Formative assessment can be fun as well as educational

J. N. Hudson1 and D. R. Bristow2

1Graduate School of Medicine, University of Wollongong, Wollongong, New South Wales, Australia; and
2Peninsula Medical School, Portland Square, Plymouth University, Plymouth, United Kingdom

Submitted 1 July 2005; accepted in final form 29 September 2005

Hudson, J. N., and D. R. Bristow. Formative assessment can be fun as well as educational. Adv Physiol Educ 30: 33–37, 2006; doi:10.1152/advan.00040.2005.—An understanding of the hormonal basis of normal growth and development, including the changes occurring at puberty, is important foundation knowledge for contemporary medical practice in most fields of medicine. A quiz, testing the important physiological concepts of growth and puberty, was designed using the format of the well-known television game “Who Wants to Be a Millionaire.” An evaluation of this formative assessment activity revealed that a cohort of first-year undergraduate medical students valued learning with peers in an enjoyable, interactive environment, where they were able to admit to uncertainties and clarify answers. It also showed that making an educational activity fun need not detract from the focus of giving feedback on learning. Formative assessment, known to produce learning gains in a range of educational settings, is an important activity in contemporary medical education. With a greater emphasis on self-directed learning and less well-defined curriculum boundaries, feedback helps students to understand and apply the important physiological concepts that underpin the practice of medicine.

Formative assessment is an important educational activity because it gives learners feedback while their learning is still taking place. According to Black and William (1), formative assessment “...encompasses all those activities undertaken by teachers, and/or their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged.” Central to all definitions of formative assessment is the concept of feedback, because it informs learners of their present state of learning, so they can plan action to close the gap between their present state of learning and their desired goal(s) (1, 13). For many students in contemporary medical curricula, where problem-based learning (PBL) is a major learning strategy, considerable anxiety can arise from uncertainty about the boundaries of their self-directed learning. Feedback can provide guidance about which concepts are core and give opportunities to modify learning during the learning process. While the key issue in formative assessment is the quality of the feedback given to learners, innovative educational settings that are also fun can add to the learning experience.

In their article on creativity in medical education, Handfield-Jones et al. (3) describe some of the advantages of using innovative techniques to make learning more fun for both teachers and learners. They explain that the use of a new technique to “liven up” a topic increases learners’ attention and concentration and can make both learners and teachers more interested and enthusiastic. They suggest that competition and games are potentially very powerful media to facilitate learning as long as they are used with care. Two health professionals from Australia have described favourable outcomes, namely, positive feedback and greater attendance at hospital educational sessions, when games as well as debates were used to “lighten up” these sessions (5).

Physiology educators, it seems, are also committed to providing enjoyable active learning environments for their students, as several authors have used games to facilitate learning of physiological concepts (4, 10, 11, 14). Howard et al. (4) reported that the game show “Survivor” torched “Who Wants to Be a Physician” in the educational games war, as it responded to student requests for multiple-choice questions (MCQs) to match the Medical Board exam format and sped up the game to cover more material in respiratory physiology. While the Survivor format was thought to promote team-based and collaborative learning, being voted out by the rest of the team could be humiliating for some students. As will be discussed below, care must be taken to ensure that all students are comfortable with the game and that competitiveness does not hinder learning. Zakaryan and colleagues (14) looked to the game “Trivial Pursuit” to design a game format for interactive review sessions in cardiac physiology. They aimed to encourage peer-to-peer instruction as well as team learning, and, as for the other physiology games, students were positive about learning this way. All these articles suggested that an interactive quiz, modelled on games or a television game show known to students, could make medical education fun as well as educational.

Therefore, time was invested to develop a similar activity at Peninsula Medical School (PMS), a relatively new school in the southwest of England, learning from the experience of earlier reports in the literature. Because the quality of feedback given to learners was thought to be key to the success of our formative assessment game activity, considerable time was invested in designing questions that were valid for medical undergraduates and matched the style used in our summative assessments. The context in which our educational game was situated will now be briefly described, before describing its features, implementation, and evaluation.

