Journal of Online Learning Research (2019) **5**(1), 25-48

(Blended) Learning: How Traditional Best Teaching Practices Impact Blended Elementary Classrooms

ELIZABETH ANTHONY
University of Notre Dame
eanthony@nd.edu

Decades of research have shown that the quality of instruction a student receives has a greater impact than any other factor on his or her learning, but the emergence of the blended learning model has caused many teachers and leaders to focus on innovation and changing practices rather than on developing the research-based teacher practices that have a demonstrated positive impact on students' learning. Through observations of and interviews with six teachers, this study shows that best teaching practices from traditional classrooms – especially (1) demonstrating flexibility and responsiveness, (2) using assessment in instruction, and (3) engaging students in learning tasks – do, in fact, have a substantial impact on student learning in blended elementary classrooms. These results should help leaders and teachers understand where to focus their time and energy in order to improve student learning with the blended learning model.

Keywords: blended learning, teacher practices, instruction

INTRODUCTION

Blended learning is a relatively new educational innovation that has swept the nation in recent years and, in the process, garnered quite a bit of both support and skepticism. Critics point to the fact that there is virtually no evidence that blended learning works in K–12 settings, but both critics and supporters agree that measures of the effectiveness of blended learning will always depend on a number of factors, such as teacher preparedness and how the model is implemented (Schwirzke, Vashaw, & Watson, 2018). The better question, then, is which specific factors influence the success of blended learning, and to what extent?

Decades of education research show that one variable has proven time and again to be the single most critical for student learning: the teacher (Sanders & Horn, 1994; Wright, Horn, & Sanders, 1997). In fact, the impact of the teacher on student learning is far greater than the impact of the school or the curriculum, and the most ordinary teaching strategies, when implemented effectively and with fidelity, can overcome many other factors on students' academic growth (Hattie, 2018; Hattie, 2009; Sanders & Horn, 1994; Wright et al., 1997). Research on the teacher's impact on student learning, however, has rarely taken place in the context of a blended learning environment (Kennedy & Ferdig, 2018).

Blended learning is defined as a formal education program in which a student has an integrated learning experience that takes place at least in part online, with some element of student control over the time, place, path, and/or pace of learning, and at least in part in a brick-and-mortar location (Horn & Staker, 2015). Because the field of blended learning, and particularly the field of K-12 blended learning, is still emerging, research on blended learning – and more specifically on the impact of teachers and specific teaching practices on student learning in blended learning environments – is still quite limited (Schwirzke et al., 2018).

This study seeks to answer the following two research questions on this topic: do best teacher practices studied in traditional classrooms distinguish high-growth blended classrooms from low-growth blended classrooms? And, if so, which of these practices have the greatest impact on growth in blended classrooms? These questions work together, first, to demonstrate the applicability of teacher practices studied in traditional classrooms to blended classrooms and, second, to identify the specific teacher practices that will have the greatest impact on student learning in blended classrooms. The field needs this research to examine where blended learning teachers and leaders should focus their time and energy in order to make blended learning work for their students.

LITERATURE REVIEW

There is widespread agreement in the literature that the single most critical influence on student learning is the quality of teaching students receive (Hattie, 2018; Hattie, 2009; Hargreaves & Fullan, 2012; Sanders & Horn, 1994; Wright et al. 1997). In fact, improved classroom instruction has been identified as the key "factor to improve student achievement gains" (Odden & Wallace, 2003, p. 64) and several years of simple, effective teaching has been shown to dramatically alter the academic trajectory of some struggling students (Marzano, 2003; Sanders & Horn, 1994; Wright et al., 1997).

John Hattie's *Visible Learning* (2009) and continued research (2018) measured the effect size of hundreds of different factors contributing to student achievement, including those from the student, home, school, curricula, teacher, and teaching approaches. Hattie categorized the effect size of each of these contributing factors as negative, low, medium, or high, with high being the zone of desired effectiveness. In his synthesis, Hattie (2009) found that only one category of contributing factors had more meta-analyses classified as highly effective than medium, low, or negative combined, and that was the category of teaching practices/approaches (Hattie, 2018; Hattie, 2009). This confirms findings from Sanders & Horn (1994), Wright et al. (1997), and others (Hargreaves & Fullan, 2012; Marzano, 2003) that teaching practices have a greater impact on student learning than other contributing factors, including but not limited to ethnicity, socioeconomic status, parent involvement, curricular resources, and school resources.

Experts do not always agree on the exact teacher practices that characterize effective teaching, but extensive research has allowed a number of practices and tools to emerge as particularly impactful. The Danielson Group's Framework for Teaching (2014) is a research-based set of components of instruction that is widely considered to be both comprehensive and sound (Little, Goe, & Bell, 2009). The Framework for Teaching (Danielson, 2014) is separated into four domains – planning and preparation, the classroom environment, instruction, and professional responsibilities – which together include 22 components and 76 smaller elements. The scope and validity of this framework make it an obvious choice for a framework of effective teaching practices, but the breadth and number of specific practices identified beg further discussion of which particular practices have the greatest impact on student learning.

Fortunately, other research indicates that certain practices are more impactful than others. Marzano (2003) showed that he and other researchers (Horn & Sanders, 1994; Wright et al., 1997) generally concur on the three, key teacher-level factors: instructional strategies, classroom management, and classroom curriculum design. These factors correspond with the Danielson Framework (2014) domains of planning and preparation, classroom

environment, and instruction. This correlation suggests that the fourth domain of the Danielson Framework (2014) – professional responsibilities – may not be as impactful for student learning as the other three.

