2013

Academic Performance, Course Completion Rates, and Student Perception of the Quality and Frequency of Interaction in a Virtual High School

Abigail Hawkins

Charles R. Graham
Brigham Young University

Richard R. Sudweeks
Brigham Young University

Michael K. Barbour
Sacred Heart University

Follow this and additional works at: http://digitalcommons.sacredheart.edu/ced_fac

Part of the Educational Assessment, Evaluation, and Research Commons, Elementary and Middle and Secondary Education Administration Commons, and the Online and Distance Education Commons

Recommended Citation

This Article is brought to you for free and open access by the Isabelle Farrington College Of Education at DigitalCommons@SHU. It has been accepted for inclusion in Education Faculty Publications by an authorized administrator of DigitalCommons@SHU. For more information, please contact ferribyp@sacredheart.edu.
Academic performance, course completion rates, and student perception of the quality and frequency of interaction in a virtual high school

Abigail Hawkins\textsuperscript{a}, Charles R. Graham\textsuperscript{b*}, Richard R. Sudweeks\textsuperscript{b} and Michael K. Barbour\textsuperscript{c}

\textsuperscript{a}Learning Design and Development, Investools® from TD Ameritrade Holding Corp., Draper, Utah, USA; \textsuperscript{b}Instructional Psychology and Technology, Brigham Young University, Provo, Utah, USA; \textsuperscript{c}Instructional Technology, Wayne State University, Detroit, Michigan, USA

(Received 17 March 2012; final version received 1 October 2012)

This study examined the relationship between students’ perceptions of teacher–student interaction and academic performance at an asynchronous, self-paced, statewide virtual high school. Academic performance was measured by grade awarded and course completion. There were 2269 students who responded to an 18-item survey designed to measure student perceptions on the quality and frequency of teacher–student interaction. Quality of interaction was subdivided into three constructs representing feedback, procedural, and social interaction. A confirmatory factor analysis helped to establish the fit of the statistical model for teacher–student interaction. Hierarchical logistical regression indicates that an increase in the quality and frequency of interaction resulted in an increased likelihood of course completion but had minimal influence on grade awarded. The estimated effect for quality and frequency composite items on completion was .83 and .56 respectively. Low practical significance of student–teacher interaction on grade awarded may be the result of mastery-based teaching approaches that skew grades for the completers toward the high end.

Keywords: virtual schooling; online K–12 learning; completion rates; interaction

Within the last 16 years virtual schooling has spread across 48 states and the District of Columbia (Watson, Murin, Vashaw, Gemin, & Rapp, 2010) in the United States (US). In 2000, there were an estimated 40,000 to 50,000 K–12 students enrolled in online programs (Clark, 2000) in the US. More recent estimates put the figure at 1,030,000 K–12 students (Picciano & Seaman, 2009), a 20-fold increase in less than a decade. The student population among virtual schools are becoming increasingly diverse. Historically, virtual school students were described as highly motivated, honors/advanced, independent learners who were more likely to attend four-year college than their face-to-face counterparts (Barbour, 2009). Today, a broader range of students is choosing virtual schooling for the purpose of credit recovery or to fulfill a graduation requirement (Watson, Gemin, & Ryan, 2008; Watson et al., 2010). The virtual school phenomenon is even extending down into
the elementary grades with more than 26 states offering at least some full- or supplemental-online learning opportunities for grades K–5 (Watson et al., 2010).

Attrition from online courses, particularly at the K–12 level, has been a significant challenge (Simpson, 2004; Smith, Clark, & Blomeyer, 2005; Zucker & Kozma, 2003). Though reasons for student persistence in an online environment are complex and multifaceted (Willging & Johnson, 2004), interaction may be a key factor to student success. In contrast to a sizable amount of interaction research in postsecondary education, there has been a paucity of empirical research into online learning at the K–12 level (Barbour & Reeves, 2009). In a review of literature, Rice (2006) found “very little research examining relationships between K–12 interaction that directly related to student performance, satisfaction, and retention in distance education contexts” (p. 439). Clearly, more research is needed to examine the relationship between student success and interaction in virtual schooling contexts to identify teacher best practices to improve student retention and academic performance.

The purpose of this study was to examine the relationship between students’ perceptions of teacher–student interaction and academic performance in a supplemental, asynchronous, self-paced, statewide virtual high school. We begin this article with an exploration of the factors related to academic success in K–12 and postsecondary online learning environments. We then describe the analysis of student perceptions of teacher–student interaction and academic performance using Pearson’s product-moment correlation coefficient and hierarchical linear modeling (HLM). Finally, we conclude by discussing two changes institutions can make to improve teacher–student interaction, along with two avenues of potential research.