At PMS, one of the major learning strategies in the first 2 years of the course is PBL. Each PBL case unit focuses on issues affecting health at different stages of the life cycle, and a variety of learning activities are offered to support the learning associated with each PBL case. These activities include plenary sessions, information technology (IT) resources, clinical skill and life science small group sessions, and community practice experiences. The activity described in this educational report took place at the Life Science Resource Centre (LSRC) as part of the “Adolescence” case unit, in the
first year of the course. For each case unit (2 weekly), there were 4 h of “practical” at the LSRC mostly comprising surface, gross, clinical, and radiological anatomy; microstructure and function; and human function. The emphasis was on application of these basic sciences to clinical practice and, where possible, to use the contact time to check on student understanding.

After the “Conception,” “Fetal,” “Infancy,” and “Childhood” case units, where students explored growth at early stages of the life cycle, one of the major learning issues associated with Adolescence was the events and hormonal basis of puberty. Because learning issue feedback suggested that individual PBL groups had tackled the issue of growth and puberty to varying degrees, the “function” session at the LSRC was used to check on, give feedback on, and promote further learning in these topics. An interactive quiz, modelled on the internationally known television game show “Who Wants to Be a Millionaire,” was used to give students a novel learning environment that was fun and reinforced important concepts in human growth and development. The quiz was the focus of the session. Its features are outlined below, using seven guidelines outlined by colleagues (3) for the use of games in education.

METHODS

Time for Preparing and Planning the Session

Writing a quiz with 15 MCQs of gradually increasing difficulty that are valid in terms of clinical practice and meet the desired learning outcomes is time consuming. However, this is offset to some degree by the ease of editing this electronic resource in subsequent years. While the initial few questions were more knowledge based, most tested the application of knowledge to clinical practice. The latter questions were of the style used to test applied medical knowledge, in “progress tests,” at PMS. Progress testing (8) is a longitudinal assessment, particularly suited to PBL curricula, where 4 tests/yr sample each student’s progress toward curriculum outcomes. Students valued practice in this style of questioning to help them prepare for the ongoing summative progress tests. Some sample questions from the quiz are shown in Table 1.

Because assessment was recognized as a powerful drive to learning, many hours were invested in blueprinting (6) and developing the quiz. A study guide, highlighting the learning outcomes and core concepts, was also prepared and posted in advance for students on our managed learning environment, “Emily.” The latter was the commercial software package “Blackboard,” a Web-based virtual learning environment that supports and manages various aspects of teaching and learning. We used “Emily” to distribute course content and learning resources and facilitate communication between staff and students at the dispersed campuses of PMS, in Devon and Cornwall. The study guides remained on “Emily” as an ongoing self-directed learning resource for students and tutors and could be reviewed and updated as required.

Review of the Characteristics of the Learners and Their Needs

The formative assessment activity had to meet the needs of a diverse group of learners, in terms of age and experience, in the first-year cohort of our undergraduate medical curriculum

