amid more and more regulations and policies, assessment stands out as a place where teachers can exercise professional judgment and enact their educational philosophies. Indeed, teachers have deeply held beliefs about assessment, usually encased in strong emotion and almost always grounded in their own experience of assessment as students (Barnes et al., 2017; Crossman, 2007). Their decisions around assessment reflect their perspectives on teaching and learning, their views on how best to prepare students for the future, and their self-concepts as educators (Xu & Brown, 2016).

Based on their unique beliefs, teachers make highly individualized decisions about assessment in the classrooms (McMillan, 2001). To the extent that assessment decisions vary from one teacher to the next, students are likely to encounter multiple approaches to assessment in a given year and across multiple years. Such variability between teachers can generate confusion about the nature and purpose of assessment, as well as the meaning of the grades assigned to assessments.

In recent years, leaders at The Jade Coast Academy noted the highly variable nature of assessment practices at the school. The purpose of this study is to better understand and analyze that variability with attention to the relationship between the beliefs that teachers have about assessment and the assessment practices that they employ in the classroom.

### **Organizational Context**

Founded in [REDACTED], The Jade Coast Academy is a K-12, co-ed day school in [REDACTED]. The school serves 1,200 students across three divisions: lower school, middle school, and upper school. Since its founding, the school has been guided by its mission to [REDACTED]. The spiritual component of the educational program draws on teachings from the [REDACTED] faith, though the school welcomes students of all faiths.

While an emphasis on [REDACTED] ties the three divisions together, each fosters a unique learning environment. [REDACTED]

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]. The

student-to-teacher ratio across all divisions is 7:1.

The Jade Coast Academy employs 180 teachers. Teachers in the middle school and upper school serve in one of several departments, including classical languages, computer science and engineering, English, history, mathematics, religion, science, performing arts, visual arts, and world languages.

Teachers at The Jade Coast Academy are supported by the school's Center for Teaching and Learning, the purpose of which is to promote sustained professional growth in teaching. The Center is staffed by two teacher-leaders whose job is to offer personalized, one-on-one support, host on-campus workshops and seminars, and curate an on-campus education resource library.

The coordinators of the Center for Teaching and Learning serve as this study's primary stakeholders. Findings from the study could inform the nature and purpose of assessment-related professional development offered by the Center. Because the coordinators collaborate closely with the school's academic administration, findings could also inform the shape and direction of assessment policies at the school.

### **Area of Inquiry**

Variability in assessment practices is a common phenomenon in educational settings (DeLuca et al., 2018). On the one hand, variability is the natural byproduct of settings that create space for and value teacher autonomy. Given the relationship between teacher autonomy and motivation, job satisfaction, and professional empowerment (Pearson & Moomaw, 2005), there is good reason for schools to respect teacher autonomy. On the other hand, variability between teachers can also raise important questions about the presence of a shared vision for teaching and learning at a school, as well as the equitable treatment of students in the same grade level with different teachers (Wiggins & McTighe, 2007).

Leaders at The Jade Coast Academy report fielding such questions about assessment at the school. The COVID-19 pandemic, which put a spotlight on teaching practices via remote instruction, has recently fueled these inquiries. However, variability in assessment practices predates the pandemic. While formal reviews have not been conducted at the school to determine the extent of the variability and its consequences, anecdotal observations by administrators, in

addition to their conversations with teachers, students, and parents, suggest that assessment practices vary in some measure by teacher.

This study proposes to understand the variability at The Jade Coast Academy by examining teacher assessment beliefs, teacher assessment practices, and the relationship between the two. To the degree that the current state of variability represents a problem to be solved, leaders at the school could use the findings to cultivate closer alignment in practices between teachers. Professional development efforts aimed at changing assessment practices typically neglect to consider teacher beliefs (Remesal, 2011). However, because of the power of beliefs to guide and shape practice, those beliefs represent a critical point of departure in terms of fully understanding the range of practices at the school, as well as designing effective professional development activities that might affect changes in practice.



## **Literature Review**

Literature was reviewed in accordance with the primary areas of inquiry: assessment beliefs, assessment practices, and the relationship between beliefs and practices. Articles and other resources were gathered through Google Scholar using keyword searches for assessment, assessment beliefs, assessment conceptions, assessment practices, and equitable assessment practices. Literature was selected by reviewing titles and abstracts with attention to relevance to the primary areas of inquiry. Reference lists of selected articles and other resources were reviewed to identify additional relevant literature. The section below reviews the extant literature by first taking up the critical connection between beliefs and practices.

### **Beliefs and Practices**

Teachers' beliefs about teaching and learning affect their classroom practices (Pajares, 1992). According to Kagan (1992), teacher beliefs exist “at the very heart of teaching” (p. 85). These beliefs serve as filters for new information and guides for future action (Fives & Buehl, 2014). Attention to teacher beliefs is important because research indicates that teacher beliefs influence not only how teachers teach but also how students perform (Brown, 2004; Muis & Foy, 2010).

The relationship between beliefs and practice is complex and iterative. Individual teacher beliefs exist in larger systems of beliefs, where the centrality of a given belief mediates its impact on practice (Fives & Buehl, 2014; Pajares, 1992). At the same time, teacher practice serves as a site of professional learning in its own right and therefore shapes beliefs (Kagan, 1992). The link between belief and practice can be both supported and interrupted by external factors, such as socio-cultural and policy contexts (Bonner, 2016; Davis & Neitzel, 2011; James & Pedder, 2006).

## Assessment Conceptions

The research focused on teacher assessment beliefs has gained momentum in the last twenty years, premised on the view that teacher assessment beliefs operate like other beliefs inasmuch they influence professional decision-making and activity in the given domain (Opre, 2015). Beliefs about assessment can frame the way teachers design assessments, interpret assessment results, and process information about alternate assessment practices and principles (Barnes et al., 2017; Brown, 2008; Vandeyar & Killen, 2007). These beliefs have cognitive and affective dimensions: the cognitive aspect relates to what teachers believe is true and false about assessment, whereas the affective aspect concerns the emotions that teachers have about assessment based on their personal and professional experiences (Xu & Brown, 2016).

Brown (2004) identified four conceptions that teachers might hold about assessment. The term conception is meant to encompass beliefs. It is situated as a more general mental structure or “framework through which a teacher views, interprets, and interacts with the teaching environment” (Brown, 2004, p. 303). Building on the research base related to assessment purposes, Brown (2004) used a combination of teacher surveys, exploratory factor analysis, and structural equation modeling to develop and validate a model that contains four assessment conceptions. The conceptions are listed below and explained in the paragraph that follows:

1. Assessment improves teacher instruction and student learning.
2. Assessment makes students accountable for their learning.
3. Assessment makes schools accountable for student learning.
4. Assessment is irrelevant to the work of teachers and the life of students.

The first conception holds that assessment provides valuable information about the quality of instruction and student learning. The information gained through assessment can be used to

provide student feedback and make instructional adjustments as necessary. The second conception positions students as responsible for their learning with assessment acting as a tool for assigning grades, categorizing students based on performance, and evaluating readiness for promotion. The third conception proposes that the role of assessment is to appraise the performance of schools and their teachers to ensure the effective use of society's resources. And the final conception maintains that assessment is not a legitimate part of the teaching and learning process because it diverts time away from instruction and lacks validity and reliability.

Teacher conceptions are socio-cultural constructions (van den Berg, 2002) and idiosyncratic expressions of values and experience (Cizek et al., 1995). Research has demonstrated that teachers can hold multiple—sometimes competing—conceptions of assessment simultaneously (Barnes et al., 2017). A teacher may therefore view assessment as both an instrument for improving student learning and as irrelevant. These apparent contradictions reflect the reality that assessment is often used in multiple ways and for multiple purposes by education systems and teachers alike (Brown, 2012). Moreover, as socio-cultural constructions, assessment conceptions are not immune to the influence of the shifting circumstances and systems in which they take shape (Vandeyar & Killen, 2007).

### **Assessment Practices**

Balancing internal assessment conceptions with external factors, teachers make individualized and idiosyncratic decisions about assessment practices (McMillan, 2005). Given the teacher-specific nature of assessment decisions, assessment practices have been shown to vary widely from teacher to teacher, even when characteristics such as setting, years of experience, and familiarity with assessment policies are taken into account (Cizek et al., 1995). Still, Vandeyar and Killen (2007) found that observations of teacher assessment practices

provided an accurate window into their assessment conceptions. Here assessment practices include both how students are assessed and how students are graded.

Assessments vary in form and function. The research literature has identified two primary types of assessment: formative and summative (Popham, 2020). Formative assessment is the process by which teachers collect evidence of learning to adjust ongoing instructional procedures. Summative assessment, by contrast, takes place when educators evaluate student learning to inform decisions about instructional activities already completed. There is an extensive base of research literature on these two assessment types, as well as the degree to which they have at times become confused and entangled in both theory and practice (Harlen & James, 1997; Looney et al., 2018).