Literature review

Historically, distance education has faced higher dropout and failure rates compared to traditional classrooms (Roblyer, Davis, Mills, Marshall, & Pape, 2008). Furthermore, it is believed that attrition rates are higher for virtual school settings compared to postsecondary online learning programs (Smith et al., 2005). Although no official attrition statistics exist for virtual schools by state or school type, individual evaluations of some K–12 online learning programs indicate that attrition ranges broadly from 10% up to 70% (Roblyer & Davis, 2008). For example, in its first year of operation, Illinois Virtual High School had only a 53% completion rate (Clark, Lewis, Oyer, & Schreiber, 2002), while the Alberta Distance Learning Center asynchronous courses had only a 47% completion rate (Elluminate Inc., 2006). Problems of attrition may be further masked by variations in how virtual schools calculate a successful completion and the length of trial periods when students are not considered to be “officially” enrolled (Hawkins & Barbour, 2010)

Factors influencing student performance in K–12 online learning

Several studies have examined the relationship between learner characteristics and student performance. Roblyer and Marshall (2003) identified learner characteristics predictive of high school student online course completion using the educational success prediction instrument (ESPRI), which predicted passing students with 100% confidence and failing with 95% confidence with 135 students across 13 virtual high schools. Additionally, they found no differences on personal characteristics (e.g., age, grade level, or gender). Successful students scored higher in self-efficacy,
individual initiative, organizational skills, and access to technology, and spent less time working outside of school. Roblyer’s (2008) replication study, using a larger sample \((n = 4100)\), also found successful students scored higher on technology access, self-efficacy, and organization. Additionally, past performance (i.e., grade point average) was a strong predictor of success. However, samples in both studies were selective (e.g., the majority were Caucasian; they were drawn from rural/suburban populations; they had a historically high pass rate; and they had time built into their schedule for their online coursework).

Teacher support in the online learning environment has also been tied to online course completion. However, the role of the teacher in the virtual school environment has expanded from what we typically understand the traditional classroom teacher’s role to be. Three critical roles have emerged: virtual school designer, teacher, and site facilitator (Davis, 2007). The role of virtual school site facilitator (i.e., the onsite local mentor and advocate), though not as thoroughly researched as the other two roles, has been tied to program and student success (Davis & Roblyer, 2005). In their evaluation of a statewide online program, Roblyer, Freeman, Stabler, and Schneidmiller (2007) found “facilitators that are directly working with students day-by-day are key to the success of the program” (p. 11). Though this issue is rarely documented, Mulcahy (2002) reported school-based teachers and principals often voluntarily provided technical and supervisory support, along with significant academic tutoring.

Factors influencing student performance in postsecondary online learning

In contrast to the limited research at the K–12 level, there is a more substantial body of literature focused on online learning in adult populations. Student success online is related to several factors. While demographic characteristics (i.e., gender (Bernard, Brauer, Abrami, & Surkes, 2004; Levy, 2007; Willging & Johnson, 2004), ethnicity (Dupin-Bryant, 2004), occupation (Willging & Johnson, 2004), and age (Levy, 2007; Willging & Johnson, 2004) were not predictive of completion rates, prior academic success (Bernard et al., 2004; Dupin-Bryant, 2004; Wang & Newlin, 2000; Wojciechowski & Palmer, 2005) and success in past online courses (Dupin-Bryant, 2004) are predictive. Additionally, affective traits (e.g., student’s locus of control (Willging & Johnson, 2004), motivation level (Jamison, 2003), and independent learning styles (Diaz, 2000) are predictive of success online. Factors external to the learner such as precourse orientation sessions (Wojciechowski & Palmer, 2005), study skills (Osborn, 2001), and strong computer skills (Dupin-Bryant, 2004) have also been associated with student success. Additionally, studies have identified interaction as a key factor for online student success (Thurmond & Wambach, 2004; Wallace, 2003).

Interaction and student performance in postsecondary online learning

Scholars have identified five types of interaction in online distance education: learner–instructor, learner–learner, learner–content, learner–interface, and vicarious interaction (Hillman, Willis, & Gunawardena, 1994; Moore, 1989; Sutton, 2001). Postsecondary interaction research has primarily focused on the quality and/or frequency of interaction in relation to three outcome variables: satisfaction, perceived learning, and academic achievement (Swan, 2001). High quality and levels of interaction have been associated with increased learner satisfaction (Jung, Choi, Lim, & Leem, 2002; Picciano, 2002; Russo & Benson, 2005; Shea, Fredericksen, Pickett,
Pelz, & Swan, 2001; Swan, 2001), perceived learning (Picciano, 2002; Rovai & Barnum, 2003; Stein, Wanstreet, Calvin, Overtoom, & Wheaton, 2005), and academic achievement (Jung et al., 2002; Picciano, 2002). Isolation and disconnection has the largest influence on students’ decision to drop out and disengage (Bocchi, Eastman, & Swift, 2004; Willging & Johnson, 2004).

Different taxonomies have emerged to examine teacher–student interaction. Heinemann (2005) identified three major types of teacher interactions: intellectual, organizational, and social. While there is substantial research into interaction in postsecondary online learning (as evidenced above), researchers caution against generalizing these findings to adolescents, who often lack the ability to regulate their own learning (Barbour & Reeves, 2009; Cavanaugh, Gillan, Kromrey, Hess, & Blomeyer, 2004; Rice, 2006). Typical of research in K–12 online learning, only a handful of studies have examined these three types of interaction.

**Interaction and student performance in K–12 online learning**

Research on teacher–student intellectual/instructional interactions has emphasized the importance of feedback on student satisfaction and persistence. Weiner’s (2003) qualitative study of a cyber charter school found limited teacher–student interaction a major concern. Students reported the lack of timely feedback was frustrating, impeded learning, and led to feeling “ignored, lonely, or lost” (p. 49). Furthermore, researchers who identified virtual schooling best practices also emphasized the importance of prompt feedback on student learning, progress, and connectedness (DiPietro, Ferdig, Black, & Preston, 2008; Ferdig, Cavanaugh, Dipietro, Black, & Dawson, 2009). However, these studies were based on student/teacher perceptions or standards analysis and did not tie the impact of interaction on actual student performance.