<table>
<thead>
<tr>
<th>Question No.</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Which ONE of the following is true of GH?</td>
</tr>
<tr>
<td>1</td>
<td>A. It acts directly on target cells to bring about its growth producing actions</td>
</tr>
<tr>
<td>2</td>
<td>B. It is regulated by one hormone from the hypothalamus</td>
</tr>
<tr>
<td>3</td>
<td>C. Bursts of GH secretion occur in response to exercise, stress, and hypoglycemia</td>
</tr>
<tr>
<td>4</td>
<td>D. Positive feedback loops participate in the regulation of GH secretion</td>
</tr>
<tr>
<td>4</td>
<td>Overproduction of GH in childhood results in which ONE of the following?</td>
</tr>
<tr>
<td>5</td>
<td>A. The condition known as gigantism</td>
</tr>
<tr>
<td>6</td>
<td>B. The short stature seen in African pygmies</td>
</tr>
<tr>
<td>7</td>
<td>C. The condition known as acromegaly</td>
</tr>
<tr>
<td>8</td>
<td>D. The condition known as dwarfism</td>
</tr>
<tr>
<td>10</td>
<td>A 12-yr-old boy, who is of short stature, comes to see you, with his mother. The lad displays some of the secondary sex characteristics appropriate for a male. He asks whether he will grow any more, and you explain that his potential for further growth can be most easily assessed by which ONE of the following?</td>
</tr>
<tr>
<td>9</td>
<td>A. An X-ray of his forearm</td>
</tr>
<tr>
<td>10</td>
<td>B. An X-ray of his knee</td>
</tr>
<tr>
<td>11</td>
<td>C. His blood GH levels</td>
</tr>
<tr>
<td>12</td>
<td>D. His blood testosterone levels</td>
</tr>
<tr>
<td>11</td>
<td>A 29-yr-old weight lifter asks if you can prescribe him GH because he has heard it can promote muscle mass. You advise against its use, explaining that if he uses GH for an extended time, or in large doses, he may suffer from which ONE of the following side effects?</td>
</tr>
<tr>
<td>12</td>
<td>A. A dramatic increase in height</td>
</tr>
<tr>
<td>13</td>
<td>B. Increased fat deposition</td>
</tr>
<tr>
<td>14</td>
<td>C. Increased likelihood of diabetes</td>
</tr>
<tr>
<td>15</td>
<td>D. An increase in his sex drive</td>
</tr>
<tr>
<td>13</td>
<td>Fred Jones, a 24-yr-old body builder, and his wife, Sue, have been trying unsuccessfully to have a second child for the past 9 mo. Both are in excellent health, and their 2-yr-old daughter was easily conceived 2 mo after Sue ceased the pill. They have intercourse 3–4 times/week without any mechanical problems, and Sue has been having regular periods, with her usual premenstrual symptoms. Unknown to his wife, Fred has been using androgen injections to increase his strength and muscle bulk. Apart from some acne and nipple tenderness, he didn’t note any other significant side effects. Fred was noted to have small soft testicles and a reduced sperm count. Which ONE of the following is most likely to describe this couple’s present failure to achieve conception?</td>
</tr>
<tr>
<td>14</td>
<td>A. The exogenous androgens that Fred is taking have shut off his hypothalamic-pituitary-gonadal axis</td>
</tr>
<tr>
<td>15</td>
<td>B. The exogenous androgens that Fred is taking have stimulated his hypothalamic-pituitary-gonadal axis</td>
</tr>
<tr>
<td>16</td>
<td>C. His testes have been subjected to high temperatures due to his gym “workouts” in tight shorts</td>
</tr>
<tr>
<td>17</td>
<td>D. Fred has always had a low sperm count and the paternity of his first child is doubtful</td>
</tr>
</tbody>
</table>

GH, growth hormone.

(a 5-yr course). Many had previous study in biomedical science at the secondary or tertiary school level before entering medical school, but some students had no prior learning in this area. When a review of the learning issues from each PBL tutorial group suggested “patchy” coverage of growth and puberty, there was a need to consolidate, or introduce to some individuals, important concepts in these topics. The quiz was seen as a vehicle to accommodate all student levels.
Ensuring the Technique Allowed the Educational Objectives to Be Met

Feedback for learning was the key educational objective for this session. The Who Wants to Be a Millionaire quiz, known to everyone by its international television presence, had many features to accommodate review of pertinent questions in a nonthreatening environment, and only small adaptations were needed to make it fit for our purpose. These included the need for both the audience and contestants to commit to an answer and the introduction of group discussion of correct/incorrect answers. This will be further discussed in Ensuring That All Learners Were Involved. The following brief summary is provided for those readers who are novices to the format of the Who Wants to Be a Millionaire quiz:

