Rubrics for clinical evaluation: Objectifying the subjective experience

Julie J. Isaacson, Annette S. Stacy

Arkansas State University, Jonesboro, P.O. Box 910, AR 72467, United States

Accepted 10 October 2008

Summary Rubrics have historically been used in secondary and higher education to evaluate specific assignments or tasks. There is little mention of rubrics in the nursing literature, particularly in the area of clinical evaluation. A strong case can be made for expanding the traditional use of a rubric to include its validity with clinical evaluation. Clinical evaluation remains a challenge, even for seasoned faculty. Faculty and students often interpret clinical course objectives differently. Coupled with this concern is the subjectivity of the evaluation. The use of “novice” clinical faculty, who inevitably struggle with discerning and justifying anything but stellar student performance, further compounds these issues. Rubrics also facilitate the grading experience for faculty and students. Faculty often find themselves making repetitive written comments to students. These comments can be incorporated into the rubric, thus shortening grading time while increasing the quality and quantity of instructor feedback. When clarified in a rubric, course objectives become “real”. Student benefits include increased critical thinking and a more realistic approach to self-evaluation. Clinical rubrics can be developed from existing course objectives. Though perhaps tedious in initial development, both faculty and student satisfaction with the clinical evaluation process can be enhanced with the use of rubrics.

Introduction

Clinical evaluation of nursing students has been a long-standing problem and a continued area of interest for nurse educators (Andre, 2000; Lasater, 2007; Mahara, 1998; Orchard, 1994; Seldomridge and Walsh, 2006; Woolley, 1977). It is important that clinical educators understand the inherent academic pitfalls of the common methods of clinical
evaluation from both the instructor and student perspective. There is more in the nursing literature about the problems of clinical evaluation than there is about effective solutions. Compounding the issue is the plethora of academic jargon used to describe evaluation tools (checklists, criterion — referenced objectives, practice performance assessment, etc). The purpose of this paper is to discuss faculty and student concerns with regard to clinical evaluation and to explore the use of rubrics as a tool to help objectify the clinical evaluation process.

Academic pitfalls

The process and outcome of clinical evaluation is a major concern for faculty worldwide. Regardless of the program design, faculty trepidations include the perceived subjectivity of the experience, the blending of educator/evaluator roles, the academic jargon of the evaluation tool, the challenges related to documentation and the overall inability to control the learning environment. This angst has led faculty to trust instinct and experience as a basis for evaluation as they attempt to discern exactly how a typical student should be expected to perform (Krithbaum et al., 1994). Though valid to some degree, lacking in this process is consistency in grading. Equally disconcerting to faculty is a reliance on novice clinical instructors and an overall tendency towards grade inflation. Student issues tend to focus on performance anxiety, use of self-evaluations, generational grading standards among clinical faculty.

Faculty concerns

Faculty typically try to objectify a largely subjective experience, thinking that subjectivity is not compatible with a good clinical evaluation. However, there is really no way to completely remove subjectivity from clinical evaluation (Heworth, 1991; Mahara, 1998). Faculties are also challenged to evaluate multiple behavioral domains at the same time, knowing that some domains lend themselves more toward subjectivity, such as therapeutic communication. Acknowledgement of the subjectivity of clinical evaluation is essential as there will always be some degree of subjectivity in the clinical evaluation process.

Clinical educators are continually faced with dual roles in the clinical setting, functioning as mentor/educator and evaluator/judge at the same time. The primary purpose of clinical instructors is that of educator, but it is equally important to maintain professional standards and to assess performance (Orchard, 1994). The challenge lies in determining when to transition from the role of mentor to that of evaluator in the clinical area. Nursing faculty is drawn to nurture and guide but at some point must become the evaluator/judge.

Nursing faculty has a tendency to go to great lengths to create course objectives and corresponding evaluation tools that are all encompassing of a broad range of experiences. In doing so, many clinical evaluation tools are notorious for their academic jargon, perhaps even to the point of being a language that only the author of the tool understands! Clinical evaluation tools can be so broad that no one really knows what the tool is measuring and students and faculty alike may experience an "academic illiteracy" with regard to clinical performance expectations (Seldomridge and Walsh, 2006). Additionally, these same tools pose documentation challenges. They may be hard to use, lengthy to grade, and time consuming to complete. They may encourage "global" ratings and use terminology not understood by faculty and students (Loustau and O’Connor, 1999). Holaday and Buckley (2008) present an interesting alternative with their "standardized clinical evaluation tool-kit". This highly individualized and comprehensive guide is also likely to pose difficulty for novice faculty as it requires a thorough understanding of 11 outcome objectives as measured by a complex rating scale. Anecdotal notes, an accepted method for providing a degree of objectivity, are time consuming and take experience to know what to write and then how to use what is written (Walsh and Seldomridge, 2005).