Furthermore, Hattie (2018; 2009) identified a number of the components and smaller elements within these three domains that are especially effective. More specifically, the following components (or smaller elements within the components) have a high effect size according to Hattie's (2018; 2009) research: setting instructional outcomes, designing coherent instruction, designing student assessments, creating an environment of respect and rapport, establishing a culture for learning, managing classroom procedures, communicating with students, engaging students in learning, using assessment in instruction, demonstrating flexibility and responsiveness, and using questioning and discussion techniques. From these 11 components, eight emerge as particularly important according to evidence from Hattie (2018; 2009), Marzano (2003), and a number of other frameworks or teacher observation tools (Highlander, 2017; LEAP, 2017; Little et al., 2009). Evidence from these five sources, which are research-based, combine to suggest specific domains, components, elements, and practices from the Danielson Framework (2014) that are especially impactful on student learning. These domains, components, elements, and practices are identified and cited in Table 1. Note that the Danielson Framework (2014), one of the most widely used teacher observation and evaluation instruments (Little et al., 2009), serves as the basis for this list so a citation indicates that the item is supported by or added from a different source.

These teacher practices are supported by a substantial amount of research conducted in traditional classrooms, but their applicability to the nontraditional settings of online and blended learning environments lacks evidence. Only a small number of studies have examined the role of the teacher in K-12 online or blended environments (Archambault, 2018; Pulham, Graham, & Short, 2015), and most of those studies have taken place in fully online environments (Kennedy & Ferdig, 2018). It is important to recognize that blended learning is not merely a combination of in-person and online learning, as it is so often defined, but also that it allows students greater autonomy over what, when, where, and/or how they learn (Patrick, Kennedy, & Powell, 2013). Blended learning entails a shift in instructional design away from traditional models with a single pathway and learning objectives that are the same for every student in the class to innovative models with multiple ways in which teachers measure and students demonstrate mastery towards individual learning goals and objectives aligned to academic standards (Patrick et al., 2013). This suggests that research in fully offline or fully online environments cannot necessarily be directly applied to the unique context of blended learning.

Table 1
Research Framework with Citations

Domain	Component	Elements	Practice
Planning and Preparation	Setting Instructional Outcomes (Hattie, 2009)	Value, sequence, and alignment Clarity Balance	Teacher identifies specific outcomes for each unit and lesson (LEAP, 2017; Marzano, 2003). Outcomes are: • Challenging • Statements of student learning, not student activity. • Standards-aligned • Measurable • Differentiated for students of varied academic levels (LEAP, 2017).
			Teacher identifies which outcomes need to be met to progress or receive recognition (LEAP, 2017).
	Designing Coherent Instruction	Learning activities Instructional materials and resources Instructional groups Lesson and unit structure	Teacher designs lessons that support instructional outcomes and reflect important concepts. Lessons: • incorporate multiple presentations of new concepts (Marzano, 2003). • allow for effective transfer of knowledge (Marzano, 2003). • explicitly connect new content to prior knowledge and skills (LEAP, 2017). • offer opportunities for student input and choice (Highlander, 2017; LEAP, 2017). • include modified or varied content according to students' interests, strengths and needs (LEAP, 2017).
			Teacher thoughtfully plans learning groups and regularly changes groups based on assessment results and lesson objective(s) (University, 2014; LEAP, 2017).
	Designing Student Assessments	Congruence with outcomes Criteria and standards Formative assessments (Hattie, 2018; Hattie, 2009)	Lesson plans indicate correspondence between assessments and instructional outcomes.
			Teacher offers a variety of assessment and performance opportunities for students to demonstrate mastery (LEAP, 2017).
			Teacher utilizes pre-assessment (LEAP, 2017).
			Teacher designs formative assessments to inform decision-making during instruction.
			Teacher utilizes an independent end-of-lesson assessment aligned to the instructional outcome University, 2014; LEAP, 2017).
			Teacher requires that students demonstrate mastery of competencies before progressing (Hattie, 2009).

Table 1, *Continued*

Domain	Component	Elements	Practice
Instruction	Communicating with Students (Hattie, 2009)	Expectations for learning (Hattie, 2009) Directions and procedures Explanations of content Use of oral and written language	Clarity of lesson purpose (objective)/intended outcome (Marzano, 2003).
			Clear directions and procedures specific to the lesson activities (Highlander, 2017).
			Absence of content errors and clear explanations of concepts and strategies (Marzano, 2003).
			Explicit instruction using modeling and think-aloud (University, 2014).
	Using Questioning and Discussion Techniques	Quality of questions (Hattie, 2018; Hattie, 2009) Discussion techniques Student participation	Teacher and students formulate questions of high cognitive challenge, i.e. setting hypotheses, identifying similarities and differences, questions with multiple correct answers or approaches (Marzano, 2003).
			Teacher effectively uses student responses and ideas.
			High levels of student participation in discussion.
	Engaging Students in Learning (Hattie, 2018; Hattie, 2009)	Activities and assignments Student groups (Hattie, 2009) Instructional materials and resources Structure and pacing (Fisher & Frey, 2013; Hattie, 2009)	Lesson includes multiple exposure to and complex interaction with knowledge, such as summarizing, note-taking, and non-linguistic representation (Marzano, 2003).
			Learning tasks: (Highlander, 2017; LEAP, 2017; Marzano, 2003) • require high-level student thinking • invite students to explain their thinking • involve multiple points of failure/perseverance • require students to transfer knowledge
			Students demonstrate enthusiasm, interest, thinking, problem solving, etc.
			Students are actively working rather than watching while their teacher works.
			Students work together to solve problems or create work products (Highlander, 2017; LEAP, 2017; Marzano, 2003).