- For each group of students, one student earns the right to become the first contestant (take the “hot seat”) by answering and explaining the answer to the first question correctly.
- Once in the hot seat, the contestant continues answering questions until they are unable to choose and explain the correct answer to a question. They are then replaced with a new contestant.
- When uncertain, contestants have three lifelines (assistance) to help obtain the correct answer. They may ask a friend in the group; ask the audience or whole group; or have two incorrect answers removed, narrowing their choice. These lifelines are available only once to each contestant.
- Prizes are available at various stages, after nominated numbers of questions are answered. Question difficulty increases as the quiz continues, culminating in the million-pound question, number 15. Needless to say, the winning student does not receive a million pounds!

Ensuring Competitiveness Doesn’t Hinder Learning

Many medical students are naturally competitive, and this can be harnessed to help drive the learning from the quiz. While the element of competitiveness, and the manner in which the quiz is conducted by the “quizmaster” (in this case, the tutor) is part of the fun of the quiz, care must be taken that all students are comfortable and enjoying this aspect. No one is asked to go in the hot seat unless willing, and focusing on why an answer is correct, rather than who gets it right or wrong, can help ensure that competition is not used in a negative way.

Making It Simple

Familiarity with the quiz format made it easy to design and administer. However, at the start of each session, the ground rules were explained to ensure all students knew how it would proceed. Only one style of question (scenario, followed by the choice of one best answer) was used, and that was in keeping with the MCQs used in the progress test at PMS. The quiz was prepared so that each question was contained on one PowerPoint slide, and each question could then be projected for easy viewing for all. The study guide was also open on PowerPoint for easy referral during the discussion to answers, if required.

Ensuring That All Learners Were Involved

Each group member had four cards, labelled A, B, C, and D, and for each question everyone had to commit to an answer by separating out the card that represented their answer to the MCQ. After the contestant in the hot seat declared their choice and why they chose it, all group members were invited with a show of hands to declare which answer they had chosen and why. This ensured that all learners were actively involved in each question. The animated discussion of why answers were correct or not made for a rich, interactive learning environment. For students who had prepared for the session, they were rewarded by success in the hot seat or by feedback on the correct answers. Those with poor knowledge, or an inability to apply it, learned from the discussion and the questions of their peers or the tutor. A major role for the quizmaster was to set a safe environment where students were able to admit uncertainty.

Evaluating the Method/Technique and Modifying It as Required

The quiz was evaluated using a questionnaire that was administered at the conclusion of the session. Using a Likert scale from 1 (totally disagree) to 5 (totally agree), students were invited to make a value judgement in relation to five positive statements (see Table 2) in relation to the quiz and its learning objectives. They were also invited to make spontaneous comments on the best feature of the learning activity and how the session may be improved to complement the quantitative information. The responses to the open-ended questions were subjected to a content analysis to determine central themes. Two coders (the authors) independently analyzed the qualitative data for the content analysis. Themes that emerged from the written data were organized into mutually exclusive categories, and two tables (Tables 3 and 4) were developed to express the themes and ideas of the participants.

Student sample. All first-year students who attended the LSRC session on Growth and Puberty during the Adolescent case unit were invited to complete the questionnaire immediately on completion of the interactive quiz. The questionnaires were administered to, and collected from, each group by the LSRC technician at each site; 108 students completed the session and the questionnaire. One questionnaire was not included in the analysis because the comments made it clear that the student took the opportunity to give feedback on another session. Thus, the total sample was 107 students.

