Journal of Information Systems Education

Volume 33
Issue 3
Summer 2022

Change in Students' Peer Evaluations of Requirements Elicitation Interviews Across the Pre-Pandemic and Pandemic-Affected Semesters

Dmytro Babik and Iryna Babik

Recommended Citation: Babik, D., & Babik, I. (2022). Change in Students' Peer Evaluations of Requirements Elicitation Interviews Across the Pre-Pandemic and Pandemic-Affected Semesters. *Journal of Information Systems Education*, 33(3), 283-299.

Article Link: https://jise.org/Volume33/n3/JISE2022v33n3pp283-299.html

Initial Submission:	30 June 2021
Minor Revision:	6 October 2021
Accepted:	18 November 2021
Published:	15 September 2022

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Journal of Information Systems Education, 33(3), 283-299, Summer 2022

Change in Students' Peer Evaluations of Requirements Elicitation Interviews Across the Pre-Pandemic and Pandemic-Affected Semesters

Dmytro Babik
Department of Computer Information Systems and Business Analytics
James Madison University
Harrisonburg, VA 22807, USA
babikdx@jmu.edu

Iryna Babik
Department of Psychological Science
Boise State University
Boise, ID 83725, USA
irynababik@boisestate.edu

ABSTRACT

Successful development of an information system to solve a business problem depends on the analyst's ability to elicit system requirements from a user. This complex competency could be trained via critical peer evaluation of the requirements elicitation (RE) interviews. In this study, 294 students across four pre-pandemic and two COVID-19 pandemic-affected semesters evaluated recorded sample RE interviews of low and high quality. A piecewise regression modeling was used to examine the change in students' evaluations separately for the pre-pandemic and pandemic-affected semesters. Current results showed that students exhibited inflated evaluation scores (relative to instructors' scores) for the high-quality, but not for the low-quality interview. While students' evaluations for the low-quality interview remained stable across the pre-pandemic semesters, a significant decrease in evaluation scores for the high-quality interview reduced the gap between the students' and instructors' evaluations. The onset of the COVID-19 pandemic brought a significant increase in students' evaluation scores, which decreased during the second pandemic-affected semester. Moreover, females inflated their evaluations compared to males, specifically for technical, rather than soft skills. Current findings shed light on several important trends in students' peer evaluations in the context of RE training and possible effects of massive learning disruptions, such as the pandemic.

Keywords: Requirements analysis & specification, Systems analysis & design, Technical skills, Soft skills, Peer evaluation, Pandemic

1. INTRODUCTION

Successfully deploying an information system to solve a business problem depends on the development team's competency of determining system requirements. Therefore, competent requirements elicitation (RE) is an important learning objective in a contemporary information systems (IS) curriculum and a vital marketable skill for IS professionals (Ezell et al., 2019). The literature, however, indicated a persisting lack of RE competence by IS program graduates and young professionals (Browne & Ramesh, 2002; Costain & McKenna, 2011; Kamthan & Shahmir, 2019; Schenk et al., 1998; Turner, 1990; Watson & Frolick, 1993; Zowghi & Coulin, 2005). Oftentimes, students attain cursory, basic understanding of RE and do well on a multiple-choice test but lack sufficient practice to effectively apply RE techniques in an organizational setting. To bridge the gap between cursory RE knowledge and demonstrable RE skills, in 2015-2020, the Department of Computer Information Systems and Business Analytics (CISBSAN) at James Madison University (JMU) implemented a multiyear, faculty-led project which integrated RE aptitude training and learning assessments into their Information System curriculum (Ezell et al., 2016). This project was grounded in rigorous methodology for curriculum improvement (Fulcher et al., 2014), particularly, through designing and applying an analytical Requirements Elicitation Interviews Assessment Rubric (REIAR) (Ezell et al., 2019; Ezell et al., 2016). The project resulted in considerable, recorded improvement of students' RE competencies, as well as the program's ability to methodically assess this learning outcome (Lending et al., 2018; Satkus, n.d.).

One of the courses in which students hone their RE interview skills is CIS 454 "Systems Analysis and Design." This course includes several team projects emulating phases of the Systems Development Life Cycle; in one of the projects, students interview stakeholders to determine requirements for a system. To help students understand the key components of the RE interview skills and performance expectations, the RE training includes two steps. First, students use the REIAR to evaluate two video-recorded, sample interviews conducted by other students. This individual exercise is followed by the team project, in which students conduct a mock RE interview. The premise of this two-step training process is for students to internalize the RE components and to calibrate their performance expectations, and then strengthen these skills through learning by doing (Costain & McKenna, 2011) and learning from mistakes. Critical assessment of own and peers' performance is an established precursor to one's own professional growth (Adachi et al., 2018; Burgess et al., 2013; Cao et al., 2019; Lundstrom & Baker, 2009). Studies in peer assessment demonstrated that not only receiving but also providing peer assessment is a learning opportunity, benefits of which include learning from seeing models of effective and ineffective performance, developing metacognition through practicing revision strategies, and improving the ability to detect, diagnose, and address problems (Li et al., 2010; Lundstrom & Baker, 2009; Patchan & Schunn, 2015).

The purpose of the current study was to explore the emerging patterns in students' evaluations of sample RE interviews in a system development project and to address the following research questions: 1) Is there a temporal trend in students' evaluations of the sample RE interviews? 2) Are there any changes in students' evaluations during the pandemic-affected semesters compared to pre-pandemic ones? Specifically, we intended to determine whether: 1) the average REIAR evaluation scores changed over time for either the low-quality or high-quality interview; 2) there was a temporal change in students' evaluations of the soft versus technical skills for each type of interview; 3) there was a temporal change in students' evaluations of the individual REIAR criteria for each interview; 4) there was a structural break in any of the evaluation trajectories from pre-pandemic to pandemic-affected semesters; 5) there were persistent differences between students' and instructors' evaluations of the same interviews; 6) there were gender differences in students' evaluations. The unit of analysis in this study was an individual student.

This paper is organized as follows. Section 2 frames our investigation in the broader context of previous research. Section 3 describes the methods of our empirical investigation. Section 4 reports the findings. Section 5 discusses the findings in the context of previous research. Section 6 outlines the conclusions and future directions.

2. REVIEW OF RELEVANT RESEARCH

2.1 The Importance of RE Interviews and Efficacy Criteria

The RE outcomes are important for successful systems development (Browne & Ramesh, 2002; Byrd et al., 1992; Havelka, 2003; Zowghi & Coulin, 2005). The RE outcomes, in turn, are determined by the quality of the RE process, which encompasses discovery and refinement of user needs through recurring and varied interactions between users and analysts (Jain et al., 2003; Marakas & Elam, 1998; Zowghi & Coulin, 2005). Despite the accepted importance of RE, the analyst teams oftentimes lack skills and training to perform an effective RE (Browne & Ramesh, 2002; Turner, 1990; Watson & Frolick, 1993). The failure of newly developed systems in up to 90% of projects could be attributed to poorly executed RE processes (Davis et al., 2006; Dennis et al., 2015; Lindquist, 2005).

Undoubtedly, any RE technique, such as document analysis, survey, or interview, has its limitations; therefore, understanding advantages and disadvantages of different techniques, and skillfully combining a variety of RE sources are critical for successful RE (Burnay, 2016; Burnay et al., 2014). One of the most effective RE techniques, widely used in practice and, unfortunately, often found to be weak in recent graduates, is the user-analyst interview (Agarwal & Tanniru, 1990; Alvarez, 2002; Browne & Rogich, 2001; Davey & Cope, 2008; Holtzblatt & Beyer, 1995; Moody et al., 1998). Aside from other limitations discussed in the literature, interviews between the user and analyst could be plagued by various cognitive and communication biases, which hinder the RE outcomes (Browne & Ramesh, 2002; Byrd et al., 1992; Gallivan & Keil, 2003; He & King, 2008; Jain et al., 2003; Pitts & Browne, 2004; Valusek & Fryback, 1985; Zhang, 2007; Zowghi & Coulin, 2005).

The impact of cognitive and communication biases can be reduced by skillfully executing the following practices: 1) opening the interview by presenting its purpose and agenda (Browne & Ramesh, 2002; Gallivan & Keil, 2003); 2) asking specific questions about the as-is and to-be systems (Browne & Ramesh, 2002); 3) visualizing various aspects of the system via modeling and prototyping techniques (Browne & Ramesh, 2002; Vijayan & Raju, 2011; Zowghi & Coulin, 2005); 4) actively listening to the user and appropriately rerouting the conversation (Pitts & Browne, 2007); 5) fostering inter-team and user-analyst relationships (Hickey & Davis, 2003); and 6) closing the interview with a proper summarization and outlining future steps (Pitts & Browne, 2004). These interview strategies were incorporated into the REIAR (Ezell et al., 2019).

2.2 Professional Factors That Affect Interview Evaluations

Although previous research, in general, showed moderate (around $r = .69$) positive correlation between student peer-evaluations and instructor evaluations (De Grez et al., 2012; Falchikov & Goldfinch, 2000; Sridharan et al., 2019), students oftentimes inflate their evaluation scores by as much as 5% (Langan et al., 2005; McCarty & Shrum, 2000; Pond et al., 1995). In addition, the spread of scores assigned by instructors tends to be twice as large as the spread of scores assigned by students; instructors are also more likely to assign scores at the extremes of the range compared to students (Freeman, 1995; Hughes & Large, 1993; Langan et al., 2008).

There are various possible explanations behind inflated student self- and peer-evaluations: lack of assessment experience, limited domain knowledge, students' "generosity" toward peers and reluctance to assign low scores (Ballantyne et al., 2002; Kruger & Dunning, 1999; Langan et al., 2008). Previous research showed that more years of academic experience (e.g., seniors versus freshmen) and more practice with the peer-evaluation process help alleviate the inflation issue (Langan et al., 2008; Sutherland & Ellery, 2004).

2.3 Psychological Factors that Affect Interview Evaluations

Previous research suggested that evaluative judgements may be affected by a variety of psychological factors, such as stress, anxiety, depression, mood, and empathy. Therefore, while establishing the background for this study, we considered the potential influence of these factors on students' evaluations of other students' performance.

Stress is an everyday component of our life. It stems from a mismatch between the person's resources and their perceptions of environmental demands (Eaton & Bradley, 2008). Transition to college often results in such mismatch, making student life quite stressful. The prevalence of stress in college students reportedly reached an alarming 27-30% (Bayram & Bilgel, 2008; Sax, 2003; Yusoff et al., 2010). Previous research reported that female students are more vulnerable to stress than their male counterparts (Bayram & Bilgel, 2008; Brougham et al., 2009; Misra & McKean, 2000; Pierceall & Keim, 2007).