In relation to procedural interactions, Roblyer (2006), who identified best practices from three successful virtual schools, found policies and practices that required teachers to track student progress and proactively reach out to inactive students via e-mail, telephone calls, and monthly student and parental consultations. While these practices appear promising within these three specific virtual schools, they were also not based on systematic research showing improved academic performance.

Social interactions in the form of self-disclosure, humor, and encouragement are also important to virtual high school student motivation and progress (DiPietro et al., 2008; Mulcahy, Dibbin, & Norberg, 2008; Nippard, 2005; Roblyer, 2006; Weiner, 2003). Mulcahy et al. (2008) found struggling students missed social interactions and felt distant from their online teachers, preferring to seek help from their face-to-face teachers. Additionally, Nippard and Murphy (2007) found a disconnect between the mediums that teachers and students used to create social exchanges, likely making the positive effects difficult to achieve. However, both studies were conducted with largely rural samples, which typically display a stronger sense of community (Kannapel & DeYoung, 1999). While these studies provide insight into the nature of interactions in K–12 virtual schooling, few connected these forms of interactions to student persistence or academic performance.

**Methodology**

The purpose of this study was to examine the relationship between teacher–student interaction and academic performance. This led to the following research question:
What is the relationship between students’ perceptions of the quality and frequency of teacher–student interaction and online course completion and academic performance?

Quality measures of teacher–student interaction were further subdivided into the three categories of feedback, procedural, and social interactions based on Heine-mann (2005). We hypothesized that both quality and frequency factors would be positively correlated with course completion and academic performance. We also hypothesized that of the three different types of interaction instructional interactions would likely have the greatest positive correlation with completion and performance outcomes. We used correlation and HLM research methods from survey data to address this research question.

Context
Utah’s Electronic High School (EHS), a primarily supplemental, statewide, self-paced asynchronous virtual school, was the research setting. With 46,089 student enrolments from February 1, 2008, to January 31, 2009, the school operated on an open-entry/exit enrolment model where students proceeded at their own pace with little, if any, student–student interaction. Students had to submit a graded assignment every 30 days, remain active in the course, and complete the course within a 6-month time period. If they violated either of these policies, they were automatically dropped from the course but could immediately reenroll at any time at no cost to the student or penalty on their academic record.

EHS teachers developed the primarily text-based curriculum using Utah’s State Core Curriculum Standards (see http://www.uen.org/core/). At the time of the study, there were 66 unique classes taught by one full- and 76 part-time teachers. This is quite common for supplemental K–12 online learning programs, as few employ full-time personnel. According to the 2008 audit, approximately 64% of teachers were also employed as teachers somewhere other than EHS (Center for Educational Leadership and Technology, 2008). Of these teachers, 90% were also employed full time elsewhere. All respondents were certified to teach in their content area. Nearly 60% of teachers taught only one class at EHS, while 28% taught two classes, and 13% taught three or more classes. The average student-to-teacher ratio was 233:1, but ranged from two students to 1726 students per section. While this may seem like a high student-to-teacher ratio, it is not uncommon for online K–12 teachers to have student-to-teacher ratios two to four times that of a traditional classroom teacher (and open enrolment/exit programs—like EHS—are also even higher than the normal K–12 online learning range).

Overall, 34% of students enrolled from February 1, 2008, to January 31, 2009, completed their courses. Completion rates by individual disciplines ranged widely from 10.27% to 62%, with a large degree of variance. Table 1 indicates the mean completion rates by discipline for students enrolled from February 1, 2008, to January 31, 2009.

The high completion rates for the courses in financial literacy, health and physical education, and driver’s education may be due to the fact that they were required for graduation.
Participants

Participants were invited from a pool of 67,759 students who were enrolled from February 1, 2008, through September 29, 2009, and met the following criteria:

- submitted a request to sit for their proctored final exam,
- earned a grade,
- did not earn a grade, but had enrolled in the class for more than six months as of May 1, 2009, and
- did not earn a grade, yet enrolled in class for three months but less than six months or completed at least 50% of the class work.

A total of 2269 surveys with full or partial data were received. This represented a 3.34% response rate. While the response rate is quite low, we did not have control over who had access to the surveys. There is some evidence to indicate that as much as 40% of enrolments at EHS represent students who have not started work on a course (Hawkins & Graham, 2010). The students who have not started or have done very little work on a course are likely underrepresented in our survey because they have little motivation to respond to a survey about a course into which they have not invested much energy.

Over two-thirds of respondents were females; 84.4% were Caucasian; and 95% identified English as their native language. Overall, it was a fairly homogenous group of respondents with a significant number of Caucasian and female respondents. Other virtual schooling studies have also found an overrepresentation of dominant cultures within the sample (Black, Thompson, Askenazi, Ferdig, & Kisker, 2010). Because EHS does not collect demographic data on incoming students, it was not possible to compare the respondent sample to the larger population.