Table 1, *Continued*

Domain	Component	Elements	Practice
Instruction	Using Assessment in Instruction	Assessment criteria Monitoring of student learning Feedback to students (Hattie, 2018; Marzano, 2003; Hattie, 2009) Student self-assessment and monitoring (Hattie, 2009)	Teacher pays close attention to evidence of student understanding.
			Teacher poses questions specifically created to elicit evidence of student understanding.
			Teacher uses a variety of techniques to assess student learning (University, 2014; Highlander, 2017).
			Teacher circulates to monitor student learning and to offer feedback.
			Teacher provides feedback that is objective and non-judgmental (LEAP, 2017).
			Teacher reinforces and recognizes effort (Marzano, 2003).
			Teacher corrects student misconceptions (University, 2014).
			Students assess their own work against established criteria.
			Students set learning goals and monitor their progress towards them (LEAP, 2017).
			Teacher poses questions specifically created to elicit evidence of student understanding.
			Teacher uses a variety of techniques to assess student learning (University, 2014; Highlander, 2017).
			Teacher circulates to monitor student learning and to offer feedback.
			Teacher provides feedback that is objective and non-judgmental (LEAP, 2017).
			Teacher reinforces and recognizes effort (Marzano, 2003).
			Teacher corrects student misconceptions (University, 2014).
			Students assess their own work against established criteria.
			Students set learning goals and monitor their progress towards them (LEAP, 2017).
	Demonstrating Flexibility and Responsiveness	Lesson adjustment Response to students (LEAP, 2017) Persistence	Teacher adjusts instruction in response to evidence of student understanding (or lack of it) (Highlander, 2017).
			Teacher offers strategic flexible modalities, groupings, and times/places for learning (LEAP, 2017).
			Teacher provides flexible time to allow learners to struggle/work on a problem or project for an extended period of time (LEAP, 2017).

The Clayton Christensen Institute (CCI) (2018) has identified seven different models of blended learning ranging from the station rotation model, which involves integrating adaptive online learning programs into a fairly typical classroom setting, to the enriched virtual model, which adds a few in-person touchpoints to an otherwise fully online learning program. The station rotation model is the most common model in elementary schools because teachers are often already familiar with having students rotate in stations or centers, which makes the transition to a blended learning model seem less daunting (CCI, 2018). The flex model, which allows students to move on fluid schedules among learning activities and modalities according to their needs, is also seen in elementary schools and is relevant to this study (CCI, 2018). The station rotation model typically involves a less drastic shift in instructional design than the flex model, but both require teachers to use student data to differentiate learning objectives and instruction, integrate in-person learning with online learning, and incorporate elements of mastery-based progression. Both might also be considered hybrid forms of innovation that attempt to deliver the best of both online instruction and traditional schools by using online content to bolster and differentiate instruction within a classroom setting (Schwirzke et al., 2018).

There is an emerging body of literature regarding effective instruction in K-12 online and blended learning (Horn & Staker, 2015; Kennedy & Ferdig, 2018; Patrick et al., 2013), including a teacher competency framework specifically for blended learning environments (Powell, Rabbit, & Kennedy, 2014), but most research has focused specifically on online learning (Pytash & O'Brien, 2018). The field is "just now maturing away from the precursory studies involving distance learning," (Lokey-Vega, 2018, p. 3) and has expanded to primarily include teaching and learning in fully online settings (Kennedy & Ferdig, 2018).

Furthermore, much of the research on blended learning has attempted to answer the question "Does blended learning work?," when research really needs to examine the conditions under which blended learning works (Schwirzke et al., 2018). The most valuable research, then, will determine the factors that correlate with student success in blended learning environments (Schwirzke et al., 2018). Having said that, there is a clear need for research that examines one of the factors that correlates most strongly with student success in traditional classrooms – teacher practices – in a blended learning setting.

METHODOLOGY

This study seeks to answer the following two questions.

1. Do best teacher practices studied in traditional classrooms distinguish high-growth blended elementary classrooms from low-growth blended elementary classrooms?
2. If so, which teacher practices studied in traditional classrooms have the greatest impact on growth in blended elementary classrooms?

These questions were investigated through a collective case study of six blended learning elementary classrooms, combining both qualitative and quantitative measures (Cresswell, 2007; Yin, 2003). The collective case study method was chosen as a means by which to examine one issue using the same procedure in multiple different contexts (Creswell, 2007), and to do so without manipulating behaviors (Yin, 2003). There are, of course, complications to this method of research, including variations in program definition and aspects of the program or environment that existed prior to or independent of the program (Yin, 2003). In this study, specific complications included variations of the blended learning model in each classroom and many factors that could contribute to student academic growth besides teacher practices. Still, the literature review indicated that even different blended learning models have much in common (CCI, 2018; Patrick et al., 2013) and teacher practices are more important than most other factors on student learning (Sanders & Horn, 1994; Wright et al., 1997), so these complications were not seen as prohibitive.

