Failing to Fail in Undergraduate Nursing: Understanding the Phenomenon

Angie Docherty

Abstract

AIM The aim of the study was to explore and understand the phenomenon of "failing to fail."
BACKGROUND Phase 1 of a mixed-methods study suggested faculty in clinical settings instructed students that should not have passed preceding placements; students in didactic settings also passed exams that merited a fail. Phase 2 explored this phenomenon.
METHOD A multisite qualitative case study targeted baccalaureate and community college faculty to support analysis using replication logic. Data collection was conducted via semistructured interview.
RESULTS Eighteen demographically diverse cases were recruited (including age, experience, and full-/part-time status). Factors supporting failing to fail included being good enough, clinical/didactic dichotomy, team grading, and being the bad guy.
CONCLUSION The consistency of enabling factors suggests a collective approach is required to address failing to fail, including pedagogical preparation and cross-school mechanisms for ensuring grading parity. Effort must address integrity and teaching excellence in all aspects of nursing education.
KEY WORDS Appraisal – Assessment – Failing to Fail – Nursing Education – Nursing Students

I n 2015, phase 1 of a mixed-methods study evidenced "failing to fail" in undergraduate nursing education (Docherty & Dieckmann, 2015). Faculty in an educational consortium in one northwestern state were invited to participate in a cross-sectional survey covering both clinical and didactic instruction. From a response rate of 34 percent, the results were stark: 67 percent of faculty had instructed students they felt should not have passed the preceding placement, 43 percent of faculty had awarded a higher grade than deserved, and 18 percent had given a passing grade for an exam that merited a fail.

Failure to fail was evident across baccalaureate and associate degree programs and across clinical and didactic settings. Failure to fail also showed no relation to faculty experience, age, and qualifications. The findings suggest some common enabling factors. These include reluctance to fail students in the latter stages of their program and, paradoxically, reluctance to fail students early in the program, on the assumption that they have time to attain the required standard. Other potential enablers include team grading norms, lack of rubric clarity, personal bias, and fear of potential litigation.

Given the potential implications for patient care and professional standards, it was imperative to expand these quantitative findings with a more nuanced exploration of grading practices. This article outlines the sequential second phase of study, which aimed to explore and understand the phenomenon of failing to fail.

About the Author *Angie Docherty, NursD, MPH, RN, is an assistant professor Oregon Health & Science University School of Nursing, Monmouth, Oregon. This research was funded by a Nursing Education Research Grant from the National League for Nursing. For more information, contact Dr. Docherty at docherty@ohsu.edu. The author has declared no conflict of interest.*

BACKGROUND

The phenomenon of failing to fail in nursing education was originally highlighted in a study funded by the UK Nursing and Midwifery Council (Duffy, 2003). In Duffy's clinically focused study, it was evidenced that educators could be reluctant to fail students when their performance suggested failure was merited. What emerged was a tendency to give students the benefit of the doubt, and unless there was clear evidence of a risk to patient safety, students generally continued to subsequent placements.

Further studies and reviews have provided international support for Duffy's findings. These include a systematic review evidencing grade inflation (Donaldson & Gray, 2012) and studies highlighting failing to fail from a clinical perspective (Jervis & Tilki, 2011; Larocque & Luhanga, 2013;Tanicala, Scheffer, & Roberts, 2011). In the Tanicala et al. study, the authors reached the conclusion that education and experience of faculty are likely to be factors that support accurate clinical grading decisions. However, the Docherty and Dieckmann (2015) study showed no relationship between these factors and grading practices.

There has been some additional exploration of the phenomenon of failing to fail. In particular, two systematic reviews have provided further evidence of this concern in health professions education. A Best Evidence Medical Education protocol was used to review 28 studies published between 2005 and 2015 from the fields of medicine, nursing, and dentistry (Yepes-Rios et al., 2016). The reviewers focused on evidence relating to experiences and perceptions of clinical grading. Factors supporting grading inaccuracies included institutional culture, lack of available remediation, impact on workload and faculty evaluation, and fear of litigation. Factors supporting rigorous grading included professional duty, institutional support, and opportunities for student remediation.

Hughes, Mitchell, and Johnston (2016) also explored the literature, focusing specifically on failing to fail in undergraduate nursing.