Delanshere and Jones (1999) outlined the belief structures that underpin the two assessment approaches. A view of learning that centers on the acquisition of facts, rules, and skills calls for summative forms of assessment that sanction and verify student learning. This view of learning aligns with Brown's (2004) conceptions of assessment that emphasize student and school accountability. Alternatively, a view of learning that foregrounds constant student development demands a formative approach to assessment that features ongoing and descriptive feedback, aligning more closely with Brown's (2004) conception of assessment as a tool for improving teaching and learning. However, because teacher beliefs are complex and multifaceted, and since teachers may hold multiple beliefs at once, researchers warn against the strict classification of teachers along dichotomous lines (Remesal, 2011). Indeed, teachers may strategically deploy formative and summative assessments at different points in a given instructional unit.

Grading is the process by which teachers assign symbols to individual assessments and composite measures of student performance on report cards. In a review of over one hundred years of grading research, Brookhart et al. (2016) noted a great deal of variation in the grades that teachers assign to students. Major sources of variation include differences in grading criteria, as well as differences in teacher severity or leniency. Additional variation can be attributed to the multidimensional nature of grades, reflecting not only achievement on assessments but also the nonachievement factors valued by a given teacher. Examining the relative weight assigned by teachers to nonachievement factors in their grading, McMillan (2001) found that teachers emphasize “academic enablers,” such as effort, work habits, attention, and participation, over other nonachievement factors, such as personality and behaviors. Taken together, the studies reviewed by Brookhart et al. (2016), including the McMillan (2001) study, portray grades as variable, multidimensional constructs that combine achievement and nonachievement factors in accordance with the individualized beliefs of the teachers who assign them.

Multidimensional grades, also known as “hodgepodge” grades, contradict the recommendations of measurement experts and professional developers, who hold that grades should be based on academic performance alone (Brookhart, 1991; Feldman, 2019; Guskey, 2000; O’Connor, 2011). Concerns about multidimensional grades have typically centered on the reliability, validity, and interpretability of grades that combine multiple factors. More recently, the inclusion of nonachievement factors has been criticized as an inequitable practice that draws on bias-prone evaluations of student behavior (Feldman, 2019). Such criticisms posit that a unidimensional grade, based just on academic performance, addresses issues of validity and equity by emphasizing content mastery and ignoring nonachievement factors.

Extensive research on assessment practices has revealed their wide-ranging impact on student achievement, metacognitive abilities, motivation, and self-perception (DeLuca et al., 2018). Such research provides an incentive for schools and other educational organizations to not only foster coherence in their assessment practices but also to do so around the practices that are most strongly associated with the positive effects above.

## Conceptual Framework

At the core of the reviewed literature is the complex and iterative relationship between assessment conceptions and practices. The proposed conceptual framework, adapted from Xu and Brown (2016), makes visible the relationship between conceptions and practice and situates that relationship in the service of assessor identity (re)construction. The construct of identity (re)construction describes the ongoing negotiation of teachers' roles vis-a-vis assessment as they reflect on their practice, interact with others, and engage with innovations in assessment. The framework visualizes the relationship of the constructs under review and provides a roadmap for changing conceptions, practice, and ultimately identity.

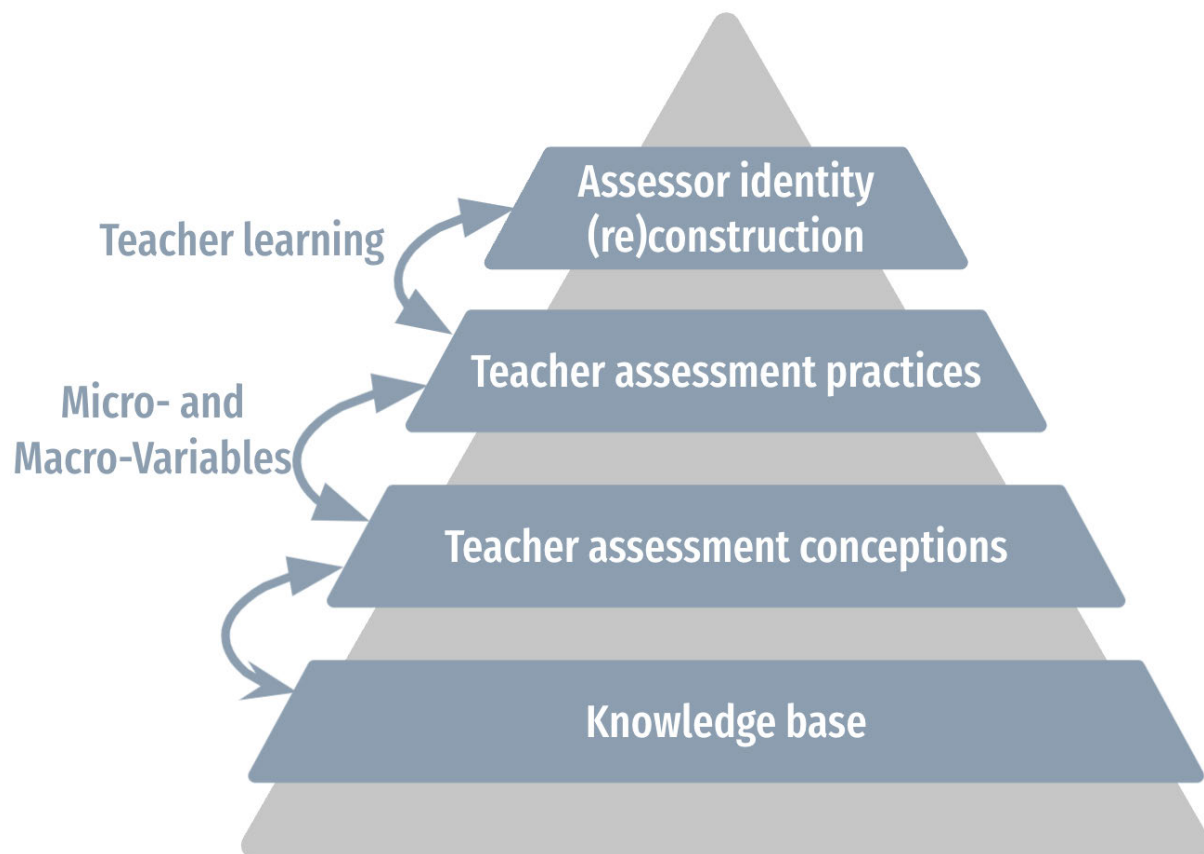


Figure 1: Conceptual Framework

The location of the knowledge base at the bottom of the framework represents its foundational relationship with the other components. Here the knowledge base represents the body of knowledge related to assessment that is available to a teacher in a given context. What teachers assimilate from the knowledge base depends on their conceptions of assessment, which act as an interpretive and guiding framework. Micro- and macro-contextual variables mediate the relationship between conceptions and practice. Whereas micro-variables represent the affordances and restraints of the immediate workplace, macro-variables refer to the influence of larger social, political, and cultural contexts. These variables re-cast assessment practice as a site of compromise, where teachers balance external factors and internal conceptions. Such compromises may even require teachers to experience a gap between their conceptions and practices. The framework proposes that teacher learning takes place when teachers reflect on their assessment decisions in the context of the given micro- and macro-variables. To the extent that such learning leads to change, teachers re(construct) their identities as assessors in the light of how they understand their assessment work and see themselves as assessors.

The double-sided arrows reflect the reciprocal nature of the relationship between the components. Assessment conceptions, informed by a base of knowledge and influenced by micro- and macro-variables, guide assessment practices, and the learning that happens around assessment practices sets the stage for the re(construction) of identities. At the same time, (re)constructed identities influence practice, and altered practices generate new insights into assessment conceptions, which renew and extend the knowledge base. The bottom-up flow is animated by research-based theories and principles, whereas the top-down flow stems from identities that are situated and developed in context.



The framework developers identified two primary ways of promoting teacher learning around assessment: a) reflective practice and b) participation in community activities (Xu & Brown, 2016). Reflection invites examination of deeply-held assessment conceptions and creates opportunities for unlearning and relearning. Community activities advance the work of reflection by exposing teachers to alternative perspectives and generating interpersonal dialogue. Taken together, reflection and community activities empower teachers to re(construct) their individual assessment identities inside of contextualized communities of practice. In-service professional development activities related to assessment are vital because most teachers lack extensive training in the area (DeLuca, 2012).

### **Project Questions**

The project sought to address the following questions:

- In what ways do assessment conceptions vary across the institution?
- In what ways do assessment practices vary across the institution?
- Does a relationship exist between assessment conceptions and practices at the institution?

## **Data Collection**

The study utilized quantitative methods to examine teacher assessment conceptions, practices, and the relationship between the two at The Jade Coast Academy. A convenience sample was taken by emailing an online survey instrument to all active teachers at the school (Appendix A). A total of 47 responses were received, amounting to a response rate of 26%.

The study replicates the design of a study completed by Calveric (2010) on assessment conceptions and practices in a sample of public school teachers in Virginia. Calveric utilized a three-section survey instrument, with the first section focused on teacher demographics, the second on assessment conceptions, and the third on assessment practices. In examining assessment practices, Calveric surveyed teachers exclusively on the types of assessments used in their classrooms. Because this study also includes grading as an area of assessment practice, this study replicates and expands on Calveric's design by using a four-section survey. The first section centers on demographics, the second assessment conceptions, the third assessment practices related to assessment type, and the fourth assessment practices related to grading.

