

Abstract

Many universities shifted how students were assessed during the COVID-19 pandemic. This movement to online learning altered the format of some assessments that were previously administered in-person and proctored. Since the start of COVID-19 in 2020, James Madison University (JMU) shifted some assessments to an unproctored internet testing (UIT) format. The bi-annual, university-wide Assessment Day was one such set of assessments that underwent the change to UIT at JMU. As we interpret scores from those UIT administrations and contemplate future changes, it is important to understand what the experience was like for the students. At the end of their battery of assessments, students were asked to share their thoughts and suggestions. The current study employed a conventional content analysis to code responses to this item for two recent Assessment Days. About 20% of students responded to the item, of which many of the comments were generally positive and said something positive about UIT specifically. Few comments were negative. This study highlights the positive impact of UIT on our campus. We aim to continue incorporating the student perspective into our assessment process.



AUTHORS

Katarina E. Schaefer, M.A.
James Madison University

Dena A. Pastor, Ph.D.
James Madison University

Samantha N. Harmon, M.A.
James Madison University

Examinee Perspectives on Unproctored Internet Testing

CCOVID-19 prompted assessment professionals in higher education to make quick decisions; decisions that would typically take months or years to finalize happened in a matter of weeks. Moreover, rather than making a single change, many universities and other higher education programs had to make multiple, drastic changes or completely reconstruct their traditional instruction and assessment processes altogether. Unprecedented times became the norm. Faculty began teaching online with varying comfort levels, assignments and assessments were modified, deadlines were extended, alternative assessments were assigned, previously proctored assessments were unproctored, and empathy toward students increased (Jankowski, 2020a, 2020b; Pastor & Love, 2020). Large-scale university admissions testing began to offer at-home testing, which had been an impossible thought only weeks before (Camara, 2020). Looking back on the 2020-2021 school year, we wonder: how do we interpret the data that came from such a hectic time?

Assessment professionals voiced their worry about the shift to online instruction and testing. Many worried that cheating would increase, students would be less motivated academically, or students would perform poorly due to increased anxiety (Jankowski, 2020b). Some speculated that the validity of the results of assessments administered after this quick transition would be lowered. However, Fulcher and Leventhal (2020) and Busby (2020) stressed that testing can and should go on despite these fears. They emphasized that it is still important to track student knowledge. Without continued testing, we would be unable to understand whether students gain, maintain, or lose knowledge due to the drastic changes that have taken place since COVID-19 began. Continued testing also

CORRESPONDENCE

Email
schae2ke@jmu.edu

provides the opportunity to explore the effects of the pandemic on the validity of assessment scores and whether the potential effects are the same for all students.

Without continued testing, we would be unable to understand whether students gain, maintain, or lose knowledge due to the drastic changes that have taken place since COVID-19 began.

Recognizing the benefits of continued assessment during the pandemic, many higher education academic and non-academic programs alike made changes to their assessment procedures. Whether assessment was administered for academic degree programs, student affairs programs, campus initiatives, etc., many institutions chose to administer assessments in a new, UIT format. At James Madison University (JMU), one such set of assessments that shifted to UIT was the biannual, low-stakes Assessment Day (Pastor & Love, 2020). These assessments have been administered both proctored and in-person for over 30 years. About 4,000 students are assessed on a typical Assessment Day throughout two three-hour sessions. Assessments are “low stakes” for students because they have no direct personal consequences to the student. Although university-wide Assessment Days are not common, smaller scale assessment of student learning is routine in higher education. Similar to other higher education programs navigating the pandemic, our university knew that these assessments could not be administered in-person and proctored.

