QANE AFI | Quality Advancement in Nursing Education / Avancées en formation infirmière

CASN ACESI

Vol. 12 | No. 1 (2026) Article 3

Two-Stage Exams in Undergraduate Nursing Education: A Descriptive Study on the Reduction of Barriers and Grade Inflation

Amy N. Rodall, College of New Caledonia

CC BY-NC-ND 4.0

Cover Page Footnote | Note de page couverture

This research was conducted on the traditional unceded lands of the Lheidli T'enneh people. Portions of this article were revised for clarity and conciseness using ChatGPT version 5 following the peer review process (Open AI, 2025). The author reviewed and verified all content for accuracy and originality. The author wishes to acknowledge the faculty members who assisted with participant recruitment and data collection for this study. Their contributions were invaluable to the successful completion of this project. Students in the program have provided valuable contributions to this research. | Cette recherche a été menée sur les terres traditionnelles et non cédées du peuple Lheidli T'enneh. Certaines sections de cet article ont été révisées pour en améliorer la clarté et la concision à l'aide de ChatGPT version 5, à la suite du processus d'évaluation par les pairs (OpenAI, 2025). L'auteure a examiné et vérifié l'ensemble du contenu afin d'en assurer l'exactitude et l'originalité. L'auteure souhaite remercier les membres du corps professoral qui ont contribué au recrutement des participantes et participants ainsi qu'à la collecte de données pour cette étude. Leur apport a été essentiel à la réalisation de ce projet. Les étudiantes et étudiants du programme ont également offert des contributions précieuses à cette recherche.

Background

Collaborative assessment has been researched for over a decade in many different disciplines of undergraduate education (Helms et al., 2019; Radford & Blount, 2023). Collaborative assessment, including two-stage exams, occurs when students complete a shared assessment and subsequently share the resulting grade (Saunders et al., 2019). The individual exam and group exams are identical. Students complete the group exam by reviewing questions and selecting answers as a group, then submitting one answer sheet. The individual exam prepares the student to participate in the group exam (Rieger & Heiner, 2014), allowing students to build on their individual results. Faculty can implement this exam style in lieu of traditional midterm assessments that include multiple-choice, true/false, and short-answer questions. Students typically complete the group exam in half the time of the individual exam (Rieger & Heiner, 2014). Two-stage exams are based on the social constructivist theory of teaching and learning (Radford & Blount, 2023; Witt et al., 2022). Students work and learn together in social groups in the second stage, building on knowledge from the initial stage. Social constructivist learning forms a foundational aspect of nursing education; collaborative learning opportunities help students build on foundational knowledge, and learning occurs in a community (Witt et al., 2022). Two-stage exams are relatively easy to implement in both small- and large-group settings (Gilley & Clarkston, 2014; Rieger & Heiner, 2014) and include both formative assessment and summative evaluation in their design (Radford & Blount, 2023; Rieger & Heiner, 2014). The immediate feedback and rich discourse between student groups allow for a deeper level of understanding (Gilley & Clarkston, 2014; Peck et al., 2013; Rieger & Heiner, 2014). Collaborative assessment is not only a method of evaluation but also a strategy for teaching and learning (Helms et al., 2019; Riley et al., 2021). Typically, the group exam score will be higher than any individual exam score (Nicol & Selvaretnam, 2022). The weighting of each exam in the overall grade varies greatly by instructor, which can be linked to potential grade inflation. The use of collaborative testing in nursing programs across Canada is increasing (Gilley & Clarkston, 2014); however, more research is needed to support its integration. This study aims to describe the reduction of learning barriers and the occurrence of grade inflation associated with two-stage exams among undergraduate nursing students.