### **Section 1 - Demographics**

The demographics section includes three questions. Teachers were asked to indicate the division in which they teach, the subject area in which they teach, and their total number of years of teaching experience.

### **Section 2 - Assessment Conceptions**

Brown's (2006) *Teachers' Conceptions of Assessment III (CoA-III) Inventory* was adapted for the second section of the survey instrument. The CoA-III includes 27 items that load onto one of four factors, thus structuring four conceptions of assessment: assessment improves

teacher instruction and student learning; assessment makes students accountable for their learning; assessment makes schools accountable for student learning; and assessment is irrelevant to the work of teachers and the life of students. Brown (2006) used confirmatory factor analysis to validate the instrument in New Zealand and Australia. However, after administering the CoA-III in other jurisdictions, including Hong Kong, Cyprus, Catalonia, Egypt, India, and Ecuador, Brown et al. (2019) found that the four-factor statistical model validated in New Zealand and Australia is not universal or generalizable. In other words, there is no global model for the structures of teacher assessment conceptions. They concluded that the inventory could only be used cross-culturally after adaptations account for local context and culture.

The researcher adjusted Brown's (2006) model to account for The Jade Coast Academy's status as an independent school that does not participate in [REDACTED] standardized testing program. Accordingly, items that originally loaded onto the factor related to school accountability were not included in the survey instrument. The result was a three-factor model featuring the following assessment conceptions: assessment improves teacher instruction and student learning; assessment makes students accountable for their learning; and assessment is irrelevant to the work of teachers and the life of students. The items were scored on a Likert-type scale from 1 to 5 (1 = strongly disagree and 5 = strongly agree). Appendix B shows the items used in the survey and their associated factors in Brown's statistical model.

### **Sections 3 and 4 - Assessment Practices: Types of Assessment and Grading**

The final two sections of the survey instrument were adapted from an assessment practices questionnaire introduced by McMillan et al. (2002). Two pilot tests were used to strengthen the validity and reliability of the questionnaire. The final version of the questionnaire included items in three categories: items assessing different types of assessment used, items

assessing different factors used to determine grades, and items assessing the cognitive level of the assessments. The items assessing different types of assessment used constitute section 3 of this study's survey instrument, and items assessing different factors used to determine grades constitute section 4 of the instrument. Certain items were either adjusted or removed to fit the independent school setting. Items assessing the cognitive level of assessments were not used. Items in section 3 are scored on a Likert-type scale from 1 to 5 (1 = not important and 5 = very important). Additionally, items in section 4 are scored on a Likert-type scale from 1 to 5 (1 = not at all and 5 = entirely).

### **Data Analysis**

Survey responses were collected in Qualtrics and transferred to R Studio for analysis. Forty-seven surveys were completed by teachers at The Jade Coast Academy, but response rates vary by question as all questions were marked optional. Descriptive statistics were used to analyze demographics by division, subject area, and years of teaching experience (Table 1).

Table 1

*Participant demographic descriptions*

Variable	<i>n</i>	%
<b>Division</b>		
Lower School	16	34.04
Middle School	13	27.66
Upper School	18	38.30
<b>Subject Area</b>		
Classics or World Languages	9	19.57
English, History, or Religion	9	19.57
Lower School Homeroom	11	23.91
Math, Science, or Computer Science and Engineering	15	32.61
Performing Arts or Visual Arts	2	4.35
<b>Years of Experience</b>		
Fewer than 4	6	12.77
Between 4 and 10	7	14.89
Between 11 and 20	18	38.30
More than 20	16	34.04

A combination of descriptive and inferential statistics were used to analyze the data in the sections of the survey that followed the demographic section. Table 2 outlines the relationship between project questions, survey sections, and data analysis methods.

Table 2

*Project Questions, Survey Sections, and Data Analysis*

Project Question	Survey Sections	Data Analysis
1. In what ways do assessment conceptions vary across the institution?	2	Descriptive Statistics (Means, Standard Deviations, Frequencies) and Inferential Statistics (ANOVA, Post hoc)
2. In what ways do assessment practices vary across the institution?	3 and 4	Descriptive Statistics (Means, Standard Deviations, Frequencies) and Inferential Statistics (ANOVA, Post hoc)
3. Does a relationship exist between assessment conceptions and practices at the institution?	2, 3, and 4	Inferential Statistics (Pearson Correlation)

The section below provides a more detailed overview of the data analysis procedures, organized by project question.

**Project Question 1:** A subscore for each of the three assessment conceptions was calculated for each respondent by taking the average of the responses to the items associated with each conception. Then, to analyze the way conceptions vary across the institution, conditional means and standard deviations were calculated for assessment conceptions by division, subject area, and years of teaching experience. Finally, to identify statistically significant differences in assessment conceptions by division, subject area, and years of teaching experience, Analysis of Variance (ANOVA) tests were conducted.

**Project Question 2:** Conditional means and standard deviations were calculated for each assessment practice by division, subject area, and years of teaching experience. To identify statistically significant differences in assessment practices by division, subject area, and years of teaching experience, Analysis of Variance (ANOVA) tests were conducted.

**Project Question 3:** A Pearson Correlation was used to measure the relationship between

assessment conceptions and assessment practices. Results of the analysis indicate the direction and strength of the relationships.

## Findings

Project Question 1: In what ways do assessment conceptions vary across the institution?

### Finding 1

Teachers at the Jade Coast Academy agreed most strongly with the conception that assessment improves teaching and learning, followed by the conception that assessment makes students accountable, and finally the conception that assessment is irrelevant (Table 3).

Table 3

#### *Descriptive Statistics for Conception Factors*

Conceptions	<i>M</i>	<i>SD</i>
Makes Students Accountable	3.00	0.59
Improves Teaching and Learning	3.90	0.49
Is Irrelevant	2.44	0.48

Table 4 provides conception factor means and standard deviations by division. A trend is observed for the accountability factor, where conception scores increase from the lower school to the upper school. The other conceptions do not reveal a clear trend by division.

Table 4

#### *Conception Factor Means and Standard Deviations by Division*

Conceptions	Lower School		Middle School		Upper School	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Makes Students Accountable	2.79	0.61	3.03	0.55	3.17	0.59
Improves Teaching and Learning	3.88	0.46	3.60	0.56	4.13	0.35
Is Irrelevant	2.39	0.38	2.63	0.63	2.34	0.42

A one-way ANOVA at alpha level .01 revealed significant differences between scores for the improvement factor by division,  $F=5.29$ ,  $p<0.01$  (Table 5). Post Hoc analyses (Tukey's HSD) revealed that improvement scores were significantly higher in the upper school ( $M = 4.13$ ) than



the middle school ( $M = 3.60$ ). No other significant differences were observed by division.

Table 5

*ANOVA for Conception Factors by Division*

Conceptions	<i>F</i>	Sig
Makes Students Accountable	1.76	0.19
Improves Teaching and Learning	5.29	0.009**
Is Irrelevant	1.47	0.24

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$ .

Table 6 provides conception factor means and standard deviations by department. The math, science, and computer science and engineering departments show the highest scores for the accountability and improvement conceptions, whereas the classics and world language departments show the highest scores for the irrelevant conception. A one-way ANOVA did not reveal any significant differences in conception scores by department.

Table 6

*Conception Factor Means and Standard Deviations by Department*

Conceptions	Classics or World Language		English, History, or Religion		Lower School Homeroom		Math, Science, or Computer Science and Engineering		Performing Arts or Visual Arts	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Makes Students Accountable	2.89	0.69	3.07	0.74	2.82	0.50	3.16	0.53	2.83	0.71
Improves Teaching and Learning	3.73	0.54	3.83	0.62	3.76	0.44	4.18	0.30	4.00	0.47
Is Irrelevant	2.70	0.60	2.56	0.41	2.47	0.40	2.17	0.43	2.36	0.30

Table 7 provides conception factor means and standard deviations by years of teaching experience. Average accountability scores increase with years of experience up to 20 years of experience, whereas improvement scores decrease with years of experience up to 20 years of experience.

Table 7

*Conception Factor Means and Standard Deviations by Years of Experience*

Conceptions	Fewer than 4		Between 4 and 10		Between 11 and 20		More than 20	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Makes Students Accountable	2.72	0.53	3.19	0.47	3.24	0.57	2.75	0.59
Improves Teaching and Learning	4.06	0.32	3.95	0.66	3.82	0.45	3.90	0.54
Is Irrelevant	2.36	0.15	2.31	0.53	2.43	0.52	2.54	0.51

A one-way ANOVA at alpha level .05 revealed significant differences between scores for the accountability factor by years of experience,  $F=2.94$ ,  $p<0.05$  (Table 8). Though the F-statistic is significant, post hoc analyses did not reveal any pairs with significant differences.

Table 8

*ANOVA for Conception Factors by Years of Experience*

Conceptions	<i>F</i>	Sig
Makes Students Accountable	2.94	0.04*
Improves Teaching and Learning	0.36	0.78
Is Irrelevant	0.44	0.72

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$ .