The inability to control the learning environment is another significant issue with student evaluation. Unlike simulations, patient census and assignments will vary from day to day. How do you evaluate the student who doesn’t have the opportunity to shine based on their clinical assignment? Perhaps patient acuity was low; patients were similar in diagnoses, etc (Walsh and Seldomridge, 2005). Similarly, the tenor of the nursing unit and the ease with which the staff accepts the students can impact the experience. How much of the student’s performance should be compensated for based on whether or not they had a good learning environment? How much did the staff impact the student experience? Many nursing programs rely heavily on "novice" clinical faculty for clinical instruction. Novice clinical faculty would include inexperienced full time academic faculty, as well as faculty whose primary job is not in academia (part time, adjunct, etc).
Novice faculty may have limited formal education and experience in evaluation of students and often lack confidence, doubting their ability to fairly evaluate students (Oermann and Gaberson, 2006; Seldomridge and Walsh, 2006). The lack of confidence stems from their experiential base. A novice faculty's main qualification for teaching may be their clinical expertise. Also, part-time faculty often has multiple roles (Sonner, 2000; Speer et al., 2000). From the institutional perspective, they function as good will ambassadors or recruiters, often establishing relationships with students that are more 'peer' like. Additionally, novice faculty tend to be younger and to relate to students in a different way. The novice faculty's experiential focus and multiple roles impact the evaluation process. Novice faculty is reluctant to give low grades because student evaluations might effect their own evaluations and ultimately rehires and/or the ability to obtain tenure (Sonner, 2000; Speer et al., 2000). Similarly, they are reluctant to give low grades to someone with whom they may ultimately work. Novice faculty draws from their clinical base when evaluating students, focusing more on 'how' to complete a task rather than 'why' the task is important. Novice faculty also value efficiency; developing skill competence may be more important than understanding the theory behind the skill (Ferguson and Calder, 1993). Using an example of administration of heparin with regard to concrete thinking versus conceptual evaluation: novice faculty may be pleased if the IV was started efficiently, the infusion rate was set appropriately and documentation written correctly. The seasoned instructor may have additional safety parameters to consider such as the rationale for heparin administration, side effects, corresponding laboratory values and patient teaching in relation to the heparin administration (Seldomridge and Walsh, 2006). Walsh et al. (2008) continues to study the merits of different styles of evaluation tools and their use by preceptors, but as of yet, have not found an acceptable solution. Academic illiteracy is often another problem for novice faculty. Course objectives may need to be translated into a language that has meaning to the clinical faculty as course objectives and performance criteria may mean different things to novice faculty. Again, the interpretation may be more concrete than conceptual.

Of paramount concern to academic nursing faculty is grade inflation (Walsh and Seldomridge, 2005). According to Speer et al. (2000), grade inflation exists when there is a disproportionate percentage of higher scores than student performances warrant. Faculty would rejoice if all students consistently excelled, but when this is not the case, the earned grade should be reflective of the performance, and not the faculty's inability to differentiate levels of performance. Yet, a disparity between didactic and clinical grades exists. Often faculty is faced with a student who performs poorly in the classroom but appears, according to standard clinical evaluation methods, to excel clinically. There are several potential reasons for this grade disparity. Face-to-face evaluations, typical of the clinical evaluation process, lead to lenient grading. It is more difficult to critique a student face-to-face. As such, faculty tends to avoid conflict by awarding higher grades (Colletti, 2000). Combining formative and summative evaluation styles also contributes to the problem (O'Connor, 2001). Faculty should continually ask themselves, how long is the learning curve for students? Unlike exams, which are equally weighted during the semester, there is a tendency to grade on growth rather than equally weighting all experiences throughout the rotation. What is graded most heavily is what occurred most recently. Discipline specific requirements are another factor. Many professional schools deem any letter grade (earned in the major) lower than a ''C'' as unsatisfactory. Therefore, faculty and students expect higher-level performance, particularly as they reach the conclusion of their program. The standard of ''C'' rather than ''D'' as minimally acceptable for work within the discipline further enhances grade inflation. The weak student gets a ''C'' and the average student gets the ''B''. The design of the evaluation tool also impacts the grading process. Equal weighting of objectives can lead to grade inflation because students can succeed overall while missing the ''big picture'' and important components of the course. Other challenges to faculty include the multiple domains which must be evaluated in a clinical setting. The psychomotor domain is easy to evaluate, but the affective and cognitive domains challenge many educators. It is far easier to evaluate sterile technique than it is to assess critical thinking. Faculty cannot observe each student all of the time and rely to some degree, on the student self-evaluation. This lack of omnipresence in the clinical setting also raises issues of doubt about the comprehensiveness of the observation, becoming another contributing factor to leniency (Walsh and Seldomridge, 2005). Most clinical courses have a limited faculty: student ratio. This small group size also impacts the evaluation process. Due to the relatively small number of students in clinical rotations, faculty get to know students and their personal circumstances more than in a classroom setting. Student effort
may take on a larger dimension. This personalization process further contributes to grade leniency (Shoemaker and DeVos, 1999).