All assessments were administered remotely and unproctored during the Fall 2020 and Spring 2021 Assessment Days to reduce exposure to COVID-19, a change from previous years. Test length and content during these administrations was identical to in-person administrations from previous years. Given the numerous differences in administration format (e.g., remote vs. in-person; proctored vs. unproctored) and context (pre-COVID-19 vs. COVID-19), we anticipated that the results from Fall 2020 and Spring 2021 would differ from those in the past. Indeed, initial reports of first-year students’ scores from Fall 2020 Assessment Day reveal that overall, students seem to have much more varied scores compared to in-person administrations (Alahamadi & DeMars, 2021). The tests considered in the study included a history test, a global issues test, and a test of scientific reasoning. On the scientific reasoning assessment, which contains more items and is more cognitively demanding than the other tests, Alahamadi & DeMars (2021) reported that first-year students did much worse than expected in Fall 2020 (during COVID-19) compared to the four previous years’ students.

Looking only at the numbers, we know scores were affected for at least some students, with a more pronounced effect for one assessment. However, though we might speculate how students were affected given the data, the only people who know the entire story are the students who experienced those assessments.

What students thought about the use of UIT for Assessment Day was particularly important because it was a considerable departure from the norm and results from previous studies were mixed.

Before COVID-19, higher education assessment professionals had already considered what it would mean for their programs to integrate the student perspective into their practice. The leading voices of diversity, equity, and inclusion in assessment have emphasized that the student perspective must be considered (Jankowski, 2020a, 2020b; Montenegro & Jankowski, 2020). Jankowski (2020a, 2020b) has emphasized that it is even more important to consider the student perspective during these unprecedented times. Such calls motivated us to obtain the students’ perspective on their assessment experience in general and, more specifically, their take on the remote administration format. What students thought about the use of UIT for Assessment Day was particularly important because it was a considerable departure from the norm and results from previous studies were mixed. Some research shows that students have generally had positive online testing experiences (e.g. Milone et al., 2017), although some report negative experiences with proctors in online testing (Karim et al., 2014). At JMU, assessments were unproctored so we expected students to have a generally positive experience, but did not know for sure. We also did not know how COVID-19 would affect their experience without asking them – so we did.

Method

Procedures & Sample

Data were collected during the Fall 2020 and Spring 2021 Assessment Days, which were forced to use UIT due to COVID-19. Incoming first-year students were assessed in Fall 2020, and sophomore students¹ with 45-70 credit hours were assessed in Spring 2021.

¹Although some students in this credit hour range are juniors, we refer to students who completed the spring assessment as sophomores throughout this article.

Students were assigned to a battery of online assessments during both Assessment Days. In a video describing the content of the last assessment, examinees were informed that the testing format differed from the typical in-person, proctored experience. Additionally, examinees were told they would be asked to describe their Assessment Day experience and provide suggestions for improvement during the last assessment. Examinees were asked to respond to the following questions at the end of the assessment: “Want to tell us about your Assessment Day requirement experience? Have suggestions for how to improve the Assessment Day requirement? If so, please share your experience and/or suggestions below.” The item did not inquire about UIT specifically to avoid leading students to mention something about UIT.

A little less than 20% of examinees (including first-years² and sophomores³) responded to the question, yielding 1,421 responses. The first-year and sophomore samples were 63% female and 77% White, with all other races and ethnicities representing less than 10% of the sample. These distributions align with those for undergraduates at the university overall during the 2020-2021 academic year (58% female, 75% White).

Analysis

Meaning was extracted from the responses through a conventional content analysis (Hsieh & Shannon, 2005), which is appropriate when the goal is to allow themes to emerge from the data. Although we anticipated some generally positive responses, we did not want to constrain the categorization of responses to our preconceived notions; instead, we wanted to allow themes to “flow from the data” (Hsieh & Shannon, 2005, p. 1279). Two authors read two different sets of comments from 50 randomly selected first-year students and separately created initial codes to begin the analysis. After discussing the initial codes, the final set of codes and their descriptions were created. Example responses for each code were identified along with responses for training purposes. The remaining author and two additional raters were trained to use the codes, with each of the five raters assigned an equal number of responses. Although not ideal, to make the workload manageable, all first-year responses were rated first (before collecting the sophomore data) and all sophomore responses were rated second. Thus, raters were aware of the class level of the students during rating.