Stress is often accompanied by anxiety and depression (Beiter et al., 2015), threatening to transform college-related worries into debilitating short- or long-term mental health conditions. Importantly, the COVID-19 pandemic caused a significant increase in the prevalence of stress, anxiety, and depression among the general population (Cooke et al., 2020; Gallagher et al., 2020; Salari et al., 2020), as well as college students (Son et al., 2020; Wang et al., 2020). As many as 71.3% of students reported an increase in their stress levels due to COVID-19, 38.5% displayed significant anxiety symptoms, while 48.1% succumbed to depression during the pandemic (Son et al., 2020; Wang et al., 2020). A variety of factors lead to increased stress levels during the COVID-19 pandemic; among them are risk of exposure and infection, social isolation, uncertainty and lack of control over the situation, financial instability, insufficient supplies, difficulty with concentration, disturbed sleep, worries about inadequate academic performance, concerns about using distance/remote learning tools, boredom, frustration, anger, and stigma (Brooks et al., 2020; Son et al., 2020).

Importantly, the valence of mood (positive versus negative) has been shown to affect people's evaluations. Previous research found that evaluative judgements tend to be congruent with the mood, be this due to elaborate cognitive processing of available information ("mood-congruent retrieval" model; Blaney, 1986; Bodenhausen & Wyer, 1987; Bower, 1991; Kahneman, 2002; Morris, 1989; Sherman & Corty, 1984; Wyer & Srull, 1986) or the lack of motivation for deep analysis and the use of "feeling heuristic" ("feeling heuristic" model; Clore et al., 1994; Forgas, 1994, 1995; Schwarz, 1990; Schwarz & Clore, 1983; Siemer & Reisenzein, 1998). According to these models, during the COVID-19 pandemic, one might expect that students would evaluate peers' interviews in a less favorable way.

More recent research, however, suggested that evaluation judgements depend not only on the valence of mood, but also on the specific type of emotion. For example, fear and sadness are typically associated with blaming situational factors and making pessimistic judgements; in contrast, anger is often related to blaming other individuals, while producing optimistic judgements of a situation and punitive judgements of other individuals (Goldberg et al., 1999; Keltner et al., 1993; Lerner & Keltner, 2000). As we mentioned above, negative emotions associated with the COVID-19 pandemic may vary from fear to anger, thus students' evaluations of others may, correspondently, shift in a positive or negative direction.

Students evaluating interviews conducted by other students, while knowing that they will be responsible for performing a similar activity in a week, might also feel empathy. We define empathy here as an ability to recognize and share another person's emotional state or situational context (Eisenberg & Strayer, 1987). Previous studies suggested that empathy is inversely related to aggressive attitudes (Cohen & Strayer, 1996). Therefore, in difficult times, students experiencing empathy towards other students may evaluate them more positively. Also, females typically show greater emotional empathy than males (Cohen & Strayer, 1996; Nwankwo, 2013), which may also result in more favorable evaluations. Indeed, previous research found that males, compared to females, tend to have higher expectations for others' performance and judge them more critically (Abad-Tortosa et al., 2017; Alagna, 1982).

2.4 Hypotheses

Previous research mostly did not provide support to the directional hypotheses, which justifies the exploratory nature of the current study. Based on the review of relevant literature, we hypothesized that:

H1: Students' evaluations were stable during the pre-pandemic semesters (Ezell et al., 2019; Ezell et al., 2016; Lending et al., 2018).

H2: There was a significant change in students' evaluations from the pre-pandemic to the pandemic-affected semesters (Brooks et al., 2020; Son et al., 2020; Wang et al., 2020).

H3: There were differences in students' evaluations of the soft versus technical skills (Ezell et al., 2019; Ezell et al., 2016; Lending et al., 2018).

H4: Students inflated their evaluation scores compared to instructors (Langan et al., 2005; McCarty & Shrum, 2000; Pond et al., 1995).

H5: There were gender differences in students' evaluations of the sample RE interviews (Abad-Tortosa et al., 2017; Alagna, 1982; Cohen & Strayer, 1996; Nwankwo, 2013).

3. METHODS

3.1 Participants

Our empirical investigation was a cross-sectional study implemented over the six consecutive semesters from fall 2018 to spring 2021. Participants were 294 students (231 males, median age 20 years) majoring or minoring in Computer Information Systems and taking the required upper-level course CIS 454 "Systems Analysis and Design" in the Department of CISBSAN at JMU. All participants enrolled during different semesters were taught by the same instructor. All data were obtained from required graded assignments of the course; students received no compensation for participating in the study.

3.2 Procedures

Students received training in the information system development under the Waterfall Model by completing three team case-based projects. Project 1 emulated the planning

phase of the Waterfall, with the written Project Plan and a presentation to the stakeholder as deliverables. Projects 2 and 3 emulated the analysis phase of the Waterfall. In Project 2, student teams, acting as system development teams, elicited system requirements from the project stakeholder. This project focused on the development of the RE interview skills. Each team prepared for the interview using information from the project case; conducted a 20-25-minute-long mock interview with the project stakeholder (role-played by the instructor) and submitted a short report summarizing collected requirements. This report included a memo, as well as functional and process models for the as-is system (use-case and high-level activity diagrams). At the conclusion of Project 2, students were provided detailed, REIAR-based feedback on their interview performance and models. In Project 3, student teams analyzed the collected requirements, compiled the System Proposal, and presented it to the stakeholder. Each project took about 2-4 weeks to complete.

In preparation for Project 2, all students were required to individually complete evaluations of two video-recorded mock RE interviews conducted by other student teams in earlier semesters. The same two sample interviews were used in the study; one interview represented overall strong performance of a team eliciting requirements; the other interview showed overall poor performance of a team. Students were blind to the quality of these interviews before completing this assignment. This interview evaluation was a take-home assignment, and students could watch and evaluate the two sample interviews in any order they liked. The same REIAR was used by students evaluating these sample interviews as by the instructor assessing team interviews in Project 2.

During the four semesters from fall 2018 to spring 2020, students had in-person instruction, whereas in the fall 2020 and spring 2021 semesters, due to the COVID-19 pandemic, all the instruction was done online in a synchronous mode.

3.3 Measures

The participants evaluated each sample interview using the REIAR (Appendix C), which consists of the following eight criteria (outcome variables of this study): 1) *Opening* – the quality of the opening phase of the interview; 2) *Closing* – the quality of the closing phase of the interview; 3) *Listening* – active listening during the interview; 4) *Relation* – relationship building with the interviewee; 5) *Teamwork* – interpersonal interactions within the interviewing team; 6) *Analysis* – analysis of the as-is (current) system; 7) *Design* – design of the to-be (proposed) system; and 8) *Visual* – the use of visual aids and models (Ezell et al., 2019).

For each interview, each criterion was evaluated on a scale from 1 to 5; 1 marking the worst (Beginner) and 5 marking the best (Outstanding) outcome (Ezell et al., 2019). Note that students were informed that, when performing the project interview, they needed to reach at least level 3 (Competent) to receive credit; this threshold may have influenced their own evaluations of the sample interviews. Importantly, the two sample RE interviews were also evaluated by four faculty from the Department of CISBSAN at JMU to examine potential differences between student and faculty evaluations.

The mean score from all eight criteria was calculated separately for the low-quality interview (*All_Low*) and the high-quality interview (*All_High*). Furthermore, the eight criteria were classified into *soft skills* (*Opening*, *Closing*, *Listening*, *Relation*, and *Teamwork*) and *technical skills* (*Analysis*, *Design*, and *Visual*). Note: specific technical skills relevant to this project included identifying and formulating a business problem, identifying functional system requirements, and visualizing those requirements using UML business-process and functional models. The variables *Soft_Low*, *Soft_High, Tech_Low*, *and Tech_High* were calculated by averaging scores across the corresponding criteria, computed separately for the low- and high-quality interviews.

The time point of the interview evaluations was coded into the *Time* variable: 0 = fall 2018; 1 = spring 2019; 2 = fall 2019; 3 = spring 2020; 4 = fall 2020; and 5 = spring 2021. In order to evaluate possible gender differences in student evaluations, we included a dummy-coded *Gender* variable (0 = males; 1 = females) to all statistical models.

3.4 Statistical Analyses

PASW Statistics software (version 18.0.3) was used for all statistical analyses. Results were considered statistically significant at $\alpha \leq .05$.

3.4.1 Change Over Time in Students' Evaluations. Suspecting that the COVID-19 pandemic was a dramatic event that could potentially affect students' evaluations of the interviews, we implemented a piecewise modelling to accurately represent change in evaluations over time (testing H1-H3). The piecewise statistical model estimated the two regression lines (the first one for pre-COVID semester – time points 0, 1, 2, 3; the second one for the COVID-affected semesters – time points 4 and 5), allowing individual intercepts and slopes for each segment of the trajectory. The final piecewise model was represented by the following equation:

$$Y_i = \beta_{01}\, Int1_i + \beta_{02}\, Int2_i + \beta_1\, Time1_i + \beta_2\, Time2_i + \beta_3\, Gender_i + \epsilon_i, \text{ where:}$$

Y_i – the student's interview evaluations (two models for the *All_Low* and *All_High* dependent variables; four models for the *Soft_Low*, *Soft_High*, *Tech_Low*, *Tech_High* dependent variables; and 16 models for the *Opening*, *Closing*, *Listening*, *Relation*, *Teamwork*, *Analysis*, *Design*, and *Visual* dependent variables for both the low- and high-quality interviews);
β_{01} – the intercept for the pre-pandemic segment of the trajectory;
$Int1_i$ – a variable coded as 1 for $Time_i \leq 3$, and 0 for $Time_i > 3$;
β_{02} – the intercept for the pandemic-affected segment of the trajectory;
$Int2_i$ – a variable coded as 0 for $Time_i \leq 3$, and 1 for $Time_i > 3$;
β_1 – the slope of change over time for the pre-pandemic segment of the trajectory;
$Time1_i$ – a variable coded as ($Time_i - 4$) for $Time_i \leq 3$, and 0 for $Time_i > 3$;
β_2 – the slope of change over time for the pre-pandemic segment of the trajectory;
$Time2_i$ – a variable coded as 0 for $Time_i \leq 3$, and ($Time_i - 3$) for $Time_i > 3$;
β_3 – difference in the intercept between the trajectories for males and females;
ϵ_i – independent error term that follows a normal distribution.

3.4.2 Comparison of Students' and Instructors' RE Interview Evaluations. The low number of instructors providing their evaluations of the sample interviews (n = 4) precluded any formal statistical analysis of these data. Therefore, visual inspection of the graphs representing evaluation scores (mean across the six time points) for students and instructors was performed (testing H4). Both summarized scores (All_Low and All_High variables) and individual criteria (Opening, Closing, Listening, Relation, Teamwork, Analysis, Design, and Visual variables) were evaluated for the low- and high-quality interviews.

3.4.3 Change in Gender Composition Over Time. Since previous research noted significant gender differences in evaluative judgements, we wanted to ensure that possible changes in student evaluations across the six semesters were not due to shifts in gender composition. We conducted Pearson chi-square analysis to test whether there was a significant difference in gender composition across the six semesters (testing H5).