**Student concerns**

Students also cite several concerns with regard to clinical evaluation. Students are often nervous when interacting with faculty in the clinical setting and profess that their performance is better when faculty are not directly working with them, proclaiming "You make me nervous"! Also, most students are unschooled in the appropriate use of self-evaluation tools, thinking that unless they indicate high-level performance, that faculty would not evaluate them highly and their grade will be affected. Students are universally grade conscious (Nagle, 1998; Sonner, 2000; Speer et al., 2000). The typical "Generation X" students have been conditioned to receive rewards and accolades, even in the face of marginal performance. There may also be pressure from parents to earn high grades. Students often resemble novice faculty in some aspects of concreteness and skills, thinking that acquiring a skill set should transfer to a good grade. They think, "look what I can do", without regard for the level of understanding or degree of support required. Lastly, students may complain that every instructor has different expectations and grades differently. It is easy to understand their confusion when one faculty member evaluates them very highly and a similar clinical performance yields an average grade with a different instructor.

**Rubrics**

The use of rubrics provides a potential solution to the subjectivity/objectivity clinical grading dilemma. According to Webster’s Unabridged Dictionary 1913, "Rubrica red earth for coloring, red chalk, the title of a law...that part of early manuscripts and typography which was colored red...the directions or rules for the conduct of service, formerly written or printed in red..." Obviously, from its earliest definitions, rubrics related to teachers’ affinity for grading with red ink! A more current definition would be "an assessment tool that uses clearly defined evaluation criteria and proficiency levels to gauge student achievement of those criteria" (Montgomery, 2000 p. 325). Evaluation rubrics have long been used in secondary and higher education to evaluate assignments or tasks, especially papers and presentations. Many nursing faculty may have used rubrics in didactic nursing courses for papers and projects but have not incorporated them into clinical evaluation. However, the standard goal oriented outcomes and criterion referenced objectives typical of clinical nursing courses are in line with rubrics and are actually a logical progression to the development of a clinical grading rubric.

**Benefits of rubrics**

There are many purported benefits of grading rubrics. Rubrics facilitate communication because feedback can be given quickly, fairly, efficiently, and individually. According to Rucker and Thompson (2003), feedback is most effective when given quickly after task completion and students are more concerned about getting information back quickly than getting helpful feedback. Rubrics foster rapid faculty review, as faculty tend to write the same comments over and over. These comments can be incorporated into the rubric as descriptors, thus making clinical educators more efficient in today’s busy clinical environment (Fitzgerald et al., 2007). For example, a nursing student who writes the nursing diagnosis "at risk for pain" or "at risk for infection" when caring for an open-heart surgical patient might receive the instructor response of "This diagnosis is not incorrect but could apply to any surgical patient. What diagnoses are specific to this patient? Continue working on individualization of nursing diagnoses". From these comments stem performance descriptors like "nursing diagnoses are individualized and specific to the stage of patient illness" versus "nursing diagnoses are generic and lack individualization". Actually describing performance behaviors will take the guesswork out of what the student needs to do to get an "A". Also, because feedback can be given in a timely manner, there is less chance that the student will keep repeating the same mistakes. Feedback that focuses on self-improvement is a form of intrinsic motivation. Other rubric benefits include allowing students to note patterns of problems and/or improvements in their work (Arter and Tighe, 2001). If they repeatedly have problems or successes in the same area, a rubric will highlight this. Rubrics allow faculty to communicate specific goals and intentions to everyone involved; everyone knows what is expected of the students, what behavior constitutes what grade, etc. Similarly, rubrics can act as mechanisms of translation for a variety of diverse groups, helping students understand what is expected and helping teachers...
understand what is not understood. This aspect of rubrics can be particularly helpful with first generation college students or with diverse populations for whom the college experience may be unique or foreign. Rubrics also add a dimension to typical student evaluations by highlighting areas of strengths and weaknesses that become evident for faculty; thus allowing modifications of what or how something was taught based on patterns that may emerge (Arter and Tighe, 2001; Stevens and Levi 2005; Truemper 2004). If everyone in class is lacking in the same area, perhaps the objective or expectation needs to modified or explained more clearly. Lastly, rubrics serve as a blueprint for grading. There is less interpretation of what level of performance transfers to which grade.