Universal Design to Reduce Barriers

Individual and systemic barriers to assessment may include the physical environment, general anxiety, mental health concerns, and testing anxiety (Burgess & Medina-Smuck, 2018; Caboral-Stevens & Fox, 2020; Epstein et al., 2021; Helms et al., 2019; Martin, 2018; Radford & Blount, 2023; Sonmez et al., 2023). Universal Design for Learning (UDL) has been adopted by many post-secondary programs as a framework for providing inclusive education that meets all learners' needs by optimizing teaching practices (Kotcherlakota et al., 2024). UDL has three core principles: multiple means of engagement, multiple means of representation, and multiple means of action and expression (CAST, 2024). A UDL assessment is based on reducing barriers and anxiety while increasing student engagement, motivation, and satisfaction (Balta et al., 2021). Rates of mental health concerns such as anxiety and depression are elevated in universities across the globe, with rates among the nursing student population reported to be as high as 34% (Epstein et al., 2021; Sonmez et al., 2023). Stress and anxiety negatively affect the health and development of the nursing student's professional identity, with exam-related stress contributing (Bangcola, 2023; Sonmez et al., 2023). Inclusivity and equity can be supported using a UDL approach and two-stage exams in undergraduate nursing programs. Racialized and Indigenous nurses are vastly underrepresented in health care and nursing education; the retention of these students after entry into a nursing program is essential to the development of a diverse workforce (Ferrell & DeCrane, 2016; Murray et al., 2016). Helms et al. (2019) report that scores from individuals to groups in collaborative assessments

were proportionately higher for Black nursing students than for white nursing students, indicating higher inclusivity. Meaders and Vega (2023) concluded that two-stage exams were of increased benefit to students of a historically excluded race or ethnicity, improving equality in assessment. While implementing two-stage exams in lieu of traditional exams integrates a UDL approach to exams by reducing barriers to assessment, adopting a holistic UDL approach to nursing education would include multiple means of assessment in which exams are potentially replaced by alternative approaches to assessment.

Benefits of Two-Stage Exams

The use of a collaborative approach to assessment and evaluation in the form of two-stage exams has many documented benefits. Individual student knowledge increases during a two-stage exam (Gilley & Clarkston, 2014), and learning occurs on a deeper level than simply rote memorization (Gilley & Clarkston, 2014; Nicol & Selvaretnam, 2022), with a high potential for learning (Martin, 2018; Rieger & Heiner, 2014). Mahoney and Harris-Reeves (2017) demonstrate that collaborative assessment improves student performance on higher-order thinking questions across the board. Students report a very positive opinion of collaborative exams (Helms et al., 2019; Rieger & Heiner, 2014; Riley et al., 2021). Faculty indicate that two-stage exams are a powerful learning opportunity that increase student motivation, teamwork, collaboration, and communication (Burgess & Medina-Smuck, 2018; Gilley & Clarkston, 2014; Helms et al., 2019; Martin, 2018; Nicol & Selvaretnam, 2022; Patiwael et al., 2021). Students are highly engaged with the content and the members of their groups (Peck et al., 2013; Rieger & Heiner, 2014). This type of exam can positively affect problem solving, communication, and group skills (Gilley & Clarkston, 2014; Radford & Blount, 2023; Riley et al., 2021). Two-stage exams enhance student knowledge by decreasing test anxiety and stress while promoting problem solving and enhancing critical thinking, critical reasoning, and communication skills (Burgess & Medina-Smuck, 2018; Caboral-Stevens & Fox, 2020; Helms et al., 2019; Martin, 2018; Radford & Blount, 2023). Active and collaborative learning provides an opportunity for students to develop and practise peer interaction (Burgess & Medina-Smuck, 2018; Nicol & Selvaretnam, 2022). There are benefits to students at both ends of the grading spectrum, as most groups will score higher than an individual grade (Gilley & Clarkston, 2014; Riley et al., 2021). Studies show that giving and receiving verbal feedback elaborates on the student understanding of the content (Nicol & Selvaretnam, 2022), while the discussion process for the second stage has been linked to higher-order thinking and content retention (Witt et al., 2022).

Pitfalls of Two-Stage Exams

Two-stage exams are not without challenges in both implementation and group dynamics. Some research indicates that two-stage exams may not lead to increased knowledge retention (Martin, 2018; Peng et al., 2023). Caboral-Stevens and Fox (2020) found that collaborative assessment increases short-term retention but not long-term retention; however, they noted increased proficiency in critical thinking and clinical reasoning. Two-stage exams are not meant to replace all standardized exams but rather used strategically to enhance collaboration, communication, and critical thinking skills (Riley et al., 2021). Mahoney and Harris-Reeves (2017) indicate the risk for unbalanced groups during two-stage exams, including “freeriding, resentment towards peers, conflict and distrust” (p. 26). Social loafing (relying on other students for a “free ride”) is described by Saunders et al. (2019) as a limitation to a collaborative approach.