This study included six blended learning classrooms from five elementary schools in the Midwest United States. Only classrooms that used a blended learning model in grades one through five for both reading and math were eligible to participate in this study. Four of the classrooms used a station-rotation model and two used a flex model, but all six qualified as blended and had relatively similar blended practices. Of the eligible classrooms, three were chosen to represent high-growth blended learning classrooms and three were chosen to represent low-growth blended learning classrooms. The rationale for comparing high-growth to low-growth classrooms (as opposed to high-performing and low-performing) was that high-growth classrooms are where the most learning occurs whereas low-growth classrooms are where the least learning occurs. The high-growth classrooms were the three classrooms in the sample with the highest growth; each of the high-growth classrooms had an average conditional growth percentile (CGP) greater than 65 on Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) assessments (NWEA, 2018). Similarly, the low-growth classrooms were three of the classrooms with the lowest growth; each of the low-growth classrooms had an average CGP of less than

40 on NWEA MAP assessments. The CGP compares the growth of a class to the growth of other classes across the country at that same level (NWEA, 2018). So, for example, a second-grade class with a conditional growth percentile of 98 grew more than 98% of other second-grade classes across the country. Table 2 shows the average CGPs for the studied classrooms.

Table 2

Conditional Growth Percentiles for Participating Classrooms

Classroom	Average CGP	Grade Level	Classification
A	98.5	2	High-Growth
B	87	5	High-Growth
C	67.5	5	High-Growth
X	39.5	3	Low-Growth
Y	38	1	Low-Growth
Z	9.5	3	Low-Growth

To create a comprehensive evaluation system, the researcher conducted three 90-120-minute observations of each classroom – one or two of which was during reading instruction and one or two of which was during math instruction. In addition, the researcher conducted one 20-40-minute interview with each teacher for a total of 18 observations and six interviews over a three-month period. Each observation took place during the class' normal reading or math block and was used to measure the actions that make a teacher highly effective. Observation is the "most common form of teacher evaluation" and is considered to be "the most direct way to measure teaching practice because the evaluator can see the full dynamic of the classroom" (Little et al., 2009, p. 6-7). During each observation, the researcher used the tool developed during the literature review to quantify the evidence (or lack of evidence) of the 47 research-based best teaching practices (Table 1). The interviews were then used as supplements to the observations to "describe the 'whys' and 'hows' of teacher performance and its impact" (Little et al., 2009, p. 12). During each interview, the researcher asked for more information about teacher practices that are difficult to observe in a class period and attempted to tap into the teachers' intentions and thought processes. This process also created opportunities for member checks, which allowed the researcher to confirm that the practices the researcher observed were interpreted correctly (Lincoln & Guba, 1985). Example research questions are included in Appendix A. The researcher is an

experienced observer of classroom instruction and very comfortable speaking with educators about their practice. To avoid bias, the researcher looked for evidence of the practices explicitly named in the framework and did not form conclusions based on prior perceptions or any sort of expectations (Yin, 2003).

Teachers received a score of -2 (substantial negative evidence), -1 (negative evidence), 0 (neutral or no evidence), 1 (positive evidence), or 2 (substantial positive evidence) for each teaching practice during each observation. The scores were based on the researcher's observations as well as background information gathered in the interview. The scores for each of the practices were totaled at the end to assign the teacher an overall score for the observation. For an individual practice, the highest possible score was 2 and the lowest possible score was -2, meaning the greatest possible difference between two scores was 4 points. For an overall observation, the highest possible score was 94 and the lowest possible score was -94, meaning the greatest possible difference between two scores was 188. The researcher compared scores for overall observations and individual practices to answer the two research questions. Though this exact tool and scoring system has not been used previously, other studies have scored teachers according to the Danielson framework's indicators, assigned teachers scores on a similar scale, and examined student achievement data for correlations between the teachers' scores and the students' achievement scores (Kane, Taylor, Tyler, & Wooten, 2011; Kimball, White, Milanowski, & Borman, 2004). The methodology for this study was informed by these previous studies but was created to be more appropriate for the number of classrooms included and to utilize the observation framework borne of the literature review.

RESULTS

Each of the six participating teachers were observed and scored on the observation tool three times. As explained in the previous section, the minimum score on the evaluation tool was -94 and the maximum score was 94, but in practice it was intentionally challenging to receive an especially low or high overall score as minimum and maximum scores on each practice were reserved for instances where there was substantial evidence in either direction. Each teacher's three observation scores were recorded and averaged to find his or her average observation score. As seen in Table 3 and Figure 1, the average observation scores almost directly correlate with the average CGPs. The two teachers (A and B) with extremely high growth in their classrooms had average observation scores substantially higher than the others, and the reverse is true for the teacher with extremely low growth (Z), whose average observation score was substantially lower than the other teachers.

Table 3
Observation Scores and Conditional Growth Percentiles for Participating Teachers

Classroom/ Teacher	Obs. 1	Obs. 2	Obs. 3	Average Observation Score	Average CGP
A	45	53	53	50.3	98.5
B	46	42	47	45	87
C	26	17	26	23	67.5
X	4	24	16	14.7	39.5
Y	17	9	38	21.3	38
Z	-46	-10	-26	-27.3	9.5