Based on the survey data, 75.1% of respondents \((n = 1705)\) reported having successfully completed the course (i.e., a grade was awarded), while 24.9% of respondents \((n = 564)\) indicated they did not complete the course. The sample was skewed toward successful students, as evidenced by the fact that overall completion rates for 2008–2009 were only 34%. However, this is typical of K–12 online learning in general, and supplemental programs in particular, which often experiences a bimodal distribution of students skewed toward the higher achieving students (Barbour, 2011). Interestingly, the grade distribution from A through D- for participants was

<table>
<thead>
<tr>
<th>Discipline</th>
<th>N</th>
<th>Percent completion</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer science education</td>
<td>448</td>
<td>10.27</td>
<td>30.39</td>
</tr>
<tr>
<td>Fine arts</td>
<td>1393</td>
<td>14.86</td>
<td>35.58</td>
</tr>
<tr>
<td>World languages</td>
<td>1756</td>
<td>18.05</td>
<td>38.47</td>
</tr>
<tr>
<td>Language arts</td>
<td>7149</td>
<td>18.73</td>
<td>39.02</td>
</tr>
<tr>
<td>Science</td>
<td>3031</td>
<td>19.80</td>
<td>39.85</td>
</tr>
<tr>
<td>Electives/Career/Technology</td>
<td>5337</td>
<td>26.72</td>
<td>44.25</td>
</tr>
<tr>
<td>Social studies</td>
<td>7082</td>
<td>26.77</td>
<td>44.28</td>
</tr>
<tr>
<td>Mathematics</td>
<td>2034</td>
<td>26.84</td>
<td>44.33</td>
</tr>
<tr>
<td>Driver’s education</td>
<td>2429</td>
<td>42.57</td>
<td>49.46</td>
</tr>
<tr>
<td>Health &amp; physical education</td>
<td>8776</td>
<td>47.76</td>
<td>49.95</td>
</tr>
<tr>
<td>Financial literacy</td>
<td>6654</td>
<td>62.41</td>
<td>48.44</td>
</tr>
<tr>
<td>Total/Average</td>
<td>46,089</td>
<td>34.18</td>
<td>47.43</td>
</tr>
</tbody>
</table>

Table 1. Mean percent completion rates by discipline \((n = 46,089)\).
reflective of the larger population (although this does not speak to students’ overall ability or grade point average—as it is common for students to score 10% or more or less in their online courses as compared to their face-to-face performance (Barbour & Hill, 2011). Possible reasons for the response bias include motivational factors (i.e., if students didn’t complete the course, one would not expect them to take time to complete the survey), the way the survey was administered (i.e., an e-mail sent out to all students as well as an integrated step in the course completion process just prior to signing up for a proctored exam), and respondents’ original motive for enrolling at EHS (i.e., original credit or accelerated graduation). In addition, the nature of volunteer subjects may also account for the higher completion rates, as volunteers generally show higher levels of achievement, motivation, and intellect than those who do not participate in nonmandatory assessments (Rosenthal & Rosnow, 1975).

**Instrument development**

Our choice to use a survey to collect data regarding student–teacher interaction was rooted in the fact that it was the only practical way to access the EHS population. Additionally, it should be noted that the nature of any K–12 online learning environment in the US involves multiple methods of interaction (Cavanaugh et al., 2009). While the learning management system might capture the interaction that occurs within that environment (Dickson, 2005), interaction also occurs through personal e-mail, telephone calls, text messaging, Skype and other voice over Internet Protocol (VOIP) tools, social networks, and in person (Barbour & Hill, 2011; Barbour & Plough, 2009, 2012). Unless researchers had teachers and students log each and every interaction, as it occurred, a survey presented the best method to collect data related to frequency of interaction.

We first sought to find an existing survey instrument that could measure quality and frequency of interaction in a virtual school setting. We reviewed 15 quantitative surveys where at least one scale measured teacher involvement, interaction, or support and found that none were a good fit for the study context (i.e., self-paced, asynchronous online secondary courses with little or no interaction with peers). Based on the literature search, a conceptual model of interaction introduced by Heinemann (2005) was then chosen to frame the study. Quality interaction items were designed to measure the three interaction types Heinemann identified:

1. **Intellectual/Instructional interactions**: exchanges related to academic feedback
2. **Organizational/Procedural interactions**: exchanges related to class logistics, procedures, and processes
3. **Social interactions**: exchanges related to support, encouragement, and connectedness.

Thirteen items were developed to measure students’ perceptions of the three quality constructs of student–teacher interactions and five items to measure students’ perceptions of the frequency of interactions. The instrument was piloted on 10 youth, with feedback used to improve the survey prior to its broader administration.

To estimate the construct validity of the instrument, we used confirmatory factor analysis (CFA). The purpose of CFA is to determine whether items purported to measure a particular construct shared a significant portion of common variance and
load to the intended construct (Hair, Black, Babin, Anderson, & Tatham, 2006). CFA can be used to assess the degree of convergent and discriminant validity. Convergent validity refers to the extent to which the items purported to assess a certain construct (i.e., frequency, social, feedback) share common variance and load to the same construct. In contrast, discriminant validity refers to the extent to which constructs are discrete factors with little overlap. Constructs that have weak correlations between each other indicate a high level of divergent validity.

CFA was performed using Mplus version 5.21. Preliminary CFA analyses were done on four proposed single-factor interaction scales: (a) frequency, (b) feedback, (c) procedural, and (d) social factor scales. Secondary analysis was conducted on the 13 quality items, where CFA was used to determine whether a first-order factor structure (i.e., interaction quality composite score) or a three-separate factor structure (i.e., instructional, procedural, social interaction composite scores) best fit the data. Reliability was estimated using Cronbach’s alpha coefficient.