Methods

Study Design

This study used a descriptive design to examine collaborative assessment and two-stage exams in undergraduate nursing education. Survey data were collected from students to explore the perceived barriers to assessment before the two-stage exam. A second survey was implemented after the exam to explore the perceived reduction of barriers, knowledge retention, feeling of preparedness for future courses/clinical rotations, and level of satisfaction with two-stage exams. In addition, a descriptive numerical analysis of exam scores was conducted and applied to different exam weighting scenarios. No inferential statistical testing was performed on these data.

Study Setting and Participants

Participants were recruited from the collaborative baccalaureate nursing program at a college in Western Canada. Participants were in years 1 and 2 of the program and enrolled in nursing-specific courses in the fall of 2024. Participating courses included foundational nursing knowledge, pathophysiology, pharmacology, and nutrition. Courses included two-stage exams as part of the assessment. Each exam was developed by nursing faculty and not reviewed or altered by the researcher. Exams included 40 to 100 short-answer, multiple-choice, and true/false questions. Data collection in the form of surveys took place during scheduled exam times. Participants completed the first stage of the exam, followed by the second stage in small groups of three to six students. The group size, composition, and timing were determined by faculty and were not factors in this study. Six faculty participated in the recruitment and data collection phases of this study. A total of 130 participants in years 1 and 2 of the collaborative baccalaureate nursing program were recruited from 130 potential students. Exam scores were collected from 130 students by participating faculty. All students participating completed the pre-exam survey, and 108 completed the post-exam survey. Years of study were not analyzed.

Inclusion and Exclusion Criteria

Inclusion criteria were enrolment in a year 1 or 2 nursing-specific course with in-person classes and a two-stage exam. Exclusion criteria included online or hybrid/flexible courses, non-nursing courses, or courses without a two-stage exam. Faculty variations in the delivery of the two-stage exam did not affect data collection. Participants' grades were not affected by the research study.

Data Collection Methods

A survey tool was used to collect descriptive data before and after the two-stage exam. The pre-exam survey collected data on perceived systemic and individual barriers to assessment and the perceived impact of these barriers. The survey listed 13 barriers to assessment (see Table 1). The post-exam survey collected data about reduction of barriers to assessment and whether students felt better prepared to apply this knowledge to future courses and/or practice (see Table 2). Survey links were provided by QR codes. Participants used their electronic devices to complete anonymous surveys using their unique participant code. Participant codes were affixed to both individual and group exams. Participating faculty provided the individual and group exam scores correlated to participant codes to the researcher.

Pre-exam Survey

The pre-exam survey collected data on systemic and individual barriers to assessment, based on 13 barriers identified from a literature review of over 30 peer-reviewed studies and supplemented with

faculty and researcher experience. A literature review was conducted to determine potential barriers to assessment using the keywords *collaborative testing*, *two-stage exams, universal design for learning, post-secondary students*, *barriers to learning*, and *grade inflation*. Key themes that emerged from the literature included anxiety, mental health concerns, test anxiety, and stress (Burgess & Medina-Smuck, 2018; Caboral-Stevens & Fox, 2020; Epstein et al., 2021; Helms et al., 2019; Martin, 2018; Radford & Blount, 2023; Sonmez et al., 2023).

Post-exam Survey

The post-exam survey examined student perception of reduction of barriers, how barriers were reduced, and whether students felt better prepared for future courses and clinical practice. Questions included Likert-scale ratings from 1 to 5 for satisfaction, participation, and knowledge mastery.

Ethical Considerations

Research ethics approval was granted from the research ethics board at the college and conducted according to institutional policy and following Tri-Council Policy Guidelines regarding ethical conduct for research involving humans. Participants were provided with an information session and informational sheet from the researcher prior to the first exam. Written informed consent was obtained by participating faculty prior to the start of the exam. Student permission to write the exam was not affected by lack of consent. Participants were not recruited from the researcher's courses. No student names were provided to the researcher. All efforts were made to protect anonymity. Participating faculty were not provided individual survey results to protect student confidentiality.