Tutors. Two tutors acted as the quizmaster (facilitator) of the four small group sessions (12–13 students/group) at each of

Table 2. Means and SDs of student responses to five questions about the quiz session

<table>
<thead>
<tr>
<th>Statement</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I enjoyed this small group quiz session on growth and puberty</td>
<td>4.56</td>
<td>0.55</td>
</tr>
<tr>
<td>The session built on my previous learning in relation to growth and puberty</td>
<td>4.10</td>
<td>0.83</td>
</tr>
<tr>
<td>Exploring the topics in a collaborative environment with a group of colleagues was useful for learning</td>
<td>4.37</td>
<td>0.64</td>
</tr>
<tr>
<td>The facilitator provided a learning environment in which I felt confident to explore my understanding of growth and puberty</td>
<td>4.40</td>
<td>0.67</td>
</tr>
<tr>
<td>The study guide “Growing Up” was useful for learning</td>
<td>3.92</td>
<td>0.86</td>
</tr>
</tbody>
</table>

The total number of students polled was 107. The Likert scale was used, where 1 is totally disagree and 5 is totally agree.
Table 3. Summary of student responses to the following question: What was the best feature of this learning activity?

| Theme of Student Responses                                      | Number of Student Responses (%)
|------------------------------------------------------------------|----------------------------------
| Fun and enjoyable learning                                      | 32 (30)                          
| Opportunity to explore and explain answers/admit uncertainty    | 16 (15)                          
| Group learning, collaborative teamwork                         | 14 (13)                          
| Humorous responses                                             | 14 (13)                          
| The style of learning (quiz)                                   | 9 (8)                            
| Relaxed and informal learning environment                      | 8 (7)                            
| Checked and gave feedback on learning                         | 8 (7)                            
| Novel/different approach to learning and testing knowledge      | 7 (6)                            
| Stimulated preparation for learning/motivated learning          | 5 (5)                            
| Interactive and supportive learning environment                 | 5 (5)                            
| Building on/using PBL                                          | 3 (3)                            
| Attributes of the tutor                                        | 2 (2)                            
| Preferred learning style                                       | 1 (1)                            
| Knowledge of topic                                             | 1 (1)                            

The number of students polled was 107. The total number of student responses was 125, because some students made more than one comment. PBL, problem-based learning.

RESULTS

Table 2 shows that the majority of students enjoyed the quiz session, agreeing that it built on their prior learning in growth and puberty and that facilitators successfully built a collaborative environment for exploration of these topics. There was a less positive response for the utility of the study guide for learning. Close examination of the raw data revealed that, although most students agreed it was useful for their learning, many checked the “don’t know” option, claiming they were unaware that it was available in advance on “Emily.”

The following student comments spontaneously offered about the best features of the session (examples from Table 3) give greater confidence to the findings summarized above:
● Group learning promoted sharing of knowledge.
● The collaborative reasoning.
● It’s good to draw on each other’s advice/knowledge.
● Working as a team to discuss why the wrong answers were incorrect.
● Getting answers wrong was okay—you could learn from these errors.
● Did not feel under pressure about knowing the answers.
● I felt at ease even though my knowledge was being tested.
● Informal way of learning and recalling knowledge.
● Learning in a fun, relaxed atmosphere is beneficial.
● Everyone was involved whilst having fun and learning.
● It was more fun than self-directed learning.
● Working in an interactive fashion and helping colleagues.
● Being able to prepare thoroughly and then test my own knowledge.
● It makes you want to go away and learn it more thoroughly.
● It stimulated learning and I feel I will remember this session.
● Revision that’s fun and memorable.
● It was surprising to see how much I knew.
● If you are a visual learner, learning in this way is very beneficial, as it will reinforce your learning.

Constructive comments to improve the session (examples from Table 4) included the following:
● Extremely helpful—could be a longer session perhaps.
● Not enough time to discuss wrong answers.
● Always have it at the end of a case unit.
● Prepare beforehand (us!).
● Bit more preparation by students.
● Please be sure everyone knows [of] the existence of a study guide.
● Greater awareness of stuff on Emily.
● I would work harder if there were Mars bars (KING SIZE) for prizes.
● More in depth questions.
● Possibly need to get through the questions faster.
● A closer resemblance to the progress test would be [of] more benefit.