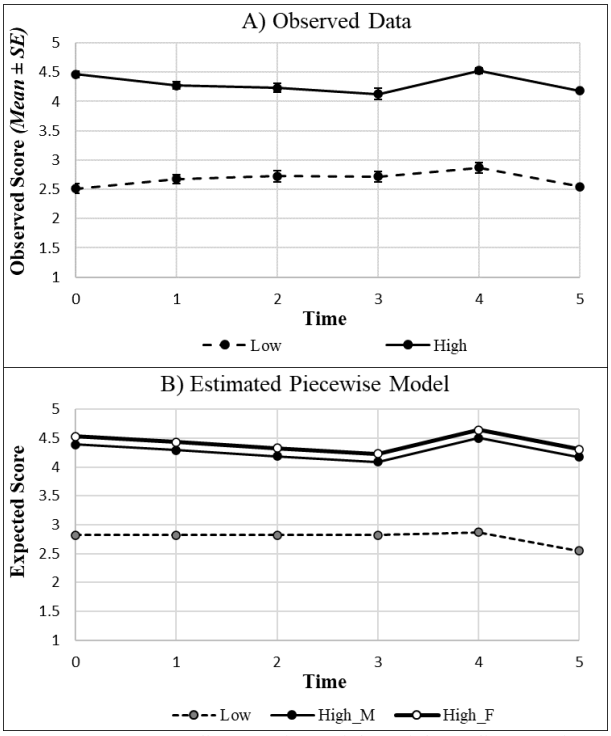
4. RESULTS

Summarized raw data for the sample composition and outcome variables from both low- and high-quality interviews across the six semesters are presented in Appendix A. Statistical parameters from the implemented piecewise models are displayed in Appendix B.

4.1 Change Over Time in Students' Evaluations Averaged across All Skills (H1-H2)

Figure 1 illustrates change in students' evaluations, averaged across all skills, of the low- and high-quality interviews over the six semesters. For the ***low-quality interview***, the piecewise regression model suggested no change in evaluation scores during the pre-pandemic semesters (p = .090), a slight increase in scores during the first pandemic-affected semester (β = 3.19, SE = 0.18, p < .0001), and a significant decrease during the second pandemic-affected semester (β = -0.32, SE = 0.11, p = .004). An independent-samples t-test was used to check whether the evaluation scores returned to the pre-pandemic level during the second pandemic-affected semester; the t-test showed no significant difference in students' overall evaluations of the low-quality interview between the last pre-pandemic and the second pandemic-affected semesters (t(99) = 1.42, p = .158). No difference between males and females was detected (p = .737).

For the ***high-quality interview***, the piecewise model showed a significant decrease over time in students' evaluations during the pre-pandemic semesters (β = -0.10, SE = 0.03, p = .001), a significant increase in evaluation scores during the first pandemic-affected semester (β = 4.83, SE = 0.14, p < .0001), and a decrease in scores during the second pandemic-affected semester (β = -0.33, SE = 0.09, p < .0001). An independent-samples t-test identified no significant difference in students' overall evaluations of the high-quality interview between the last pre-pandemic and the second pandemic-affected semesters (t(75) = -0.47, p = .637). Interestingly, females, on average, evaluated the high-quality interview higher than males across both pre-pandemic and pandemic-affected semesters (β = 0.14, SE = 0.07, p = .039).



Note: Low = Low-Quality Interview; High = High-Quality Interview; M = Males; F = Females

Figure 1. Observed Average Scores (Mean ± SE) (A) and Estimated Piecewise Models (B) for the Students' Evaluations of the Low- and High-Quality Interviews

4.2 Change Over Time in Students' Evaluations of the Soft versus Technical Skills (H3)

Figure 2 represents the change in students' evaluations of the soft versus technical skills demonstrated during the low- and high-quality interviews across the six semesters. For the ***soft skills in the low-quality interview***, the piecewise regression model suggested no change in evaluation scores during the pre-pandemic semesters (p = .185), a slight increase in scores during the first pandemic-affected semester (β = 2.91, SE = 0.19, p < .0001), and a significant decrease from the first to the second pandemic-affected semesters (β = -0.37, SE = 0.12, p = .003). No differences were observed between males and females (p = .979).

For the ***soft skills in the high-quality interview***, the piecewise model suggested a steady decrease in the evaluation scores during the pre-pandemic semesters (β = -0.10, SE = 0.03, p = .003), a steep increase in scores during the first pandemic-affected semester (β = 4.91, SE = 0.15, p < .0001), and a significant decrease from the first to the second pandemic-affected semesters (β = -0.34, SE = 0.09, p < .0001). No differences were observed between males and females (p = .137).

For the ***technical skills in the low-quality interview***, the piecewise model showed no change in students' evaluation scores during the pre-pandemic semesters (p = .079), a significant increase during the first pandemic-affected semester (β = 3.65, SE = 0.21, p < .0001) and no change between the two pandemic-affected semesters (p = .062).

Also, no differences between males and females were detected ($p = .439$).

For the ***technical skills in the high-quality interview***, the piecewise model suggested a significant decrease in the evaluation scores across the pre-pandemic semesters ($\beta = -0.11$, $SE = 0.04$, $p = .005$), a steep increase in scores during the first pandemic-affected semester ($\beta = 4.69$, $SE = 0.17$, $p < .0001$), and a decrease from the first to the second pandemic-affected semesters ($\beta = -0.31$, $SE = 0.10$, $p = .003$). Importantly, females significantly inflated their evaluations scores compared to males ($\beta = 0.19$, $SE = 0.08$, $p = .017$).

4.3 Change Over Time in Students' Evaluations of the Individual Criteria (H1-H3)

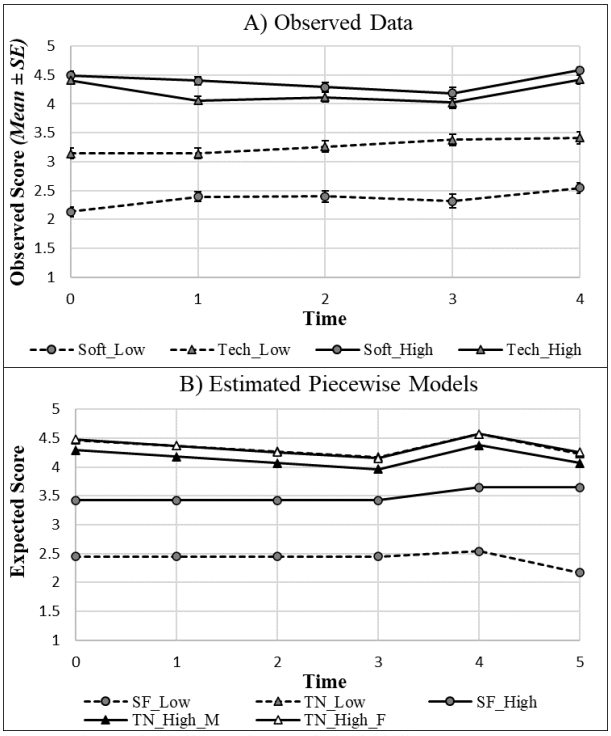
For the ***low-quality interview*** (see Appendix B), there was a significant increase in evaluation scores during the first pandemic-affected semester for all the measured criteria. Students' evaluations of the *Closing*, *Analysis*, *Design*, and *Visual* criteria remained constant during both the pre-pandemic and pandemic-affected semesters. Scores for *Listening* and *Relation* remained constant during the pre-pandemic semesters but decreased from the first to the second pandemic-affected semesters. Scores for *Teamwork* increased during the pre-pandemic semesters and remained unchanged from the first to the second pandemic-affected semesters.

For the ***high-quality interview***, again, there was a significant increase in evaluation scores during the first pandemic-affected semester for all the measured criteria. Students' evaluations for the *Closing* criterion did not change during the pre-pandemic semesters or between the pandemic-affected semesters. Scores for *Opening* and *Relation* did not change during the pre-pandemic semesters but decreased from the first to the second pandemic-affected semesters. Scores for *Listening*, *Teamwork*, *Analysis*, and *Design* decreased during the pre-pandemic semesters, as well as from the first to the second pandemic-affected semesters. In contrast, scores for *Visual* did not change during the pre-pandemic semesters but decreased from the first to the second pandemic-affected semesters. Finally, females' evaluations of the *Listening* and *Visual* criteria were higher than those from males.

4.4 Comparison of Students' and Instructors' RE Interview Evaluations (H4)

Although summarized scores (*All_Low* and *All_High* variables) for both interviews were, on average, higher among the students compared to the instructors, this difference was very small for the low-quality interview and quite substantial for the high-quality interview. Thus, on average, students adequately assessed the low-quality interview, but assigned inflated evaluations to the high-quality interview (Figure 3A). Moreover, the inflation of student evaluation scores in the high-quality interview was equally pronounced in both soft and technical skills (Figure 3B).

For the ***low-quality interview***, students assessed criteria of the *Opening*, *Listening*, and *Analysis* higher than instructors. By contrast, instructors gave higher scores than students for *Closing* and *Relation* outcomes. The criteria of *Teamwork*, *Design*, and *Visual* were assessed by students and instructors quite similarly (Figure 3C).



Note: Soft = Soft Skills; Tech = Technical Skills;
Low = Low-Quality Interview; High = High-Quality Interview;
M = Males; F = Females

Figure 2. Observed Average Scores (*Mean* ± *SE*) (A) and Estimated Piecewise Models (B) for the Students' Evaluations of the Soft vs. Technical Skills

For the ***high-quality interview***, students assessed *Opening*, *Relation*, *Teamwork*, and *Analysis* criteria considerably higher than instructors. *Closing* received slightly higher scores from instructors compared to students. Students and instructors had comparable scores on *Listening*, *Design*, and *Visual* outcomes (Figure 3D).

When we compared student and instructor evaluations, we noticed consistent gender differences (Figure 4). For the high-quality interview, while evaluating both soft and technical skills, female students inflated their evaluation scores compared to male students, who, in turn, inflated their scores compared to instructors.

4.5 Change in Gender Composition Over Time (H5)

Pearson chi-square analysis showed no differences in gender composition of the sample across the six semesters: χ^2 (5, $N = 294$) = 5.21, $p = .391$. Thus, the observed changes in student evaluations could not be attributed to this factor.

5. DISCUSSION

The goal of the current study was to explore students' evaluations of the requirements elicitation interviews and determine possible: 1) change over time in students' evaluations of the low- versus high-quality interviews (**H1-H2**); 2) change over time in students' evaluations of soft versus technical skills for the two types of interviews (**H3**); 3) change over time in students' evaluations of the individual

criteria for the interviews (**H1-H3**); 4) differences between students' and instructors' evaluations (**H4**); and 5) potential gender differences in student evaluations of RE interviews (**H5**).