All four raters independently coded 100 of the same randomly selected first-year student comments to compute intercoder reliability. O’Connor and Joffe (2020) report that all raters typically code between 10%-25% of the same data to estimate intercoder reliability. For this study, all raters independently coded roughly 7% of the same comments. After all responses were coded, the responses associated with each code were reviewed by the study authors, resulting in the creation of subcategories and the merging of two initial codes. The number of responses classified into each code and subcategory was then tallied.

Table 1 contains the codes, representative examples of text to describe each code, and two indices of intercoder reliability calculated using all four raters. These code descriptions were used to train all raters. Although “Assessment Content” and “Assessment Format” were merged during the review process, estimates of intercoder reliability were calculated separately for these codes. In addition to the percent agreement index, Gwet’s AC1 is provided. Gwet’s AC1 differs from percent agreement because it corrects for chance agreement and is preferable to many alternatives (Gwet, 2014). Intercoder reliability indices were favorable, with values $> .92$ for 11 of the 13 codes.

²The responses from only those examinees who completed testing by the extended deadline were used in this study. Of the 3,847 incoming first-years required to participate in Assessment Day and assigned to complete the assessment used in this study, 3,408 completed the assessment on which the item was administered by the extended deadline. Out of these 3,408 examinees, 718 provided responses to the item. Thus, 21% (718/3408) of the incoming first-years who completed the assessment by the extended deadline provided a response.

³The responses from only those examinees who completed testing by the Assessment Day deadline were used in this study. Out of the 3,524 sophomores required to participate in ADay, 3,174 completed the assessment on which the item was administered by the deadline. Of these 3,174 examinees, 703 provided responses to the item. Thus, 22% (703/3174) of the sophomores who completed the assessment by the deadline provided a response.

Table 1
Codes, Definitions, and Average Intercoder Reliability Estimates

Code	Representative Examples of Text	Gwet's AC ₁	%	
Positive	Overall liked their experience. Felt excited to start school. Felt like they knew what to anticipate. Had a predominately good experience. Didn't feel overwhelmed.	0.71	85.2%	
Neutral	Overall didn't have strong feelings one way or another about the assessment. Felt like they had enough time to complete it. Said "It was ok." It was uneventful.	0.82	87.0%	
Negative	Overall didn't like something about ADay. Had a mainly bad experience. Their assessments took too long. Felt overwhelmed. They didn't care about this.	0.93	94.2%	
Online Positive	Liked the online format. Liked that they could spend as much time as they wanted on the assessments. Didn't feel overwhelmed specifically because it wasn't in person. They don't have to explicitly mention the online or remote format.	0.92	94.8%	
Online Negative	Didn't like the online format. Would rather be in person.	0.98	98.2%	
Communication/ A-Day Purpose	Felt like they didn't receive enough information about ADay. Something they said could have been changed if they'd read the emails/received more emails. Would like to know more about why it's important. Would want to know why they should feel motivated to do the assessments.	0.96	96.2%	
Low Motivation	Didn't feel motivated to do well. Didn't try their hardest.	0.96	96.2%	
Stressed	Said they had a lot going on at the time. This added a lot to their plate. They were dealing with lots of stress (COVID-19 related or not).	0.98	98.0%	
Performance concern	Don't think they did well. They think something affected their performance today. Didn't feel prepared. They want to know their scores.	0.94	94.7%	
Assessment Format/Content	Comment on a specific aspect of the test. Offers suggestion to the format. Said something was too long. Wished there were less of a type of question (multiple choice, short answer, etc.). Comment on the content of the text related to how questions were asked, what questions were asked, or the difficulty. Mentioned grammar or spelling mistakes.	Format	0.92	92.8%
		Content	0.97	97.8%
Flag	Student brings up something concerning.	0.96	96.2%	
Other	Noteworthy information in response not captured by other codes.	0.92	94.8%	

Note. % = percent agreement.

The percentage of comments classified as conveying something positive about UIT was 15.6% and 28.7% for first-year and sophomore students, respectively.

