

How Do Online Course Design Features Influence Student Performance?

Abstract

Given the rapid growth in online coursework within higher education, it is important to establish and validate quality standards for these courses. While many online learning quality rubrics do exist, thus far there has been little empirical evidence establishing a clear link between specific course design features and concrete, student-level course outcomes. In the current study, the authors develop an online course design assessment rubric that includes four areas, and explore the impact of each area on student end-of-semester performance in 23 online courses at two community colleges. The results indicate that the quality of interpersonal interaction within a course relates positively and significantly to student grades. Additional analyses based on course observation and interview data suggest that frequent and effective student–instructor interaction creates an online environment that encourages students to commit themselves to the course and perform at a stronger academic level.

Keywords: computer-mediated communication; pedagogical issues; postsecondary education; teaching/learning strategies; distance education and telelearning

1. Introduction

Online coursework¹ has become increasingly popular in postsecondary education, with an estimated 5 to 7 million students now enrolling in at least one online course each year (Allen & Seaman, 2015). A large body of research has compared outcomes between online and traditional face-to-face college courses; however, results have been mixed across studies, with some finding positive results for online learning and others finding negative results (e.g., see Bernard et al., 2004; Jahng, Krug, & Zhang, 2007; Sitzmann, Kraiger, Stewart, & Wisher, 2006; U.S. Department of Education, 2010; Zhao, Lei, Yan, Lei, & Tan, 2005;). Negative results are more common among students who come from disadvantaged backgrounds or have lower levels of preparation (Figlio, Rush, & Yin, 2013; Kaupp, 2012; Xu & Jaggars, 2014), resulting in overall negative effects within the community college population (Johnson & Mejia, 2014; Xu & Jaggars, 2011; Xu & Jaggars, 2013). These mixed findings suggest that the effectiveness of online learning may vary according to how the online course is designed and taught, particularly among less-privileged student populations. Accordingly, this paper focuses on the community college context to address two research questions: (1) Are there particular online course design features that significantly impact students' course performance?; and (2) Among those design features that have a significant impact on student course learning outcomes, what practices and techniques are typical of "higher quality" versus "lower-quality" online courses?

Our first research question aims to identify the relative importance of four online course design and instructional features that the literature suggests may influence students' course learning outcomes: (1) organization and presentation, (2) learning objectives and assessments, (3) interpersonal interaction, and (4) use of technology. While these online course design features have been widely covered in theoretical discussions, empirical studies, and various online quality rubrics, research has not yet established a clear empirical link between any of these specific indicators and concrete, student-level outcomes, which makes it difficult for colleges to choose among the wide

¹ In this study, "online coursework" refers to semester-length asynchronous fully-online college courses, which are the most typical type of online courses offered by community colleges (Jaggars & Xu, 2010; Xu & Jaggars 2011b).

variety of options for measuring course quality or designing effective online courses (e.g., see Benson, 2003). Our second question aims to identify specific practices and techniques that are typically used in effective online courses and can be replicated to improve student learning outcomes.

To achieve these goals, we draw from the literature on online course quality to create a simple holistic rubric that contains the four design feature subscales mentioned above. We then use the rubric guidelines to assess 23 online courses taught at two community colleges in spring 2011, and relate the rubric scores to students' learning outcomes through a multilevel model approach that controls a variety of student-level and course-level baseline characteristics. We further complement the quantitative results with in-depth course observation data and interview data with both course instructors and students enrolled in these online courses.

We focus on the community college setting in part because online learning enrollments have increased particularly quickly at community colleges (Choy, 2002; Parsad & Lewis, 2008), where many students need the flexibility of online coursework in order to balance school with work or family demands (Jaggars, 2014); and in part because community college students may particularly struggle in online courses due to relatively low levels of self-directed learning skills (Xu & Jaggars, 2014). Although this study examines online courses at just two community colleges, the challenges faced by instructors to develop high-quality semester-long online courses, and faced by students to succeed in a fully online environment, are similar to many other institutions. Thus, lessons learned from this study are likely to be useful to other college administrators, online course instructors, and policymakers.

Overall, our findings indicate that while well-organized courses with well-specified learning objectives are certainly desirable, these design features do not significantly predict student learning outcomes per se. Among the four design features examined, only the quality of interpersonal interaction within a course relates positively and significantly to student grades. Additional analyses based on course observation and interview data suggest that frequent and effective student-instructor interaction creates an online environment that encourages students to commit themselves to the course and perform at a stronger academic level. The importance of an engaged instructor in online

learning is in line with Holmberg's (1995) theory that instructors must create a personal relationship with students in order to motivate them to succeed. Our results, therefore, lend support to creating more interactive opportunities within fully-online courses.

2. Existing Studies on Online Course Design Features

The literature on online course design features includes four distinct strands: (1) practitioner-oriented literature, (2) surveys, (3) controlled studies, and (4) course quality rubrics, which pull together work from the first three strands. In this section, we briefly describe the types of work conducted in each area; in subsequent sections, we summarize themes and findings across areas.

2.1 Four Strands of Research

First, the large practitioner-oriented literature includes theory-based frameworks, case studies of successful courses, papers on perceived best practices, and syntheses or reviews of such work. For example, Moore's theory of transactional distance (2013) sketches out three key design features that define the nature of an online course: structure, or the extent to which curricular materials and assessments are strictly regulated and sequenced; dialogue, or the extent to which teacher and student engage in constructive interaction that builds the student's knowledge; and autonomy, or the extent to which the student has freedom to choose what to learn, how to learn, and how much to learn. Transactional distance theory does not define quality per se, but it does suggest that a course with greater transactional distance – for example, a low degree of dialogue – will be more challenging for less autonomous learners. As another example from the practitioner literature, Grandzol and Grandzol's (2006) much-cited review of best practices in online education pulled together information from accreditation standards, learning theories, prior reviews of best practices, and the very few available studies that attempted to empirically validate best practices.

Second, several surveys have captured students' and instructors' opinions regarding the elements that characterize high-quality online courses (MacDonald, Stodel, Farres, Breithaupt, & Gabriel, 2001; Smissen & Sims, 2002; Keeton, 2004; Young, 2006;

Ralston-Berg, 2010, 2011). Typically, such surveys have focused on a small set of courses or students within a single institution. For example, in one of the largest single-institution studies (Smitsen & Sims, 2002), researchers convened focus groups to generate a list of quality features, then surveyed 231 staff members and 893 students in order to prioritize the resulting list. In two surveys that encompassed more than a single institution, Frederickson, Pickett, Shea, Pelz, and Swan (2000) surveyed students across the State University of New York system to examine factors that contribute to self-perceived learning and satisfaction in online courses; similarly, Ralston-Berg (2010) surveyed students from 31 institutions across 22 states regarding 13 areas that were thought to potentially influence the quality of an online course.

Third, some researchers have moved beyond survey work, designing experiments or other control group studies that attempt to assess the causal effects of specific design features of online learning on students' attitudes or outcomes. Some have focused on the presence or absence of specific technology tools (such as discussion boards, online quizzes, and embedded video), some have focused on the degree of interaction among students or with instructors, and some have focused on the degree of learner control or learner reflection. The most rigorous studies in these areas were synthesized in a recent review by the U.S. Department of Education (2010).

Fourth, several educational associations have drawn from the literature (particularly the first two strands described above) to create rubrics assessing the quality of online programs or courses. These rubrics each focus on slightly different sets of quality characteristics. For example, the Institute for Higher Education Quality (Merisotis & Phipps, 2000) designed 24 quality benchmarks clustered along seven categories to measure the quality of distance education courses. The Council of Regional Accrediting Commissions developed guidelines for online degree and certificate programs (Middle States Commission on Higher Education, 2002), which they divided into five quality components. The Sloan Consortium also created a framework of five broad categories—the “Five Pillars”—for assessing the quality of online learning (Moore, 2005). Perhaps the most widely adopted rubric in terms of online course quality is “Quality Matters,” developed by MarylandOnline (Quality Matters Program, 2014). The Quality Matters rubric includes eight general standards and 41 specific benchmarks, which were designed

by faculty with the goal of evaluating online courses and improving student learning in distance education.