Their final analysis included 24 studies covering the period 1995 to 2014. The authors concluded that failing students is a difficult and emotional experience that requires faculty confidence and institutional support. In both reviews, most of the nursing literature stemmed from the United Kingdom and Canada; in total, only six papers (one not published beyond conference abstract) focused on education in the United States. Also of note across both reviews was the presence of only one cited paper (Walsh & Seldomridge, 2005) that explored both clinical and didactic grading. Most studies focused on challenges related to either clinical or didactic grading, which seems contrary to the drive to strengthen the theory-practice interface. It also increases the risk that, as a nursing education sector, we miss intrinsic grading challenges at the institutional and structural level (Docherty & Dieckmann, 2015). Both reviews called for further research in this area.

The primary aim of this phase 2 study was to expand our understanding of grading challenges across didactic and clinical settings to determine what factors support or limit an environment of failing to fail in schools of nursing. Primarily, the focus was on exploring areas where faculty passed students who merited a fail (whether pass/fail or graded assignments). Attention was also paid to the issue of awarding higher grades than merited.

METHOD

This multisite, qualitative case study returned to the phase 1 study population between November 2015 and June 2016. The population consisted of more than 200 undergraduate faculty from a consortium of five university campuses and nine community colleges. A shared curriculum provided standardization of student learning outcomes and expectations across all programs.

The conceptual framework was based on the Morse Notation System (Morse & Niehaus, 2009). Specifically, the design was a sequential QUAN *qual,* where the focus of phase 2 was to use qualitative methodology to add context and expand understanding of the findings from the phase 1 quantitative survey. Case study methodology was chosen for the qualitative component given the need to understand the "why" and "how" of personal actions rather than the actions themselves (Polit & Beck, 2012). The goal was to recruit a purposive sample of 15 to 20 participants or "cases" from across the target population to ensure sufficient case data for replication analysis (Yin, 2009). Recruitment was through email, and participation was voluntary. The study was ethically approved by the university institutional review board.

Procedure

Data collection was primarily via semistructured interview. Structured questions were derived from an expansion of specific questions from the phase 1 survey and from a review of free-text responses obtained in that survey. An embedded approach collected data on a number of subunits within each case, including age, nursing education and experience, formal or informal educator preparation or instruction, grading experience in clinical and didactic settings, and current and prior grading practices. Data were also collected on level of perceived collegial and institutional support, personal experience of failing to fail, and, if applicable, perceived factors that supported or reduced the failing to fail phenomenon. Participants could add further unstructured information if they felt anything had not been captured via the a priori questions. The interviews were conducted by the researcher face to face, via telephone, or via videoconference, depending on interviewee preference and proximity. All interviews were digitally recorded and transcribed verbatim.

Data Analysis

Data analysis, completed by the researcher, used replication logic supported by Nvivo10 software. The primary unit of analysis was the case, with the logic being level of replication between cases, that is, the degree to which each case was similar or contrasting to other cases (Yin, 2009). On a case-by-case basis, raw data were grouped into university or community college cases. This was to facilitate individual case analysis and replication logic within and across case groupings to determine if factors associated with the educational setting were of relevance to the phenomenon under study. Data were then explored and coded under the a priori question categories, followed by a process of case-by-case pattern matching to determine the level of replication. Additional coding was clustered throughout the analysis prior to being grouped into key themes.

RESULTS

The target recruitment was met with 18 cases participating, 13 from the university campuses and five from the community colleges. This recruitment reflected the 2:1 consortium ratio. Cases were diverse in age, experience, and full-time/part-time status. The minimum qualification was a master's degree; five cases (one from the community college setting) had doctoral qualifications.

The key finding from the qualitative data was that failing to fail is a real phenomenon across educational settings and across clinical and didactic assessment. Furthermore, faculty know when they or their colleagues fail to fail. The overarching evidence is presented below. The findings are then grouped thematically into factors that may support or limit the potential for faculty to fail to fail. These themes include the good enough approach; the clinical versus didactic dichotomy; student stage, early or late; team grading, good or bad; and being the bad guy. In all cases, themes and, where relevant, subthemes are directly supported by the voice of the participants (university cases U1-U13 and community college cases C1-C5).

The analysis suggested that faculty, across instructional settings, grade with reflection and insight into their grading practices and generally strive to be pragmatic in their approach:

(C1) We take very seriously here, our responsibility to graduate someone who we believe we'd be comfortable with if they were taking care of us in the future.

(U2) It's like not managing the symptom in the nursing clinical world. Our job is to manage symptoms and if we don't manage symptoms, we're not doing our work.

However, failing to fail was evident across both case groupings, either as a personal action or as an action witnessed in current or prior educational settings:

(U1) I felt like she learned very little, and she progressed very little, and she really was unsafe in my opinion. But I was told not to fail her, and that they would continue to work with her.