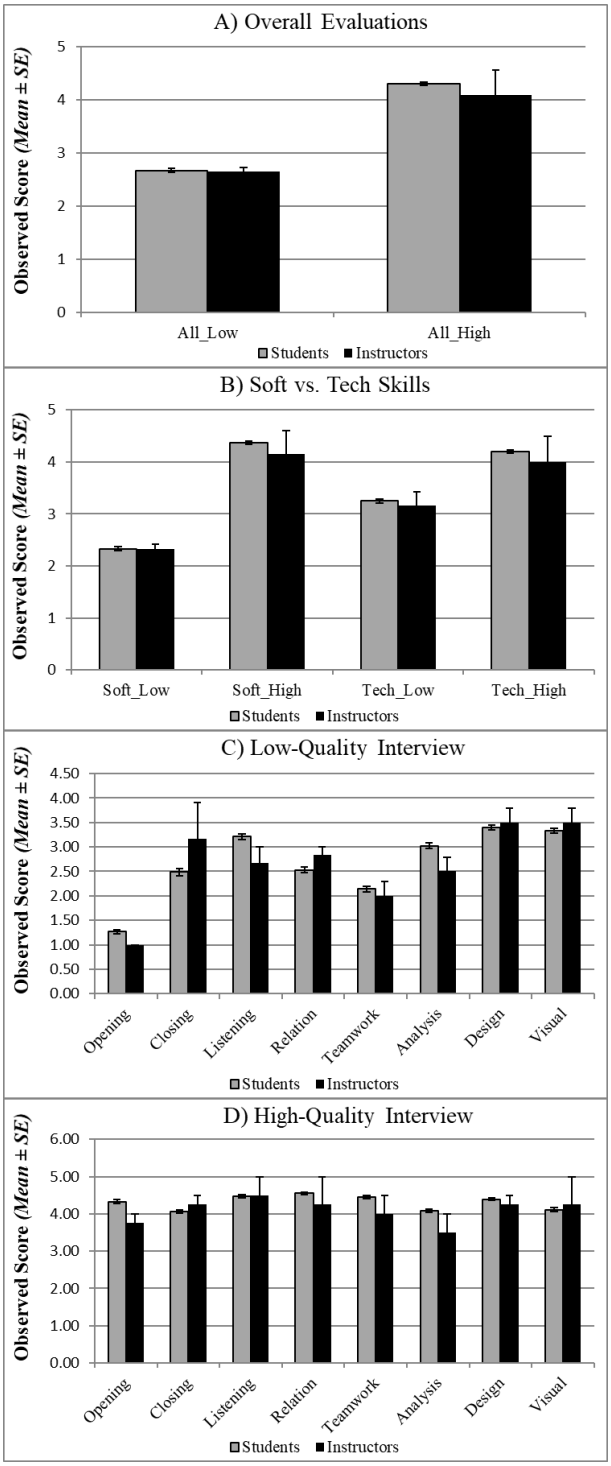
5.1 Change Over Time in the Low-Quality versus High-Quality Interview Evaluations (H1-H2)

Current results suggested partial support to hypotheses **H1** and **H2** that students' evaluations would remain stable during the pre-pandemic semesters but may shift considerably between the pre-pandemic and pandemic-affected semesters. Over the pre-pandemic semesters, students' evaluation scores remained stable for the low-quality interview and decreased steadily for the high-quality interview. At the onset of the pandemic, similar changes in the trajectories were observed for both types of interviews: during the first pandemic-affected semester, there was an inflation of students' evaluation scores, whereas during the second pandemic-affected semester the scores dropped significantly, back to the pre-pandemic level. Thus, the estimated models suggested a significant disruption during the pandemic-affected semesters. Based on previous research, we suggest that students' inflation of evaluations during the first pandemic-affected semester resulted from negative psychological effects associated with the COVID-19 pandemic, specifically stress, fear, sadness, and empathy (Cohen & Strayer, 1996; Goldberg et al., 1999; Keltner et al., 1993; Lerner & Keltner, 2000). A significant drop in evaluation scores during the second pandemic-affected semester may indicate students' adaptation to the negative factors associated with the pandemic.

5.2 Student Evaluations of the Soft versus Technical Skills (H3)

In support to the hypothesis **H3**, we found that the trajectories of change in student evaluations differed between the soft and technical skills. Students' evaluations of the soft skills remained constant across the pre-pandemic semesters, whereas evaluations of the technical skills decreased across the pre-pandemic semesters for both types of the interview. Again, there was a significant inflation of evaluation scores during the first pandemic-affected semester for both soft and technical skills in both types of the interview; the second pandemic-affected semester brought a significant decrease in all the skills and interviews, except soft skills in the high-quality interview, for which the evaluation scores remained inflated as much as during the first pandemic-affected semester.

Thus, separating the set of evaluated criteria into soft versus technical allowed us to pinpoint the location of change. With each passing semester, students were more critical while evaluating technical skills in both types of the interview. This can be attributed to the increased instructor's attention to mastering technical skills (such as correct use of the UML syntax and semantic accuracy of the models); this shifted attention was based on the past observations of weakening students' technical skills and the program-wide decision to bring them back in focus. Previous research showed that technical skills are important for success in IT professions (Medlin et al., 2001; Merhout et al., 2009). The trend toward more critical evaluation of the technical skills during the pre-pandemic semesters indicates the strength of the training program, while a significant inflation of evaluation scores during the first pandemic-affected semester may suggest the disruptive effect of the COVID-19 pandemic on student learning.



Note: Soft = Soft Skills; Tech = Technical Skills;
Low = Low-Quality Interview; High = High-Quality Interview

Figure 3. Comparison of the Observed Overall Evaluations (A), Soft vs. Technical Skills (B), as well as Scores for Individual Criteria in the Low-Quality (C) and High-Quality (D) Interviews Between Students and Instructors

5.3 Student Evaluations of the Individual REIAR Criteria (H1-H3)

While looking at the change in students' evaluations of the individual criteria, we noticed that the stability in evaluation scores during the pre-pandemic semesters for the low-quality interview was due to the contribution of all the criteria except *Opening* and *Teamwork*, which showed an upward trend. Similarly, not all the criteria exhibited a steady decrease across the pre-pandemic semesters in the high-quality interview: *Opening*, *Closing*, *Relation*, and *Visual* showed no trend.

Furthermore, although all the criteria in both the low- and high-quality interviews showed a significant inflation during the first pandemic-affected semester, a decrease in scores during the second pandemic-affected semester was not observed in *Closing*, *Teamwork*, *Analysis*, *Design*, and *Visual* criteria for the low-quality interview, as well as in *Closing* for the high-quality interview. Thus, analysis of the individual evaluation criteria, rather than aggregated measures, may shed some light on the areas of strengths and weaknesses in student evaluations, as well as areas most affected by the pandemic.

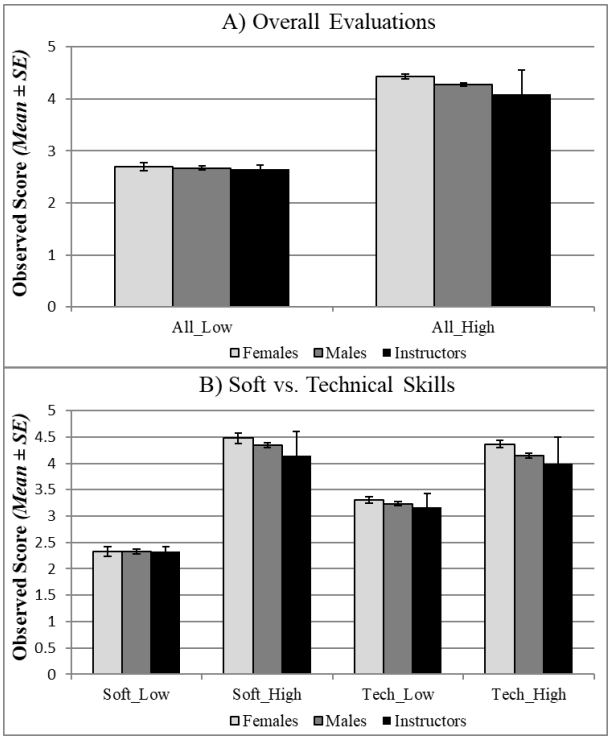
5.4 Comparison of Students' and Instructors' RE Interview Evaluations (H4)

Current results provided support to hypothesis **H4** that students inflated their evaluation scores compared to instructors. Indeed, on average, students evaluated the high-quality interview higher than instructors; however, very small difference was found between students' and instructors' evaluations for the low-quality interview. For the high-quality interview, students inflated their scores for both soft and technical skills; in both cases, female students assigned higher evaluation scores than male students, who, in turn, assigned higher scores compared to instructors. When we compared evaluation scores for the individual criteria between students and instructors, the picture became more complicated: in both low- and high-quality interviews, some criteria were assessed higher by students compared to instructors, whereas other criteria received higher scores from instructors compared to students. In both low- and high-quality interviews, students evaluated the *Opening*, *Teamwork*, and *Analysis* criteria higher than instructors, whereas instructors assigned higher scores than students to the *Closing* criterion.

Importantly, at the beginning of the study, students assigned higher scores than instructors to the skills in the high-quality interview; however, a decrease in scores across the pre-pandemic semesters reduced the gap between students' and instructors' evaluations: male students reached the instructor's level of evaluations by the fourth semester, while female students still exhibited inflated scores. We may conclude that during the pre-COVID period, the increasing quality of RE training delivery promoted shared understanding (between the instructor and students) of interview skills and expected proficiency, and enabled students to evaluate even the high-quality interviews more adequately with each passing semester. However, during the semesters affected by the COVID-19 pandemic, this positive training and learning trend was disrupted.

5.5 Gender Differences in Student Evaluations (H5)

In partial support to hypothesis **H5**, significant gender differences were found in students' evaluations of the high-quality interview, but not the low-quality interview. On average, females tended to assign higher scores than males across all six semesters. Furthermore, gender differences appeared only in the evaluation of the technical skills, rather than the soft skills, and only for the high-quality interview. While looking at the individual evaluation criteria, we found that females assigned higher scores than males only for the *Listening* and *Visual* criteria.



Note: Soft = Soft Skills; Tech = Technical Skills;
Low = Low-Quality Interview; High = High-Quality Interview

Figure 4. Comparison of the Observed Overall Evaluations (A) and Soft vs. Technical Skills (B) Between Female Students, Male Students, and Instructors

These findings align well with previous research suggesting that males, compared to females, may be more critical while evaluating others (Abad-Tortosa et al., 2017; Alagna, 1982). In addition, females are more susceptible to the effects of stress (Bayram & Bilgel, 2008; Brougham et al., 2009; Misra & McKean, 2000; Pierceall & Keim, 2007), while the resulting fear, sadness, and helplessness may trigger more positive evaluations (Goldberg et al., 1999; Keltner et al., 1993; Lerner & Keltner, 2000). Females' tendency to empathize more than males (Cohen & Strayer, 1996; Nwankwo, 2013) may also stimulate less critical evaluations of others exhibited by females (Abad-Tortosa et al., 2017).

Importantly, the above-mentioned gender differences in evaluative judgements could potentially affect students' evaluations in this study. For example, a shift in the gender composition of the sample towards higher proportion of females during the first pandemic-affected semester compared to previous semesters could have resulted in the inflation of student evaluations. Additional analysis showed that this was not the case: there was no significant difference in gender

composition across the six semesters. Thus, the observed shift towards inflation of student evaluations during the first pre-pandemic semester may be attributed to the effects of the COVID-19 pandemic.