Across the four strands of the literature, sources differed widely in their conceptions of the key elements of course quality. However, when viewed in broad strokes, most sources seemed to agree on four general areas of quality: (1) the extent to which the course interface is well organized and easy to navigate; (2) the clarity of learning objectives and performance standards; (3) the strength and diversity of interpersonal interaction; and (4) the extent to which technology is effectively used. In the following sections, we discuss the theoretical, empirical, and quality-framework literature that exists within each of these four areas.

2.2 Key Evidence on Online Course Design Features

Course Organization and Presentation

Across the set of quality rubrics, Quality Matters most strongly emphasizes the importance of course organization and presentation.² For example, the Quality Matters standards specify that “instructions make clear how to get started and where to find various course components,” and that “learners are introduced to the purpose and structure of the course” (Quality Matters Program, 2014). In the practitioner literature, Grandzol and Grandzol (2006) also suggest that a consistent and clear structure—including navigational documents and instructions that explicitly instruct students in terms of where to go and what to do next—is vital to student success.

Several surveys have also emphasized the importance of a well-organized course structure with intuitive navigation. A study of two online criminal justice courses (Fabianic, 2002) indicates that students regard ease of navigation as a key quality criterion. An institutional survey (Young, 2006) found that students appreciated instructors who made a strong effort to create a thoughtful course that was well organized and carefully structured. In larger-scale survey work, Smissen and Sims (2002) found that “ease of use” (an intuitive, user-friendly interface and navigation) was one of the three

² Found under the first Quality Matters general standard, entitled “Course Overview and Introduction.”

most important quality criteria identified by students, faculty, and staff; similarly, when Ralston-Berg (2010) asked students to rate the most important factors that contribute to their success in online courses, these factors included clear instructions regarding how to get started, how to find various course components, and how to access resources online. Beyond survey work, however, little empirical research has been conducted in this area.

Learning Objectives and Assessments

Most online course quality rubrics highlight the importance of clearly stated and well-aligned learning objectives, a close relationship between course objectives and assessments, and specific and transparent grading criteria. For example, the Institute for Higher Education Quality (2000) specifies that students should be provided with supplemental course information outlining course objectives, concepts, and ideas. Quality Matters (Quality Matters Program, 2014) also includes a long list of specific standards³ in this area that states that learning objectives should be measurable, clearly communicated, and consistent across learning modules; that assessments should be in line with the stated learning objectives and the level of the course; and that students should have clear instructions in terms of how they are to be evaluated.

In the theoretical literature, Naidu (2013) argued that while carefully designed learning goals are important in all educational settings, they may be particularly critical in distance education, given that students are often “studying independently and without a great deal of opportunity to interact with their peers and tutors” (p. 269). In transactional distance theory, the relationship between clearly-specified learning goals and quality is more ambiguous: a course with overly rigid goals and assessments is thought to create more “distance” between the teacher and learner, and thus to require learners to exercise more autonomy, which could be challenging for some students (Moore, 2013).

Data from surveys and qualitative work, however, suggest that clearly stated and sequenced learning objectives, relevant assessments, and a transparent grading policy all contribute to student satisfaction. When Ralston-Berg (2011) asked students to rate the importance of 68 online course benchmark items, four of the students’ top 10 selections

³ Found under the general standards “Learning Objectives (Competencies)” and “Assessment and Measurement.”

were related to course objectives and their measurement.⁴ Respondents also specified that a quality online course should have “information presented in a logical progression, broken down into lessons that are spaced apart properly” and that the course content should be “straightforward and what will be on the tests.” In a qualitative study, Hara and Kling (1999) provided an example of how unclear course objectives can negatively affect student performance. In an attempt to provide students with flexibility (and a lesser degree of transactional distance), the instructor profiled by Hara & Kling (1999) did not specify course objectives or expectations for assignments. However, many students did not consider this an advantage; rather, several were frustrated by their uncertainty regarding the instructor’s expectations.

Interpersonal Interaction

Nearly every published online quality framework has emphasized the importance of interpersonal communication and collaboration. The Middle States guidelines (2002), for example, explicitly state that courses and programs should be designed to promote “appropriate interaction” between the instructor and students, as well as among students. Most frameworks also detail specific best practices under the general heading of interpersonal interaction. The Quality Matters guidelines, for instance, include four items under the general standard of “course activities and learner interaction,” as well as two other items regarding self-introductions by the instructor and students (Quality Matters Program, 2014).⁵

Online learning theories have also strongly emphasized the critical role of interpersonal interaction (e.g., Anderson, 2003; Moore, 2013; Moore & Kearsley, 1996; Scardamalia & Bereiter, 2006), which is thought to have two key impacts on student learning. First, theorists and practitioners believe that collaborative work can help build a learning community that encourages critical thinking, problem solving, analysis, integration, and synthesis; provides cognitive supports to learners; and ultimately

⁴ The four specific items were: “How work will be evaluated,” “Grading policy,” “Assessments measure learning objectives,” and “Assessment time, appropriate to content.”

⁵ The two items are “the self-introduction by the instructor is appropriate and available online,” and “learners are asked to introduce themselves to the class,” both of which are listed under the general standard “Course Overview and Introduction.”

promotes a deeper understanding of the material (Fulford & Zhang, 1993; Kearsley, 1995; Moore & Kearsley, 1996; Friesen & Kuskis, 2013; Picciano, 2001; Salmon, 2002, 2004; Scardamalia & Bereiter, 2006; Sherry, 1995). Second, interpersonal interaction may help reduce transactional distance, and strengthen students' psychological connection to the course by enhancing "social presence"—the degree to which a person is perceived as a "real person" in mediated communication (e.g., Gunawardena & Zittle, 1997; Moore, 2013; Shearer 2013; Short, Williams & Christie, 1976; Young, 2006).

Survey research has bolstered the notion that effective student–instructor and student–student interactions are critical to effective online learning (e.g., Fredericksen et al., 2000; Smissen & Sims, 2002; Baker, 2003; Ralston-Berg, 2010, 2011). Perhaps more importantly, a passel of empirical studies have also focused on interpersonal interaction, including both student–instructor interaction (e.g., Arbaugh, 2001; Picciano, 2002; Young, 2006) and student–student interaction (e.g., Bangert, 2006; Matthew, Felvegi, & Callaway, 2009; Balaji & Chakrabarti, 2010). Bernard et al. (2009) conducted a meta-analysis of 74 studies of interaction in online learning⁶ and concluded that increased interpersonal interaction within the framework of the course, either with the instructor or with student peers, positively affects student learning.

More recently, theorists and researchers have begun to move beyond examining the extent of such interaction to investigating its quality. This line of questioning suggests that the mere existence of communication and collaboration is not sufficient to improve student learning. Rather, communication and collaboration must have a clear purpose and facilitate content delivery (Baran & Correia, 2009; Naidu, 2012). In support of this notion, one qualitative study (Ho & Swan, 2007) found that the quality of participation in online discussions (defined in terms of whether the posting was a new contribution, was reflective of the student's opinions, and was supported by sufficient evidence where necessary) predicted students' own grades in the course. Another study (Balaji & Chakrabarti, 2010) asked students to rate their course's online discussion forum in terms of the perceived quality, interactivity, and participative nature of the discussion.

⁶ Each study compared a treatment versus control condition, where the treatment involved a stronger element of interpersonal interactivity and the control included a lesser or non-existent level of interactivity.