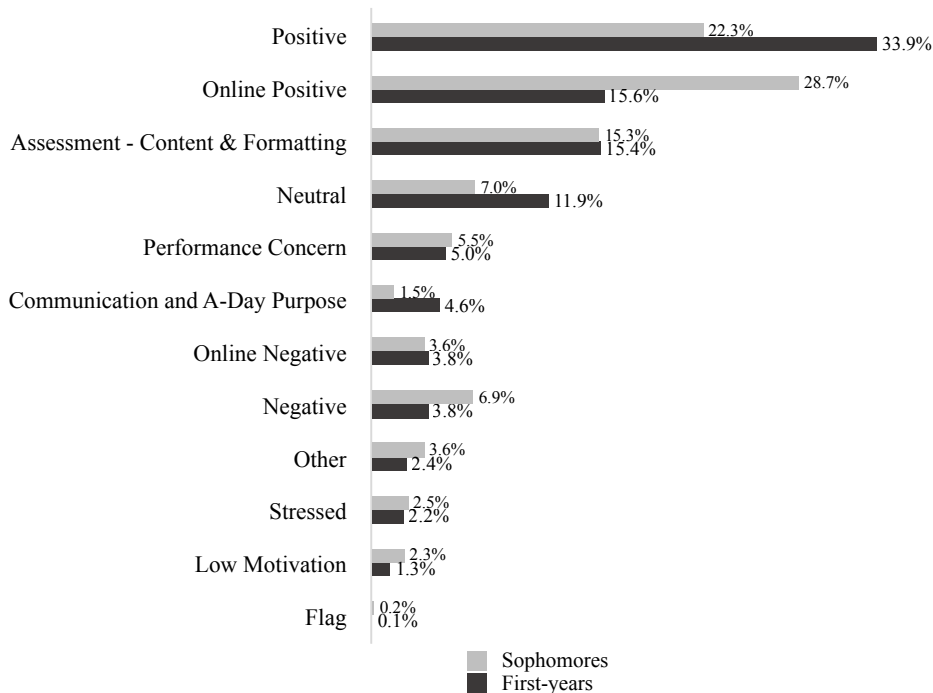
Although all codes and subcategories are informative, the most relevant for understanding students' experiences with UIT is "Online Positive" and "Online Negative." For this reason, we begin by considering the percentage of comments classified according to these two codes and whether these percentages differed across first-year and sophomore students. Additionally, general "Positive" and "Negative" codes are discussed for comparison. We then consider the subcategories of "Online Positive" and "Online Negative" to better understand students' specific positive and negative comments regarding UIT.

The percentage of comments classified by code is shown in Figure 1 separately for first-year and sophomore students. Comments classified as "Online Positive" said something generally positive concerning the remote testing format. For example, a comment that was coded as "Online Positive" for a sophomore student stated, "Taking the Assessment Day requirement remotely was stress free and more impactful." The percentage of comments classified as conveying something positive about UIT was 15.6% and 28.7% for first-year and sophomore students, respectively. The only other code capturing a larger percentage of student responses was the "Positive" code, which captured general positive comments about the testing experience (not necessarily related to UIT). A comment that was coded as "Positive" for one sophomore student stated, "It went well." In contrast, very few comments (approximately 4%) were classified as "Online Negative" across first-year and sophomore

student comments. Comments classified as “Online Negative” mentioned feeling displeased with a remote Assessment Day or mentioned that Assessment Day should be in-person in the future. For example, one sophomore student said, “Having Assessment Day online is not good. in person [sic] is better.” Additionally, only 3.8% of first-years and 6.9% of sophomores said something that fell into the general “Negative” code. These students said something generally negative about Assessment Day (not necessarily related to UIT). For example, one first-year student said, “It was boring and tedious, there are much better ways to spend time on campus, like studying our courses or making friends, rather than sitting in a room answering an assessment survey.”

In contrast, very few comments (approximately 4%) were classified as “Online Negative” across first-year and sophomore student comments.