Data Analysis

Surveys

Pre- and post-exam survey data were analyzed for trends and frequencies using Microsoft Forms to complete a basic results analysis. Mean scores were calculated to assess barriers of assessment and potential reductions of those barriers. Likert-scale questions were rated on a scale of 1 to 5, where 1 indicated *not at all* and 5 indicated *to the greatest extent*.

Quantitative Data

Exam scores of 130 students across two exams were recorded for a total of 258 incidences. Raw data were collated and sorted by individual exam letter grades. Grade inflation was calculated based on the increase from the individual exam to the two-stage exam score, where the individual exam score is multiplied by a theoretical weight (%) and the group exam score is multiplied by the remaining theoretical weight (%). That result was added and compared to the initial individual exam score to determine the net increase in grade based on the two-stage exam. Theoretical grade increases were estimated using four common weighting scenarios: 75/25, 80/20, 85/15, 90/10, based on faculty feedback and a study by Rieger and Heiner (2014) that explores the student perspective and faculty implementation of two-stage exams. Descriptive analyses were conducted (see Table 3).

Results

Pre-exam Survey

A total of 130 participants were invited to participate in the surveys. All 130 participants participated in the pre-exam survey, a response rate of 100%. The majority (69%) identify with at least one systemic or individual barrier to assessment. The most common barriers are test anxiety (59.2%), anxiety, stress, or mental health concerns (46.1%), and attention-deficit/hyperactivity disorder (ADHD), attention deficit disorder (ADD) or Autism spectrum disorder (ASD), whether diagnosed or suspected (26.9%). Of the 90 participants who identify as having barriers to assessment, 14% indicated significant impact on exam scores and 34% indicated a moderate impact, while 38% indicated a minor impact.

Table 1

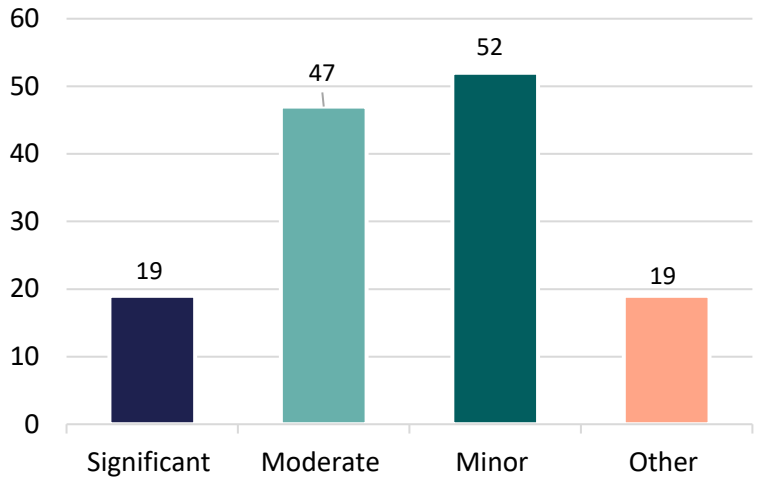
Systemic and/or Individual Barriers to Assessment (n = *130)*

Barrier	***n***	**%**
Dyslexia, reading comprehension, or other	4	3.1
ADHD, ADD, ASD (diagnosed or suspected)	35	26.9
English as a second or alternative language	11	8.5
Anxiety, stress, or mental health concerns	60	46.1
Text anxiety	77	59.2
Language comprehension	4	3.4
Concussion or post-concussion symptoms	3	2.3
Difficulty with focus during exams	40	30.8
Environment (too hot, too cold, too loud)	24	18.5
Outside influences related to social or personal concerns	19	14.6
Trauma related to previous exams or experiences	6	4.6
Trauma unrelated to educational experiences	6	4.6
Other	21	16.1

Note. ADHD = attention-deficit/hyperactivity disorder; ADD = attention deficit disorder; ASD = Autism spectrum disorder.

Figure 1

Impact of Barriers (n = *130)*



Post-exam Survey

Of the 130 participants recruited, 108 completed the post-exam survey. Most participants (86.1%) indicated that two-stage exams helped to reduce barriers to assessment. The top three ways that two-stage exams helped to reduce barriers were by helping to revisit the question with the group (67.6%), allowing a second chance to answer the question (65.7%), and helping to understand the question as it was written (64.8%). A total of 86% of participants reported that they found a reduction of barriers with two-stage exams, while 90% (*n* = 97) indicated that they feel better prepared to apply this knowledge in the future courses and/or the practice setting. Knowledge mastery was rated as 4.47 on a five-point Likert scale where *n* = 107. Participant satisfaction was 4.44 on a five-point Likert scale where *n* = 108. Participants reported 4.65 on their level of participation with this exam style and 73% rating it 5/5 where *n* = 107.