Perceived quality of discussion was positively related to students' participation and interaction, as well as to self-perceived learning.

Technology

Most online learning quality rubrics incorporate standards regarding the availability of technology and its ease of use. For example, the Quality Matters rubric specifies that course technologies should be current and that students should have ready access to required technologies (Quality Matters Program, 2014). Survey studies support this general notion; two of the students' top 10 selections in the Ralston-Berg (2011) study were related to easily accessible and downloadable technology.⁷ Grandzol and Grandzol's (2006) review of best practices also indicates that students prefer to interact with course content via "current" technologies (such as PowerPoints with voiceover narration) rather than by reading a textual explanation.

An emerging literature has focused not only on the availability of technology but also on how it is employed. As Fulton noted, "dazzling technology has no value unless it supports content that meets the needs of learners" (2001, p. 22). Indeed, a recent review of the effectiveness of specific technology tools that are frequently incorporated into online courses (such as discussion boards, online quizzes, and embedded video) found that the mere presence of such tools did not seem to affect student learning (U.S. Department of Education, 2010). Rather, technology tools that did promote student learning were those that provided students with more control over their interactions with media (e.g., Balaji & Chakrabarti, 2010) and those that encouraged learner reflection (e.g., Roschelle et al., 2010). These results suggest that simply incorporating current technology into a course is not sufficient to improve student success; rather, the technology must also be thoughtfully leveraged to support student learning in the service of defined learning objectives.

2.3 The Practical Usefulness of Existing Online Rubrics

⁷ The two items are "Technology available, easily downloaded" and "Course components web-based or easily downloaded."

Overall, the existing quality rubrics' standards dovetail fairly well with the theoretical, survey, and empirical research literature; all four strands of the literature tend to agree that a high-quality online course should be well organized, with clearly specified learning objectives, an appropriate level of interpersonal interaction, and the use of current technologies. However, the existing rubrics have two key limitations.

First, specific rubric items have been validated only in terms of student and faculty ratings, perceptions, and opinions (e.g., Ralston-Berg, 2011), rather than in terms of student outcomes. Thus it is unclear whether a course deemed "high-quality" by a specific rubric will have stronger student success rates than would a "low-quality" course.

Second and perhaps more importantly, the rubrics' grading criteria tend to focus on the presence or absence of surface-level characteristics. For example, while the literature on interpersonal interaction and technology usage are increasingly beginning to emphasize the quality of these activities and tools, most existing rubrics merely indicate whether or not they are available. This tactic is understandable; it is much easier to quickly and reliably assess "course quality" if the grader has only to mark the presence or absence of various characteristics. Yet this method may not provide much insight into whether the course provides a high-quality learning experience. For example, an instructor may succeed in creating a highly engaging and interactive learning environment without necessarily using every strategy on a quality checklist. Such an instructor would receive a lower score than a second teacher who mechanically adhered to each item on the checklist, even if the second instructor's course seemed sterile, boring, and impersonal. A deeper and more nuanced examination of quality, however, may seem infeasible to many colleges, which are concerned about the staffing and resource requirements of peer evaluation.

Overall, our review of existing online quality rubrics suggested that it may be helpful to create a rubric that assesses not only the presence of specific quality elements but also how they are leveraged in the service of learning – using a quick and efficient assessment method. Moreover, it would be particularly helpful to validate each rubric area against student outcomes.

3. Methods

This particular analysis is part of a larger study of online learning (see Jaggars, 2014) which collected data at two community colleges within the same state. Researchers examined the course offerings for the spring 2011 semester and selected the most popular introductory academic subjects offered online (freshman composition, sociology, psychology, humanities, music, information technology, business, mathematics, and chemistry) as the focus of the research. Below we describe in additional detail only those aspects of the study unique to this analysis.

3.1 Analysis Sample

Instructors teaching an online course within the selected subject areas were invited to participate in the study, resulting in 19 faculty participants who taught 23 courses (some faculty taught multiple courses) in a total of 35 online course sections during spring 2011. After the semester was complete, the state system provided us with anonymized information on the $N = 678$ students who completed at least one of the 35 course sections, including transcript information and demographic characteristics. End-of-semester course grades were converted to a 0–4 scale, with 0 representing an *F* and 4 representing an *A*.

Across the 678 students, 90% were enrolled in only one of the 23 courses, 10% were enrolled in two, and less than 1% were enrolled in three. Students were primarily female (76%), either White (56%) or Black (34%), and under age 25 (54%). Most were on a transfer-oriented track (62%), with fewer in a career–technical track (32%) and a remainder of students who had an unknown or undeclared status. Overall, 87% were continuing students (i.e., had already been at community college for at least one semester), 69% had previously taken an online course, and 51% were attending school full time during the semester under study, with the sample carrying an average load of 10.46 credits. Continuing students had earned an average of 28 credits prior to the semester under study, with an average GPA of 2.74. In the courses included in the current study, the average end-of-semester grade was 2.32 (with continuing and new students earning fairly similar grades, 2.32 versus 2.30). Among a subsample of students enrolled

in these courses who were interviewed ($N = 43$, see details below), three-quarters were employed and approximately one-third mentioned dealing with childcare responsibilities.

3.2 Assessing Each Area of Quality

To assess the design features of each course, we developed a rubric that addresses four areas:

- *Organization and presentation*, which examines ease of navigation, as well as clarity and organization of materials;
- *Learning objectives and assessments*, which evaluates whether the course clearly outlines course-level and unit-level objectives, along with clear expectations for assignments;
- *Interpersonal interaction*, which assesses the effectiveness of interpersonal interaction in reinforcing course content and objectives; and
- *Use of technology*, which examines the effectiveness of the chosen technology to support learning objectives.

The Appendix includes a more detailed description of the quality expectations for each of the four areas. For the sake of concision, we will hereafter refer to the four areas as “organization,” “objectives,” “interaction,” and “technology.”

Our primary aim for the rubric was to help the rater think through a complex set of quality characteristics within each area (in a more deeply-reflective fashion than is required by a yes/no checklist), while capturing the rater’s judgment quickly and efficiently (similar to the time required by a yes/no checklist). To assess a course’s quality in each area, one member of the seven-person research team logged into the course website several times each semester to view ongoing activities⁸ (e.g., announcements, discussion board postings, chat sessions), collected course documents (including but not limited to course syllabi, assignments, and other written materials), provided detailed comments regarding the extent to which the course met expectations

⁸ From a practical perspective, all these activities were archived and could have been viewed at the end of the semester if necessary. We chose to monitor the courses throughout the semester in order to inform our ongoing interviews and other research activities.

for each of the four areas, and then provided a numeric rating for each area. A second researcher then reviewed the original documentation and provided their own rating.

In order to provide a final set of numeric ratings for each course, we needed to create scoring guidelines that would ensure relatively high agreement between different raters. In pilot attempts at scoring, our team found it difficult to agree on ratings using a fine-grained scale (e.g., a 5-point scale). Using a 3-point scale, rater disagreements were rare for the areas of objectives, interaction, and technology, but remained common for the area of organization. Such cases were resolved in conference between the two raters and the research director. The final rating scale for each area ranged from 1 (lower quality) to 3 (higher quality), as described in the Appendix.

For courses with multiple sections taught by the same instructor, each section was fairly identical to the others; instructors typically used exactly the same materials and structures across sections, and levels or types of interaction did not vary markedly across sections. Accordingly, each instructor's specific course (rather than the specific section) was treated as the unit of analysis.