Project Question 2: In what ways do assessment practices vary across the institution?

## **Finding 2**

The findings below detail the ways that assessment practices vary across the institution.

Assessment practices are organized in two categories: assessment type and grading

### **Assessment Type**

Teachers rated the importance of each of the ten assessment types in the context of their larger assessment toolkit. Table 9 provides school-wide means and standard deviations for assessment type scores, where a score of 1 indicates ‘Not Important’ and a score of 5 indicates ‘Very Important.’ Teachers at The Jade Coast Academy identified authentic assessments as the most important assessment type in their assessment toolkit, whereas assessments provided by publishers represent the least important assessment type. Scores for major examinations showed the highest standard deviation.

Table 9

#### *Descriptive Statistics for Importance of Assessment Types*

Assessment Type	<i>M</i>	SD
Major examinations	2.96	1.21
Oral presentation	3.29	0.91
Objective assessments	2.8	1.00
Performance assessments	3.89	0.97
Publisher assessments	1.8	0.96
Short answer	3.7	0.81
Projects in teams	3.38	0.90
Projects by self	3.84	0.73
Quizzes	3.53	0.86
Authentic assessments	4.09	1.07

Table 10 provides assessment type means and standard deviations by division. Notably, scores

for both authentic assessments and publisher assessments decrease from the lower school to the upper school.

Table 10

*Assessment Type Means and Standard Deviations by Division*

Assessment Type	Lower School		Middle School		Upper School	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Major examinations	2.73	0.96	2.31	1.11	3.65	1.22
Oral presentation	3.40	0.91	3.08	1.19	3.35	0.70
Objective assessments	3.00	0.76	2.46	1.05	2.88	1.17
Performance assessments	3.87	0.92	3.85	0.90	3.94	1.14
Publisher assessments	2.13	0.83	1.85	1.14	1.47	0.87
Short answer	3.36	0.93	3.69	0.85	4.00	0.61
Projects in teams	3.53	0.83	3.38	0.96	3.24	0.97
Projects by self	4.00	0.65	3.77	0.83	3.76	0.75
Quizzes	3.67	0.72	3.46	0.97	3.47	0.94
Authentic assessments	4.40	0.91	4.31	0.95	3.65	1.22

A one-way ANOVA at alpha level .01 revealed significant differences between scores for the major examinations by division,  $F=5.83$ ,  $p<0.01$  (Table 11). Post Hoc analyses (Tukey's HSD) revealed that major examination scores were significantly higher in the upper school ( $M = 3.65$ ) than the middle school ( $M = 2.31$ ). No other significant differences were observed by division.

Table 11

*ANOVA for Assessment Practices by Division*

Assessment Type	<i>F</i>	Sig.
Major examinations	5.83	0.006**
Oral presentation	0.48	0.62
Objective assessments	1.08	0.35
Performance assessments	0.04	0.96
Publisher assessments	1.97	0.15
Short answer	2.51	0.09
Projects in teams	0.42	0.66
Projects by self	0.49	0.63
Quizzes	0.26	0.78
Authentic assessments	2.45	0.10

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$

Table 12 provides assessment type means and standard deviations by department. Publisher assessments, the least important school-wide assessment type, are considered most important by teachers of lower school homeroom of all the departments. Authentic assessments, the most important school-wide assessment type, are also considered most important by teachers of lower school homeroom of all the departments.

Table 12

*Assessment Type Means and Standard Deviations by Department*

Assessment Type	Classics or World Language		English, History, or Religion		Lower School Homeroom		Math, Science, or Computer Science and Engineering		Performing Arts or Visual Arts	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Major examinations	2.00	1.20	3.33	1.41	2.90	0.88	3.47	1.06	2.50	0.71
Oral presentation	3.00	1.31	3.33	0.87	3.70	0.82	3.13	0.74	3.00	1.41
Objective assessments	2.13	1.13	2.67	1.41	3.00	0.82	3.07	0.80	3.00	0.00
Performance assessments	3.75	1.39	4.11	1.17	4.00	0.94	3.67	0.72	4.50	0.71
Publisher assessments	1.50	1.07	1.33	0.71	2.20	0.79	2.07	1.10	1.50	0.71
Short answer	3.38	0.74	3.78	0.83	3.67	0.71	4.00	0.65	2.50	2.12
Projects in teams	2.63	1.06	3.78	0.97	3.60	0.84	3.47	0.64	2.50	0.71
Projects by self	3.25	0.71	4.22	0.97	4.00	0.67	3.87	0.52	3.50	0.71
Quizzes	3.63	0.92	3.11	0.93	3.90	0.57	3.60	0.91	2.50	0.71
Authentic assessments	3.75	1.28	3.89	1.54	4.50	0.71	4.20	0.77	3.50	2.12

Three of the ten assessment types showed significant differences by department (Table 13).

Performing arts and visual arts were excluded from the analysis because of their insufficient sample size. The three significant ANOVAs are presented in decreasing order of significance.

The first significant ANOVA at alpha level .05 revealed differences between scores for major examinations by department,  $F=3.21$ ,  $p<0.05$ . Post hoc analyses (Tukey's HSD) revealed that major examination scores were significantly higher in math, science, and computer science and engineering ( $M = 3.47$ ) than in classics and world language ( $M = 2.00$ ). The second significant ANOVA at alpha level .05 revealed differences between scores for projects in teams by department,  $F= 3.01$ ,  $p<0.05$ . Post hoc analyses (Tukey's HSD) revealed that team project scores were significantly higher in English, history, and religion ( $M = 3.78$ ) than in classics and world language ( $M = 2.63$ ). The third significant ANOVA at alpha level .05 revealed differences

between scores for projects by self (individual projects) by department,  $F=2.94$ ,  $p<.05$ . Post hoc analyses (Tukey's HSD) revealed that individual project scores were significantly higher in English, history and religion ( $M = 4.22$ ) than in classics and world language ( $M = 3.25$ )

Table 13

*ANOVA for Assessment Type by Department*

Assessment Type	<i>F</i>	Sig.
Major examinations	3.21	0.03*
Oral presentation	1.09	0.365
Objective assessments	1.67	0.19
Performance assessments	0.45	0.72
Publisher assessments	1.94	0.14
Short answer	1.35	0.27
Projects in teams	3.01	0.04*
Projects by self	2.94	0.05*
Quizzes	1.40	0.26
Authentic assessments	0.91	0.44

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$

Table 14 provides assessment type means and standard deviations by years of experience. No significant differences were observed by years of experience.

Table 14

*Assessment Type Means and Standard Deviations by Years of Experience*

Assessment Type	Fewer than 4		Between 4 and 10		Between 11 and 20		More than 20	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Major examinations	2.50	0.55	3.43	1.40	2.94	1.24	2.94	1.34
Oral presentation	3.00	0.89	3.43	1.13	3.63	0.72	3.00	0.97
Objective assessments	3.17	0.75	2.71	1.50	2.81	0.91	2.69	1.01
Performance assessments	3.83	1.17	4.43	1.13	4.06	0.57	3.50	1.10
Publisher assessments	1.83	0.75	1.57	0.98	1.88	1.09	1.81	0.98
Short answer	3.17	1.17	4.00	0.82	4.00	0.63	3.47	0.74
Projects in teams	3.00	0.63	3.57	0.98	3.75	0.68	3.06	1.06
Projects by self	3.67	0.82	4.14	0.69	4.00	0.52	3.63	0.89
Quizzes	3.50	1.05	3.29	0.95	3.75	0.68	3.44	0.96
Authentic assessments	4.33	1.21	3.43	1.72	4.25	0.77	4.13	0.96

Standard deviations were analyzed to identify areas of high variability institution-wide and within groups. Major examination scores showed the highest standard deviation institution-wide, as well as within the lower school, the upper school, teachers with between 11 and 20 years of experience, and teachers with more than 20 years of experience.

### **Grading**

Teachers marked the extent to which they base their semester grades on each of the sixteen grading areas, where a score of 1 represents “Not at all” and a score of 5 represents “Entirely.” Table 15 provides school-wide means and standard deviations for grading areas scores. Results show that teachers base their semester grades most heavily on learning mastery, whereas non-academic extra credit weighs least heavily in semester grades. Scores for zeros for missing assignments showed the highest standard deviation.



Table 15

*Descriptive Statistics for Grading Areas*

Grading Areas	<i>M</i>	SD
Disruptive behavior	1.88	0.95
Improvement in performance	3.23	0.88
Effort	3.12	0.89
Ability level	3.27	0.83
Work habits and neatness	2.65	0.83
Grades of other teachers	1.47	0.73
Homework completion	2.63	0.99
Homework quality	2.72	0.84
Academic performance	3.84	0.53
Performance compared to class	2.02	0.85
Performance compared to prior classes	1.67	0.80
Learning mastery	3.93	0.59
Participation and attention	2.95	0.94
Zeros for missing assignments	2.33	1.18
Extra credit (non-academic)	1.23	0.56
Extra credit (academic)	1.88	0.89

Table 16 provides grading area means and standard deviations by division. Three of the sixteen practices show trends across the divisions: disruptive behavior scores decrease from the lower school to the upper school, homework completion scores increase from the lower school to the upper school, and homework quality scores increase from the lower school to the upper school.