Figure 2

Two-Stage Exams Reducing Barriers (n = *108)*

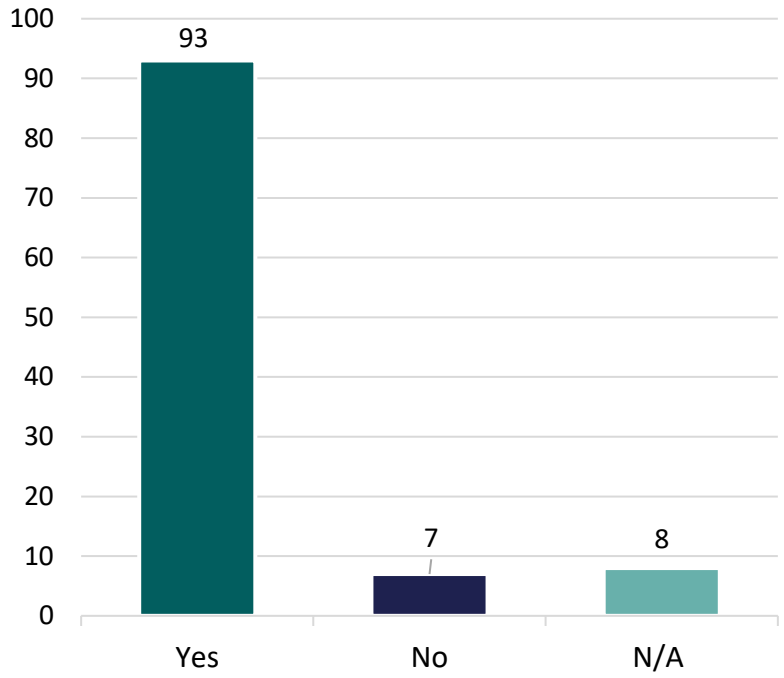


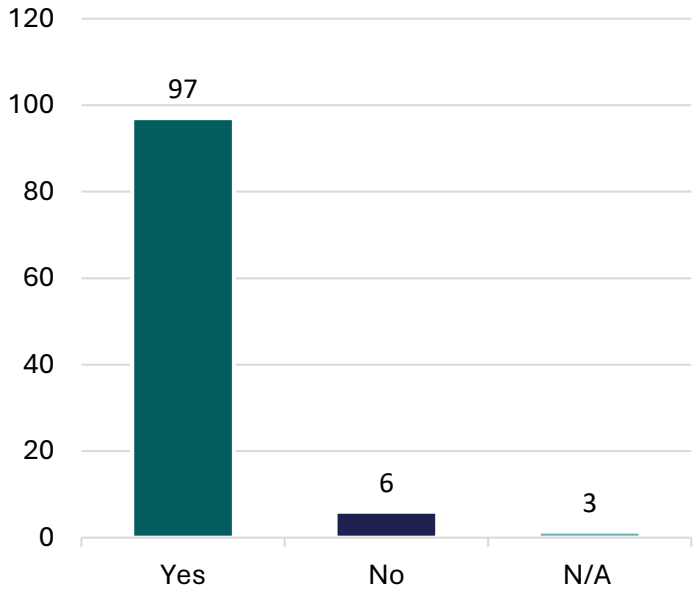
Table 2

Two-Stage Exams Reducing Barriers (n = *108)*

How did it help reduce barriers?	***n***	**%**
Understand the questions as it was written	70	64.8
Understand the terminology or language used	43	39.8
Reduce anxiety	64	59.3
Increase focus and attention when reviewing	40	37.0
Allow a second chance to answer the question	71	65.7
Revisit the question with insight from the group	73	67.6
Have other group members explain the concepts	56	51.9
N/A	2	1.9
Other	2	1.9

Figure 3

Preparedness to Apply Knowledge (n = *106)*



Grade Inflation

Grade increases were proportional to the weighting of the exam. An expected grade inflation was determined based on theoretical weighting scenarios of 75/25, 80/20, 85/15, 90/10. The grade inflation can range from 0.33% for an A+ grade weighted at 90/10 to 9.63% for a D grade based on a 75/25 scenario. This data can be used by faculty when planning and implementing a two-stage exam to decide the weight of the assessment. Little research exists about determining weighting of two-stage exams.