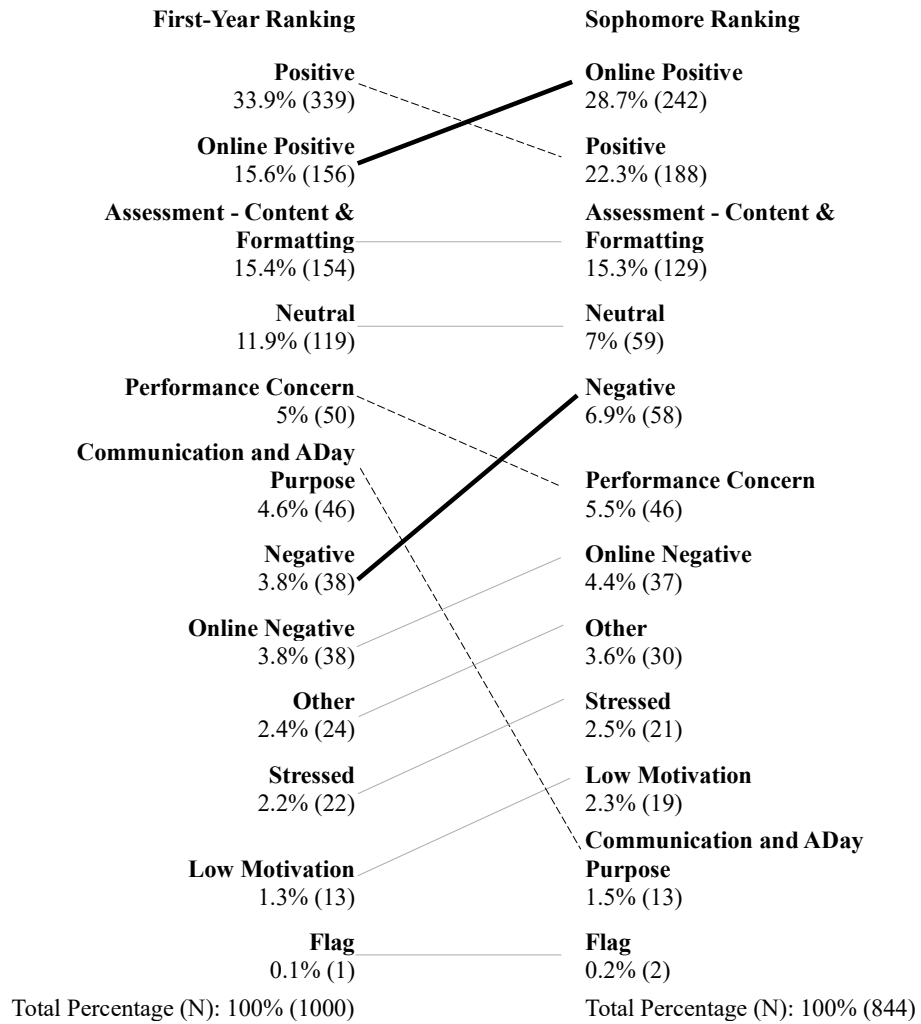
Figure 1.
Percentage of Comments Classified by Code for First-year and Sophomore Students



Although most comments were classified into each code at similar rates for first-year and sophomore students, there were some differences. Figure 2 presents the rank-ordering of codes separately for first-year and sophomore students according to the percentage of comments classified by each code. Lines are provided within the Figure to illustrate differences in rank-ordering of codes between the two groups. Notable differences include the rank ordering of the “Positive” and “Online Positive” codes. A larger percentage of comments were classified as “Online Positive” for sophomore students (28.8%) relative to first-year students (15.6%), while “Positive” was higher for first-year students (33.9%) relative to sophomore students (22.3%). Another notable difference was the rank ordering of “Negative.” This code was higher for sophomore students (6.9%) compared to first-year students (3.8%). The rank-order and percentage of comments classified as “Online Negative” stayed relatively similar for the two groups.