3.3 Additional Qualitative Data

As part of the larger study, the research team also conducted in-depth interviews with 24 instructors (including all 19 who taught one of the courses used in this analysis) and 47 students (43 of whom were enrolled in one of these courses). Interviews were conducted using semi-structured protocols focused on experiences and perceptions related to online learning, and for students, particularly on their learning experiences in the online courses in which they were currently enrolled. All interviews were transcribed and then coded using ATLAS.ti qualitative data analysis software, using an array of codes related to different research topics of interest (for other analyses produced using these data, see Bork & Rucks-Ahidiana, 2013; Jaggars, 2014). For the current analysis, we used four codes (one for each of the four course design features) to flag instructors' reflections and students' experiences relevant to the given area. We also pulled data on the observed characteristics of each course, using the raters' course descriptions. In Section 5, we utilize both interview and observation data to explore the qualitative differences between courses that received a high versus a low rating in a given area.

4. Quantitative Results

4.1 Descriptive Statistics for Each Area

Table 1 presents the rubric score means and intercorrelations across the 23 online courses observed. On a descriptive basis, courses tended to have somewhat higher ratings in terms of interaction and objectives, and somewhat lower ratings in terms of technology and organization. The four ratings tended to be moderately intercorrelated, with the strongest correlation being between technology and objectives.

Table 1
Means and Correlations for the Four Quality Ratings ($N = 23$ courses)

Subscale	Mean (SD)	Organization	Objectives	Interaction
Organization	1.78 (0.60)			
Objectives	2.00 (0.74)	0.31		
Interaction	2.04 (0.77)	0.42*	0.32	
Technology	1.78 (0.80)	0.47*	0.62**	0.46*

4.2 Predicting Course Grades

We first correlated each of the four ratings with student end-of-semester grades on a bivariate basis ($N = 678$ students). Overall, only interaction ($r = 0.15, p < .001$) and technology ($r = 0.12, p < .01$) were significantly associated with student grades; the correlations for organization ($r = -0.05, n.s.$) and objectives ($r = 0.05, n.s.$) were negligible. In order to more precisely estimate the relationship between the course ratings and student outcomes, however, it is important to control for student-level characteristics that could obscure (or magnify) the impact of each area of quality on student outcomes. Given that the ratings were measured at the course level and that other student characteristics were measured at the individual student level, we employed a multilevel model (also known as a random-effects model, mixed model, or hierarchical linear model) in order to include each predictor at the appropriate level of analysis.

In a multilevel model, the variance in the outcome is partitioned into two pieces: “course-level variation” (τ_{00}) and “student-level variation” (σ^2). In rough conceptual terms, one can think of course-level variation as the differences in average student grades between courses (for example, in one course, the average grade might be a B+ while in another course it might be a C+), and student-level variation as the differences in grades between students in a given course (for example, in a course with an average grade of C, some students will do better than a C while others will do worse). Typically, a course-level predictor explains only course-level variation, while a student-level predictor can explain both student-level and course-level variation.⁹ As a consequence, our effective sample size for assessing the predictive capability of the course-level quality ratings was quite small, at $N = 23$.

A first step in the multilevel modeling process is to determine whether there is a significant degree of variation across courses in terms of student grades—that is, if there is no such variation, then there is nothing to explain. For this study, a basic unconditional-means model suggested that there was a significant degree of variation in student grades across the sample of online courses ($\tau_{00} = 0.48$, $SE = 0.17$, $p < .01$), as well as a significant degree of variation in grades across students within any given course ($\sigma^2 = 1.79$, $SE = 0.10$, $p < .001$).

Our second step was to add student-level predictors to the model (including whether the student had taken a prior online course, prior credits accrued, current-semester credit load, aged 25 or older, White, female, and type of academic track) to help remove unexplained variation at both the student and course level. Overall, the set of predictors explained 6% of the within-course student-level variation and 8% of the course-level variation.

In a third step, we added the four course-level quality ratings as predictors, which improved the explanation of the course-level variation substantially, to 23%. However, subsequent tests performed with each of the four ratings on a separate basis indicated that

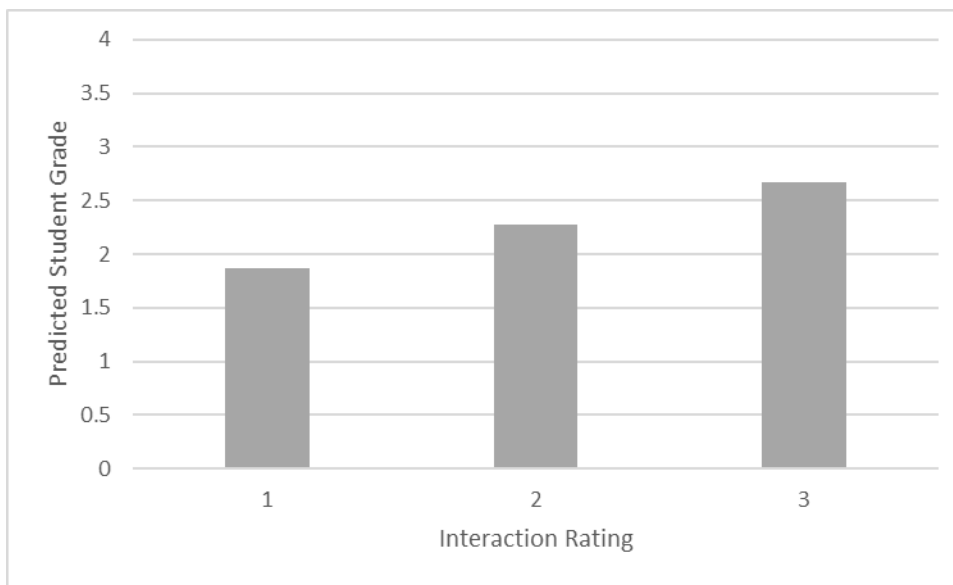
⁹ More precisely, the estimate for any course-level predictor is adjusted for differences between courses in terms of the student-level variables. For this study, all student-level predictors were grand-mean-centered. For more details on the application of multilevel models in the educational context, the reader may find it useful to consult Raudenbush and Bryk (2002), Kreft and deLeeuw (1998), and Singer (1998).

all of this additional variance was attributable solely to the area of interaction.¹⁰

Accordingly, in our final model, we discarded the other three ratings and focused on the interaction rating.¹¹

The final model explained 23% of the variance in course-level grades and indicated that the interaction rating had a significant positive impact on student grades ($b = 0.40$, $SE = 0.19$, $p < .05$). As shown in Figure 1, the average student (i.e., a student with a mean score on each of the covariates included in the model) taking a course with the typical interaction rating of 2 would have an estimated course grade of 2.27, or a C-. The results of the final model suggest that if the same student took a course with an interaction rating of 3, his or her estimated course grade would increase to 2.67, or a C+.

Figure 1. Predicted Student Course Grade by Interaction Rating



4.3 Potential Limitations and Sensitivity Analysis With Prior GPA

¹⁰ That is, when each of the other three ratings were added separately to the model, each explained 0% additional course-level variance over and above the student predictors.

¹¹ Ideally, we would also explore the contribution of course subject (e.g., English, Chemistry) to the explanation of variance, given that different disciplines may be more or less difficult for the typical student, or more or less oriented toward interpersonal interaction or the use of technology tools. In our dataset, we had too few degrees of freedom to allow inclusion of course subject; however, we did not note obvious differences between subject areas in terms of course ratings. For example, despite the common stereotype that physical science courses are depersonalized, the most highly-interactive class we observed was a Chemistry course.