Table 16

*Grading Area Means and Standard Deviations by Division*

Grading Area	Lower School		Middle School		Upper School	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Disruptive behavior	2.50	0.94	1.67	0.89	1.53	0.80
Improvement in performance	3.64	0.84	2.92	1.00	3.12	0.78
Effort	3.57	0.85	2.75	1.14	3.00	0.61
Ability level	3.36	0.84	2.92	1.08	3.47	0.52
Work habits and neatness	3.07	0.73	2.25	0.87	2.59	0.80
Grades of other teachers	1.50	0.94	1.17	0.39	1.65	0.70
Homework completion	2.14	1.03	2.67	1.23	3.00	0.61
Homework quality	2.29	0.91	2.92	0.90	2.94	0.66
Academic performance	3.50	0.52	4.00	0.60	4.00	0.35
Performance compared to class	1.93	0.83	1.83	0.94	2.24	0.83
Performance compared to prior classes	1.71	0.91	1.25	0.62	1.94	0.75
Learning mastery	4.07	0.47	3.67	0.65	4.00	0.61
Participation and attention	3.29	0.91	2.67	1.15	2.88	0.78
Zeros for missing assignments	1.50	0.94	2.83	1.03	2.65	1.17
Extra credit (non-academic)	1.14	0.53	1.50	0.80	1.12	0.33
Extra credit (academic)	1.64	0.92	2.17	0.58	1.88	1.05

Four of the sixteen grading areas showed significant differences by division (Table 17). The four significant ANOVAs are presented in decreasing order of significance. The first significant ANOVA (at alpha level .01) revealed differences between scores for zeros for missing assignments by division,  $F=6.39$ ,  $p<0.01$ . Post hoc analyses (Tukey's HSD) revealed that zeros for assignment scores were significantly higher in both the middle school ( $M = 2.83$ ) and the upper school ( $M = 2.65$ ) than the lower school ( $M = 1.50$ ). The second significant ANOVA (at alpha level .01) revealed differences between scores for disruptive behavior by division,  $F=5.27$ ,  $p<0.01$ . Post hoc analyses (Tukey's HSD) revealed that disruptive behavior scores were significantly higher in the lower school ( $M = 2.50$ ) than the upper school ( $M = 1.53$ ). The third

significant ANOVA (at alpha level .05) revealed differences between scores for academic performance by division,  $F=4.97$ ,  $p<0.05$ . Post hoc analyses (Tukey's HSD) revealed that academic performance scores were significantly higher in both the middle school ( $M = 4.00$ ) and the upper school ( $M = 4.00$ ) than the lower school ( $M = 3.50$ ). The fourth significant ANOVA (at alpha level .05) revealed differences between work habits and neatness scores by division,  $F=3.54$ ,  $p<0.05$ . Post hoc analyses (Tukey's HSD) revealed that work habit scores were significantly higher in the lower school ( $M = 3.07$ ) than the middle school ( $M = 2.25$ ).

Table 17

*ANOVA for Grading Area by Division*

Grading Area	<i>F</i>	Sig.
Disruptive behavior	5.27	0.009**
Improvement in performance	2.53	0.09
Effort	3.19	0.05
Ability level	1.60	0.214
Work habits and neatness	3.54	0.04*
Grades of other teachers	1.57	0.22
Homework completion	3.11	0.05
Homework quality	2.95	0.06
Academic performance	4.97	0.01*
Performance compared to class	0.89	0.42
Performance compared to prior classes	2.82	0.07
Learning mastery	1.76	0.19
Participation and attention	1.48	0.239
Zeros for missing assignments	6.39	0.004**
Extra credit (non-academic)	1.92	0.16
Extra credit (academic)	1.09	0.35

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$

Table 18 provides grading area means and standard deviations by department. ANOVA calculations revealed one significant difference by department: disruptive behavior ( $F=4.03$ ,

$p < 0.05$ ). Post Hoc analyses (Tukey's HSD) revealed that disruptive behavior scores were significantly higher in lower school homeroom ( $M = 2.44$ ) than math, science, computer science, and engineering ( $M = 1.33$ ).

Table 18

*Grading Area Means and Standard Deviations by Department*

Grading Area	Classics or World Language		English, History, or Religion		Lower School Homeroom		Math, Science, or Computer Science and Engineering		Performing Arts or Visual Arts	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Disruptive behavior	1.63	0.92	2.22	1.09	2.44	0.88	1.33	0.62	3.00	0.00
Improvement in performance	3.00	0.93	3.33	0.71	3.56	1.01	3.00	0.93	4.00	0.00
Effort	2.75	1.16	3.33	0.50	3.44	1.01	2.87	0.83	4.00	0.00
Ability level	3.57	0.79	3.11	0.78	3.33	1.00	3.22	0.80	3.00	1.41
Work habits and neatness	2.13	0.99	2.78	0.44	3.11	0.93	2.53	0.83	3.00	0.00
Grades of other teachers	1.25	0.46	1.44	0.73	1.44	1.01	1.60	0.74	1.50	0.71
Homework completion	2.88	1.25	2.56	1.01	2.44	1.01	2.73	0.88	2.00	1.41
Homework quality	2.63	0.92	3.11	0.60	2.44	0.88	2.73	0.96	2.50	0.71
Academic performance	4.13	0.64	3.89	0.33	3.56	0.53	3.93	0.46	3.00	0.00
Performance compared to class	1.63	0.52	2.22	1.09	2.22	0.83	2.07	0.88	1.50	0.71
Performance compared to prior classes	1.50	0.76	1.56	0.73	1.78	0.97	1.87	0.83	1.00	0.00
Learning mastery	3.88	0.64	3.89	0.78	3.89	0.33	4.00	0.53	4.00	1.41
Participation and attention	2.50	1.20	3.44	0.53	3.22	1.09	2.67	0.82	3.50	0.71
Zeros for missing assignments	2.38	1.30	2.78	1.20	1.44	1.01	2.60	1.06	2.00	1.41
Extra credit (non-academic)	1.25	0.71	1.33	0.71	1.00	0.00	1.27	0.59	1.50	0.71
Extra credit (academic)	2.38	0.92	1.89	0.93	1.44	0.88	1.93	0.88	1.50	0.71

Table 19 provides grading area means and standard deviations by years of experience. Trends across experience levels were observed in four grading areas. The extent to which teachers base their semester grades on a) disruptive behavior decreases with increasing years of experience, b) effort decreases with increasing years of experience, c) participation and attention decreases with

increasing years of experience, and d) academic extra credit increases with increasing years of experience.

Table 19

*Grading Area Means and Standard Deviations by Years of Experience*

Grading Area	Fewer than 4		Between 4 and 10		Between 11 and 20		More than 20	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Disruptive behavior	3.00	0.63	2.14	1.07	1.73	0.88	1.47	0.74
Improvement in performance	3.83	0.41	3.14	1.07	3.27	0.70	3.00	1.07
Effort	3.83	0.41	3.14	1.07	3.13	0.64	2.80	1.08
Ability level	3.17	0.75	3.14	1.21	3.47	0.52	3.15	0.99
Work habits and neatness	3.00	0.00	2.43	0.79	2.80	0.86	2.47	0.99
Grades of other teachers	1.83	1.17	1.29	0.49	1.47	0.74	1.40	0.63
Homework completion	2.33	0.82	2.00	1.15	3.07	1.03	2.60	0.83
Homework quality	2.83	0.41	2.43	1.40	2.73	0.80	2.80	0.77
Academic performance	3.17	0.41	4.00	0.00	3.80	0.56	4.07	0.46
Performance compared to class	1.83	0.75	1.86	0.90	2.13	0.83	2.07	0.96
Performance compared to prior classes	1.33	0.82	1.86	0.90	1.60	0.74	1.80	0.86
Learning mastery	4.00	0.63	4.14	0.90	3.80	0.56	3.93	0.46
Participation and attention	3.50	0.55	3.14	0.69	2.87	0.99	2.73	1.10
Zeros for missing assignments	1.67	1.03	2.57	1.51	2.93	1.10	1.87	0.92
Extra credit (non-academic)	1.17	0.41	1.29	0.76	1.20	0.56	1.27	0.59
Extra credit (academic)	1.50	0.84	1.86	0.90	1.87	0.99	2.07	0.88

Three of the sixteen grading areas showed significant differences by years of experience (Table 20). The three significant ANOVAs are presented in decreasing order of significance. The first significant ANOVA (at alpha level .01) revealed differences between scores for academic performance by years of experience,  $F=5.88$ ,  $p<0.01$ . Post hoc analyses (Tukey's HSD) revealed that academic performance scores were significantly higher for teachers with between 4 and 10 years of experience ( $M = 4.00$ ), teachers with between 11 and 20 years of experience ( $M = 3.90$ ),

and teachers with more than 20 years of experience ( $M = 4.07$ ) than teachers with fewer than 4 years of experience ( $M = 3.17$ ). The second significant ANOVA (at alpha level .01) revealed significant differences between scores for disruptive behavior by years of experience,  $F=5.15$ ,  $p<0.01$ . Post hoc analyses (Tukey's HSD) revealed that disruptive behavior scores were significantly higher for teachers with fewer than 4 years of experience ( $M = 3.00$ ) than for teachers with between 11 and 20 years of experience ( $M = 1.73$ ) and teachers with more than 20 years of experience ( $M = 1.47$ ). The third significant ANOVA (at alpha level .05) revealed differences between scores for zeros for missing assignment by years of experience,  $F=3.20$ ,  $p<0.05$ . Though the F-statistic is significant, post hoc analyses did not reveal any pairs with significant differences.