(C5) Well, [failing to fail] has happened many times actually. And we wonder how come they got to where they're at.

The Good Enough Approach
There was evidence to support faculty adopting a "benefit of the doubt" or "good enough" approach to their clinical grading and assessment of students that stopped short of failing to fail:

(C1) There have been students who've graduated who I would say were minimally competent. They were not at the strength that we would like to see them, but they were minimally competent.

(C3) I've seen some students that, they're maybe not fit for let's say a hospital practice, but they demonstrate very good skills for a practice that is maybe a clinical practice.

The decisive factor appeared to be related to clinical safety:

(U3) But the really big things that are critical to being a practicing nurse, if they're not going to be able to be safe, then they shouldn't graduate from our program.

(U12) [Failing] was supported because it was a situation of safety. And the language related to the expectation for safety was very clear. So there really wasn't a problem.

The remainder of the results section explores thematic findings that may mitigate or contribute to the presence of failing to fail. In each theme, there was no notable distinction between university and community college settings. However, as the first theme suggests, there may be discreet perceptions between clinical and didactic grading.

Clinical Versus Didactic Dichotomy
Faculty reported a perception that didactic or theory grading was often less subjective and could potentially be less problematic as a means to fail students:

(C1) You know, in some senses, it's easier to fail them academically because it seems like it's more cut and dry…if they don't have the knowledge to pass those exams then they fail.

(U6) They were hoping that they would fail the theory portion of the course so that it wouldn't matter that they had passed the clinical, they would end up failing anyway.

Student Stage: Early or Late
One of the contributory factors, particularly in relation to clinical failing, was the student's stage in the program, with a perception that it was more challenging to address problems later in the program. However, there was a paradoxical finding that avoiding early problems may not be the best solution over time.

(C4) We also have a little bit of a feeling of how did they get that far? They should have been identified earlier.

(U1) If it's the first term, then yes. If it's the first term, I'm thinking maybe they just need to have more practice with this specific competency. So it really does depend on where they're at in the program.

(U3) It gets harder, every year. Because we have a history then of having invested in them as being successful, and they had that understanding, and they've been financially and cognitively and emotionally involved in the program…. You know, it's like, we failed them back here.

(U8) I don't know that it's ever easy, but certainly, you know, when somebody gets to the end, you know they're two terms away from graduating, that really is hard.

The failure to address learning deficits when they first emerge was noted within faculty teams and was considered problematic for faculty and students:

(U2) I said it already, giving the benefit of the doubt, just drives me nuts. It drives me nuts. Because look what happens. We perpetuate irritation.

(U3) It means somebody along the way hasn't attended to what's going on with this student. And I think, you know, we do them a disservice.

Team Grading: The Good and Bad

Grading as part of a team was perceived as a process that could both support and limit failing to fail. In general, there was some acceptance of the subjectivity of grading processes:

(U3) I think even with the best of intentions, people evaluate differently. If they're taking a Likert test, you know, some people mark all the 5s, and some people mark all the 3s, and so faculty grade in that way too.

However, there was evidence that some faculty perceived a degree of peer pressure:

(U6) And I've had pressure from other faculty…Oh, you know, they've had these things going on in their life and I know that if only they didn't have that then they'd be able to do it.

(U11) I think while I was being mentored, I don't know if I would say there was pressure, but there probably was an expectation that the person who was the primary teacher for the course set the standard.

There was evidence to suggest this "peer pressure" could be internalized as well as overt:

(U3) When I turn in all my grades and all the other students in the class got higher than my students, and then I feel like I graded too hard, and then I will go in then and change that.

Positive aspects, where an intentional approach to comparing and reviewing grading was embedded into grading systems, appeared to build confidence and a shared responsibility:

(U12) It was a collective will, because then we also learned to use it, where when somebody had to give a grade that was lower, like a 75, they felt confident saying, "The faculty has graded this, not just me." And one person felt so supported by it, she would write on the paper, This paper had a second reader.

(U13) The first person that grades, and then I can go in there and look

and see how that person graded and we'll talk about, you know, we take off so many points for this and so many points for that, so we have interrater reliability.

However, even with this team effort, there were occasions when pressure was felt from colleagues:

(U3) And sometimes, everybody got 100% except the ones I graded, and then that didn't feel very good, and we had discussions, and I felt at that point in time I needed to give all of mine 100% too after looking at the others so that there was equity among the cohort.