The conclusions of our analysis are limited by several methodological issues. First, while the students' records contained useful demographic information, other important student characteristics were unobserved and could potentially cause both higher levels of course-level interaction and stronger grades. For example, a particular course section might be filled with students who were naturally quite interactive, motivated, and self-directed; these students would increase the overall level of interaction in the course while also earning higher grades. Similarly, an instructor might reward a class full of naturally-interactive students with higher grades, due to a perception that such students are more motivated or harder-working. We believe an unobserved student-level disposition toward interactivity is unlikely to completely explain the significant relationship between course-level interactivity and student-level grades because – as we will discuss in more detail below – course-level variation in interactivity seemed driven primarily by instructor actions, rather than by student initiative. To assess the extent to which student-level unobservables might explain away the relationship between course interactivity and student grades, however, we conducted a sensitivity check using students' prior GPA, among the subset of 591 students who had attended the college for at least one semester. The unconditional-means model for the reduced sample had similarly-sized variance components at the course level ($\tau_{00} = 0.50$, $SE = 0.18$, $p < .05$) and the student level ($\sigma^2 = 1.75$, $SE = 0.10$, $p < .001$). Including prior GPA together with the other student-level predictors in the second step resulted in much stronger explanatory power, explaining 37% of the course-level variance. Adding the interaction rating in the third step of the model improved the proportion of explained variance by about 6 percentage points to 43%. The final model indicated a marginally significant positive effect for the interaction rating ($b = 0.30$, $SE = 0.17$, $p < .10$).

Second, our rating scales may not have captured the most important variation within each design area. In particular, the weak predictive power of course organization and structure could be related to its measurement. As noted above, grading for this subscale was the most difficult for our research team to agree upon. There were a few courses that were clearly poorer in terms of organization and navigability and a few that were clearly better, but most were a mix, combining some confusing organizational features with others that were more clear. These courses tended to earn a 2 on our grading

scale. It is possible that some of these organizational features were more important than others. Negligible associations with course outcomes for both the organization and the learning objective scales could also be due to restricted range. While there was certainly some variability in each scale across courses, it is possible that only a minimal threshold is necessary for each area and that variability beyond that threshold is unimportant. And in terms of interaction, we could not assess the quantity, timeliness, and quality of private personal communications such as emails and face-to-face office meetings. However, as we will discuss below, we conducted interviews with a subset of students enrolled in the 23 target courses, and students' reports regarding the quantity and quality of private communications were consistent with the public communications we were able to observe and assess.

5. Qualitative Data on Interpersonal Interaction

Overall, the areas of organization and objectives appeared to have no relationship with student grades. Interaction had the strongest relationship, while technology had only a raw relationship that did not endure after controlling for student-level characteristics. To shed more light on the meaning of the interpersonal interaction ratings, this section uses course observation and interview data to describe in more detail courses that earned high-quality ratings in terms of interpersonal interaction. Below we discuss both student–instructor and student–student interaction. Overall, however, the 23 observed courses varied more obviously and strongly in terms of student–instructor interaction than in terms of student–student interaction; accordingly, differentiation in ratings between the courses tended to be driven most strongly by student–instructor interaction.

5.1 Student–Instructor Interaction

In courses with high interpersonal interaction ratings, instructors tended to post frequently, invite student questions through a variety of modalities, respond to student queries quickly, and solicit and incorporate student feedback. These elements are discussed in more detail below.

First, high-interaction instructors posted announcements on a regular basis to remind students about requirements for assignments, coming deadlines, newly posted documents, examinations, and other logistic issues. In courses where the instructor made

limited announcements, students were more likely to express dissatisfaction with the course. For example, one student complained:

I understand it's distance learning, but here I am just studying away, and then I have to—to me, it's just too lenient. Like, he doesn't say "This is due, this is due, this is due," like the other class. ... It's very lenient, but he wants it all done by the end.

Second, students in high-interaction courses reported that their instructors responded to questions in a timely manner (typically, within 24 hours). These instructors also tended to provide multiple ways for students to communicate with the instructor, including email, telephone, discussion board postings, synchronous chatting, and in-person office hours. Struggling students particularly appreciated opportunities to interact face-to-face; for example, one student noted,

Actually with this test that we have coming up, I plan on finding out where her office is and try to get some extra help rather than emailing her and just getting a response. It's more helpful.

Another student noted that both in-person office hours and video chats were helpful:

But I think once a month we should all meet with the instructor. ... Because you know, you can talk to a person over the computer all day long. But there's something about sitting in front of that person that does it for me. ... You know, I just don't want you to tell me something to shut me up. I want to see what your face looks like when you give me the answer.

High-interaction instructors were also more likely to ask for student feedback and seem responsive to that input. One student, when asked to rate his experience in the online course from one (least satisfactory) to five (most satisfactory), explained:

I'm going to say 5, and I'm going to say why. ... She has sent us emails twice, you know, [saying that] "People didn't do well, what can I do?" ... Like the one suggestion, she's going to implement it next semester. But I say "excellent" because she already does a good job, and she

wants to know what she can do to fix it. To me that says a lot.

The strategies above seemed to help students to feel that the instructor cared about the course and students' performance in the course, which in turn helped students personalize the instructor, feel connected to the course, and strengthen their motivation to learn and succeed. Effective teacher interaction and the sense that the teacher actually "cares" seemed to carry a lot of weight in students' assessments. One student noted:

I admire people like that; you're just not a number. ... That they really care about your learning, whether it's face-to-face or online. ... Maybe you are a little bit more personal [in face-to-face classes] because you are face-to-face, and you can put a face, and they can really "size you up." ... You may lose that little piece [in an online course], but I'm not looking for that. I just need you to talk to me when I need you and I tell you I'm struggling. Or if I'm struggling or just have a question or whatever, and that helped me.

It seemed that students could easily distinguish between instructors who cared and those who did not; as one student explained:

Some teachers, they just don't care. Like you're on your own, like you're online, like it doesn't matter. You're getting a grade. You do my work online and call it a day. ... Some teachers, they want you to understand the difference in online and in-class, and they want you to be helped, and they know that you can't get to class. Some teachers, they try to understand, and some teachers they just don't care, like they're getting paid regardless.

Several students made explicit the link between teacher interaction and caring. One student who felt that the instructor really cared described her experiences with the instructor during office hours:

I actually study with her. It's hard to find teachers like that. ... That will study with you for like an hour-and-a-half. On campus, yeah, office hours. ... So we'll set it up in a chat or through email. I'll go over there to her and sit down and go over, if I have questions. ... She's very personable.

Another student in a humanities course appreciated the helpfulness that came through in the narrative videos her instructor had created. When asked whether she had a sense that the instructor cared, she responded:

Yes, because he has videos and stuff, so I can see or hear what he expects in his videos, like he's trying to help me, basically. ... And I see that he has examples, and I see that if we don't know what we're doing, that we can click on what someone else said and read what they said.

Another student was able to sense the instructor's caring through live chat and the discussion board:

She really has a passion for it. ... She was like, "You don't get it; well, what can I [do]?" She really wants you to understand it. ... She took the time to learn how to use the live chat and the communication discussion board, you know, audio. ... She's just a "one of a kind" type of instructor. She's someone I'll never forget just because of her passion and how much she cares.

In a similar vein, high-interaction instructors often brought up in their interviews that online students may feel lonely, frustrated, isolated, or less motivated than students in traditional face-to-face classes. As a result, these instructors felt that it was particularly important to make their students feel that they cared and were actively engaged in the course. One accounting instructor, when discussing a previous course that did not go well, explained:

So I think all of that, in addition to the mere nature of it being distance learning, was contributing to not a very good success rate. ... So because of that, that's how I kind of ventured out in the technology piece. We've got to do something else; we've got to make them feel like they're here. ... We've got to make them feel like we care. ... You know, we've got to make them feel like we're engaged with them. ... Rather than "Okay, here are all of these things to use, you know, go at it."