Table 20

*ANOVA for Grading Area by Years of Experience*

Grading Area	<i>F</i>	Sig.
Disruptive behavior	5.15	0.004**
Improvement in performance	1.30	0.29
Effort	2.00	0.129
Ability level	0.42	0.74
Work habits and neatness	0.90	0.45
Grades of other teachers	0.66	0.58
Homework completion	2.24	0.10
Homework quality	0.33	0.8
Academic performance	5.88	0.002**
Performance compared to class	0.27	0.85
Performance compared to prior classes	0.62	0.61
Learning mastery	0.55	0.65
Participation and attention	1.07	0.37
Zeros for missing assignments	3.20	0.03*
Extra credit (non-academic)	0.08	0.97
Extra credit (academic)	0.54	0.652

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$

Standard deviations were analyzed to identify areas of high variability institution-wide and within groups. Scores for zeros for missing assignments showed the highest standard deviation institution-wide, as well as within the upper school, the classics and world languages departments, the English, history and religion departments, the math science, and computer science and engineering departments, the performing arts and visual arts departments, teachers with between 4 and 10 years of experience, and teachers with between 11 and 20 years of experience.

Project Question 3: Does a relationship exist between assessment conceptions and practices at the institution?

### Finding 3

Correlational analysis found twelve moderate-to-strong relationships between assessment conceptions and practices at The Jade Coast Academy. The two groups of assessment practices (assessment type and grading area) were treated separately. Table 21 provides the correlation coefficients for the relationships between assessment conceptions and assessment types.

Table 21

*Correlation of Assessment Conception Subgroups and Assessment Types*

Assessment Type	Makes Students Accountable	Improves Teaching and Learning	Is Irrelevant
Major examinations	0.26	0.54**	-0.40**
Oral presentation	0.29	0.11	-0.30*
Objective assessments	0.42**	0.34*	-0.28
Performance assessments	0.23	0.38*	-0.12
Publisher assessments	0.09	0.13	-0.32*
Short answer	0.37*	0.27	-0.24
Projects in teams	-0.11	-0.21	-0.05
Projects by self	-0.09	-0.05	-0.06
Quizzes	0.17	0.21	-0.19
Authentic assessments	-0.21	-0.21	-0.01

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$ .

The following significant relationships were identified, listed in decreasing order of significance:

- A strong positive relationship between the improvement conception and major examinations ( $r = 0.54$ ). The stronger the improvement conception, the more important major examinations become in the teacher's assessment toolkit.
- A moderate positive relationship between the accountability conception and objective



assessments ( $r = 0.42$ ). The stronger the accountability conception, the more important objective assessments become in the teacher's assessment toolkit.

- A moderate negative relationship between the irrelevant conception and major examinations ( $r = -0.40$ ). The stronger the irrelevant conception, the less important major examinations become in the teacher's assessment toolkit.
- A moderate positive relationship between the improvement conception and performance assessments ( $r = 0.38$ ). The stronger the improvement conception, the more important performance assessments become in the teacher's assessment toolkit.
- A moderate positive relationship between the accountability conception and short answer assessments ( $r = 0.37$ ). The stronger the accountability conception, the more important short answer assessments become in the teacher's assessment toolkit.
- A moderate positive relationship between the improvement conception and objective assessments ( $r = 0.34$ ). The stronger the improvement conception, the more important performance assessments become in the teacher's assessment toolkit.
- A moderate negative relationship between the irrelevant assessment conception and publisher assessments ( $r = -0.32$ ). The stronger the irrelevant conception, the more important publisher assessments become in the teacher's assessment toolkit.
- A moderate negative relationship between the irrelevant assessment conception and oral presentations ( $r = -0.30$ ). The stronger the irrelevant conception, the more important oral presentations become in the teacher's assessment toolkit.

Table 22 provides the correlation coefficients for the relationships between assessment conceptions and grading areas

Table 22

*Correlation of Assessment Conception Subgroups and Grading Areas*

Grading Area	Makes Students Accountable	Improves Teaching and Learning	Is Irrelevant
Disruptive behavior	0.06	-0.1	-0.04
Improvement in performance	0.00	-0.15	0.07
Effort	-0.08	0.04	-0.07
Ability level	-0.10	-0.13	0.31*
Work habits and neatness	-0.03	-0.13	-0.02
Grades of other teachers	0.22	0.17	-0.02
Homework completion	0.37*	0.11	-0.07
Homework quality	0.28	-0.04	0.02
Academic performance	0.23	-0.14	0.40**
Performance compared to class	0.03	-0.01	-0.03
Performance compared to prior classes	-0.05	0.14	-0.03
Learning mastery	-0.16	0.22	-0.18
Participation and attention	-0.14	-0.13	-0.03
Zeros for missing assignments	0.41**	-0.02	-0.17
Extra credit (non-academic)	0.25	-0.27	0.03
Extra credit (academic)	0.04	-0.19	0.02

*Note.* \* indicates  $p < .05$ . \*\* indicates  $p < .01$ .

The following significant relationships were identified, listed in decreasing order of significance:

- A moderate positive relationship between the accountability conception and zeros for missing assignments ( $r = 0.41$ ). The stronger the accountability conception, the greater the extent to which teachers base their semester grades on zeros for missing assignments.
- A moderate positive relationship between the irrelevant conception and academic performance ( $r = 0.40$ ). The stronger the irrelevant conception, the greater the extent to which teachers base their semester grades on academic performance.
- A moderate positive relationship between the accountability conception and homework

completion ( $r = 0.37$ ). The stronger the accountability conception, the greater the extent to which teachers base their semester grades on homework completion.

- A moderate positive relationship between the irrelevant conception and ability level ( $r = 0.31$ ). The stronger the irrelevant conception, the greater the extent to which teachers base their semester grades on ability level.

## Summary

The study found several significant differences in conceptions (Project Question 1) and practices (Project Question 2) across the institution. Table 23 summarizes the observed differences in conceptions.

Table 23

### *Significant Differences in Assessment Conceptions*

Conception	Difference By	Sig.	Higher Score ( $M_{\text{Higher}}$ )	Lower Score ( $M_{\text{Lower}}$ )
Improves Teaching and Learning	Division	0.009	Upper school (4.13)	Middle school (3.60)

Table 24 summarizes the observed differences in assessment practices.

Table 24

*Significant Differences in Assessment Practices*

Assessment Practice	Difference By	Sig.	Higher Score (M <sub>Higher</sub> )	Lower Score (M <sub>Lower</sub> )
Academic performance (Grading area)	Years of experience	0.002	Between 4 and 10 (4.00), Between 11 and 20 (3.90), More than 20 (4.07)	Fewer than 4 (3.17)
Zeros for missing assignments (Grading area)	Division	0.004	Middle school (2.83) and upper school (2.65)	Lower school (1.50)
Disruptive behavior (Grading area)	Years of experience	0.004	Fewer than 4 (3.00)	Between 11 and 20 (1.73), More than 20 (1.47)
Major examinations (Assessment type)	Division	0.006	Upper school (3.65)	Middle school (2.31)
Disruptive behavior (Grading area)	Division	0.009	Lower school (2.50)	Upper school (1.53)
Academic Performance (Grading area)	Division	0.01	Middle school (4.00) and upper school (4.00)	Lower school (3.50)
Disruptive behavior (Grading area)	Department	0.01	Lower school homeroom (2.44)	Math, science, computer science, and engineering (1.33)
Major examinations (Assessment type)	Department	0.03	Math, science, computer science, and engineering (3.47)	Classics and world language (2.00)
Projects in teams (Assessment type)	Department	0.04	English, history, and religion (3.78)	World language (2.63)
Work habits and neatness (Grading Area)	Division	0.04	Lower school (3.07)	Middle school (2.25)

Projects by self (Assessment type)	Department	0.05	English, history, and religion (4.22)	World language (3.52)
---------------------------------------	------------	------	--	--------------------------

The study also found several significant relationships between conceptions and practices (Project Question 3). Table 25 summarizes the significant relationships.