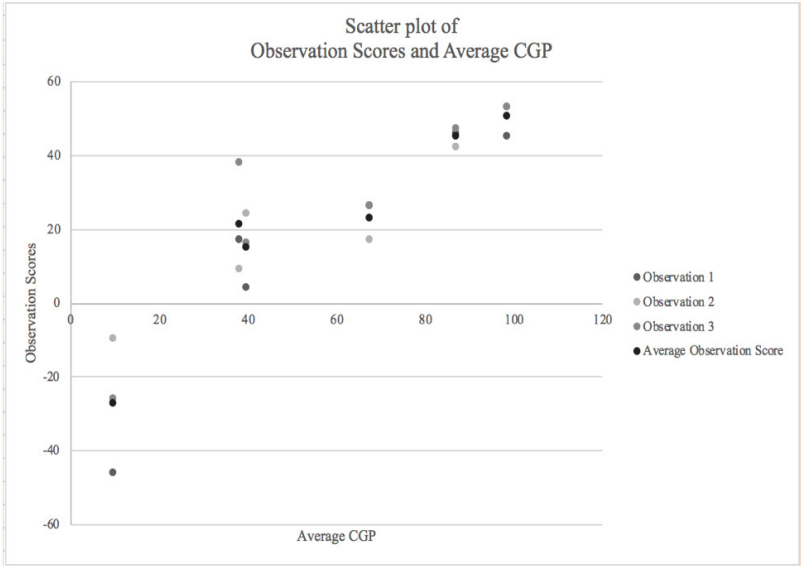

Figure 1. Scatterplot of each participating teacher's observation scores and average Conditional Growth Percentile.

The results for C, X, and Y were slightly more complicated. C was classified as "high-growth" with an average CGP of 67.5, and C's average observation score was higher than those of X and Y, but the differences between C's average observation score and the average observation scores of the next two teachers below (X and Y), both classified as "low-growth,"

were not commensurate with the differences between C's CGP and the CGPs of X and Y. In other words, the difference in growth between C and X/Y was substantially greater than the difference in their observation scores. Furthermore, though X had a slightly higher CGP than Y, Y actually received a higher average observation score than X. In fact, Y's average observation score was fewer than two points below C's average observation score, and C was classified as high-growth whereas Y was the teacher with the second-to-lowest growth.

The scores for individual practices were compared as well to answer the second research question. Just as each teacher received an average overall observation score, each teacher also received an average observation score for each practice, which was the average of the teacher's scores on that practice from all three observations. The minimum score for any given practice was -2 and the maximum was 2 but, again, scores of -2 or 2 were intentionally reserved for substantial evidence. To determine the differences in practices between high-growth and low-growth classrooms, an average practice score was calculated for each group. This score was the average observed score for the practice across all observations for the three teachers in each group. For each practice, the difference between the high-growth group's average score and low-growth group's average score was examined. Table 4 reports the average scores and difference for each practice. All scores in the tables are rounded to the nearest tenth but the differences were calculated using unrounded numbers.

Table 4

Average Scores of Observed Practices and Differences between High-Growth and Low-Growth Classes

Component	Practice	High-Growth Average	Low-Growth Average	Difference
Setting Instructional Outcomes	Teacher identifies specific outcomes for each unit and lesson. Outcomes are:	0.8	0.6	0.2
	• Challenging	1	0.4	0.6
	• Statements of student le arning, not student activity.	0.9	0.7	0.2
	• Standards-aligned	1	0.8	0.2
	• Measurable	0.7	0.6	0.1
	• Differentiated for students of varied academic levels.	0.3	-0.7	1
	Teacher identifies which outcomes need to be met to progress or receive recognition.	0.2	0	0.2

Table 4, *Continued*

Component	Practice	High-Growth Average	Low-Growth Average	Difference
Designing Coherent Instruction	Teacher designs lessons that support instructional outcomes and reflect important concepts. Lessons:	1.1	0.3	0.8
	• incorporate multiple presentations of new concepts.	0.4	-0.3	0.8
	• allow for effective transfer of knowledge.	0.7	0	0.7
	• explicitly connect new content to prior knowledge and skills.	0.8	0.2	0.6
	• offer opportunities for student input and choice.	0.3	-0.6	0.9
	• include modified or varied content according to students' interests, strengths and needs.	0.8	-0.4	1.2
	Teacher thoughtfully plans learning groups and regularly changes groups based on assessment results and lesson objective(s).	1.1	0.7	0.4
Designing Student Assessments	Lesson plans indicate correspondence between assessments and instructional outcomes.	0.3	0	0.3
	Teacher offers a variety of assessment and performance opportunities for students to demonstrate mastery.	-0.3	-0.4	0.1
	Teacher utilizes pre-assessment.	0.7	0	0.7
	Teacher designs formative assessments to inform decision-making during instruction.	1.1	0.2	0.9
	Teacher utilizes an independent end-of-lesson assessment aligned to the instructional outcome.	0.6	0	0.6
	Teacher requires that students demonstrate mastery of competencies before progressing.	0.1	-0.4	0.5

Table 4, *Continued*

Component	Practice	High-Growth Average	Low-Growth Average	Difference
Communicating with Students	Clarity of lesson purpose (objective)/intended outcome.	0.6	1	-0.4
	Clear directions and procedures specific to the lesson activities.	1	0	1
	Absence of content errors and clear explanations of concepts and strategies.	0.9	-0.2	1.1
	Explicit instruction using modeling and think-aloud.	1	0.1	0.9
Using Questioning and Discussion Techniques	Teacher and students formulate questions of high cognitive challenge, i.e. setting hypotheses, identifying similarities and differences, questions with multiple correct answers or approaches.	0.8	-0.1	0.9
	Teacher effectively uses student responses and ideas.	0.6	0.2	0.3
	High levels of student participation in discussion.	0.8	0.6	0.2
Engaging Students in Learning	Lesson includes multiple exposure to and complex interaction with knowledge, such as summarizing, note-taking, and non-linguistic representation.	1.2	-0.2	1.4
	Learning tasks:			
	• require high-level student thinking	1.2	-0.2	1.4
	• invite students to explain their thinking	0.9	-0.3	1.2
	• involve multiple points of failure/perseverance	0.4	0.1	0.3
	• require students to transfer knowledge	0.9	0.3	0.6
	Students demonstrate enthusiasm, interest, thinking, problem solving, etc.	0.8	0.4	0.3
	Students are actively working rather than watching while their teacher works.	1.4	0.3	1.1
	Students work together to solve problems or create work products.	0.4	-0.1	0.6