When effort was made to standardize processes, such as using grading rubrics, the results were not always as objective as may be desired:

(U5) When I have the ability to look within a course and see grading, it frustrates me when I see rubrics not being followed and grades being given to students that I don't think they have deserved on the merit of the evidence they submitted.

(U7) And it tends to go back to how I look at the rubric and how others interpret the rubric, and it brings up issues. Is the rubric really supporting to all of us in grading as objectively as we can?

In didactic settings, an important factor may be whether faculty used anonymous grading processes or knew whose paper they were grading:

(C1) Our instructors have implemented the anonymous, where they just have the student ID number on the assignment, and they have found that very interesting, to then identify who the student is after they have finished grading.

(U3) I took all the identifiers off, it was a lot of work, and sent them out and then got the grades back, and the disparity in the grades, using the same rubric, as amazing.

(U12) Well people found that they were biased towards their own clinical group. If they were marking a paper, and they knew the student, even if they didn't hit the criteria, they could say, "Well, I know they did this in clinical," but it wasn't in the paper.

Being the Bad Guy

It was evident in the data that some faculty, across all settings, had a reluctance to fail to students. This occasionally related to lacking the energy or "bandwidth" to face potential or actual consequences:

(C4) Uh, this is very devastating in the fact that most every single time this happens the student grieved through the college process, which kind of seemed to turn the whole situation around and the faculty then is on the hot seat, not the student.

(U3) If you give a student a poor grade on something, they come back and they want to talk with you about it and you know, even if it's a B, which I think is a great grade, it's like, there's a hassle factor that goes with

that. And I'm not sure that faculty feel that it's worth it.

(U8) Part of that was, during that time, I really, I just didn't have the emotional bandwidth to do it. I knew I should. I wanted to. I talked to colleagues about it, but time just went by, and here we are week 10 and that's when it happened.

There was some evidence to suggest that the risk of losing their perceived standing with students impacted faculty decision making:

(C5) Sometimes it's a new instructor, who hasn't quite got the feel for it yet, or is too afraid to be critical of a student, or I don't know if I should say critical, if that's a good term, but you know, they want to be nice, they want to be liked, or something or other.

(U6) And I guess the other reason, the unspoken one was just aversion to doing it.

In general, faculty experienced administrative support for difficult grading decisions, but there was evidence that "higher-order" processes could impact final outcomes. There was also evidence that these processes may not always be fully understood by faculty.

(C2) I've seen situations where faculty had not crossed their t's and dotted their i's and we have a poor paper trail. And then the frustration is then those students grieve that and then the dean may make an overriding decision, and that doesn't feel well.

(U1) They're having to juggle many different things and they don't know

exactly what we've been doing here. So they haven't been with us, every month, every step of the way, like we have with the student.

(U2) If we have legal practice leapfrogging policies and trumping policies because we are afraid of litigation, we then have litigation trumping policy, or worse yet, litigation creating policy, that then steps outside of course outcomes.

Ultimately, although there was evidence to support failing to fail and evidence pointing to a number of contributing factors, there was evidence to support the effort that many faculty put toward grading and student success in the long term.

(C3) I think that, again, the patient should be any instructor's center of foci, and really thinking about what impact any nursing student, any new nurse will have on our patient.

(U5) Counseling on feedback directly related to course outcomes that the student wasn't achieving — the student came to her own understanding that she was in completely the wrong line of work, and that was actually a very successful removal from the program.

DISCUSSION

The introspective data, obtained through 18 interviews, suggest one overarching point: faculty are aware of the responsibilities of the accuracy of their grading, both in terms of student success and public safety, and they strive to honor this responsibility. However, the data also suggest two additional points: 1) there are a number of factors, positive and negative, that impact grading practices, and 2) when the negative factors are prominent, the risk of failing to fail can become the reality. In this study, the factors that contribute to a failure

to fail included personal factors such as emotional ability and lack of confidence, team factors such as peer pressure, and institutional factors such as administrative and legal requirements.

The phase 2 qualitative findings support the phase 1 quantitative findings and provide a more detailed understanding of these contributory factors. The results also support the findings of the most recent reviews discussed above. Importantly, in the Yepes-Rios et al. (2016) review, the findings from the one US-focused nursing paper (Debrew & Lewallen, 2014) are congruent with the findings of this study.

The Debrew and Lewallen (2014) study was clinically focused, but it is evident from our findings that failing to fail is not a phenomenon specific to any one instructional setting. This was noted in the systematic review by Hughes et al. (2016), which illustrated grading challenges across settings. There is, however, a persistent and long-standing perception in the cited evidence (Duffy, 2003; Paskausky & Simonelli, 2014; Walsh & Seldomridge, 2005) that students are more likely to fail in didactic settings. These findings, which span UK and US education, are supported in the current study.