Table 25

*Significant Relationships between Assessment Conceptions and Assessment Practices*

Conception	Assessment Practice	Correlation Coefficient	Strength
Improves teaching and learning	Major examinations (Assessment type)	0.54	Strong
Makes students accountable	Objective assessments (Assessment type)	0.42	Moderate
Makes students accountable	Zeros for missing assignments (Grading area)	0.41	Moderate
Is irrelevant	Academic performance (Grading area)	0.40	Moderate
Is irrelevant	Major examinations (Assessment type)	-0.40	Moderate
Improves teaching and learning	Performance assessments (Assessment type)	0.38	Moderate
Makes students accountable	Short answer assessments (Assessment type)	0.37	Moderate
Makes students accountable	Homework completion (Grading area)	0.37	Moderate
Improves teaching and learning	Objective assessments (Assessment type)	0.34	Moderate
Is irrelevant	Publisher assessments (Assessment type)	-0.32	Moderate
Is irrelevant	Ability level (Grading area)	0.31	Moderate

Is irrelevant	Oral presentations (Assessment type)	-0.30	Moderate
---------------	---	-------	----------

Table 26 indicates the assessment type and the grading area with the highest standard deviation school-wide. Also indicated are the other groups within which the two scores showed the highest standard deviation.

Table 26

*Assessment Types and Grading Areas with the Highest Standard Deviation*

Variable	SD
<b>Major Examinations (Assessment Type)</b>	
School-Wide	1.21
Lower School	0.96
Upper School	1.22
Between 11 and 20 Years	1.24
More than 20 Years	1.34
<b>Zeros for Missing Assignments (Grading Area)</b>	
Upper School	1.17
Classics or World Languages	1.30
English, History, or Religion	1.20
Math, Science, or Computer Science and Engineering	1.06
Performing Arts or Visual Arts	1.41
Between 4 and 10 Years	1.51
Between 11 and 20 years	1.10

## Recommendations

### Recommendation 1: Develop and articulate an assessment policy

Assessment conceptions shape practices, but they do not act alone. A range of contextual variables, such as institutional policies, norms, and expectations, influence classroom assessment decisions at a school by setting the boundaries for acceptable and unacceptable practices (Xu & Brown, 2016). Because The Jade Coast Academy does not have a global assessment policy, local norms and expectations—amongst teachers who regularly collaborate, for instance—are likely the strongest moderators of assessment decisions at the school. To address school-wide variability in assessment practices, the Center for Teaching and Learning should collaborate with administrators to develop and articulate an assessment policy. Such a policy could include both a statement of philosophy on assessment and an outline of assessment expectations.

In crafting a school-wide policy, attention should be paid to the intentional differences in pedagogy and curriculum that exist between divisions and even departments. Leaders should review the findings to differentiate between the practices whose high-variability reflects those intentional differences and the practices whose high-variability does not. For instance, the statistically significant difference in the importance of major examinations between the upper and middle schools is likely an intentional difference that accounts for the distinct developmental stages of the two student populations. The assessment policy could recognize such a difference as intentional. On the other hand, the high standard deviation of scores related to giving zeros for missing assignments may reflect variability based on teacher preference alone. In cases such as this, leaders could use the assessment policy to reduce unintended variability.

Teachers may perceive an assessment policy as a threat to professional autonomy, especially if the expectations of the policy are incongruous with their conceptions of assessment

(Xu & Brown, 2016). Because assessment represents a final frontier of professional autonomy for many teachers (Feldman, 2019), leaders at The Jade Coast Academy should carefully consider the balance between autonomy and alignment in classroom practices.

Indeed, the policy may not represent a serious incursion on professional autonomy if it builds on and reinforces existing assessment conceptions. Brown's (2004) research suggests that leaders should introduce policy in a manner that maximizes the association between the most prominent conception at the school (i.e., assessment improves teaching and learning) and the relevant components of the new policy. To the degree that the assignment policy does not overlap entirely with existing conceptions, leaders at the school should pair the policy with appropriate professional development and support, as outlined in the recommendation below.

### **Recommendation 2: Engage teachers in reflection and community activity around assessment**

Professional development can lead to changes in assessment conceptions and practices, but change is often gradual (Bonner, 2016). To support deep learning around assessment, the Center for Teaching and learning should engage teachers in professional development activities sustained over time. Xu and Brown (2016) identified two modes of professional development that support teacher learning: a) reflective practice and b) participation in community activities (See Conceptual Framework). Both could prove useful in the work of the Center for Teacher and Learning with teachers.

In inviting teachers to reflect, the Center for Teaching and Learning should start by making assessment conceptions explicit. Learning is unlikely if teachers do not come to identify their conceptions and understand the way those conceptions operate in their daily practice



(Brown, 2004; Remesal, 2011). With their conceptions made explicit, teachers can reflect on the assumptions that underpin conceptions, the decisions that flow from them, and the experiences and emotions to which they are attached. Such reflection, to the degree that it produces cognitive conflict, can lead to unlearning and relearning about assessment (Xu & Brown, 2016).

Community activities should create opportunities for teachers to engage in conversation about their assessment practices. Because assessment is a socio-cultural activity, teachers stand to learn from the differing conceptions and practices of their peers (Looney et al., 2018). Given the significant differences observed between the practices of less experienced teachers (i.e., fewer than four years of experience) and more experienced teachers, the Center for Teaching and Learning could use community activities to help new teachers acculturate to the school and the proposed assessment policy.

### **Recommendation 3: Examine and harness the relationships between assessment conceptions and practices**

The study found twelve significant relationships between assessment conceptions and practices. Some of the relationships, such as the relationship between the improvement conception and major examinations, do not align with expectation and should be examined.

Brown's (2004) research suggests that high-stakes assessments, such as major examinations, are more typically associated with the accountability conception than the improvement conception. The relationship between the improvement conception and major examinations at The Jade Coast Academy could represent a valid, idiosyncratic relationship at the school, or it could represent a shortcoming of the survey design. By making conceptions explicit and opening professional conversation about assessment practices, the Center for Teacher and Learning could gain insight

into the nature of the relationship at the school.

The observed correlations between conceptions and practices do not imply causation in one direction or the other. However, the literature on the subject suggests a reciprocal relationship between conceptions and practices moderated by contextual variables (Fives & Buehl, 2014; Kagan, 1992; Xu & Brown, 2016). The implication is that changes in conceptions can produce changes in practice and vice versa. The Center for Teaching and Learning, in collaboration with administrators at the school, should take these relationships into account when introducing new policy and designing professional development. Suppose that leaders wished to decrease the use of zeros for missing assignments. A reasonable place to start would be to observe the positive relationship between the accountability conception and the practice of giving zeros for missing assignments. To affect change, then, leaders might design professional development activities to strengthen the improvement conception and attenuate the accountability conception at the school.

### **Conclusion**

The study examined classroom assessment conceptions, practices, and the relationship between the two at The Jade Coast Academy, an independent school in [REDACTED]. Informed by research that posits a link between conceptions and practices, the study utilized an online survey to identify significant differences in both conceptions and practices by division, department, and years of teaching experience. Quantitative analysis found multiple significant differences, some of which reflect expected variability between unique areas of the school, and others of which point to unintentional misalignment. Recommendations for addressing the variability in assessment practices at the school include developing and articulating an

assessment policy, engaging teachers in reflection and community activity around assessment, and examining and harnessing the relationship between assessment conceptions and practices.

### References

- Barnes, N., Fives, H., & Dacey, C. M. (2017). U.S. teachers' conceptions of the purposes of assessment. *Teaching and Teacher Education, 65*, 107–116.  
<https://doi.org/10.1016/j.tate.2017.02.017>
- Bonner, S. M. (2016). Teachers' Perceptions About Assessment: Competing Narratives. In G. Brown & Harris, L (Eds.), *Handbook of human and social conditions in assessment* (pp. 21–39). London University Press.
- Brookhart, S. M. (1991). Grading Practices and Validity. *Educational Measurement: Issues and Practice, 10*(1), 35–36. <https://doi.org/10.1111/j.1745-3992.1991.tb00182.x>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research, 86*(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Brookhart, S. M., Walsh, J. M., & Zientarski, W. A. (2006). The Dynamics of Motivation and Effort for Classroom Assessments in Middle School Science and Social Studies. *Applied Measurement in Education, 19*(2), 151–184.  
[https://doi.org/10.1207/s15324818ame1902\\_5](https://doi.org/10.1207/s15324818ame1902_5)
- Brown, G. (2004). Teachers' Conceptions of Assessment: Implications for Policy and Professional Development. *Assessment in Education Principles Policy and Practice, 11*, 301–318. <https://doi.org/10.1080/0969594042000304609>
- Brown, G. (2006). Teachers' Conceptions of Assessment: Validation of an Abridged Version.