Table 3

Summary of Grade Inflation Averages (n = *258)*

Weighted percentage (%) increase (individual/group mark)				
Letter grade	**75/25**	**80/20**	**85/15**	**90/10**
A+	0.83	0.67	0.50	0.33
A	1.82	1.46	1.09	0.73
A–	2.16	1.72	1.29	0.86
B+	2.88	2.30	1.73	1.15
B	3.79	3.03	2.28	1.52
B–	4.86	3.88	2.91	1.94
C+	5.69	4.55	3.41	2.27
C	6.55	5.24	3.93	2.62

C–	8.33	6.66	5.00	3.33
D	9.63	7.70	5.78	3.85

Discussion

Barriers to Assessment

A high proportion of participants in this study identified barriers to assessment, including anxiety, mental health, learning disabilities, language comprehension, and trauma. These barriers are common themes in the literature (Burgess & Medina-Smuck, 2018; Caboral-Stevens & Fox, 2020; Epstein et al., 2021; Helms et al., 2019; Martin, 2018; Radford & Blount, 2023; Sonmez et al., 2023). The UDL Guidelines 3.0 describe how barriers at the individual, institutional, and system levels can negatively affect learning and assessment, while emphasizing the value of collaborative and interdependent learning (CAST, 2024). In this study, rates for test anxiety were 59.2%, and overall stress, anxiety, and mental health concerns were 46.1%. Of the participants, 69% reported that these barriers affect their exam scores. Prior to this study, minimal research existed on the use of collaborative testing or two-stage exams to reduce systemic and individual barriers to assessment and evaluation. Results indicate that many undergraduate nursing students report barriers to assessment and those barriers may affect exam scores. Reducing barriers to assessment by implementing two-stage exams in nursing education can potentially mitigate the negative effect on exam scores.

Benefits of Two-Stage Exams

The post-exam survey shows that the majority of participants perceived increased knowledge mastery. Many participants in this study reported a reduction in barriers to assessment and felt better prepared and knowledgeable about the topic. High levels of reported satisfaction and participation suggest increased student engagement, which may facilitate deeper learning. A two-stage exam reframes the assessment from a regurgitation of knowledge and information to an active, engaging, and critical inquiry of the questions and answers, wherein the students receive immediate feedback and an opportunity to review and revise questions (Gilley & Clarkston, 2014). Proposed solutions for alleviating stress, mental health concerns, and the associated decrease in enrolment include developing a collaborative and supportive environment (Sonmez et al., 2023). Using two-stage exams in undergraduate nursing education incorporates collaboration and communication skills into an assessment technique, creating a supportive assessment strategy (Rieger & Heiner, 2014).

Grade Inflation and Two-Stage Exams

The literature review identifies a gap in the research on potential grade inflation based on the weighting of two-stage exams. Currently, no best-practice standard exists for the weighting of the two components, individual and group. A study by Rieger and Heiner (2014) describes a range of weighting options but lacks a detailed analysis based on letter grades. This study finds that students with an A+ grade on the individual attempt receive less grade inflation than those with a lower grade. This is proportionate to the weighting of the group exam. The grade inflation ranges from 0.33% to 0.83% for an A+ individual grade. Students with a lower individual grade have a higher percentage of grade inflation, which is again proportionate to the weighting scale used. For example, a C– individual grade can expect a grade inflation of 3.33% to 8.33%. While these basic calculations stand to reason, the descriptive analysis of calculating 258 different exam scores with correlating group scores gives the expected grade inflation stronger

validation when sorted by letter grade. The results are comparable to the study by Reiger and Heiner (2014) that calculates the average increase of a two-stage exam weighted at 85/15 to be 3.3%. An understanding of expected grade inflation provides practical guidance for faculty when determining exam weighting. Higher rates of grade inflation can be decreased by scaling the group exam to a lower percentage. While it is intuitive that students who score lower on the individual exam will have a greater increase in exam score after the group exam, having a comprehensive understanding of the potential increase in exam score based on weighting of the exams can help inform faculty assessment design. A theme that emerged from this study is whether the increased mark should be categorized as grade inflation. The assumption of grade inflation may oversimplify the phenomenon; further research is needed to clarify whether increased grades reflect knowledge gains rather than simple inflation.