Table 4, *Continued*

Component	Practice	High-Growth Average	Low-Growth Average	Difference
Using Assessment in Instruction	Teacher pays close attention to evidence of student understanding.	1.6	-0.7	2.2
	Teacher poses questions specifically created to elicit evidence of student understanding.	1.6	0.2	1.3
	Teacher uses a variety of techniques to assess student learning.	0.7	-0.7	1.3
	Teacher circulates to monitor student learning and to offer feedback.	2	0.1	1.9
	Teacher provides feedback that is objective and non-judgmental.	1.3	0.3	1
	Teacher reinforces and recognizes effort.	1	1.2	-0.2
	Teacher corrects student misconceptions.	1	0.4	0.6
	Students assess their own work against established criteria.	0.4	-0.2	0.6
	Students set learning goals and monitor their progress towards them.	0.8	-0.1	0.9
Demonstrating Flexibility and Responsiveness	Teacher adjusts instruction in response to evidence of student understanding (or lack of it).	1.2	-0.2	1.4
	Teacher offers strategic flexible modalities, groupings, and times/places for learning.	1.1	-0.3	1.4
	Teacher provides flexible time to allow learners to struggle/work on a problem or project for an extended period of time.	1.3	-0.7	2

As seen in Table 4, the teachers in the high-growth classrooms were observed to implement almost all of these effective practices more, on average, than teachers in the low-growth classrooms. There are three components that high-growth teachers were observed to implement far more often than their low-growth counterparts, and those are, in descending order, (1) demonstrating flexibility and responsiveness, (2) using assessment in instruction, and (3) engaging students in learning. Practices in these components were those that presented the starkest contrast between high-growth

and low-growth classrooms and therefore are likely the practices that most related to the differences in academic growth between the two categories of classrooms. There were only two practices in which the low-growth teachers actually scored higher than the high-growth teachers: (1) clarity of lesson purpose/intended outcome and (2) teacher reinforces and recognizes effort.

DISCUSSION

Though these results are from a relatively small sample in an exploratory, observational study, they provide promising initial answers to both research questions. First, the overall average scores indicate that research-based, best teacher practices from traditional classrooms are, in fact, likely candidates for best teacher practices in blended learning classrooms as well. This conclusion is significant in the blended learning space, where there is ample talk of technology and innovation but relatively little talk of what teacher instruction should look like, especially vis-a-vis what we already know about sound pedagogy. In addition, these results indicate that the tool developed for this study could be used to observe and evaluate instructional practices relevant to blended learning teachers even though it is not blended learning-specific.

It is important to recognize potential limitations to this study. First, and most obviously, observations are subjective. Different researchers might interpret observations in different ways, and there is always a potential for bias in observations (Yin, 20013). Here, the use of an a priori framework, coupled with interviews after observations was used to limit this bias (Lincoln & Guba, 1985), but these biases can never be completely removed.

Second, teachers C, X, Y and Z were in their first year of blended instruction, which means that their average CGPs reflected only one semester of growth, a semester wrought with change and learning. When the research observations took place in the second semester, these teachers were continuing to learn and develop, which may have impacted their observation scores. Furthermore, all three of these teachers nearly fall into the middle third of growth percentiles, as shown in Figure 2 below. It would have been ideal for the study to include teachers with equal amounts of experience, growth data from multiple years of blended instruction, and greater differences in their growth (perhaps with high-growth being only the 75th percentile or above), but recruiting teachers meeting all of these criteria was not possible. Further research could examine these same questions with an even more well-matched sample to ensure that the results still stand, but the results from this study indicate a nontrivial contrast between the highest-growth teachers and the lower-growth teachers, which supports the hypothesis that best teaching practices in traditional settings distinguish high-growth blended learning classrooms from low-growth blended learning classrooms.

Figure 2. Distribution of Conditional Growth Percentiles of Participating Classrooms.

Third, and finally, this study focused entirely on the face-to-face aspect of the blended environment, which is shared with a traditional environment. It was not examining the factors which are unique to a blended environment, namely the technology. There were likely behaviors that went unnoticed because they are unique to the blended environment and not included on the tool developed to examine the face-to-face practices. As a result, this tool would be insightful to others looking to observe and evaluate instructional practices, but more research is needed to examine the effectiveness of the practices and aspects of the classroom that are unique to blended teaching.

Of all the best teaching practices identified through the literature review, three categories of practice emerged as particularly meaningful in this study: demonstrating flexibility and responsiveness, using assessment in instruction, and engaging students in learning. Research has shown that demonstrating flexibility and responsiveness is important for teachers in traditional classrooms, but it makes sense that this skill might be even more important in a blended learning setting in which students are often working independently, at their own paces, and in many different contexts. In contrast to the low-growth teachers in this study, who struggled to relinquish traditional means of control over their classrooms, teachers like the high-growth teachers from this study who have the capacity to allow for strategically flexible modalities, settings, and paces for learning seem to thrive in a blended learning setting. One simple but important example of this flexibility is in seating options and arrangements. Specifically, in the highest-growth classrooms, students were scattered all over the room (or even throughout the hallways) as they were given the freedom to choose a place and space for learning that made them most comfortable. In contrast, students in the lower-growth classrooms sat in assigned seats at desks for the majority of the lesson. This relatively small step in flexibility on the part of the teacher is likely indicative of a greater overall ability to provide a responsive learning environment for students.