5.6 Limitations and Strengths of the Current Study

Only four instructors evaluated RE interviews for this study; such limited data did not permit a more rigorous statistical analysis comparing student and instructor evaluations. Moreover, all the instructors who provided their RE interview evaluations for this study were males. Acknowledging gender differences in evaluative judgements, future research should replicate current results while considering RE interview evaluation scores from both male and female instructors.

Furthermore, in the current study, student evaluations from only two semesters were potentially affected by the COVID-19 pandemic. More longitudinal data covering the COVID-19 pandemic is needed to re-evaluate the emerging trends detected in the current study. Also, during the two pandemic-affected semesters, the instruction mode was changed from in-person to synchronous online. One might argue that the changes in student evaluations we attributed to COVID-19 pandemic could be due to the change in the instruction mode. Although it is impossible to separate the two effects, we propose that the change in the instruction mode had very little effect on student evaluations since the interview evaluation format and procedures did not change with the onset of the pandemic – students were expected to watch and evaluate the interviews on their own in the comfort of their homes during both pre-pandemic and pandemic-affected semesters. Thus, we propose that the observed disruption in student evaluations was due to the effects of COVID-19 pandemic rather than the change in the instruction mode.

On the positive side, the piecewise modeling implemented in the current study allowed us to chart trajectories of change in students' evaluations of the RE interviews separately for the pre-pandemic and pandemic-affected semesters. One might argue that the COVID-19 pandemic was a disruptive event that could potentially influence students' evaluations in multiple ways, and the data generated during the pandemic should be discarded. To such readers, we suggest to consider only the pre-pandemic segment of the trajectory and disregard the pandemic-affected segment. Others might argue that the COVID-19 pandemic may have had no effect on students' learning and evaluative judgments, and the regression analysis should have modelled only one trajectory across all six semesters. To examine this option, we ran an additional analysis of change across all six semesters in average scores that showed no change over time for either the low-quality (p = .335) or the high-quality (p = .234) interviews, meaning that the identified effects (a significant decrease in scores across pre-pandemic semesters and an increase in scores at the onset of the pandemic) cancelled each other to produce an erroneous appearance of no trend of change over time.

Moreover, some might argue that a t-test could suffice to compare evaluation scores aggregated across the pre-pandemic semesters versus the pandemic-affected semesters. We would like to note that studying change over time, rather than combining the data across multiple semesters, allowed us to identify several interesting and important trends: 1) an increase in the instruction effectiveness across multiple semesters; 2) an inflation of evaluation scores as a result of negative psychological effects and disruption of learning processes at the onset of the pandemic; and 3) a steep decrease, back to the pre-pandemic values, during the second pandemic-affected semester due to adaptation to the negative conditions.

6. CONCLUSIONS

The current study provided a comprehensive account of the change in students' peer evaluations of the low- and high-quality RE interviews over six semesters, including two semesters during the COVID-19 pandemic. Exploring the data at different levels of analysis (low- versus high quality interview, soft versus technical skills in each interview, and eight individual evaluation criteria for each interview) provided important insights into the complexity of learning trends within the data. We found that students' evaluations for the high-quality interviews were originally inflated compared to instructors' ones, but with each semester, students' evaluations were becoming more critical and approximated the instructors' evaluation scores after four pre-pandemic semesters. This trend indicated the ability of the program to coach students' RE interview skills and to promote more critical outlook. However, during the first semester affected by the COVID-19 pandemic, students significantly inflated their evaluations of RE interviews. This change could have stemmed from negative psychological effects associated with the pandemic.

We also found significant gender differences in students' perceptions of effective technical skill application. Specifically, females tend to assign significantly higher scores than males in the evaluation of the technical skills in a high-quality interview; this result may be indicative of females' difficulty to recognize more subtle nuances in the application of technical skills in the medium- to high-level performance. Further research should examine the ways technical skills are taught to and learned by males and females in the IS discipline. Our results may be of interest and practical use to the instructors and course designers involved in integrating RE training in their IS courses. Future research should further evaluate the long-term effects associated with the pandemic disruption in students' lives and academic practices. In particular, it is important to investigate possible interventions and techniques that could mitigate the negative factors discussed in this paper affecting the RE process.

7. ACKNOWLEDGMENTS

We would like to thank: the Editors and anonymous reviewers for their valuable comments and recommendations; the College of Business at JMU for supporting this work; JMU CISBSAN faculty and students who participated in data collection; Drs. Diane Lending, Thomas Dillon, and Jeremy Ezell for providing feedback on the early versions of this study; JMU graduate research assistant Nicolau Nguimbi for his help with data processing.

8. REFERENCES

Abad-Tortosa, D., Alacreu-Crespo, A., Costa, R., Salvador, A., & Serrano, M. Á. (2017). Sex Differences in Autonomic Response and Situational Appraisal of a

Competitive Situation in Young Adults. *Biological Psychology*, 126, 61-70. https://doi.org/10.1016/j.biopsycho.2017.04.008

Adachi, C., Tai, J. H.-M., & Dawson, P. (2018). Academics' Perceptions of the Benefits and Challenges of Self and Peer Assessment in Higher Education. *Assessment & Evaluation in Higher Education*, 43(2), 294-306. https://doi.org/10.1080/02602938.2017.1339775

Agarwal, R., & Tanniru, M. R. (1990). Knowledge Acquisition Using Structured Interviewing: An Empirical Investigation. *Journal of Management Information Systems*, 7(1), 123-140. https://doi.org/10.1080/07421222.1990.11517884

Alagna, S. W. (1982). Sex Role Identity, Peer Evaluation of Competition, and the Responses of Women and Men in a Competitive Situation. *Journal of Personality and Social Psychology*, 43(3), 546-554. https://doi.org/10.1037/0022-3514.43.3.546

Alvarez, R. (2002). Confessions of an Information Worker: A Critical Analysis of Information Requirements Discourse. *Information and Organization*, 12(2), 85-107. https://doi.org/10.1016/S1471-7727(01)00012-4

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing Procedures for Implementing Peer Assessment in Large Classes Using an Action Research Process. *Assessment & Evaluation in Higher Education*, 27(5), 427-441. https://doi.org/10.1080/0260293022000009302

Bayram, N., & Bilgel, N. (2008). The Prevalence and Socio-demographic Correlations of Depression, Anxiety and Stress Among a Group of University Students. *Social Psychiatry and Psychiatric Epidemiology*, 43(8), 667-672. https://doi.org/10.1007/s00127-008-0345-x

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The Prevalence and Correlates of Depression, Anxiety, and Stress in a Sample of College Students. *Journal of Affective Disorders*, 173, 90-96. https://doi.org/10.1016/j.jad.2014.10.054

Blaney, P. H. (1986). Affect and Memory: A Review. *Psychological Bulletin*, 99(2), 229-246. https://doi.org/10.1037/0033-2909.99.2.229

Bodenhausen, G. V., & Wyer, R. S. (1987). Social Cognition and Social Reality: Information Acquisition and use in the Laboratory and the Real World. In H.-J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social Information Processing and Survey Methodology* (pp. 6-41). Springer. https://doi.org/10.1007/978-1-4612-4798-2_2

Bower, G. H. (1991). Mood Congruity of Social Judgments. In *Emotion and Social Judgments* (pp. 31-53).

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The Psychological Impact of Quarantine and How to Reduce it: Rapid Review of the Evidence. *The Lancet*, 395(10227), 912-920. https://doi.org/10.1016/S0140-6736(20)30460-8

Brougham, R. R., Zail, C. M., Mendoza, C. M., & Miller, J. R. (2009). Stress, Sex Differences, and Coping Strategies Among College Students. *Current Psychology*, 28(2), 85-97. https://doi.org/10.1007/s12144-009-9047-0

Browne, G. J., & Ramesh, V. (2002). Improving Information Requirements Determination: A Cognitive Perspective. *Information & Management*, 39(8), 625-645. https://doi.org/10.1016/S0378-7206(02)00014-9

Browne, G. J., & Rogich, M. B. (2001). An Empirical Investigation of User Requirements Elicitation: Comparing the Effectiveness of Prompting Techniques. *Journal of Management Information Systems*, 17(4), 223-249. https://doi.org/10.1080/07421222.2001.11045665

Burgess, A. W., Roberts, C., Black, K. I., & Mellis, C. (2013). Senior Medical Student Perceived Ability and Experience in Giving Peer Feedback in Formative Long Case Examinations. *BMC Medical Education*, 13(1), 79. https://doi.org/10.1186/1472-6920-13-79

Burnay, C. (2016). Are Stakeholders the Only Source of Information for Requirements Engineers? Toward a Taxonomy of Elicitation Information Sources. *ACM Transactions on Management Information Systems*, 7(3), 1-29. https://doi.org/10.1145/2965085

Burnay, C., Jureta, I., & Faulkner, S. (2014). What Stakeholders Will or Will Not Say: A Theoretical and Empirical Study of Topic Importance in Requirements Engineering Elicitation Interviews. *Information Systems*, 46, 61-81. https://doi.org/10.1016/j.is.2014.05.006

Byrd, T. A., Cossick, K. L., & Zmud, R. W. (1992). A Synthesis of Research on Requirements Analysis and Knowledge Acquisition Techniques. *MIS Quarterly*, 16(1), 117-138. https://doi.org/10.2307/249704

Cao, Z., Yu, S., & Huang, J. (2019). A Qualitative Inquiry into Undergraduates' Learning from Giving and Receiving Peer Feedback in L2 Writing: Insights from a Case Study. *Studies in Educational Evaluation*, 63, 102-112. https://doi.org/10.1016/j.stueduc.2019.08.001

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective Causes and Consequences of Social Information Processing. In *Handbook of social cognition: Basic processes; Applications, Vols. 1-2, 2nd ed* (pp. 323-417). Lawrence Erlbaum Associates, Inc.

Cohen, D., & Strayer, J. (1996). Empathy in Conduct-disordered and Comparison Youth. *Developmental Psychology*, 32(6), 988-998. https://doi.org/10.1037/0012-1649.32.6.988

Cooke, J. E., Eirich, R., Racine, N., & Madigan, S. (2020). Prevalence of Posttraumatic and General Psychological Stress During COVID-19: A Rapid Review and Meta-analysis. *Psychiatry Research*, 292, 113347. https://doi.org/10.1016/j.psychres.2020.113347

Costain, G., & McKenna, B. (2011). Experiencing the Elicitation of User Requirements and Recording them in Use Case Diagrams through Role-Play. *Journal of Information Systems Education*, 22(4), 367-380.

Davey, B., & Cope, C. (2008). Requirements Elicitation—What's Missing. *Issues in Informing Science and Information Technology*, 5. https://doi.org/10.28945/1027

Davis, A., Dieste, O., Hickey, A., Juristo, N., & Moreno, A. M. (2006). Effectiveness of Requirements Elicitation Techniques: Empirical Results Derived from a Systematic Review. *14th IEEE International Requirements Engineering Conference (RE'06)*, 179-188. https://doi.org/10.1109/RE.2006.17

De Grez, L., Valcke, M., & Roozen, I. (2012). How Effective Are Self- and Peer Assessment of Oral Presentation Skills Compared with Teachers' Assessments? *Active Learning*

in Higher Education, 13(2), 129-142. https://doi.org/10.1177/1469787412441284

Dennis, A., Wixom, B. H., & Tegarden, D. (2015). *Systems Analysis and Design: An Object-Oriented Approach with UML* (5th ed.). John Wiley & Sons.