- Psychological Reports*, 99, 166–170. <https://doi.org/10.2466/PR0.99.5.166-170>
- Brown, G. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Nova Science Publishers, Inc.
- Brown, G. (2012). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45–70.  
<https://doi.org/10.18296/am.0097>
- Brown, G., Gebril, A., & Michaelides, M. P. (2019). Teachers' Conceptions of Assessment: A Global Phenomenon or a Global Localism. *Frontiers in Education*, 4, 16.  
<https://doi.org/10.3389/feduc.2019.00016>
- Calveric, S. (2010). Elementary Teachers' Assessment Beliefs and Practices. [*Doctoral Dissertation, Virginia Commonwealth University*], *VCU Scholars Compass*.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' Assessment Practices: Preparation, Isolation, and the Kitchen Sink. *Educational Assessment*, 3(2), 159–179.  
[https://doi.org/10.1207/s15326977ea0302\\_3](https://doi.org/10.1207/s15326977ea0302_3)
- Davis, D. S., & Neitzel, C. (2011). A Self-Regulated Learning Perspective on Middle Grades Classroom Assessment. *The Journal of Educational Research*, 104(3), 202–215.
- DeLuca, C. (2012). Preparing Teachers for the Age of Accountability: Toward a Framework for Assessment Education. *Action in Teacher Education (Association of Teacher Educators)*, 34(5/6), 576–591. <https://doi.org/10.1080/01626620.2012.730347>
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: A large-scale survey. *Assessment in Education: Principles, Policy & Practice*, 25(4), 355–375.  
<https://doi.org/10.1080/0969594X.2016.1244514>

- Feldman, J. (2019). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms* (First edition). Corwin, a SAGE publishing company.
- Fives, H., & Buehl, M. M. (2014). Exploring Differences in Practicing Teachers' Valuing of Pedagogical Knowledge Based on Teaching Ability Beliefs. *Journal of Teacher Education*, 65(5), 435–448. <https://doi.org/10.1177/0022487114541813>
- Guskey, T. R. (2000). Grading Policies that Work Against Standards...and How to Fix Them. *NASSP Bulletin*, 84(620), 20–29. <https://doi.org/10.1177/019263650008462003>
- Harlen, W., & James, M. (1997). Assessment and Learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365–379. <https://doi.org/10.1080/0969594970040304>
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109–138. *Curriculum Journal*, 17, 109–138. <https://doi.org/10.1080/09585170600792712>
- Kagan, D. M. (1992). Implication of Research on Teacher Belief. *Educational Psychologist*, 27(1), 65–90. [https://doi.org/10.1207/s15326985ep2701\\_6](https://doi.org/10.1207/s15326985ep2701_6)
- Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, 25(5), 442–467. <https://doi.org/10.1080/0969594X.2016.1268090>
- McMillan, J. H. (2001). Secondary Teachers' Classroom Assessment and Grading Practices. *Educational Measurement: Issues and Practice*, 20(1), 20–32. <https://doi.org/10.1111/j.1745-3992.2001.tb00055.x>
- McMillan, J. H. (2005). Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice. *Educational Measurement:*

- Issues and Practice*, 22(4), 34–43. <https://doi.org/10.1111/j.1745-3992.2003.tb00142.x>
- McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary Teachers' Classroom Assessment and Grading Practices. *The Journal of Educational Research*, 95(4), 203–213.
- Muis, K. R., & Foy, M. J. (2010). The effects of teachers' beliefs on elementary students' beliefs, motivation, and achievement in mathematics. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal Epistemology in the Classroom* (pp. 435–469). Cambridge University Press. <https://doi.org/10.1017/CBO9780511691904.014>
- O'Connor, K. (2011). *A repair kit for grading: 15 fixes for broken grades* (2nd ed). Pearson.
- Opre, D. (2015). Teachers' Conceptions of Assessment. *Procedia - Social and Behavioral Sciences*, 209, 229–233. <https://doi.org/10.1016/j.sbspro.2015.11.222>
- Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, 62(3), 307–332. <https://doi.org/10.3102/00346543062003307>
- Pearson, L. C., & Moomaw, W. (2005). The Relationship between Teacher Autonomy and Stress, Work Satisfaction, Empowerment, and Professionalism. *Educational Research Quarterly*, 29(1), 38–54.
- Popham, W. J. (2020). *Classroom assessment: What teachers need to know* (Ninth edition). Pearson.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472–482. <https://doi.org/10.1016/j.tate.2010.09.017>
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of*

- classroom assessment*. State University of New York Press.
- van den Berg, R. (2002). Teachers' Meanings Regarding Educational Practice. *Review of Educational Research*, 72(4), 577–625. <https://doi.org/10.3102/00346543072004577>
- Vandeyar, S., & Killen, R. (2007). Educators' conceptions and practice of classroom assessment in post-apartheid South Africa. *South African Journal of Education*, 27, 101–115.
- Wiggins, G. P., & McTighe, J. (2007). *Schooling by design: Mission, action, and achievement*. Association for Supervision and Curriculum Development.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>

## Appendix A

### Online Survey Instrument

#### Section 1

1. In which division do you primarily teach?

- Lower School
- Middle School
- Upper School

2. In which department do you primarily teach?

- Lower School Homeroom
- Classics or World Languages
- English, History, or Religion
- Math, Science, or Computer Science and Engineering
- Performing Arts or Visual Arts
- Not Applicable

3. For how many years have you taught?

- Fewer than 3
- Between 4 and 10
- Between 11 and 20
- More than 20



## Section 2

Please respond to the statements below using your own beliefs about assessment.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Assessment provides feedback to students about their performance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Assessment results should be treated cautiously because of measurement error	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Assessment is a way to place students into categories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Assessment is integrated with teaching practice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Assessment establishes what students have learned	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Assessment is unfair to students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Assessment informs students of their learning needs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Assessment is a way to determine how much students have learned from teaching	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Assessment forces teachers to teach in a way that is contradictory to their beliefs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Assessment is assigning a grade or level to student work	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Assessment information modifies ongoing teaching of students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Assessment results are filed and ignored	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

13. Assessment determines if students are qualified to advance to the next course	0	0	0	0	0
14. Assessment measures students' higher order thinking skills	0	0	0	0	0
15. Assessment helps students improve their learning	0	0	0	0	0
16. Assessment interferes with teaching	0	0	0	0	0
17. Assessment is an imprecise process	0	0	0	0	0
18. Assessment allows different students to get different instruction	0	0	0	0	0
19. Assessment has little impact on teaching	0	0	0	0	0

### Section 3

Give a rating for each of the items below based on your opinion about assessment practices. Consider the importance of each of the practices in the context of your larger assessment 'toolkit'.

	Not Important	Slightly Important	Fairly Important	Quite Important	Very Important
1. Major examinations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Oral presentations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Objective assessments (e.g., multiple choice, matching, short answer)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Performance assessments (e.g, a speech, paper, or presentation)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Assessments provided by publishers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Short answer or other essay-type questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Projects completed by teams of students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Projects completed by individual students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Quizzes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Authentic assessments (e.g., real world performance tasks)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Section 4**

To what extent do you base your semester grades on the following?

	Not At All	Very Little	Some	Extensively	Entirely
1. Disruptive student behavior	0	0	0	0	0
2. Improvement of performance since the beginning of the year	0	0	0	0	0
3. Student effort—how much students tried to learn	0	0	0	0	0
4. Ability levels of students	0	0	0	0	0
5. Work habits and neatness	0	0	0	0	0
6. Grade distributions of other teachers	0	0	0	0	0
7. Completion of homework (not graded)	0	0	0	0	0
8. Quality of completed homework (graded)	0	0	0	0	0
9. Academic performance, as opposed to other factors, such as behavior	0	0	0	0	0
10. Performance compared with other students in the class	0	0	0	0	0
11. Performance compared with students from previous years	0	0	0	0	0

12. Specific learning objectives mastered	0	0	0	0	0
13. The degree to which students pay attention, participate in class, or both	0	0	0	0	0
14. Inclusion of zeros for missing assignments	0	0	0	0	0
15. Extra credit for non-academic performance (e.g., bringing in items for food drive)	0	0	0	0	0
16. Extra credit for academic performance	0	0	0	0	0

## Appendix B

### Conception Items and Factors

#	Item	Factor
1	Assessment provides feedback to students about their performance	Assessment improves teacher instruction and student learning
2	Assessment results should be treated cautiously because of measurement error	Assessment is irrelevant to the work of teachers and the life of students
3	Assessment is a way to place students into categories	Assessment makes students accountable for their learning
4	Assessment is integrated with teaching practice	Assessment improves teacher instruction and student learning
5	Assessment establishes what students have learned	Assessment improves teacher instruction and student learning
6	Assessment is unfair to students	Assessment is irrelevant to the work of teachers and the life of students
7	Assessment informs students of their learning needs	Assessment improves teacher instruction and student learning
8	Assessment is a way to determine how much students have learned from teaching	Assessment improves teacher instruction and student learning
9	Assessment forces teachers to teach in a way that is contradictory to their beliefs	Assessment is irrelevant to the work of teachers and the life of students
10	Assessment is assigning a grade or level to student work	Assessment makes students accountable for their learning
11	Assessment information modifies ongoing teaching of students	Assessment improves teacher instruction and student learning
12	Assessment results are filed and ignored	Assessment is irrelevant to the work of teachers and the life of students
13	Assessment determines if students are qualified to advance to the next course	Assessment makes students accountable for their learning
14	Assessment measures students' higher order thinking skills	Assessment improves teacher instruction and student learning
15	Assessment helps students improve their learning	Assessment improves teacher instruction and student learning
16	Assessment interferes with teaching	Assessment is irrelevant to the work of teachers and the life of students
17	Assessment is an imprecise process	Assessment is irrelevant to the work of teachers and the life of students

18	Assessment allows different students to get different instruction	Assessment improves teacher instruction and student learning
19	Assessment has little impact on teaching	Assessment is irrelevant to the work of teachers and the life of students