Table 2 contains the subcategories for comments coded “Online Positive” for both first-year and sophomore students. Only subcategories that were larger than 10% are listed. Recall that out of the total comments, about 16% (156) of first-year student comments and about 29% (242) of sophomore student comments contained text that was classified as “Online Positive.” Across both groups, most students fell into the top two subcategories - “online eases stress/anxiety” and “ease of online use.” Students who reported “online eases stress/anxiety” said something about the online aspect of testing that helped them feel less stressed or anxious. Students who reported “ease of online use” mentioned how testing was easy to do online. For example, one first-year student said, “The virtual Assessment Day test ran smoothly and I enjoyed being able to complete this task on my own time. In addition, the questions were

Figure 2
Ranking of Codes by Group



Note. Codes are rank ordered for the first-year student group and sophomore group separately.

clear and easy to understand. Based on my experience, I would recommend Assessment Day to be virtual permanently.” Of the “Online Positive” comments, a higher proportion of first-year students (27.4%) compared to sophomore students (20.4%) fell into “online eases stress/anxiety” over “ease of online use.” In contrast, a higher proportion of sophomore students (27.6%) compared to first-year students (19.0%) fell into “ease of online use” over “online eases stress/anxiety.” Both first-year students (14.2%) and sophomore students (18.1%) appreciated having extended time to complete their assessments and/or complete these assessments independently. First-year students (11.1%) and sophomore students (15.8%) said they would prefer online to in-person testing.

Table 2
Online Positive subcategories

Subcategory	First-Years			Sophomores		
	Count	%	Rank	Count	%	Rank
Being online eases stress/anxiety	62	27.4%	1	62	20.4%	2
Ease of online use	43	19.0%	2	84	27.6%	1
Extended time/on their own time	32	14.2%	3	55	18.1%	3
Prefer online to in-person	25	11.1%	4	48	15.8%	4

Note. Some comments fell into more than one subcategory.

Table 3 contains the subcategories for text coded “Online Negative” for both first-year and sophomore students. Only subcategories that were larger than 10% are listed. Recall that out of the total comments, only about 4% (38) of first-year student comments and about 4% (37) of sophomore student comments contained text that was classified as “Online Negative.” The majority of comments within this code fell into four subcategories: “prefer in-person,” “had an issue with the test,” “ability to focus,” and “motivation issues.” For example, one first-year student said: “I know that Covid [sic] has had a big impact on the way assessments are taken, however, I feel that the online environment is not a great way to conduct the assessments because it is much easier to just skip through and not put forth your best effort. I had trouble remaining focused and motivated to complete the assessment. I feel that being in person would have been better.” Of the “Online Negative” comments, the largest proportion of first-year student comments in this category said that assessments should be conducted in-person instead of online (27.7%), followed by comments that described testing issues completing the assessments online (19.1%). These subcategories were ranked differently for sophomores. The largest percentage of sophomores (31.7%) felt they had issues focusing with assessments being conducted online, followed by those saying they struggled with motivation to complete the assessment due to the online versus in-person administration (24.4%).

Of the roughly 20% of students who elected to provide feedback on their assessment experience, substantially more students said something positive about the online administration format than something negative.

Table 3
Online Negative subcategories

Subcategory	First-Years			Sophomores		
	Count	%	Rank	Count	%	Rank
Prefer in-person	13	27.7%	1	8	19.5%	3
Had an issue with the test	9	19.1%	2	6	14.6%	4
Ability to focus	6	12.8%	3	13	31.7%	1
Motivation issues	6	12.8%	4	10	24.4%	2

Note. Some comments fell into more than one subcategory.

Discussion

Of the roughly 20% of students who elected to provide feedback on their assessment experience, substantially more students said something positive about the online administration format than something negative. Specifically, almost 30% of sophomore student comments and 16% of first-year student comments conveyed something positive about UIT. We were encouraged to see the large number of positive statements surrounding UIT, particularly because students were not explicitly asked to address the online administration format in their feedback. Additionally, we were satisfied with the amount of generally positive comments we received about Assessment Day from first-year (34%) and sophomore (22%) students coupled with the low amount of generally negative feedback from first-year (4%) and sophomore (7%) students.