Data collection
The study items were incorporated into an existing EHS class evaluation survey and administered in two ways. In February 2010, the principal e-mailed an invitation and survey link to 46,089 participants enrolled in courses from February 1, 2008, to January 31, 2009, with reminders sent to nonrespondents 4 weeks later, and the survey was closed following another 4 week period. The survey was e-mailed or made available to all enrolled students regardless of how much, if any, coursework they had completed. The second method of administration was its integration into the course completion process as a part of the class evaluation that students completed after finishing their coursework, but prior to taking the proctored exam. A total of 21,670 students had access to this route of completing the survey via its integration into the course.

The survey administration was part of the EHS course evaluation and the EHS administrators did not provide us with demographic data on all 67,759 students who had access to the survey, only with data on survey respondents. Student responses were mapped back to the EHS student performance database, linking responses to actual course grade, date of enrolment, date of completion (if applicable), course name, and instructor. Data were then de-identified.

Data analysis
Three procedures were used to examine the relationship between quality and frequency interaction variables with academic performance, as measured by grade awarded and course completion status: Pearson’s product-moment correlation coefficient, HLM, and hierarchical logistical regression. HLM was used as a secondary analysis to examine the dependent variable grade awarded. Since the second dependent variable, completion, was dichotomous, we conducted hierarchical logistical regression using Mplus version 5.21. Both regression statistics were conducted to estimate the degree of dependence among students with the same teacher and the consequences of violating the assumption of independence.
Results

We begin this section with a discussion of the CFA to determine whether items measured intended constructs and shared a significant portion of common variance. This is followed by a presentation of the descriptive statistics to provide the reader with an understanding of the participants and response counts. This is concluded by a discussion of the results of the Pearson’s product-moment correlation coefficient and HLM, which was designed to tell us the strength of the relationship between students’ perceptions of the quality and frequency of interaction and academic performance.

CFA

Model fit was determined using the chi-square test for difference testing (CHI), comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). To determine reasonably good fit, we were guided by CFI and TLI scores close to .95 or greater, and RMSEA scores close to .08 or below (Brown, 2006). Table 2 indicates the fit statistics on the preliminary analysis of the four proposed first-order factor scales.

For all elements of the model, preliminary analysis found the model fit indices were within the acceptable range for CFI and TLI and not RMSEA, indicating the items loaded to the intended factors. Secondary analysis using CFA on the quality items was used to determine whether quality items loaded to a single quality construct or loaded best to three separate constructs: feedback, procedural, and social. Table 3 depicts the fit statistics for quality items loading to a single-order factor structure, second-order factor structure, and three separate factor structure (i.e., feedback, procedural, and social).

Again, the fit statistics were within the acceptable range for CFI and TLI as outlined by Brown (2006). It should be noted that the fit statistics were the same for the second-order and three-separate order factor structures. Thus, to determine definitively which model for quality factor structures was the best fit, we used the CHI

Table 2. Fit statistics on preliminary analysis of first-order factor scales.

<table>
<thead>
<tr>
<th>Model</th>
<th>No. of items</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency first-order factor structure</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td>0.07</td>
</tr>
<tr>
<td>Feedback first-order factor structure</td>
<td>4</td>
<td>0.99</td>
<td>0.99</td>
<td>0.19</td>
</tr>
<tr>
<td>Procedural first-order factor structure</td>
<td>4</td>
<td>1.00</td>
<td>1.00</td>
<td>0.11</td>
</tr>
<tr>
<td>Social first-order factor structure</td>
<td>5</td>
<td>0.99</td>
<td>0.98</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Table 3. Fit statistics on secondary analysis of models for quality items (n = 13).

<table>
<thead>
<tr>
<th>Quality factor models</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-order factor structure</td>
<td>0.95</td>
<td>0.98</td>
<td>0.20</td>
</tr>
<tr>
<td>Second-order factor structure</td>
<td>0.96</td>
<td>0.99</td>
<td>0.18</td>
</tr>
<tr>
<td>Three-separate factor structure (i.e., feedback, procedural, social)</td>
<td>0.96</td>
<td>0.99</td>
<td>0.18</td>
</tr>
</tbody>
</table>
for difference testing comparing the first-order and three-separate factor structure models. This resulted in CHI values of 471.202, $df = 2$, and $p = <.00001$, indicating that the three-separate factor model fit the data significantly better than the first-order factor. Figure 1 depicts the path diagram for the proposed three-separate factor structure model along with the correlations between constructs and items.

The high factor loadings indicated that items loaded to the intended constructs and had high convergent validity. However, the high correlations between feedback, procedural, and social constructs indicated low discriminant validity. Practically speaking, participants responded similarly on all three constructs implying the constructs these items measured were distinguishable only theoretically and not statistically.

The high correlations between each construct and the fact that they pointed to a common theoretical construct (i.e., quality) led us to test a second-order factor structure with feedback, procedural, and social factors linked to an overarching theoretical construct we refer to as interaction quality (see Figure 2).

Once more, this model demonstrated high correlations between the first order factor (i.e., quality) and second order factors (i.e., feedback, procedural, social). These correlations indicated that the second order factors were part of a larger, theoretical construct we refer to as interaction quality.

---

**Figure 1.** Three-separate factor structure model for quality items. This figure illustrates the path diagram for the three-separate factor structures and construct correlations.