5.2 Student–Student Interaction

Across the 23 courses included in the study, students did not seem particularly engaged with one another. Almost all of the courses included a student discussion board,

but in general, student postings were perfunctory. Most courses required a minimum number of student postings, but the content of these posts tended to be as superficial as possible, such as “I agree,” or “good job.” When asked about student interactions, several instructors said that although they encouraged students to use the discussion board to communicate with their peers, the students seldom did so. One instructor said:

That’s a piece of it that I’m sort of “up in the air” on, and don’t really know how to bring it together. I used to have discussion boards out there for “chapter one discussion board,” “chapter two discussion board,” where they could go out there and post something about the chapter, and either I would respond or someone else would respond. . . . No one ever did. . . . So the point with this conversion over the spring, I didn’t even bother. . . . Because it hadn’t worked in the past.

In contrast, in a few of the high-interaction online courses, students posted on the discussion board more regularly and thoughtfully. In these courses, postings were not only compulsory and graded, but instructors also provided a clearer sense of their expectations in terms of what constituted a “high-quality” initial post or response to a post. For example, one instructor noted in his syllabus: “Discussion board posts and replies will be closely evaluated for depth of thought and insight into the question. Replies to peers must be thoughtful, and should not simply agree with the author, or state ‘me too’.” He also made explicit that each post would be assessed in terms of focus, specificity, support, thoughtfulness, and use of language, with each counting for 2 points. Taking “focus” as an example, a student would receive 2 points for making “vividly clear references to specific readings,” 1 point for making “some reference to readings,” and 0 for making “no reference to readings.” Moreover, required responses to other students’ initial postings were also graded in terms of the extent to which they addressed points or questions raised by the initial post and drew upon readings to validate their response. The instructor felt that the grading rubric helped encourage discussions that could build “a real, personal connection” among students in an online class, “like there is in a real, traditional classroom”:

So trying to build community any way you can, I mean it helps all sorts of things. It helps the response in your courses; it helps retention for the college. It just has so

many impacts, so many different layers of impact on the success of the student.

Despite encouragement from instructors such as this one, most student respondents seemed disinterested in interacting with their online peers. When asked whether it was important to have interactions with other online students, one student from this instructor's class responded that he does not "need it for that class." Most students interpreted their experiences with online course discussions as a forced artificial communication that neither resembled spontaneous personal connections in a face-to-face classroom nor led to active learning. The following comment vividly illustrates students' perceptions of peer interactions in an online environment:

Face-to-face with my math class, at times we joke with each other and stuff. And then someone needed some notes, so she just asked. She sent an email, but then she asked. We have more of an interaction because we see each other at least twice a week. Online, we only interact because that's what we're told to do. If it was just the teacher giving me an assignment and just for me and whatever, I probably wouldn't interact with my fellow students. ... As far as actual, direct communication or anything with them, it is different. I only do it because that's what the assignment says to do.

In addition to a general indifference to online peer interactions, some students reported negative experiences with required group work in their online courses. One student provided an example:

In one of my other online classes I had to do some group work and stuff, and I just, I don't like doing that. I don't mind doing group work in a classroom setting, but online, it's just much more difficult.

When asked why it is more difficult to complete group work online, this student listed two major reasons: problems finding a common time for group work and a lack of commitment from group members.

If we're not together in classroom every day or every other day, that makes it really difficult for us all to find time to work together and to, you know, if one person isn't going to do any work, then you guys have to pick up the slack and

then they still get credit for it because you work online so nobody knows, you know, who's doing what. . . . So it just, it's just not a good system that way, and I just, they haven't done any kind of group work this semester in my online classes, which I'm thankful of because it's just extremely difficult and not really worth it basically.

6. Discussion and Conclusion

Overall, our findings indicate that while well-organized courses with well-specified learning objectives may be desirable, these qualities may not have an impact on student grades per se. In addition, while students appreciated courses that leveraged appropriate learning technologies (in contrast to courses that were extremely reading-heavy), that factor also did not necessarily impact student grades. Only the area of interpersonal interaction predicted student grades in the course.

While it is difficult to draw definitive conclusions due to our small sample size, it seems that courses in which the instructor posted frequently, invited student questions through a variety of modalities, responded to student queries quickly, solicited and incorporated student feedback, and (perhaps most importantly) demonstrated a sense of “caring” created an online environment that encouraged students to commit themselves to the course and perform at a stronger academic level. This finding aligns with work by Young (2006), which suggests that students view an effective online instructor as one who is actively involved, provides feedback to students, adapts to students' needs, and encourages students to interact with their classmates, their instructor, and the course material. The importance of an engaged instructor also accords with Holmberg's (1995) belief that instructors must create a personal relationship with students in order to motivate them to succeed.

Qualitative data imply that students in our sample valued student–instructor interaction more strongly than student–student interaction. In fact, many students viewed peer interactions in their online courses as an imposed requirement, rather than as a helpful and necessary component of learning. The general indifference to peer interaction among our sample of students could be due in part to their non-traditional characteristics. Most had professional or familial obligations that constrained their schedule and may

have made it difficult for them to participate in collaborative tasks. This observation echoes some scholars' assertions that imposed peer interactions in distance courses may interfere with student autonomy in managing the time, place, and pace of their learning, and may not necessarily benefit students (e.g., May, 1993; Ragoonaden & Bordeleau, 2000).

In general, across the literature on interpersonal interaction in the online setting, there is some disagreement as to whether student–instructor or student–student interaction is more critical to learning. Some work aligns with our observation that student–instructor interaction appears to be more salient. For example, a survey of students (Fredericksen et al., 2000) found that student–instructor interaction was rated more strongly than student–student interaction as contributing to perceived learning in distance education. Similarly, Sener (2001) found that the “tutorial model” of online instruction (featuring a strong reliance on student–instructor interaction without requiring student–student interaction) had student success rates comparable to courses adopting a “mandatory interaction model” (which required students to interact with both their instructor and with each other), suggesting that the addition of student–student interaction was unnecessary. In contrast, however, Bernard et al.'s (2009) meta-analysis found a stronger impact on student learning in studies of online interventions that increased student–student interaction (effect size = 0.49) compared with those that increased student–instructor interaction (effect size = 0.32).

One prominent theory of online learning (Anderson, 2003) may reconcile these disparate results with its suggestion that high-quality learning can occur when supported by *either* strong student–instructor or strong student–student interaction. Anderson then goes even farther to suggest that perhaps *neither* type of interpersonal interaction need be present, if course technology is sufficiently robust to support high student–content interaction (2003, p. 4):

Deep and meaningful formal learning is supported as long as one of the three forms of interaction (student–teacher; student–student; student–content) is at a high level. The other two may be offered at minimal levels, or even eliminated, without degrading the educational experience. High levels of more than one of these three modes will likely provide a more satisfying educational experience,

though these experiences may not be as cost or time effective as less interactive learning sequences.

In our study, the effective use of technology to support learning objectives was initially associated with course grades, yet the association disappeared after student characteristics were controlled. This finding may suggest that the learning technologies incorporated into these courses were insufficiently robust to promote strong student-content interaction. Course technologies were not necessarily entirely disconnected from student grades, however, as some technologies helped boost the course's interpersonal interaction rating. For example, video- or audio-taped narrations helped instructors establish a sense of personal presence.

Overall, students in our study seemed most concerned with their connection to their instructor, which could be due in part to the profile of the community college population. In an in-depth qualitative examination of community college students, Cox (2009) found that many were disheartened or even paralyzed by the fear that they could not succeed in college. Cox noted that:

Some students who started the semester with overwhelming fears were able to manage the anxiety and go on to complete their courses successfully. Doing so, however, required active intervention from someone at the college—someone who could reassure students about their academic competence and ability to succeed. For the students in these interviews, professors could play that role, especially when they were able to “come down to students’ level” or if they were the kind of instructor who “really relates.” (p. 159)

Cox's findings align with our observation that the sense of caring, which was communicated through interpersonal interaction, seemed particularly salient to students in their conversations about course quality.