Further inspecting the responses coded “Online Positive” revealed several reasons for students’ favorable attitudes toward UIT. Both first-year and sophomore students said they had lower stress or anxiety due to UIT. Additionally, both first-year and sophomore students cited the ease of the online assessments as a positive aspect of UIT. These subcategories are meaningful because they represent a substantial number of students. The few students who commented negatively about the online experience cited technical difficulties, difficulties focusing on the test, or trouble feeling motivated. It is essential to understand why students had negative comments about the online format. While positive comments may support the continued use of this type of remote testing, negative comments identify areas for improvement to the assessment process. Still, we must keep in mind that few students provided negative comments about UIT or general negative comments, and the reasons for those comments were not unanimous.

When reflecting on the comments about online testing we received, it is also worthy to note the kinds of comments we did not receive. No student mentioned trouble with internet connectivity or lack of access to a device to use for testing. The lack of such comments might also be specific to our university and a function of the characteristics of our students (e.g., socioeconomic status) and campus (e.g., availability of on-campus testing lab). Additionally, a lack of technology or technology issues may have prevented students from completing their assessments altogether. Typically students are required to complete their Assessment Day tests; however, due to COVID-19, students who had not completed their tests by the final deadline were not forced to complete them. However, noncompletion may not be a major issue because most students (around 90% for both groups) completed their assessments.

Because the codes were developed for the first-year student group, we considered whether the number of comments in each category and subcategory differed between the two groups. We felt that it was important to do so because the context for the two administrations differed and because previous research indicates test-taking motivation differs by class level (Pastor et al., 2019; Thelk et al., 2009; Wise, 2006; Wise & DeMars, 2010). In general, the results were similar between the two groups: both first-year and sophomore students reported more positive than negative comments about UIT. A noteworthy difference was that although positive UIT comments outnumbered negative UIT in both groups, the proportion of sophomore students that specifically cited UIT as a positive experience was higher than the proportion of first-year students. In other words, sophomore students particularly liked the online format, more so than first-year students. A critical difference between the two groups is that sophomore students had previously experienced an in-person Assessment Day. Contrary to first-year students, sophomores were able to compare in-person Assessment Day to remote Assessment Day. This difference could be why they seem more likely to cite UIT as the reason for their positive Assessment Day experience.

There were several limitations to this study. As mentioned previously, the codes were created using only first-year student comments so the raters were aware of the year of the students. This process is not ideal because raters’ coding may be biased by knowing the year of the student. However, splitting comments this way eased the weight of creating codes for all the comments at one time. Another limitation is that only 20% of all students responded to the open-ended item. Those who chose to respond may have had a different perspective than those who chose not to respond in ways that limit our ability to generalize these results to all students who participated in Assessment Day. Second, social desirability may or may not have been a factor in these results. Although some students may have provided less than genuine positive responses in an attempt to “look good,” the number is likely small because the assessments were low-stakes for students and answering the question was optional. Additionally, Caputo (2017) noted that social desirability might account for less than 10% of the variance in self-report measures. For that reason, we are not too worried about social desirability in this study. Finally, these results may be specific to our university, our students, and our UIT procedures. The generalizability of these results may be limited due to these settings.

Despite the limitations, these results are encouraging for the continued use of UIT for assessment. Although this study focuses on comments pertaining to UIT, the collection of comments will help us understand what the experience was like for students and inform improvements to future assessment in higher education. The act of asking and sifting through

Although this study focuses on comments pertaining to UIT, the collection of comments will help us understand what the experience was like for students and inform improvements to future assessment in higher education.

the responses brought us closer to the students, allowing us to see things from their point of view – a perspective we value but do not always actively seek out. Decisions about UIT use may still need to consider the effects of COVID-19, cost, accessibility, and the quality of the data (Jankowski, 2020b; Montenegro & Jankowski, 2020). Further, decisions of its continued use should also weigh the student's perspective. Incorporating the student perspective at our institution revealed a positive experience, which others have cited as a vital aspect of a high-quality UIT program (Beatty et al., 2009). This finding is encouraging for our institution and others who would like to use UIT. We aim to continue incorporating the student perspective into the assessment process to ensure that UIT continues to facilitate a positive experience for all students at JMU.