Strengths and Limitations

Strengths of this study include a robust cohort of undergraduate nursing students in years 1 and 2 of a collaborative nursing program. Several key barriers to assessment in this demographic have been identified. The data provide evidence that two-stage exams reduce barriers to assessment. Grade inflation is descriptively quantified, creating an easy-to-follow guideline for nurse educators to use in weighting two-stage exams. Limitations include a single semester for data collection, a single undergraduate nursing program, and a lack of qualitative feedback from the participants after the two-stage exam. The results are primarily based on participant self-reporting. The participant population is not representative of the general population of this geographic area, in that female students were overrepresented and Black, Indigenous, and racialized students were underrepresented. The concept of grade inflation in terms of two-stage exams needs further exploration to examine whether the students' increased grade can be attributed to increased knowledge mastery of the topic rather than inflated marks as an outcome of two-stage exams.

Future Research

Future research on two-stage exams around knowledge mastery and retention would be beneficial to understand the lasting impact of two-stage exams. Qualitative data around the student experience would be beneficial to support integration from a student viewpoint. Anecdotal evidence from this study showed interesting changes to the way that participants prepared and studied for a two-stage exam, with some students reporting preparation strategies aimed at deeper comprehension and peer teaching rather than rote memorization. Quantifying the participant perception of increased knowledge mastery would clarify the mixed narrative in the current literature as to whether two-stage exams improve long-term knowledge retention.

Conclusion

The results of this study show the number of barriers that undergraduate nursing students are experiencing with assessment. A reduction of these barriers can be expected when implementing a two-stage exam. Overall grade inflation can be quantified with an expected percentage increase to overall grade, filling a void in the literature. Using a UDL approach to undergraduate nursing education by implementing two-stage exams may reduce barriers and support knowledge acquisition among nursing students.

References

Balta, J. Y., Supple, B., & O'Keeffe, G. W. (2021). The Universal Design for Learning framework in anatomical sciences education. *Anatomical Sciences Education*, *14*(1), 71–78. https://doi.org/10.1002/ase.1992

Bangcola, A. (2023). Ways of coping and mental health among nursing students transitioning from online learning to in-person classes in a university setting. *The Malaysian Journal of Nursing*, *15*(1), 70–78. https://doi.org/10.31674/mjn.2023.v15i01.008

Burgess, A., & Medina-Smuck, M. (2018). Collaborative testing using quizzes as a method to improve undergraduate nursing student engagement and interaction. *Nursing Education Perspectives*, *39*(3), 178–179. https://doi.org/10.1097/01.nep.0000000000000223

Caboral-Stevens, M., & Fox, D. P. (2020). The use of collaborative testing with baccalaureate nursing students. *Teaching and Learning in Nursing*, *15*(1), 37–41. https://doi.org/10.1016/j.teln.2019.09.007

CAST. (2024). *The UDL guidelines*. https://udlguidelines.cast.org

Epstein, I., Khanlou, N., Ermel, R. E., Sherk, M., Simmonds, K. K., Balaquiao, L., & Chang, K.-Y. (2021). Students who identify with a disability and instructors' experiences in nursing practice: A scoping review. *International Journal of Mental Health and Addiction*, *19*(1), 91–118. https://doi.org/10.1007/s11469-019-00129-7

Ferrell, D. K., & DeCrane, S. K. (2016). S.O.S. (students' optimal success): A model for institutional action to support minority nursing students. *Journal of Cultural Diversity*, *23*(2), 39–45.