Rubric development

Once the decision to use a rubric in a clinical course has been made, it is fairly easy to meld existing course objectives and assignments into the components of a rubric. Existing course objectives become the descriptors for the clinical rubric. The assignment description is often taken directly from the syllabus. The scale includes terms used to describe the level of performance and describes how well or poorly the assignment was performed. The skills/dimensions are the labels or subheadings breaking the assignment down into component parts and the performance descriptors are objective descriptions of leveled behaviors from good to bad or high to low. The scale labels selected will frame the whole experience for the student so it is important to choose labels carefully. They need to have meaning for the student and be direct enough so the student understands what they actually mean. Terms selected will vary to meet the needs of the experience. Highest level terms might include exemplary, sophisticated, advanced, distinguished, excellent, outstanding, or exceptional. Mid range levels could be described as competent, proficient, intermediate, average, progressing, capable, adept, or run of the mill, and low level terms could include marginal, novice, developing, emerging, unsatisfactory, unacceptable, substandard, or disappointing (Stevens and Levi, 2005).

In their book, Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning, Stevens and Levi (2005) outline the process for developing individualized rubrics. Reflection is an important step in the process. Reflecting forces a thorough assessment of the purpose of the assignment. Based on these reflections, a listing of behaviors follows. Usually listing is one of the lengthiest tasks in the development of a rubric. Behaviors are initially described at their highest and lowest levels of expectations. Next any middle level behaviors can be discerned. After completion of this step, the behaviors are grouped and labeled, becoming the dimensions of the rubric. Once these steps are completed, the assignment can be applied to a grid or table format.

Fig. 1 illustrates a senior level nursing clinical course objective using the steps outlined above. This particular course has 8 clinical objectives (overall descriptions) related to pathophysiology, nursing process, communication, teaching/learning, research, leadership, collaboration, and professionalism. The example used is from a "task description" for pathophysiology. In the example provided, the overall description of the assignment is the course objective as listed in the syllabus. For this rubric, the terms selected for the scale ranged from exemplary to unsatisfactory. Each performance label was also given a numerical point value to allow for instructor expertise and flexibility in grading. The area of focus or overall skill involved in the assignment is pathophysiology. This is a very detailed course objective and in its entirety has three dimensions: pathophysiology, laboratory interpretation and medication administration, plus 11 leveled behaviors. The example illustrates one level of performance descriptors.

Fig. 2 is an example from the course research objective and Fig. 3 denotes a portion of the professionalism objective.

Rubrics also allow for the flexibility of weighted course objectives. In the examples provided the same objectives are used for senior students whose

<table>
<thead>
<tr>
<th>Exemplary</th>
<th>Proficient</th>
<th>Marginal</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 10 9</td>
<td>8 7 6</td>
<td>5 4 3</td>
<td>2 1 0</td>
</tr>
</tbody>
</table>

Patho

<table>
<thead>
<tr>
<th>Exemplary</th>
<th>Proficient</th>
<th>Marginal</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independently individualizes patho to patient condition</td>
<td>With assistance individualizes patho to patient condition</td>
<td>Struggles to individualize patho to patient condition</td>
<td>Unable to individualize patho to patient condition</td>
</tr>
</tbody>
</table>

**Figure 1** Applies cultural, behavioral, role and pathophysiologic theories utilizing the nursing process in acute care settings.
clinical focus changes based on their clinical site. By simply changing the weighted percentage of that portion of the rubric, that same tool can be used. For example, when seniors are in the critical care rotation, the overall task descriptions of pathophysiology and nursing process are heavily weighted. When they rotate to management focused rotations, the heavier weighting shifts to leadership and collaboration.

Conclusion

The process of sound clinical evaluation is a complex issue. Rubrics provide one possible solution to the concerns of faculty and students related to effective clinical evaluation. Rubrics are also very adaptable; rubrics are what you make them. They can be as specific or as general as needed. They can be evenly or unevenly weighted. They can be checklists or grading tools. They can be all encompassing or encourage lengthy comments. Although tedious to develop, the process itself is very educational for those involved and the end product is an overall improved assignment for faculty and students alike. Faculty become vested in all aspects of their assignments when they walk through the process of rubric development and students know what to expect and what is expected of them. Rubrics are a valid answer to the concerns of faculty and students related to clinical evaluation.

References