**Figure 2.** Second-order factor structure model for quality items. This figure illustrates the path diagram for second-order factor structure and construct correlations.
Reliability

Cronbach’s alpha for the interaction quality and quantity composites were .94 and .85 respectively. The reliability estimates for the three quality factors were .86 for feedback, .91 for procedural, and .92 for social. These were strong reliability levels as they fell well above the acceptable minimum value of alpha at .70 (Hair et al., 2006).

Descriptive statistics

The mean composite scores for interaction quality and frequency items were 3.2 (on a 4-point scale) and 2.43 (on a 5-point scale) respectively. Table 4 provides the mean scores for the factors analyzed in the survey.

Scores for quality of interaction were higher than frequency of interaction, which may have been a reflection of EHS’s independent study model.

There was a fairly normal distribution of scores for the quality composite factor. Quality items were on 4-point scale where 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree. The distribution of scores for the frequency composite of items collapsed into a 5-point scale where 1 = hardly ever, 2 = once a month, 3 = twice a month, 4 = once a week, and 5 = twice a week. Unlike that of the quality variable, the scores were negatively skewed. Frequencies of interaction varied based on the type of interaction. Table 5 indicates student-reported frequencies of interaction for feedback, procedural, and social interactions. While students reported low frequencies of interaction in all three areas, students indicated that they interacted more frequently over procedural issues than social or instructional matters.

Table 4. Factor composite means, standard deviations, and variances.

<table>
<thead>
<tr>
<th>Factors</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality composite</td>
<td>2202</td>
<td>3.20</td>
<td>0.59</td>
<td>0.35</td>
</tr>
<tr>
<td>Feedback</td>
<td>2123</td>
<td>3.38</td>
<td>0.66</td>
<td>0.43</td>
</tr>
<tr>
<td>Procedural</td>
<td>2117</td>
<td>3.34</td>
<td>0.61</td>
<td>0.37</td>
</tr>
<tr>
<td>Social</td>
<td>2122</td>
<td>2.96</td>
<td>0.66</td>
<td>0.43</td>
</tr>
<tr>
<td>Frequency composite</td>
<td>2182</td>
<td>2.43</td>
<td>0.91</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Note. The number of students (N) varies as not all of the respondents answered all of the questions.

Table 5. Reported frequency of student–teacher interaction by type.

<table>
<thead>
<tr>
<th>Interaction type</th>
<th>Hardly ever</th>
<th>Once a month</th>
<th>Twice a month</th>
<th>Once a week</th>
<th>Twice a week</th>
<th>Total respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback/Instructional</td>
<td>43.4%</td>
<td>17.3%</td>
<td>14.4%</td>
<td>18.1%</td>
<td>6.7%</td>
<td>2136</td>
</tr>
<tr>
<td>Procedural</td>
<td>37.3%</td>
<td>19.4%</td>
<td>15.5%</td>
<td>20.2%</td>
<td>7.7%</td>
<td>2147</td>
</tr>
<tr>
<td>Social</td>
<td>61.7%</td>
<td>11.6%</td>
<td>8.3%</td>
<td>12.8%</td>
<td>5.6%</td>
<td>2133</td>
</tr>
</tbody>
</table>

Note. The number of students (n) varies as not all of the respondents answered all of the questions.
Initial data analysis: Pearson’s product-moment correlation coefficient

Mean quality and frequency composite scores were correlated by academic performance as measured by grade awarded and course completion. Letter grades were recoded on a scale of 1 to 13 with a 1 as an A+, 2 as an A, 3 as an A-, etc. Table 6 indicates the correlation coefficients, N, statistical significance, and the measure of the strength of the association ($r^2$).

Correlations between perceived quality and frequency factors by grade awarded and course completion were very weak to negligible at both the composite and factor levels. Quality had higher correlations than frequency. However, quality factors (by grade awarded and by completion) only explained 2.1% and 7.2% of the variance respectively. Again, this was a very weak correlation.

Further analysis of correlations at the item level showed items correlated to completion had higher coefficient values than items correlated to grade awarded, though all items were .3 or lower. We thought that there might be differences in how lower-performing students (i.e., credit recovery) versus higher performing students (i.e., original credit) would perceive teacher–student interactions. However, correlations comparing these student types again found very weak to negligible relationships between academic performance and quality/frequency variables. In summary, correlations between the variables were almost nonexistent. These findings were unexpected and led us to either (1) question their hypothesis that quality and quantity of interactions would correlate with student performance in this context or (2) consider that perhaps respondents were not independent of one another, a primary assumption for the Pearson’s product-moment correlation coefficient.

Secondary data analysis: HLM and hierarchical logistical regression

The unexpected results from the correlational analysis led us to conduct HLM to identify the relationship between variables, believing that there may be dependencies
within groups of students who had the same teacher. The primary question in conducting HLM and hierarchical logistical regression was to determine if students with the same teacher were responding similarly in terms of the nature and frequency of interaction. Table 7 displays the mean, standard deviation, and intra-class correlation for the dependent variables: grade awarded and completion.

The mean for grade awarded was 10.04, which translated to letter grade was an A-. Completion mean is reported as a proportion with 74% of respondents successfully completing the course. Intra-class correlations were used to measure the degree of interdependence among observations (i.e., student respondents) with the same teacher. The high intra-class correlation values were evidence that there were dependencies in student responses grouped around the teacher (Raudenbush & Bryk, 2001). These dependencies would explain the weak to negligible correlation values from the initial analysis using Pearson’s product-moment correlation coefficient, which masked the actual relationship between interaction type and academic performance.