Interestingly, the process for assessing and grading students in clinical settings varies. In the United Kingdom, supervision and assessment of students are undertaken by trained clinical staff; in the United States, supervision and assessment are undertaken by clinical faculty. Yet, both instructional formats seem subject to the subjective/objective challenge suggesting clinical grading, at large, is an area where there is scope to develop more objective measures. The United Kingdom has led the way in the use of Objective Structured Clinical Examinations and Assessments as a means to counteract the subjective nature of grading clinical performance, and there is now increasing use of this methodology in the United States (Najjar, Docherty, & Miehl, 2016).

Objective Structured Clinical Examinations and Assessments, as well as other simulation-based testing models, provide a means to instill a degree of standardization and rubrics into grading process to enhance validity and reliability. However, the findings here suggest that rubrics, in their own right, may not be the complete solution to grading challenges. The evidence suggests that faculty may still incorporate known or unknown bias into their grading, particularly when they have a professional relationship with the student they are grading. Anonymous grading appears to mitigate against this bias to some extent, and there may be an argument that anonymized grading should be a standard institutional requirement for written assignments and examinations.

The evidence suggests that faculty do find grading a challenging part of their role. The persistence of personal, collegial, and institutional factors, and in this study, the evidence across instructional settings, shines an uncomfortable light on an educational problem. These new findings support the recent work by Kardong-Edgren, Oermann, Rizzolo, and Odom-Maryon (2017), which reviewed rater reliability in high-stakes simulation testing. Kardong-Edgren et al. determined that, although faculty may be content experts, they are not necessarily evaluative experts. Their conclusion, that some faculty should not be used for high-stakes testing, raises an important issue: can we do more to explore how we transition expert clinicians to become expert educators and expert evaluators? It is increasingly evident that nursing education, at both the undergraduate and graduate levels, does not always prepare one to become a nurse educator. We may need to take a reflective look at the preparation and support we give new educators as well as the continuing education and support we provide for those already in post.

As part of any educator support, it may be important to ensure that educators, across all settings, are fully cognizant with standard institutional requirements to support legal challenges to grading and assessment decisions. Although faculty report a fear of being involved in litigation (which can be strong enough to impact their own grading decisions), they also report frustration when they perceive that their grading decisions are overruled through legal or administrative processes. In these cases, it may be that assessment documentation did not reach a required legal standard, but this rationale should be communicated clearly to faculty and used as the starting point for strengthening assessment processes.

Ultimately, inconsistency in grading is not a single institution or consortium problem. It is also not a problem unique to nursing education (Yepes-Rios et al., 2016). But the willingness to explore areas where we acknowledge personal or collegial practice is falling short of desired standards illustrates that nursing faculty are ready for change. It may be that, to attain the desired and expected standard, institutional organizations and nursing education, in general, need to work together to raise accountability across the sector. We need to look at cross-sector standards and practices that strengthen quality of grading and, by default, ensure the quality of graduates. This effort will be challenging, but it will ensure nursing education is committed to leading the drive toward grading parity and integrity.

LIMITATIONS

Although this study had strengths, such as the ability to obtain data from cases across multiple sites, the ability to minimize the confounder of curriculum variance, and the ability to triangulate quantitative and qualitative data, there were a number of limitations. The most important limitation, not always acknowledged in qualitative research, is the potential for unconscious or implicit researcher bias (Morse, 2015). In this study, there was a single researcher, and the study was driven by a personal recognition of grading challenges and the desire to explore grading practices.

It is important to state that this study was not looking for evidence of failing to fail. The phase 1 study and others already suggest that failing to fail is a reality. Therefore, this phase 2 study started from that perspective and was designed to explore what may be contributing factors. In this respect, data collection and analysis happened through a reflexive process. A neutral tone was adopted at all times during data collection, and the analysis was guided initially by the a priori coding. At all times, both enablers and barriers to failing to fail were of importance.

A related limitation was the potential for participant self-selection bias. It is possible that those who volunteered to participate in the study felt strongly regarding the problem of grading inconsistencies and challenges. However, this is the nature of qualitative research and the desire for a purposive sample. A further limitation related to the study being limited to one northwestern state, although institutions across the state were included. This is countered somewhat by prior evidence that failing to fail may be a national and international problem.