AUTHORS NOTE

Special thanks are given to our interns, Bree Pifer and Tanna Walters, who assisted in the long process of reading and coding hundreds of student comments. This paper would not be possible without them.

References

- Alahmadi, S. & DeMars, C. E. (2022). Large-scale assessment during a pandemic: Results from James Madison University's remote Assessment Day. *Research and Practice in Assessment*, 12(1), 5-15.
- Beatty, J. C., Dawson, C. R., Fallaw, S. S., & Kantrowitz, T. (2009). Recovering the scientist-practitioner model: How IOs should respond to unproctored internet testing. *Industrial and Organizational Psychology*, 2(1), 58-63. <https://doi.org/10.1111/j.1754-9434.2008.01109.x>
- Busby, A. K. (2020). Resilient assessment during COVID-19. *Assessment Update*, 32(6), 1-16. <https://doi.org/10.1002/au.30231>
- Camara, W. (2020). Never let a crisis go to waste: Large scale assessment and the response to COVID 19. *Educational Measurement: Issues and Practice*, 39(3), 10-18. <https://doi.org/10.1111/emip.12358>
- Caputo, A. (2017). Social desirability bias in self-reported well-being measures: Evidence from an online survey. *Universitas Psychologica*, 16(2). <https://doi.org/10.11144/Javeriana.upsy16-2.sds>
- Fulcher, K. H., & Leventhal, B. C. (2020). James Madison University: Assessing and planning during a pandemic. *Assessment Update*, 32(6), 4-5. <https://doi.org/10.1002/au.30233>
- Gwet, K. L. (2014). Handbook of inter-rater reliability, 4th edition. Gaithersburg, MD: Advanced Analytics.
- Hsieh, H.-F. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288. <https://doi.org/10.1177/1049732305276687>
- Jankowski, N. A. (2020a). Guideposts for assessment during COVID-19. *Assessment Update*, 32(4), 10-11. <https://doi.org/10.1002/au.30222>
- Jankowski, N. A. (2020b, August). *Assessment during a crisis: Responding to a global pandemic*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
- Karim, M. N., Kaminsky, S. E., & Behrend, T. S. (2014). Cheating, reactions, and performance in remotely proctored testing: An exploratory experimental study. *Journal of Business and Psychology*, 29, 555-572. <https://doi.org/10.1007/s10869-014-9343-z>
- Milone, A. S., Cortese, A. M., Balestrieri, R. L., & Pittenger, A. L. (2017). The impact of proctored online exams on the educational experience. *Curr Pharm Teach Learn*, 9(1), 108-114. <https://doi.org/10.1016/j.cptl.2016.08.037>
- Montenegro, E., & Jankowski, N. A. (2020, January). *A new decade for assessment: Embedding equity into assessment praxis* (Occasional Paper No. 42). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1-13. <https://doi.org/10.1177/1609406919899220>
- Pastor, D. A., & Love, P. (2020, Fall). University-wide assessment during Covid-19: An opportunity for innovation. *Intersection: A Journal at the Intersection of Assessment and Learning*, 2(1).
- Pastor, D. A., Ong, T. Q., & Strickman, S. N. (2019). Patterns of solution behavior across items in low-stakes assessments. *Educational Assessment*, 24(3), 189-212. <https://doi.org/10.1080/10627197.2019.1615373>
- Thelk, A. D., Sundre, D. L., Horst, S. J., & Finney, S. J. (2009). Motivation matters: Using the student opinion scale to make valid inferences about student performance. *Journal of General Education*, 58(3), 129-151. <https://doi.org/10.1353/jge.0.0047>
- Wise, S. L. (2006). An investigation of the differential effort received by items on a low-stakes computer-based test. *Applied Measurement in Education*, 19(2), 95-114. https://doi.org/10.1207/s15324818ame1902_2
- Wise, S. L., & DeMars, C. E. (2010). Examinee noneffort and the validity of program assessment results. *Educational Assessment*, 15, 27-41. <https://doi.org/10.1080/10627191003673216>