To examine interaction and grade awarded, we used HLM. Table 8 displays the estimates of fixed effects. In this analysis, grade awarded is the dependent variable and quality and frequency of interaction are independent variables.

In all instances, the perceived quality and frequency of interaction associated with the grade awarded were significant at \( p = .05 \) level. The estimate effect column in Table 8 indicates that for every one unit increase in the perception of the 4-point quality of interaction scale, there was a .27 unit increase in the 12-point grade (e.g., B- to B is one unit difference. Note that EHS does not award failing grades, thus, a 12-point grade scale). In other words, a four unit difference in perception of interaction quality (i.e., the full scale) only accounted for a one unit grade difference. The practical significance of this finding is minimal.

To examine interaction and completion, we used hierarchical logistical regression, as the dependent variable was dichotomous in nature. Table 9 displays the estimates of fixed effects. In this analysis, completion is the dependent variable and quality and frequency of interaction are independent variables.

The perceived quality and frequency of interaction associated with the completion were significant at \( p = .05 \) level. For completers and noncompleters, there were significant differences in scores on the quality composite construct. The scores on

### Table 7. Dependent variable mean, standard deviation, and intra-class correlations.

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Mean</th>
<th>SD</th>
<th>Intraclass correlation*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade awarded</td>
<td>10.04</td>
<td>1.62</td>
<td>0.21</td>
</tr>
<tr>
<td>Completion</td>
<td>0.74</td>
<td>0.44</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Note. *Computed from the random intercept null model.

### Table 8. Grade awarded estimates of fixed effects.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality composite</td>
<td>.27</td>
<td>.07</td>
<td>3.84</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>Feedback</td>
<td>.23</td>
<td>.07</td>
<td>3.43</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>Procedural</td>
<td>.36</td>
<td>.07</td>
<td>5.08</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>Social</td>
<td>.14</td>
<td>.06</td>
<td>2.21</td>
<td>.027*</td>
</tr>
<tr>
<td>Frequency composite</td>
<td>−.09</td>
<td>.04</td>
<td>−2.27</td>
<td>.023*</td>
</tr>
</tbody>
</table>

Note. *Correlations significant at the .05 level.
quality composite by completion had an estimate effect at 0.83. Thus, for every one unit increase on the 4-point interaction scale for quality items the log odds of completing the course increased by .83. Additionally, for every one unit increase on the 5-point frequency scale, the log odds of completing the course increased .56. Thus, a one unit difference in perception of quality of interaction is nearly the difference between a non-completer and completer. Additionally a two unit difference on the frequency scale differentiates noncompleters from completers.

Discussion

There were three main findings resulting from this study: (1) methodological insights, (2) key differences between completers and noncompleters on quality and frequency of interaction, and (3) little practical significance on differences between high and low performing students' perceptions on quality and frequency of interaction.

First, there were methodological understandings that arose from this study that may be just as important as the findings related to virtual schooling. As with many other studies measuring the relationship between interaction and academic performance or perceived satisfaction (e.g., Fredericksen, Pickett, Shea, Pelz, & Swan, 2000; Restauri, 2006; Swan, 2001), we used a correlation statistic to measure this relationship and found a weak relationship. Since respondents who shared a common teacher may have answered similarly, the assumption of independence was violated, which is a major premise of correlation coefficients. Consequently, we conducted HLM and hierarchical logistical regression, resulting in more meaningful insights on student perceptions of interaction in relation to academic performance. Future research that draws participants from across multiple courses and instructors should account for the violation of response independence and nest students by teacher using HLM. Another methodological insight future researchers should heed that resulted from the CFA was significantly high correlations between feedback, procedural, and social factors, which meant that students generally responded similarly to the procedure, feedback, and social interaction questions. Students split by grade responded differently on the three constructs. However, since there were relatively few C and D students in the population one could still get the interaction effect at the end without significantly changing the correlations when examining the entire population.

Second, there were unique findings between completer and noncompleter responses. First, the results of this study indicate that the perceived quality and quantity of interaction mattered to student completion rates. Higher quality interaction and more frequent interaction scores increased the log odds of completion significantly. In other words, students who completed the course perceived greater interaction and quality of interaction than noncompleters. This finding supports the higher education research on the importance of interaction to remain engaged
(i.e., Bocchi et al., 2004; Willging & Johnson, 2004) and gives statistical backing
to the handful of qualitative studies pointing to the importance of interaction in
virtual schooling environments (i.e., DiPietro et al., 2008; Ferdig et al., 2009;
Mulcahy et al., 2008; Nippard, 2005; Roblyer, 2006; Weiner, 2003).

Finally, in contrast to interaction impacting completion, there was no practi-
cally significant effect of interaction on the grade awarded. There were several
possible reasons for this outcome. First, the results may be due to the fact that
there was little variation in grades awarded among respondents. The mean grade
was an A-, with 76% of respondents receiving an A compared to only 0.6%
receiving a D and 4.1% receiving a C. Thus, it would be difficult to detect differ-
ences in perceptions on such a small number of respondents. Another possible
reason for the absence of an effect was that students who completed the course
were satisfied solely with completion irrespective of the grade awarded, because
in many cases completion with any nonfailing grade was all that was required for
graduation. For the bulk of students, completion may have been more of an
important issue than the grade awarded. Teacher interaction impacted completion
but not necessarily grade awarded at a significant level. Finally, it was common
practice for teachers at EHS to allow students to resubmit work for an improved
grade. One could speculate that this increased interaction moved some students
from noncompletion to passing the course and they perceived high interaction as
a result of multiple resubmissions.