Demonstrating flexibility and responsiveness goes hand-in-hand with using assessment in instruction, as effective teachers not only integrate multiples forms of assessment throughout the lesson but also adjust their

instruction in response to evidence of student understanding (or lack thereof) in real time. The highest-growth teachers used the time they had with their students to carefully examine the students' work and parse out what students understood versus what they did not understand. These teachers integrated data from multiple sources to create a holistic picture of student understanding and then formulated small groups or created activities in response to this data. The low-growth teachers, in contrast, often asked the right questions but may have failed to listen to the answers or incorporate this information into their instruction. It was much more common for low-growth teachers to demonstrate only a general understanding of students' progress and understanding rather than the more nuanced understanding the high-growth teachers demonstrated. Research has shown that formative assessment is one of the single most impactful teacher practices (Hattie, 2018, 2009; Marzano, 2003), and this study adds support to that claim.

Finally, the specific practices in the "engaging students in instruction" component that differentiated high-growth classrooms from low-growth classrooms might also be described as "rigor," as these practices demand that students wrestle with complex knowledge, explain their thinking, and analyze and synthesize information rather than regurgitate it. In one of the highest-growth classrooms, the teacher almost never let a student provide an answer without justifying it. This practice of asking students to explain their thinking pushed students to think more deeply about the material and, as a result, revealed misconceptions or shallow knowledge that would not have otherwise been evident. Many online learning programs are critiqued for not requiring students to do very much high-level thinking, but these high-growth teachers demonstrated that they did not rely on someone or something else to do this important work for them. Instead, they took it upon themselves to challenge their students to understand material more deeply, and it is this deep conceptual understanding that was ultimately reflected in such astounding growth.

Given the evidence for the importance of each and every practice on the observation tool, it is somewhat surprising that low-growth teachers outperformed high-growth teachers in any practices. The researcher was especially surprised to find that low-growth teachers outperformed high-growth teachers on the first practice – clarity of lesson purpose/intended outcome. The highest-growth teacher (A) provided an exemplar of this practice through a clearly defined objective for each activity taking place in the classroom. Further, these objectives were differentiated based on student need. Her careful creation and communication of objectives could be a wonderful example for all classes, but particularly for the other two high-performing teachers (B and C) who rarely oriented the day's lesson towards a larger learning goal. It is likely that B and C, who both used a flex

model, struggled to articulate the purpose and relevance of the many different learning objectives they were simultaneously managing in their classrooms; even if they generally knew students' learning objectives, they did not take the time to clearly outline them. In contrast, the low-growth teachers (X, Y, and Z) probably found it easier to implement this practice because objectives in their classrooms were less explicitly differentiated. Blended learning does not require that teachers create different learning objectives for each student, but classrooms that are more responsive to students' needs typically include differentiated objectives. Having a single lesson objective for all students may be easier for teachers to manage and therefore communicate, but an undifferentiated learning objective also likely stifles growth. The gold standard is clearly the classroom in which objectives are differentiated and communicated clearly so that each student may understand the unique purpose and relevance of what he or she is learning (Danielson, 2014; Hattie, 2009; LEAP, 2017; Marzano, 2003), and the highest-growth teacher (A) exemplified this impeccably.

Finally, low-growth teachers also had higher scores in the practice of "recognizing students' effort," which is interesting but unsurprising given what was observed in these classrooms. The low-growth teachers all presented substantial evidence of this practice by saying things like "that's not quite right, but I like where your brain's at!" or "look at how many hands are up, I love that so many scholars are excited to share the answer!" These interactions are positive and desirable, but their absence from the high-growth classrooms was not negative. Rather, the high-growth teachers did not seem to need to recognize students' effort as it was simply expected that students would give their best effort all the time. There was a less observable culture of high expectations surrounding effort in the high-growth classrooms, and teachers did not need to cultivate this culture with constant positive statements. Students in the high-growth classrooms seemed to understand that their effort would be reflected in their progress in both online and offline learning. The high-growth teachers used their feedback and recognition to address student understanding and performance, which is a practice that is far more likely to contribute to student growth.

This difference between low-growth and high-growth classrooms could also have to do with the aforementioned differences in students' ages across both groups. Still, even the highest-growth teacher, who taught in second grade, focused the majority of her feedback on learning rather than effort. This evidence suggests, then, that teachers should invest in the creation of a culture for learning that includes best effort as an implicit expectation and focus the majority of their feedback on student learning instead of effort.

CONCLUSION

Blended learning is often described as an innovation that will "disrupt" traditional forms of education and completely transform learning as we know it (Horn & Staker, 2015), but this research suggests that there might be more similarities between blended and traditional instruction than some may think. Though teachers in online and blended settings do require some distinct skills, many of the best teacher practices from traditional classrooms apply to blended settings as well (Archambault, 2018; Pulham et al., 2015). This study has demonstrated that best teaching practices studied in traditional classroom settings do distinguish high-growth blended learning classrooms from low-growth blended learning classrooms, which suggests that best teacher practices from traditional settings continue to be relevant in blended settings as well. The blended learning model may require teachers to be more flexible and responsive to students, to integrate multiple data sources into their constant stream of formative assessment, and to deliberately incorporate more rigorous learning activities to achieve optimal growth – but these are familiar teacher practices applied in a new setting. In short, this study is a first attempt to demonstrate that efforts to identify aspects of outstanding instruction in blended learning environments should not disregard what we know to be true in other classroom settings; rather, the field should recognize that many of the best teaching practices in blended learning elementary classrooms are the same practices that decades of research have shown are the most impactful on student learning, simply applied in a new context. It might be wise, then, to marry the focus on new strategies and models with an ongoing focus on how the application of evidence-based practices may be unique in innovative contexts. Further research should examine this same question on a broader scale involving more classrooms and more researchers. Employing multiple trained researchers in the study would mitigate the potential for bias or gaps in knowledge from significantly impacting the study.