CONCLUSION

This second phase of a sequential mixed-methods study supports the phase 1 findings that failing to fail and grading inconsistencies are evident across institutional settings. The qualitative phase allowed for a detailed exploration of contributing factors and confirmed shared

enablers that cross-cut university and community settings, clinical and didactic instruction, and new and experienced educators.

This consistency of problem demands a consistency of solution. Sector-wide, open, and transparent discussion may be required to air and address this shared problem. Future research should potentially include the role of pedagogical preparation and the potential for cross-school mechanisms for ensuring grading parity. Given the implications for the nursing profession and our practice partners, effort must be made to ensure integrity and teaching excellence in all aspects of nursing education. In closing, final acknowledgement is given to those faculty who agreed to an honest and frank exposure of their grading practices.

REFERENCES

Debrew, J. K., & Lewallen, L. P. (2014). To pass or to fail? Understanding the factors considered by faculty in the clinical evaluation of nursing students. *Nurse Education Today, 34*, 631-6. doi:10.1016/j.nedt.2013.05.014

Docherty, A., & Dieckmann, N. (2015). Is there evidence of failing to fail in our schools of nursing? *Nursing Education Perspectives, 36*(4), 227-231. doi:10.5480/14-1485

Donaldson, J. H., & Gray, M. (2012). Systematic review of grading practice: Is there evidence for grade inflation? *Nurse Education in Practice, 12*(2), 101-114.

Duffy, K. (2003). *Failing students: A qualitative study of factors that influence the decisions regarding assessment of students' competence in practice.* Retrieved from www.nmc-uk.org/Documents/Archived percent20Publications/1Research percent20papers/Kathleen_Duffy_Failing_Students2003.pdf

Hughes, L. J., Mitchell, M., & Johnston, A. N. (2016). 'Failure to fail' in nursing—A catch phrase or a real issue? A systematic integrative literature review. *Nurse Education in Practice, 20*, 54-63. doi:10.1016/j.nepr.2016.06.009

Jervis, A., & Tilki, M. (2011). Why are nurse mentors failing to fail student nurses who do not meet clinical performance standards? *British Journal of Nursing, 20*(9), 582-587. doi:10.12968/bjon.2011.20.9.582

Kardong-Edgren, S., Oermann, M. H., Rizzolo, M., & Odom-Maryon, T. (2017). Establishing inter- and intrarater reliability for high-stakes testing using simulation. *Nursing Education Perspectives, 38*(2), 63-68. doi:10.1097/01.NEP.0000000000000114

Larocque, S., & Luhanga, F. (2013). Exploring the issue of failure to fail in a nursing program. *International Journal of Nursing Education Scholarship, 10*(1), 1-8. doi:10.1515/ijnes-2012-0037

Morse, J. M. (2015). Critical analysis of strategies for determining rigor in qualitative inquiry. *Qualitative Health Research, 25*(9), 1212-1222. doi:10.1177/1049732315588501

Morse, J. M., & Niehaus, L. (2009). *Mixed method design: Principles and procedures.* Walnut Creek, CA: Left Coast Press.

Najjar, R., Docherty, A., & Miehl, N. (2016). Psychometric properties of an objective structured clinical examination tool. *Clinical Simulation in Nursing, 12*(3), 88-95. doi:10.1016/j.ecns.2016.01.003

Paskausky, A. L., & Simonelli, M. C. (2014). Measuring grade inflation: A clinical grade discrepancy score. *Nurse Education in Practice, 14*(4), 374-379. doi:10.1016/j.nepr.2014.01.011

Polit, D. F., & Beck, C. T. (2012). *Nursing research: Generating and assessing evidence for nursing practice* (9th ed.). Philadelphia, PA: Lippincott, Williams & Wilkins.

Tanicala, M. L., Scheffer, B. K., & Roberts, M. S. (2011). Defining pass/fail nursing student clinical behaviors phase I: Moving toward a culture of safety. *Nursing Education Perspectives, 32*(3), 155-161. doi:10.5480/1536-5026-32.3.155

Walsh, C., & Seldomridge, L. (2005). Clinical grades: Upward bound. *Journal of Nursing Education, 44*(4), 162-168.

Yepes-Rios, M., Dudek, N., Duboyce, R., Curtis, J., Allard, R., & Varpio, L. (2016). The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42. *Medical Teacher, 38*(11), 1092-1099. doi:10.1080/0142159X.2016.1215414

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). London, United Kingdom: Sage.

For more than 169 additional continuing education articles related to Education topics, go to NursingCenter.com/CE.