There were several methodological limitations to this study. First, the survey
was administered via e-mail in March 2010 to 46,089 participants who were
enrolled from February 1, 2008, to January 31, 2009 (i.e., the remaining 21,670
students were given the opportunity to complete the survey as a part of the end-
of-course procedures). Researchers indicate that recall bias increases when there is
a delay between the time of the experience and recall of said experience (Stynes
& White, 2006). The gap in time from when students experienced the course and
when they recalled their interaction experiences was problematic. Second, the low
response rate (i.e., 3.34%) indicates a likelihood of nonresponse bias or the likeli-
hood that respondents and nonrespondents differ significantly. There is evidence
to suggest that as much as 40% of enrolled EHS students have not started work
on a course. It is likely that students who enrolled but never started work on a
course or who have done very little work on a course are significantly underrep-
resented in the results.

According to Rosenthal and Rosnow’s (1975) review of literature on studies
relying on volunteers as subjects, the authors found that volunteers were “likely
to show higher levels of achievement than their less achievement-motivated col-
leagues” (p. 40). While they found 15 studies showing no relationship between
volunteering and intelligence, they found 20 studies that indicated volunteers were
significantly more intelligent than nonrespondents. Essentially, those who volun-
teeed to complete a nonmandatory survey may not be representative of the larger
population. There is evidence of nonresponse bias in the fact that 75% of those
who responded to the survey for this study had completed their online course
compared to only 31% of the general population at EHS during the time of the
study. Due to the fact that participants were minors and dispersed geographically,
and the difficulty of obtaining Institutional Review Board (IRB) approval through
the State and Brigham Young University, beyond a follow-up survey, we did not
build in other means to reach out to the nonrespondents.
Conclusions and implications

The quality and frequency of interaction had a significant impact on student completion but not on grade awarded. Increased levels of the quality and frequency of interaction resulted in increased student completion. However, there was no difference on grade awarded, a result likely due to the limited variation in grades awarded, attitude toward completion, and resubmission practices at EHS.

Based on the results of this study, along with the limitations of EHS’s instructional model, there are two implications for practitioners in this and similar environments. First, interaction matters in terms of both the quality and frequency of interaction. Students who completed the course perceived their interactions with teachers more positively than the noncompleting students. Teachers should continue to maintain a high quality and frequency of interaction with students, particularly those at risk of dropping out. One way to achieve this is to place a greater emphasis on interaction at the beginning of a course time when all students are most likely to be engaged in the course. Teachers should interact with students on multiple occasions in the first day or so of enrolment because this may be when they are most primed to engage in the course. These interactions would include (a) an introduction to the learning management system navigation, along with how to participate in the course and interact with the content (i.e., procedural); (b) general information about the course content and learning goals (i.e., instructional); and (c) an introduction to the teacher as a person and warm welcome to the class (i.e., social). Second, teachers should take proactive measures to reach out to students regardless of their progress in the course. The increased interaction may be enough to move students from the noncompletion status to completion status.

In addition to these implications for practice, there are also two areas of future inquiry. First, perceptions of behaviors are often very different from actual behaviors (Fishman, Marx, Best, & Tal, 2003). Mining data from learning management systems, which contain all teacher–student interactions (i.e., teacher–student e-mails, student question postings, assignment submission and feedback), would be a simple way to determine the relationship between frequency of interaction and academic performance (Schneider, Krajcik, & Blumenfeld, 2005). Furthermore, content analysis of such interactions with a clear coding system could also reveal the quality of the interaction in a less subjective manner. Second, the fact that the overall student perceptions of the quality and frequency of interaction were minimal may be due to the nature of the specific context where the study occurred (i.e., a self-paced, asynchronous, open-entry virtual school with large class sizes, where failing grades are not awarded). Due to the wide variation in virtual school types, similar studies with other virtual schools may yield different, and potentially more positive, results for both grade awarded and completion.

Notes on contributors

Abigail Hawkins is a Senior Manager of Learning Design at TD Ameritrade’s Investools®. Her research interests include blended learning, K–12 online learning, and learning analytics. In 2011, she completed her Ph.D. in Instructional Psychology and Technology from Brigham Young University (dissertation title: “We’re Definitely on our Own”: Interaction and Disconnection in a Virtual High School).

Charles R. Graham studies technology-mediated teaching and learning, focusing on the design and evaluation of blended and online learning environments. He also researches the
use of technology to enhance traditional teaching and learning. Charles co-edited *The Handbook of Blended Learning*. His current research publications can be found at https://sites.google.com/site/charlesrgraham/

Richard R. Sudweeks is a professor in the Instructional Psychology and Technology Department in the School of Education at Brigham Young University. He teaches courses in psychometrics and research design. His research interests focus on longitudinal analysis of student achievement and the challenges of estimating teachers’ effects on their students’ learning.

Michael K. Barbour is an assistant professor of instructional technology at Wayne State University. For more than a decade his research has focused on the effective design, delivery, and support of K–12 online learning, particularly for students located in rural jurisdictions, in Canada, the United States, and New Zealand.

References