While our research was a good first step toward examining the predictive power of course quality rubrics, additional research is urgently needed, given the increasingly widespread use of such rubrics in online faculty training and evaluation. In particular, it would be quite helpful for colleges that currently use existing course quality rubrics to conduct and publish studies of the predictive validity of those rubrics in terms of current course student satisfaction and learning outcomes, as well as subsequent outcomes such

as next-semester persistence and follow-up course grades. The results of our study suggest that these rubrics' measurement interpersonal interaction should include not just a checklist in terms of whether certain types of interactions are present in the course, but also a more nuanced assessment of how the teacher communicates interpersonal presence and caring. Similarly, quality ratings for technology may wish to focus on not just the usage of "current" technologies, but how these technologies are used to support student interaction, confidence, motivation, and learning. As the research evidence on course quality rubrics begins to mount, practitioners can focus on refining rubrics for improved predictive validity, as well as improving the quality of online course design and delivery itself.

Acknowledgement: This research was funded by the Bill & Melinda Gates Foundation. We wish to thank other members of the research team who contributed to the data collection and coding for this analysis: Nikki Edgecombe, Melissa Barragan, Rachel Hare Bork, Thao Tran, and Zawadi Rucks-Ahidiana.

References

- Allen, I. E., & Seaman, J., with Hill, P., & Poulin, R. (2015). *Grade Level: Tracking Online Education in the United States*. Washington, DC: Pearson, Babson Survey Research Group, Online Learning Consortium, and Tyton Partners.
- Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *International Review of Research in Open and Distance Learning*, 4(2), 9–14.
- Arbaugh, J. B. (2001). How instructor immediacy behavior affects student satisfaction and learning in web-based courses. *Business Communication Quarterly*, 64(2), 42–54.
- Baker, J.D. (2003). Faculty and Student Perspectives on Educational Technology and Online Learning. Presented to the Association for Advancement of Computing in Education, Society for Information Technology and Teacher Education 13th International Conference, Albuquerque, MN.
- Balaji, M. S., & Chakrabarti, D. (2010). Student interactions in online discussion forum: Empirical research from ‘media richness theory’ perspective. *Journal of Interactive Online Learning*, 9(1), 1–22.
- Bangert, A. (2006). Identifying factors underlying the quality of online teaching effectiveness: An exploratory study. *Journal of Computing in Higher Education*, 17(2), 79–99.
- Baran, E., & Correia, A.-P. (2009). Student-led facilitation strategies in online discussions. *Distance Education*, 30(3), 339–363.
- Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L.,... Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, 74(3), 379–439.
- Benson, A. (2003). Dimensions of quality in online degree programs. *The American Journal of Distance Education*, 17(3), 145–159.
- Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M.A., & Bethel, E. C. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research*, 79, 1243–1289.
- Bork, R. H., & Rucks-Ahidiana, Z. (2013). *Role ambiguity in online courses: An analysis of student and instructor expectations* (CCRC Working Paper No. 64). New York, NY: Community College Research Center.

- Cox, R. D. (2009). *The College Fear Factor: How Students and Professors Misunderstand One Another*. Harvard University Press.
- Fabianic, D. (2002). Online instruction and site assessment. *Journal of Criminal Justice Education, 13*, 173–186
- Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it Internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics, 31*(4), 763 – 784.
- Fredericksen, E., Pickett, A., Shea, P., Pelz, W., & Swan, K. (2000). Student satisfaction and perceived learning with online courses: Principles and examples from the SUNY learning network. *Journal of Asynchronous Learning Networks, 4*(2).
- Friesen, N., & Kuskis, A. (2013). Modes of interaction. In M. G. Moore (Ed.), *Handbook of Distance Education*, 3rd ed, pp. 351-371. New York, NY: Routledge.
- Fulford, C.P., & Zhang, S. (1993). Perceptions of interaction: The critical predictor in distance education. *The American Journal of Distance Education, 7*(3), 8–21.
- Fulton, K (2001, March/April) From Promise to Practice: Enhancing Student Internet Learning. MultiMedia Schools. Retrieved on April 5th, 2013 from <http://www.infotoday.com/MMSchools/MMStocs/mar01toc.htm>
- Grandzol, J. R., & Grandzol, C. J. (2006). Best practices for online business education. *The International Review of Research in Open and Distance Learning, 7*(1).
- Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American Journal of Distance Education, 11*(3), 8–26.
- Hara, N., & Kling, R. (1999). Students’ frustrations with a web-based distance education course. *First Monday, 4*(12). Retrieved on April 05, 2013 from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/rt/prINTERfriendly/710/620>
- Ho, C. H., & Swan, K. (2007). Evaluating online conversation in an asynchronous learning environment: An application of Grice’s cooperative principle. *Internet and Higher Education, 10*, 3–14.
- Holmberg, B. (1995). The evolution of the character and practice of distance education. *Open Learning, 10*(2), 47-53.
- Jaggars, S. S. (2013). Online learning in community colleges. In M. G. Moore (Ed.), *Handbook of Distance Education*, 3rd ed, pp. 594-608. New York, NY: Routledge.

- Jaggars, S. S. (2014). Choosing between online and face-to-face courses: Community college student voices. *American Journal of Distance Education*, 28(1), pp. 27 – 38.
- Jahng, N., Krug, D., & Zhang, Z. (2007). Student achievement in online distance education compared to face-to face education. *European Journal of Open, Distance and E-Learning*. Retrieved on April 05, 2013 from <http://www.eurodl.org/>
- Johnson, H., & Mejia, M. C. (2014). *Online learning and student outcomes in California's community colleges*. San Francisco: Public Policy Institute of California.
- Kaupf, R. (2012). Online penalty: The impact of online instruction on the Latino-White achievement gap. *Journal of Applied Research in the Community College*, 19(2), 8–16.
- Kearsley, G. (1995). The nature and values of interaction in distance education. In *Third Distance Education Research Symposium*. University Park, PA: American Center for the Study of Distance Education.
- Keeton, M. (2004). Best on-line instructional practices: Report of phase I of an ongoing study. *Journal of Asynchronous Learning Networks*, 8 (2), 75-100.
- Kreft, I. G. G. & Leeuw, J. D. (1998). *Introducing Multilevel Modeling*. Sage Publications Ltd
- MacDonald, C. J., Stodel, E., Farres, L., Breithaupt, K., & Gabriel, M. A. (2001). The demand-driven learning model: A framework for Web-based learning. *The Internet and Higher Education*, 4, 9–30.
- Matthew, K. I., Felvegi, E., & Callaway, R. A. (2009). Wiki as a collaborative learning tool in a language arts methods class. *Journal of Research on Technology in Education*, 42(1), 51–72.
- May, S. (1993). Collaborative learning: more is not necessarily better. *American Journal of Distance Education*, 7(3), 39–49.
- Merisotis, J. P., & Phipps, R. A. (2000). *Quality on the line: Benchmarks for success in internet-based distance education*. Washington, DC: The Institute for Higher Education Policy. Retrieved on April 05, 2013 from <http://defiant.corban.edu/jjohnson/Pages/Teaching/qualityonline.pdf>
- Middle States Commission on Higher Education. (2002). *Best Practices For Electronically Offered Degree and Certificate Program*. Retrieved on April 05, 2013 from: http://continuingstudies.wisc.edu/campus-info/toolkit/online_article1.pdf