Eaton, R. J., & Bradley, G. (2008). The Role of Gender and Negative Affectivity in Stressor Appraisal and Coping Selection. *International Journal of Stress Management*, 15(1), 94-115. https://doi.org/10.1037/1072-5245.15.1.94

Eisenberg, N., & Strayer, J. (1987). *Empathy and its Development*. Cambridge University Press.

Ezell, J. D., Lending, D., Dillon, T. W., May, J., Hurney, C. A., & Fulcher, K. H. (2019). Developing Measurable Cross-Departmental Learning Objectives for Requirements Elicitation in an Information Systems Curriculum. *Journal of Information Systems Education*, 30(1), 27-41.

Ezell, J. D., Lending, D., Kruck, S. E., Dillon, T. W., & May, J. L. (2016). A Plan to Improve Learning of Requirements Elicitation in an IS Curriculum. *Proceedings of the 2016 ACM SIGMIS Conference on Computers and People Research - SIGMIS-CPR '16*, 31-31. https://doi.org/10.1145/2890602.2890621

Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3), 287-322.

Forgas, J. P. (1994). The Role of Emotion in Social Judgments: An Introductory Review and an Affect Infusion Model (AIM). *European Journal of Social Psychology*, 24(1), 1-24. https://doi.org/10.1002/ejsp.2420240102

Forgas, J. P. (1995). Mood and Judgment: The Affect Infusion Model (AIM). *Psychological Bulletin*, 117(1), 39-66. https://doi.org/10.1037/0033-2909.117.1.39

Freeman, M. (1995). Peer Assessment by Groups of Group Work. *Assessment & Evaluation in Higher Education*, 20(3), 289-300.

Fulcher, K. H., Good, M. R., Coleman, C. M., & Smith, K. L. (2014). A Simple Model for Learning Improvement: Weigh Pig, Feed Pig, Weigh Pig. Occasional Paper #23. In *National Institute for Learning Outcomes Assessment*. National Institute for Learning Outcomes Assessment. https://eric.ed.gov/?id=ED555526

Gallagher, M. W., Zvolensky, M. J., Long, L. J., Rogers, A. H., & Garey, L. (2020). The Impact of Covid-19 Experiences and Associated Stress on Anxiety, Depression, and Functional Impairment in American Adults. *Cognitive Therapy and Research*, 44(6), 1043-1051. https://doi.org/10.1007/s10608-020-10143-y

Gallivan, M. J., & Keil, M. (2003). The User-Developer Communication Process: A Critical Case Study. *Information Systems Journal*, 13(1), 37-68. https://doi.org/10.1046/j.1365-2575.2003.00138.x

Goldberg, J. H., Lerner, J. S., & Tetlock, P. E. (1999). Rage and Reason: The Psychology of the Intuitive Prosecutor. *European Journal of Social Psychology*, 29(5-6), 781-795. https://doi.org/10.1002/(SICI)1099-0992(199908/09)29:5/6<781::AID-EJSP960>3.0.CO;2-3

Havelka, D. (2003). A User-Oriented Model of Factors that Affect Information Requirements Determination Process Quality. Information Resources Management Journal, 16(4), 15-32. www.igi-global.com/article/user-oriented-model-factors-affect/1242

He, J., & King, W. R. (2008). The Role of User Participation in Information Systems Development: Implications from a Meta-Analysis. *Journal of Management Information Systems*, 25(1), 301-331. https://doi.org/10.2753/MIS0742-1222250111

Hickey, A. M., & Davis, A. M. (2003). Elicitation Technique Selection: How do Experts do it? *Proceedings. 11th IEEE International Requirements Engineering Conference*, 169-178. https://doi.org/10.1109/ICRE.2003.1232748

Holtzblatt, K., & Beyer, H. R. (1995). Requirements Gathering: The Human Factor. *Communications of the ACM*, 38(5), 31-32. https://doi.org/10.1145/203356.203361

Hughes, I. e., & Large, B. j. (1993). Staff and Peer-group Assessment of Oral Communication Skills. *Studies in Higher Education*, 18(3), 379-385. https://doi.org/10.1080/03075079312331382281

Jain, H., Vitharana, P., & Zahedi, F. (2003). An Assessment Model for Requirements Identification in Component-based Software Development. *The Data Base for Advances in Information Systems*, 34(4), 48-63. https://doi.org/10.1145/957758.957765

Kahneman, D. (2002). Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice. *Nobel Prize Lecture*, 8 (December), 449-489.

Kamthan, P., & Shahmir, N. (2019). On Software Projects in Academia and Industry from a Perspective of Software Engineering Education. *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, 34-39. https://doi.org/10.1109/CSCI49370.2019.00013

Keltner, D., Ellsworth, P. C., & Edwards, K. (1993). Beyond Simple Pessimism: Effects of Sadness and Anger on Social Perception. *Journal of Personality and Social Psychology*, 64(5), 740-752. https://doi.org/10.1037/0022-3514.64.5.740

Kruger, J., & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. https://doi.org/10.1037/0022-3514.77.6.1121

Langan, A. M., Shuker, D. M., Cullen, W. R., Penney, D., Preziosi, R. F., & Wheater, C. P. (2008). Relationships Between Student Characteristics and Self-, Peer and Tutor Evaluations of Oral Presentations. *Assessment & Evaluation in Higher Education*, 33(2), 179-190. https://doi.org/10.1080/02602930701292498

Langan, A. M., Wheater, C. P., Shaw, E. M., Haines, B. J., Cullen, W. R., Boyle, J. C., Penney, D., Oldekop, J. A., Ashcroft, C., Lockey, L., & Preziosi, R. F. (2005). Peer Assessment of Oral Presentations: Effects of Student Gender, University Affiliation and Participation in the Development of Assessment Criteria. *Assessment & Evaluation in Higher Education*, 30(1), 21-34. https://doi.org/10.1080/0260293042003243878

Lending, D., Fulcher, K., Ezell, J. D., May, J. L., & Dillon, T. W. (2018). Example of a Program-Level Learning Improvement Report. *Research & Practice in Assessment*, 13, 34-50.

Lerner, J. S., & Keltner, D. (2000). Beyond Valence: Toward a Model of Emotion-specific Influences on Judgement and Choice. *Cognition and Emotion*, 14(4), 473-493. https://doi.org/10.1080/026999300402763

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or Assessee: How Student Learning Improves by Giving and Receiving Peer Feedback. *British Journal of Educational Technology*, 41(3), 525-536. https://doi.org/10.1111/j.1467-8535.2009.00968.x

Lindquist, C. (2005, November 15). *Fixing the Software Requirements Mess*. CIO. https://www.cio.com/article/2448110/fixing-the-software-requirements-mess.html

Lundstrom, K., & Baker, W. (2009). To Give is Better Than to Receive: The Benefits of Peer Review to the Reviewer's Own Writing. *Journal of Second Language Writing*, 18(1), 30-43. https://doi.org/10.1016/j.jslw.2008.06.002

Marakas, G. M., & Elam, J. J. (1998). Semantic Structuring in Analyst Acquisition and Representation of Facts in Requirements Analysis. *Information Systems Research*, 9(1), 37-63. https://doi.org/10.1287/isre.9.1.37

McCarty, J. A., & Shrum, L. J. (2000). The Measurement of Personal Values in Survey Research: A Test of Alternative Rating Procedures*. *Public Opinion Quarterly*, 64(3), 271-298. https://doi.org/10.1086/317989

Medlin, B. D., Dave, D. S., & Vannoy, S. A. (2001). Students' Views of the Importance of Technical and Non-Technical Skills for Successful IT Professionals. *Journal of Computer Information Systems*, 42(1), 65-69. https://doi.org/10.1080/08874417.2001.11647040

Merhout, J. W., Havelka, D., & Hick, S. N. (2009). Soft Skills versus Technical Skills: Finding the Right Balance for an IS Curriculum. *Proceedings of the Fifteenth Americas Conference on Information Systems, San Francisco, California*, 9.

Misra, R., & McKean, M. (2000). College Students' Academic Stress and its Relation to Their Anxiety, Time Management, and Leisure Satisfaction. *American Journal of Health Studies*, 16(1), 41-51.

Moody, J. W., Blanton, J. E., & Cheney, P. H. (1998). A Theoretically Grounded Approach to Assist Memory Recall during Information Requirements Determination. *Journal of Management Information Systems*, 15(1), 79-98. https://doi.org/10.1080/07421222.1998.11518197

Morris, W. N. (1989). *Mood: The Frame of Mind*. Springer-Verlag. https://doi.org/10.1007/978-1-4612-3648-1

Nwankwo, B. E. (2013). Role of Gender, Emotional Empathy, Interpersonal Attraction on Moral Judgement. *IFE PsychologIA : An International Journal*, 21(2), 264-276. https://doi.org/10.10520/EJC141121

Patchan, M. M., & Schunn, C. D. (2015). Understanding the Benefits of Providing Peer Feedback: How Students Respond to Peers' Texts of Varying Quality. *Instructional Science*, 43(5), 591-614. https://doi.org/10.1007/s11251-015-9353-x

Pierceall, E. A., & Keim, M. C. (2007). Stress and Coping Strategies Among Community College Students. *Community College Journal of Research and Practice*, 31(9), 703-712. https://doi.org/10.1080/10668920600866579

Pitts, M. G., & Browne, G. J. (2004). Stopping Behavior of Systems Analysts During Information Requirements Elicitation. *Journal of Management Information Systems*, 21(1), 203-226. https://doi.org/10.1080/07421222.2004.11045795

Pitts, M. G., & Browne, G. J. (2007). Improving Requirements Elicitation: An Empirical Investigation of Procedural Prompts. *Information Systems Journal*, 17(1), 89-110. https://doi.org/10.1111/j.1365-2575.2006.00240.x

Pond, K., Ul-Haq, R., & Wade, W. (1995). Peer Review: A Precursor to Peer Assessment. *Innovations in Education and Training International*, 32(4), 314-323. https://doi.org/10.1080/1355800950320403

Salari, N., Hosseinian-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, S., Mohammadi, M., Rasoulpoor, S., & Khaledi-Paveh, B. (2020). Prevalence of Stress, Anxiety, Depression Among the General Population during the COVID-19 Pandemic: A Systematic Review and Meta-analysis. *Globalization and Health*, 16(1), 57. https://doi.org/10.1186/s12992-020-00589-w

Satkus, P. (n.d.). *Interview with Dr. Dillon and Dr. Lending*. https://www.jmu.edu/assessment/featuredStories/2018/DillonLending.shtml

Sax, L. J. (2003). Our Incoming Students: What Are They Like? *About Campus*, 8(3), 15-20. https://doi.org/10.1177/108648220300800305

Schenk, K. D., Vitalari, N. P., & Davis, K. S. (1998). Differences between Novice and Expert Systems Analysts: What Do We Know and What Do We Do? *Journal of Management Information Systems*, 15(1), 9-50. https://doi.org/10.1080/07421222.1998.11518195

Schwarz, N. (1990). Feelings as Information: Informational and Motivational Functions of Affective States. In *Handbook of motivation and cognition: Foundations of social behavior, Vol. 2.* (pp. 527-561). The Guilford Press.