Gilley, B. H., & Clarkston, B. (2014). Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. *Journal of College Science Teaching*, *43*(3), 83–91. https://doi.org/10.2505/4/jcst14_043_03_83

Helms, K. D., Keith, L. A., & Walker, L. P. (2019). Evaluating student learning through collaborative testing in a psychiatric mental health course. *Journal of Doctoral Nursing Practice*, *12*(1), 3–9. https://doi.org/10.1891/2380-9418.12.1.3

Kotcherlakota, S., Leeseberg Stamler, L., Clark, A., & Howell, M. C. (2024). Implementing universal design for learning in nursing education: Faculty perspectives. *Teaching and Learning in Nursing*, *19*(1), e138–e144. https://doi.org/10.1016/j.teln.2023.10.003

Mahoney, J. W., & Harris-Reeves, B. (2017). The effects of collaborative testing on higher order thinking: Do the bright get brighter? *Active Learning in Higher Education*, *20*(1), 25–37. https://doi.org/10.1177/1469787417723243

Martin, A. (2018). A quantitative framework for the analysis of two-stage exams. *International Journal of Higher Education*, *7*(4), 33–54. https://doi.org/10.5430/ijhe.v7n4p33

Meaders, C. L., & Vega, Y. (2023). Collaborative two-stage exams benefit students in a biology laboratory course. *Journal of Microbiology & Biology Education*, *24*(1). https://doi.org/10.1128/jmbe.00138-22

Murray, T. A., Pole, D. C., Ciarlo, E. M., & Holmes, S. (2016). A nursing workforce diversity project: Strategies for recruitment, retention, graduation, and NCLEX-RN success. *Nursing Education Perspectives*, *37*(3), 138–143.

Nicol, D., & Selvaretnam, G. (2022). Making internal feedback explicit: Harnessing the comparisons students make during two-stage exams. *Assessment & Evaluation in Higher Education*, *47*(4), 507–522. https://doi.org/10.1080/02602938.2021.1934653

Patiwael, J. A., Douma, A. H., Bezakova, N., Kusurkar, R. A., & Daelmans, H. E. M. (2021). Collaborative testing in physical examination skills training and the autonomous motivation of students: A qualitative study. *BMC Medical Education*, *21*(1), Article 224. https://doi.org/10.1186/s12909-021-02618-7

Peck, S. D., Stehle Werner, J. L., & Raleigh, D. M. (2013). Improved class preparation and learning through immediate feedback in group testing for undergraduate nursing students. *Nursing Education Perspectives*, *34*(6), 400–404. https://doi.org/10.5480/11-507

Peng, X., Jiang, L., Cao, D., Chen, G., Li, D., Teng, P., & Li, J. (2023). Teacher feedback-based collaborative testing improves students' knowledge gaps of parasitology. *Advances in Physiology Education*, *47*(2), 259–264. https://doi.org/10.1152/advan.00230.2022

Radford, M., & Blount, C. (2023). The impact of collaborative testing on teamwork and collaboration in nursing students. *International Journal of Nursing Education*, *15*(3), 34–36. https://doi.org/10.37506/ijone.v15i3.19569

Rieger, G. W., & Heiner, C. E. (2014). Examinations that support collaborative learning: The students' perspective. *Journal of College Science Teaching*, *43*(4), 41–47. https://doi.org/10.2505/4/jcst14_043_04_41

Riley, E., McCormack, L., Ward, N., Renteria, F., & Steele, T. (2021). 2-stage collaborative testing results in improved academic performance and student satisfaction in a prelicensure nursing course. *Nurse Educator*, *46*(4), 261–265. https://doi.org/10.1097/nne.0000000000000934

Saunders, A., Say, R., Visentin, D., & McCann, D. (2019). Evaluation of a collaborative testing approach to objective structured clinical examination (OSCE) in undergraduate nurse education: A survey study. *Nurse Education in Practice*, *35*, 111–116. https://doi.org/10.1016/j.nepr.2019.01.009

Sonmez, Y., Akdemir, M., Meydanlioglu, A., & Aktekin, M. R. (2023). Psychological distress, depression, and anxiety in nursing students: A longitudinal study. *Healthcare*, *11*(5), 636. https://doi.org/10.3390/healthcare11050636

Witt, A. J., Gerdin, B. J., De Haan, J. A., & Bergman, L. K. (2022). A comparison of collaborative testing and alternative test-taking methods on content retention in baccalaureate nursing students. *Nursing Education Perspectives*, *43*(4), 243–245. https://doi.org/10.1097/01.nep.0000000000000856