- Moore, M. G. (1994). Autonomy and interdependence. *The American Journal of Distance Education* 8 (2), 1-5.
- Moore, J. (2005). The Sloan Consortium Quality Framework and the five pillars. Newburyport, MA: The Sloan Consortium. Retrieved on June 14th, 2012 from <http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CKgBEBYwAQ&url=http%3A%2F%2Fsloanconsortium.org%2Fpublications%2Fbooks%2Fqualityframework.pdf&ei=klbfT6uXHtCK6QHA3oyuCw&usg=AFQjCNHJzBCmTVatNTnxVAU2rY6-cspeZA>
- Moore, M. G. (2013). The theory of transactional distance. In M. G. Moore (Ed.), *Handbook of Distance Education*, 3rd ed. New York, NY: Routledge.
- Moore, M. G., & Kearsley, G. (1996). *Distance education: A systems view*. Belmont, CA: Wadsworth Publishing Company.
- Naidu, S. (2013). Instructional design models for optimal learning. In M. G. Moore (Ed.), *Handbook of distance education* (3rd ed.), pp. 268-281. New York, NY: Routledge.
- Parsad, B., & Lewis, L. (2008). *Distance education at degree-granting postsecondary institutions: 2006–07* (Report No. NCES 2009-044). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Picciano, A. G. (2001). *Distance learning: Making connections across virtual space and time*. Upper Saddle River, NJ: Prentice-Hall.
- Picciano, A. G. (2002). Beyond student perceptions: Issues of interaction, presence, and performance in an online course. *Journal of Asynchronous Learning Networks*, 6(1), 21–39.
- Quality Matters Program. (2014). Quality Matters Higher Education Rubric, fifth Edition. Retrieved January 10, 2016 from <https://www.qualitymatters.org/node/2305/download/QM%2520Standards%2520with%2520Point%2520Values%2520Fifth%2520Edition.pdf>
- Ragoonaden, K., & Bordeleau, P. (2000). Collaborative learning via the internet. *Educational Technology & Society*, 3(3). Retrieved April 5, 2013 from http://www.ifets.info/others/journals/3_3/d11.html
- Ralston-Berg, P. (2010). *Do quality standards matter to students?* Paper presented at the 2nd Annual Quality Matters Conference, Oak Brook, IL.
- Ralston-Berg, P. (2011). *What makes a quality online course?* Paper presented at the 3rd Annual Quality Matters Conference, Baltimore, MD.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, California, CA.

- Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal*, 47(4), 833–878.
- Salmon, G. (2002). *E-tivities: The key to active online learning*. London, United Kingdom: Kogan Page.
- Salmon, G. (2004). *E-moderating: The key to teaching and learning on-line*. London, United Kingdom: Kogan Page.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.
- Sener, J. (2001). Bringing ALN into the mainstream: NVCC case studies. In Bourne, J., and Moore, J. (Eds.), *On-Line Education, Volume 2: Learning Effectiveness, Faculty Satisfaction, and Cost Effectiveness*. (pp.7-29). Needham, MA: Sloan Center for OnLine Education.
- Shearer, R. (2013). Theory to practice in instructional design. In M. G. Moore (Ed.), *Handbook of distance education* (3rd ed.), p. 251 - 267. New York, NY: Routledge.
- Sherry, L. (1995). Issues in distance learning. *International Journal of Educational Telecommunications*, 1(4), 337–365.
- Short, J.A., Williams, E., and Christie, B. (1976). *The Social Psychology of Telecommunications*. London: Wiley.
- Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24 (4), 323-355.
- Smissen, I., & Sims, R. (2002). Requirements for online teaching and learning at Deakin University: A case study. In A. Treloar & A. Ellis (Eds.), *AusWeb02 : The Web enabled global village : Proceedings of the 8th Australian World Wide Web Conference*. Lismore, N.S.W.: Southern Cross University. Retrieved April 5, 2013 from <http://dro.deakin.edu.au/view/DU:30013885>
- U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC.
- Xu, D. & Jaggars, S. S. (2011). The effectiveness of distance education across Virginia's community colleges: Evidence from introductory college-level Math and English courses. *Educational Evaluation and Policy Analysis*, 33 (3), 360-377

- Xu, D., & Jaggars, S. S. (2014). Performance gaps between online and face-to-face courses: Differences between types of students and academic subject areas. *The Journal of Higher Education*, 85(5), pp. 633 – 659.
- Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. *Economics of Education Review*, 37, pp. 46 – 57.
- Young, S. (2006). Student views of effective online teaching in higher education. *American Journal of Distance Education*, 20(2), 65–77.
- Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, 107(8), 1836–1884.

Appendix: Online Course Quality Rubric

Below we present our online course quality rubric. We welcome colleges to use and adapt the rubric for their own purposes. We would issue the caution, however, that we regard the rubric as a work in progress – as noted above, it may require real modification and improvement in order to more effectively predict student grades. For each subscale, we include an overall description of “high quality” in that area. Descriptions include examples of practices that the literature suggests indicates high quality. However, rather than checking off each specific practice, the reviewer considers whether the course seems to adhere to the latent conception of quality indicated by the example practices. The description is followed by a rating scale used to record the reviewer’s overall sense of the level of quality in that area.

A.1 Organization and Presentation

The course has an easy-to-navigate interface that is generally self-explanatory and helps students to identify and manage course requirements. For example: locations of course materials are clearly labeled and consistently organized; the course homepage highlights materials and resources that are central to course learning objectives, and minimizes the presence of extraneous or redundant materials; and links to internal and

external web-based materials are integrated effectively with other course materials and kept up to date. The course includes a clear step-by-step guide to locating key course documents, assignments, and assessments.

1 – Unclear navigation in presentation of course and no instructions of how to approach navigation.

2 – Clear navigation in presentation of course, but no instructions of how to approach navigation.

3 – Clear navigation in presentation of course with step-by-step instructions of how to approach course navigation.

A.2 Learning Objectives and Alignment

Learning objectives and performance standards for the course as well as for each instructional unit are clear, so that students have information about what they need to know and will be asked to do. For example: objectives are outlined on the course site and syllabus; connections among learning objectives (i.e., how pieces of content relate to one another) are articulated to generate a more explicit rationale for and coherence across instructional activities; learning objectives are specific and transparent, detailing how student performance will be measured overall in the course as well as in each unit; and grading criteria are clear to students and reiterate performance expectations.

1 – Unclear course-level or unit-level objectives, along with unclear expectations for assignments.

2 – Some course-level objectives, unit-level objectives, or expectations for assignments are clear, and others are not.

3 – Course-level objectives, unit-level objectives, and expectations for assignments are clear and well-aligned with one another.

A.3 Interpersonal Interaction

Course includes plentiful opportunities for students to meaningfully interact with the instructor, and with other students, in ways that enhance knowledge development and create productive relationships among learners. For example: instructor feedback to

students is specific, actionable, and timely, clearly indicating what students are doing well and what opportunities are available for improvement; instructors use strategies to increase “instructor presence,” allowing students to become familiar with the instructor’s personality; student-student interactions are embedded in thoughtfully designed instructional activities that are relevant and engaging to students and that advance specified learning objectives; the types and nature of interactivity are determined by the desired learning goal, not by arbitrary criteria for collaboration or communication; interactions facilitate knowledge and skill application, not recitation.

1 – Little or no meaningful interpersonal interaction.

2 – Moderate meaningful interaction with instructor and/or amongst students.

3 – Strong meaningful interaction with instructor and amongst students.

A.4 Technology

Instructor leverages technology in support of pedagogical goals, allowing students to engage course content in ways that support the achievement of specific, measureable learning objectives. “Technology” as defined here does not include text-based readings, lecture notes, or slide presentations, but may include instructor- and publisher-created audio/video presentation tools, communication software or strategies, online information resources/archives, instructional software, etc. Technologies facilitate the diversification of instructional activities; however, rather than simply integrating a large number of technologies into the course, these technologies are used to help the instructor effectively meet particular pedagogical goals.

1 – Little or no use of technology.

2 – Limited set of technological tools, and/or use of tools but with little concrete linkage to course objectives.

3 – Consistent use of a variety of tools concretely linked to course objectives.