Schwarz, N., & Clore, G. L. (1983). Mood, Misattribution, and Judgments of Well-being: Informative and Directive Functions of Affective States. *Journal of Personality and Social Psychology*, 45(3), 513-523. https://doi.org/10.1037/0022-3514.45.3.513

Sherman, S. J., & Corty, E. (1984). Cognitive Heuristics. In *Handbook of Social Cognition* (In R. S. Wyer & T. K. Srull (Eds.), Vol. 3, pp. 189-286). Erlbaum.

Siemer, M., & Reisenzein, R. (1998). Effects of Mood on Evaluative Judgements: Influence of Reduced Processing Capacity and Mood Salience. *Cognition and Emotion*, 12(6), 783-805. https://doi.org/10.1080/026999398379439

Son, C., Hegde, S., Smith, A., Wang, X., & Sasangohar, F. (2020). Effects of COVID-19 on College Students' Mental Health in the United States: Interview Survey Study. *Journal of Medical Internet Research*, 22(9), e21279. https://doi.org/10.2196/21279

Sridharan, B., Tai, J., & Boud, D. (2019). Does the Use of Summative Peer Assessment in Collaborative Group Work Inhibit Good Judgement? *Higher Education*, 77(5), 853-870. https://doi.org/10.1007/s10734-018-0305-7

Sutherland, L., & Ellery, K. (2004). Involving Students in the Assessment process: Research Article. *Perspectives in Education*, 22(1), 99-110.

Turner, J. (1990). *A Comparison of the Process of Knowledge Elicitation with that of Information Requirements Determination* (SSRN Scholarly Paper ID 1289074). Social Science Research Network. https://papers.ssrn.com/abstract=1289074

Valusek, J. R., & Fryback, D. G. (1985). Information Requirements Determination: Obstacles within, among and between Participants. *Proceedings of the Twenty-First Annual Conference on Computer Personnel Research*, 103-111. https://doi.org/10.1145/16687.16700

Vijayan, J., & Raju, G. (2011). A New approach to Requirements Elicitation Using Paper Prototype. *International Journal of Advanced Science and Technology*, 28, 9-16.

Wang, X., Hegde, S., Son, C., Keller, B., Smith, A., & Sasangohar, F. (2020). Investigating Mental Health of US College Students During the COVID-19 Pandemic: Cross-Sectional Survey Study. *Journal of Medical Internet Research*, 22(9), e22817. https://doi.org/10.2196/22817

Watson, H. J., & Frolick, M. N. (1993). Determining Information Requirements for an EIS. *MIS Quarterly*, 17(3), 255-269. https://doi.org/10.2307/249771

Wyer, R. S., & Srull, T. K. (1986). Human Cognition in its Social Context. *Psychological Review*, 93(3), 322-359. https://doi.org/10.1037/0033-295X.93.3.322

Yusoff, M. S. B., Abdul Rahim, A. F., & Yaacob, M. J. (2010). Prevalence and Sources of Stress among Universiti Sains Malaysia Medical Students. *The Malaysian Journal of Medical Sciences*, 17(1), 30-37.

Zhang, Z. (2007). Effective Requirements Development – A Comparison of Requirements Elicitation Techniques. *British Computer Society*.

Zowghi, D., & Coulin, C. (2005). Requirements Elicitation: A Survey of Techniques, Approaches, and Tools. In A. Aurum & C. Wohlin (Eds.), *Engineering and Managing Software Requirements* (pp. 19-46). Springer. https://doi.org/10.1007/3-540-28244-0_2

AUTHOR BIOGRAPHIES

Dmytro Babik is an assistant professor of computer information systems at James Madison University, Harrisonburg, VA. He received his PhD in Information Systems and PhD minor in Educational Research Methodology from the University of North Carolina at Greensboro in 2015. He teaches courses in Systems Analysis and Design, Enterprise Architecture, and Agile Development. His research focuses on interdisciplinary social learning, wicked problem solving, critical and systems thinking, and peer assessment.

Iryna Babik is an assistant professor in the Department of Psychological Science at Boise State University. She received her PhD in Psychology and PhD minor in Educational Research Methodology from the University of North Carolina at Greensboro in 2014, conducted post-doctoral research at the University of Delaware, and joined Boise State University in 2019. She teaches courses in Research Methods and Advanced Statistical Analyses. Her research focuses on text analysis, intergroup biases, and teamwork in multicultural teams.

APPENDICES

Appendix A. Summarized Data (Mean ± SE) for Each Outcome Variable

Variables	Students							Instructors
	Fall 2018 Time = 0	Spring 2019 Time = 1	Fall 2019 Time = 2	Spring 2020 Time = 3	Fall 2020 Time = 4	Spring 2021 Time = 5	2018 - 2021	2018 - 2020
Sample size	47	44	44	43	58	58	294	4
% of males	70.21	75	84.09	76.74	77.59	86.21	78.57	100
All_Low	2.51 ± 0.08	2.67 ± 0.08	2.72 ± 0.09	2.71 ± 0.10	2.87 ± 0.09	2.55 ± 0.07	2.71 ± 0.04	2.65 ± 0.08
All_High	4.46 ± 0.06	4.27 ± 0.07	4.23 ± 0.08	4.13 ± 0.09	4.52 ± 0.05	4.18 ± 0.06	4.34 ± 0.03	4.09 ± 0.47
Soft_Low	2.13 ± 0.08	2.39 ± 0.09	2.40 ± 0.11	2.32 ± 0.11	2.54 ± 0.09	2.18 ± 0.08	2.36 ± 0.04	2.33 ± 0.09
Soft_High	4.50 ± 0.07	4.40 ± 0.07	4.29 ± 0.08	4.19 ± 0.09	4.59 ± 0.05	4.23 ± 0.07	4.41 ± 0.05	4.15 ± 0.45
Tech_Low	3.14 ± 0.09	3.14 ± 0.10	3.26 ± 0.10	3.38 ± 0.10	3.42 ± 0.11	3.16 ± 0.10	3.27 ± 0.03	3.17 ± 0.25
Tech_High	4.41 ± 0.07	4.06 ± 0.08	4.12 ± 0.08	4.03 ± 0.11	4.42 ± 0.08	4.09 ± 0.07	4.22 ± 0.04	4.00 ± 0.50
Opening_Low	1.06 ± 0.04	1.34 ± 0.10	1.27 ± 0.10	1.38 ± 0.11	1.40 ± 0.10	1.16 ± 0.06	1.29 ± 0.04	1.00 ± 0.00
Closing_Low	2.45 ± 0.15	2.41 ± 0.16	2.58 ± 0.18	2.46 ± 0.17	2.62 ± 0.16	2.38 ± 0.17	2.51 ± 0.07	3.17 ± 0.73
Listening_Low	3.11 ± 0.14	3.36 ± 0.15	3.10 ± 0.16	3.00 ± 0.17	3.59 ± 0.13	3.05 ± 0.13	3.25 ± 0.07	2.67 ± 0.33
Relation_Low	2.21 ± 0.12	2.73 ± 0.14	2.70 ± 0.14	2.45 ± 0.14	2.84 ± 0.14	2.26 ± 0.09	2.60 ± 0.06	2.83 ± 0.17
Teamwork_Low	1.83 ± 0.13	2.11 ± 0.13	2.32 ± 0.12	2.27 ± 0.15	2.28 ± 0.13	2.03 ± 0.09	2.16 ± 0.06	2.00 ± 0.29
Analysis_Low	2.94 ± 0.12	3.03 ± 0.12	3.05 ± 0.13	3.19 ± 0.14	3.06 ± 0.13	2.91 ± 0.13	3.05 ± 0.06	2.50 ± 0.29
Design_Low	3.28 ± 0.11	3.14 ± 0.12	3.32 ± 0.11	3.44 ± 0.13	3.71 ± 0.12	3.41 ± 0.12	3.39 ± 0.06	3.50 ± 0.29
Visual_Low	3.21 ± 0.13	3.26 ± 0.13	3.41 ± 0.15	3.50 ± 0.13	3.47 ± 0.13	3.16 ± 0.13	3.38 ± 0.06	3.50 ± 0.29
Opening_High	4.47 ± 0.10	4.36 ± 0.12	4.31 ± 0.11	4.24 ± 0.14	4.46 ± 0.11	4.14 ± 0.12	4.38 ± 0.05	3.75 ± 0.25
Closing_High	4.13 ± 0.12	4.13 ± 0.13	3.97 ± 0.13	3.91 ± 0.14	4.18 ± 0.10	4.03 ± 0.11	4.07 ± 0.05	4.25 ± 0.25
Listening_High	4.70 ± 0.09	4.56 ± 0.08	4.34 ± 0.09	4.14 ± 0.14	4.73 ± 0.07	4.33 ± 0.08	4.51 ± 0.04	4.50 ± 0.50
Relation_High	4.66 ± 0.07	4.45 ± 0.08	4.34 ± 0.11	4.53 ± 0.11	4.78 ± 0.05	4.47 ± 0.09	4.57 ± 0.04	4.25 ± 0.75
Teamwork_High	4.52 ± 0.09	4.50 ± 0.09	4.50 ± 0.09	4.12 ± 0.14	4.79 ± 0.06	4.21 ± 0.10	4.50 ± 0.04	4.00 ± 0.50
Analysis_High	4.34 ± 0.11	4.02 ± 0.10	4.03 ± 0.10	3.90 ± 0.13	4.28 ± 0.09	3.93 ± 0.92	4.13 ± 0.05	3.50 ± 0.50
Design_High	4.58 ± 0.07	4.25 ± 0.10	4.27 ± 0.09	4.23 ± 0.11	4.60 ± 0.07	4.34 ± 0.08	4.40 ± 0.04	4.25 ± 0.25
Visual_High	4.31 ± 0.12	3.91 ± 0.10	4.05 ± 0.14	3.95 ± 0.13	4.39 ± 0.12	4.00 ± 0.11	4.14 ± 0.06	4.25 ± 0.75
Note: All = All Skills; Soft = Soft Skills; Tech = Technical Skills; Low = Low-Quality Interview; High = High-Quality Interview								