References

Archambault, L. (2018). Introduction, part III. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning (2nd ed.)* (219-220). Pittsburgh, PA: ETC Press.

Clayton Christensen Institute (CCI) (2018). Blended learning models. *Blended learning universe.* Retrieved November 14, 2017 from: https://www.blendedlearning.org/models/

Creswell, J. (2007). *Qualitative inquiry and research design* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Danielson, C. (2014). The framework for teaching evaluation instrument (2013 ed.). *The Danielson Group*. Retrieved November 14, 2017 from: https://www.danielsongroup.org/framework/

Fisher, D., & Frey, N. (2013). *Better learning through structured teaching: A framework for the gradual release of responsibility.* Alexandria, VA: ASCD.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school.* New York, NY: Teachers College Press.

Hattie, J. (2018). 250+ influences on student achievement. *Visible learning plus.* Retrieved August 6, 2018 from: https://us.corwin.com/sites/default/files/250_influences_-_7.18.18.pdf

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York, NY: Routeledge.

Highlander Institute (2017). *Highlander walkthrough tool domains and indicators.* Retrieved November 14, 2017 from: https://highlanderinstitute.org/

Horn, M. B. & Staker, H. (2015). *Blended: Using disruptive innovation to improve schools.* San Francisco, CA: Josey-Bass.

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources, 46*, 587-613.

Kennedy, K. & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning (2nd ed.).* Pittsburgh, PA: ETC Press.

Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education, 79*(4), 54-78. Retrieved from http://www.danielsongroup.org/wp-content/uploads/2018/02/Kimball_White_MIlanowski_BormanPeabodyJournal.pdf

LEAP Innovations (2017). *LEAP learning framework for personalized learning.* Retrieved November 14, 2017 from: http://leaplearningframework.org/

Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

Little, O., Goe, L., & Bell, C. (2009). A practical guide to evaluating teacher effectiveness. *National Comprehensive Center for Teacher Quality*. Retrieved November 14, 2017 from https://gtlcenter.org/sites/default/files/docs/practicalGuide.pdf

Lokey-Vega, A. (2018). Introduction, part I. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning (2nd ed.)* (3-5). Pittsburgh, PA: ETC Press.

Marzano, R. J. (2003). *What works in schools*. Alexandria, VA: ASCD.

Northwest Evaluation Association (NWEA) (2018). *MAP growth.* Retrieved April 20, 2018 from https://www.nwea.org/map-growth/

Odden, A., & Wallace, M. J. (2003). *Leveraging teacher pay*. Education Week, 22(43), 64.

Patrick, S., Kennedy, K., & Powell, A. (2013). Mean what you say: Defining and integrating personalized, blended and competency education. *iNACOL*. Retrieved November 14, 2017 from: https://www.inacol.org/resource/mean-what-you-say-defining-and-integrating-personalized-blended-and-competency-education/

Pytash, K.E. & O'Byrne, W. I. (2018). Research on literacy instruction and learning in virtual, blended, and hybrid environments. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning (2nd ed.)* (303-320). Pittsburgh, PA: ETC Press.

Powell, A., Rabbitt, B., & Kennedy, K. (2014). *iNACOL blended learning teacher competency framework*. Retrieved August 6, 2018 from: https://www.inacol.org/wp-content/uploads/2015/02/iNACOL-Blended-Learning-Teacher-Competency-Framework.pdf

Pulham, E., Graham, C. R., & Short, C. R. (2018). Generic vs. modality-specific competencies for k-12 online and blended teaching. *Journal of Online Learning Research, 4*(1), 33–52. Retrieved from https://www.learntechlib.org/j/JOLR/v/4/n/1/

Sanders, W.L. & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*, 299-311.

Schwirzke, K., Vashaw, L., & Watson, J. (2018). A history of K-12 online and blended instruction in the United States. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning (2nd ed.)* (7-20). Pittsburgh, PA: ETC Press.

University of Texas System/Texas Education Agency (2014). *Instructional decision-making procedures: Ensuring appropriate instruction for struggling students in grades K-12.* Retrieved from: http://buildingrti.utexas.org/sites/default/files/booklets/K-12_IDM_procedures.pdf.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

Yin, R. K. (2003). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage Publications.

APPENDIX

APPENDIX A
INTERVIEW QUESTIONS
TEACHER INTERVIEW PROTOCOL

Can you please tell me about your process for unit and lesson planning?
- How do you choose which concepts to cover?
- What role do data play in this process?
- Are there specific elements that you consider essential to every lesson?
- Which aspects of your lesson are pre-planned and which are made up in the moment?

What information is important for you to gather about your students (academic and otherwise)? How do you gather it? Can you give some examples of how this information impacts your instruction?

How do you determine learning groups?

What role does assessment play in your instruction?
- Can you describe how you balance more formal assessments with informal checks for understanding?
- What role does student self-assessment play?

How is student work differentiated in your classroom?