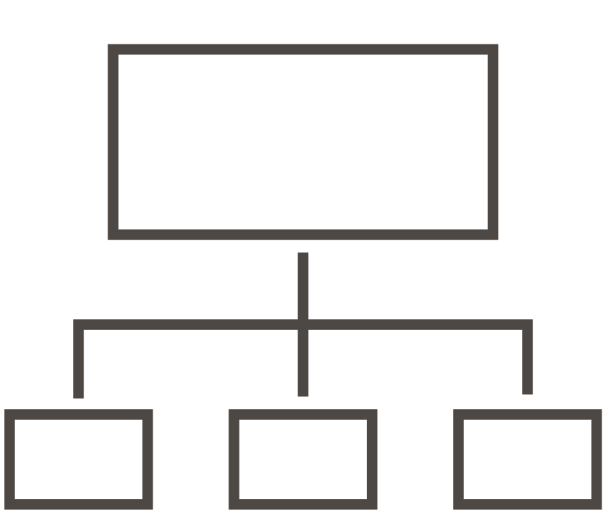
Appendix B. Statistical Parameters for the Analyzed Piecewise Models

Evaluated Skills	Statistical Parameters	Regression Equation
Average Across All Skills		
All skills: Low-quality interview	$F(3,291) = 1191.73$, $p < .0001$; $R^2 = .95$	All_Low = 2.82 Int1 + 3.19 Int2 – 0.32 Time2
All skills: High-quality interview	$F(4,290) = 5032.79$, $p < .0001$; $R^2 = .99$	All_High = 3.99 Int1 + 4.83 Int2 – 0.10 Time1 – 0.33 Time2 + 0.14 Gender
Soft vs. Technical Skills		
Soft skills: Low-quality interview	$F(3,291) = 750.65$, $p < .0001$; $R^2 = .93$	Soft_Low = 2.45 Int1 + 2.91 Int2 – 0.37 Time2
Soft skills: High-quality interview	$F(4,290) = 4563.18$, $p < .0001$; $R^2 = .99$	Soft_High = 4.07 Int1+ 4.91 Int2 – 0.10 Time1 – 0.34 Time2
Tech skills: Low-quality interview	$F(2,292) = 1239.83$, $p < .0001$; $R^2 = .96$	Tech_Low = 3.42 Int1 + 3.65 Int2
Tech skills: High-quality interview	$F(5,289) = 3318.99$, $p < .0001$; $R^2 = .98$	Tech_High = 3.85 Int1 + 4.69 Int2 – 0.11 Time1 – 0.31 Time2 + 0.19 Gender
Individual Criteria in the Low-Quality Interview		
Opening	$F(4,290) = 246.12$, $p < .0001$; $R^2 = .81$	Opening_Low = 1.51 Int1 + 1.67 Int2 + 0.09 Time1 – 0.25 Time2
Closing	$F(2,292) = 269.63$, $p < .0001$; $R^2 = .82$	Closing_Low = 2.54 Int1 + 2.87 Int2
Listening	$F(3,291) = 596.71$, $p < .0001$; $R^2 = .91$	Listening_Low = 2.98 Int1 + 4.08 Int2 – 0.52 Time2
Relation	$F(3,291) = 450.72$, $p < .0001$; $R^2 = .89$	Relation_Low = 2.71 Int1 + 3.42 Int2 – 0.58 Time2
Teamwork	$F(3,291) = 349.08$, $p < .0001$; $R^2 = .86$	Teamwork_Low = 2.52 Int1 + 2.52 Int2 + 0.15 Time1
Analysis	$F(2,292) = 669.89$, $p < .0001$; $R^2 = .92$	Analysis_Low = 3.22 Int1 + 3.18 Int2
Design	$F(2,292) = 932.56$, $p < .0001$; $R^2 = .94$	Design_Low = 3.46 Int1 + 4.00 Int2
Visual	$F(2,292) = 750.80$, $p < .0001$; $R^2 = .93$	Visual_Low = 3.58 Int1 + 3.76 Int2
Individual Criteria in the High-Quality Interview		
Opening	$F(3,291) = 1653.95$, $p < .0001$; $R^2 = .97$	Opening_High = 4.16 Int1 + 4.76 Int2 – 0.32 Time2
Closing	$F(2,292) = 1415.65$, $p < .0001$; $R^2 = .96$	Closing_High = 3.82 Int1 + 4.40 Int2
Listening	$F(5,289) = 3072.35$, $p < .0001$; $R^2 = .98$	Listening_High = 3.92 Int1 + 5.06 Int2 – 0.18 Time1 – 0.38 Time2 + 0.24 Gender
Relation	$F(3,291) = 3521.94$, $p < .0001$; $R^2 = .98$	Relation_High = 4.36 Int1 + 5.07 Int2 – 0.31 Time2
Teamwork	$F(4,290) = 2604.06$, $p < .0001$; $R^2 = .98$	Teamwork_High = 4.09 Int1 + 5.36 Int2 – 0.12 Time1 – 0.57 Time2
Analysis	$F(3,291) = 1925.81$, $p < .0001$; $R^2 = .97$	Analysis_High = 3.72 Int1 + 4.58 Int2 – 0.13 Time1 – 0.33 Time2
Design	$F(4,290) = 3097.82$, $p < .0001$; $R^2 = .98$	Design_High = 4.05 Int1 + 4.81 Int2 – 0.10 Time1 – 0.24 Time2
Visual	$F(4,290) = 1477.29$, $p < .0001$; $R^2 = .96$	Visual_High = 3.77 Int1 + 4.69 Int2 – 0.36 Time2 + 0.27 Gender
Note: See the regression equation in section 3.4.1 of the paper for coding of the variables Int1, Int2, Time1, and Time2		

Appendix C. Requirements Elicitation Interview Assessment Rubric (REIAR) (Adopted with permission from Ezell et al., 2019)

	Beginner 1	Developing 2	Competent 3	Excellent 4	Outstanding 5
Relationship Building Appropriate greeting (stand up, shake hands, introduce self, ask how the other is doing), eye contact, attentive, positive affirmation.	Interaction marred by one or more of the following: rude or condescending behavior, chronic lack of eye contact, chronic checking of phone, showing an overall lack of attention or interest.	Demonstrates some aspects of competent relationship building but may be inconsistent (e.g., inconsistent eye contact or short periods of inattention).	Appropriate greeting. Questioner engages in appropriate eye contact. Displays positive affirmation.	Meets criteria for *Competent* AND is natural or smooth. Positive body language.	Meets criteria for *Excellent* AND there is a sense of an extraordinary professional relationship.
Opening Provide an organizational frame for the client, agenda, purpose; goals to accomplish in the interview.	Provides no initial organizational frame for the client. At this level, student typically begins interaction by launching into specific questions.	Provides some frame (e.g., starts out with some organizational sentences). May stay too broad (e.g., "we are here to do requirements elicitation for your project") or provide some, but not all, of agenda, purpose, goals to accomplish.	Provides a complete organizational frame for the interview (agenda, purpose, goals to accomplish).	Meets criteria for *Competent* AND asks questions to determine type of client AND gets confirmation of frame from client AND adjusts accordingly.	Meets criteria for *Excellent* AND delivers it smoothly. "Clear", "compelling", "engaging" are the words that come to mind.
Active Listening Pay attention, provide feedback, summarize or paraphrase ideas, remember past answers, ask for appropriate clarification.	Demonstrates minimal active listening techniques. E.g., a questioner focused on questioning rather than on answers; or asking rapid questions without regard to prior conversation. May not listen to answers or talk over answers.	Demonstrates some active listening techniques. Questions and answers are marred by some of the following: double-barreled questions, allowing client to not answer questions, asking questions that have already been answered, forcing client to give opinion when the client does not know an answer.	Uses active listening techniques (feedback, recaps, clarifications). Makes sure questions are answered, questions build on prior answers.	Meets criteria for *Competent* AND confirms understanding of the answer. Flexible in questions asked by adapting discussion dynamically based on understanding client's responses.	Meets criteria for *Excellent* AND asks questions deliberately to gauge client type and gears entire style toward the client. Checks in frequently to ascertain common understanding.
Analyzing Current (As-Is) System Understand the current situation (e.g., process, system, data, artifact). Inquire what is good and what is bad about the current situation, process, system, or artifacts as	No attempt to investigate the current situation. At this level, the student often starts by asking what the client wants; not what exists now.	Articulates the current situation. May be disorganized or out of context.	Mutual communication about the current situation. Asks what is good and what is bad about the current situation.	Meets criteria for *Competent* AND adds mutual discovery that *assists* the discussion.	Meets criteria for *Excellent* AND uses visualization to guide the discussion. Examples of this may include an interactive exploration of the topic, mutual

appropriate.					discovery, or an iterative process.
Designing Proposed (To-Be) System Discuss the design of proposed (To-Be) system with the client as part of the interview.	No attempt to include the client in the design.	Asks client about the To-Be system using primarily closed-ended questions OR tells client what improvements will be and asks for opinion.	Works with client to design To-Be system. Team and client work out design together. Uses open-ended questions and an interactive process.	Meets criteria for *Competent* AND client and team design together with appropriate mutual visualization, mutual discovery, and iteration.	Meets criteria for *Excellent* AND iteration is adaptive, probing, and explorative, with value added in each iteration. Keeps in mind the scope of the project or phase.
Visualization Use appropriate and applicable visuals (process models, functional models, structural models, interface structure, mock-ups, as-is or to-be reports, visual mapping, etc.) to aid relevant aspects of meeting. Use visuals to understand scope. Effectively integrate visuals into discussion.	Does not use visuals. Does not have or request a copy of current reports, screens.	Uses visuals that do not assist in discovering the requirements OR do not reflect client input in visuals. May refer to current artifacts or to-be artifacts.	Uses visuals to guide discovery of requirements.	Meets criteria for *Competent* AND uses draft or template visuals to guide relevant aspects of meeting. Client's input leads to a dynamic development of visuals during meeting.	Meets criteria for *Excellent* AND drawings are visible to all and all are welcome to contribute. Examples of this may include a mutual exploration of the topic, mutual discovery, or an iterative process.
Team Work To the client, the team appears natural and appropriate. Roles and responsibilities (e.g., questioner and note taker) appear natural (roles may shift over interview and not each team member needs to ask a question). Team members provide different points of view, leader keeps team on track, and inter-team communication aids elicitation.	Each team member is operating on their own. May demonstrate visible dysfunction. Team members do not listen to each other.	Duties separated, with team members having different roles OR team listens to each other and works together well BUT not both.	Each team member has a role that they explain to the client. Roles are then demonstrated over the interview. Team listens to each other and works together well.	Meets the requirements for *Competent* AND team members refer to each other and add to what each other says in an appropriate way. Roles feel organic and natural.	Meets criteria for *Excellent* AND whole team performance feels strategic. Group synergy is better than sum of the individuals. The group develops and designs together, sharing different points of view.
Closing Recap, plan next step, ask final questions.	Ends interview when done with questions.	Attempts a closing but marred by one of the following: excessively long recap, closing focuses on the relational aspects and not the substance of the interview, closing focused on the agenda not the findings.	Recap of key points is on track and generally at the right level. Asks if any important issues were not discussed. Outlines future steps.	Meets criteria for *Competent* AND recap includes the ways requirements fit into the scope of project or project phase.	Meets criteria for *Excellent* AND uses artifacts created in the interview to guide the closing.



Information Systems & Computing Academic Professionals

Education Special Interest Group

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the ***Journal of Information Systems Education*** have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 2574-3872