DISCUSSION

Feedback has been called the life blood of learning and is thought to be particularly beneficial if provided frequently and under conditions that are stress free and conducive to learning (12). This study has shown that first-year undergraduate medical students in a PBL-based curriculum valued a formative assessment activity that was fun, nonthreatening, and gave them feedback on their learning. Students highlighted the benefit of having the chance to express and clarify misunderstandings.

This learning activity was offered early in the implementation of a new PBL curriculum because it was noted that students (and some staff) had a level of anxiety about student learning. Students received some feedback on their performance after each of the three monthly, summative progress tests but were requesting additional, and more detailed, feedback on how their medical knowledge was developing. More formative assessment seemed in order. Indeed, formative assessment has been described by a first-year student at an Australian medical school as a “safety net in a self-directed learning course” (12).

Feedback was also important for staff to review the teaching and learning activities in the new curriculum as well as to diagnose any misconceptions in relation to the physiological concepts underpinning the session. As Michael et al. (9) have demonstrated, students may have “undiagnosed” misconceptions from previous learning experiences or resources, and instructors

Table 4. Summary of student responses to the following question: How may the session be improved?

| Theme of Student Responses                                      | Number of Student Responses (%)
|------------------------------------------------------------------|----------------------------------
| Request for more sessions/longer time                           | 29 (27)                          
| No improvement to session required                              | 15 (14)                          
| Humorous responses                                              | 12 (11)                          
| Level of difficulty/content of questions                        | 8 (7)                            
| Study guide availability and accessibility                      | 5 (5)                            
| Session timing/number of questions                              | 4 (4)                            
| Need for students to do more study/revision/preparation for the session | 4 (4)                            

The number of students polled was 107. The total number of student responses was 77.
need to use contact time with students to detect and correct these misunderstandings. Michael and colleagues (9) proposed the creation of active learning environments to deal with this issue, and the interactive quiz was one such opportunity for students and teachers to discover any misconceptions. The quiz also showed that, although some students had learning gaps in relation to human growth and development, they were able to use the session to help close the gap between their present state of learning and the desired goal(s), and many were already competent in applying their knowledge of important concepts in growth and puberty. In addition, several students noted the learner’s own role and responsibility in achieving their learning goals.

The use of electronic “keypads” would have facilitated analysis of group answers to each question in this type of formative assessment quiz and perhaps further encourage involvement of all students. Because the considerable resources needed to implement this were not available, the answer cards provided a simple and effective alternative to ensure that all students were engaged in the activity, a feature that was critical to the success of the quiz.

To improve the response rate, students were asked to complete the evaluation questionnaire at the end of the quiz, before they moved on to the next activity in the life science resource session. This resulted in an excellent response rate in that all students completed the quantitative responses, and a high proportion of students offered spontaneous responses to the open-ended questions. Almost all commented on the best feature(s) of the learning activity, and 77% suggested how the activity could be improved. The failure of the remaining students to suggest improvements may have been due to satisfaction with the activity or that they were unable to think of a constructive suggestion in the time available. It was pleasing, however, that a larger number than usual took the time to give qualitative feedback.

Because of the human nature of researchers, coding errors in content analyzes can only be minimized, not eliminated. According to Johnson and LaMontagne (7), improved interrater reliability is obtained by the use of two independent raters, who each code at least 10–15% of the same responses. In this report, two raters independently coded 100% of the same student responses to improve the reproducibility of coding. Student quotations are included to improve dependability of the data analysis. The use of positive statements in the short questionnaire that generated the quantitative data make it potentially vulnerable to a response bias, but the qualitative data do add to the convergent validity of these quantitative findings. The qualitative data were also useful to reveal the different perspectives of individual students. Important among these were comments about the managed learning environment, “Emily.” Mastery of the latter was important at PMS, because it was the medium of communication between staff and students at dispersed learning sites. Case unit learning materials, including LSRC learning objectives and study guides, were posted in advance for each of the sessions. It was somewhat worrying that 5 mo into the course, some students were not able to find such resources on “Emily.” This was a useful outcome of the evaluation because steps could then be taken to remedy this with a further hands-on session on the use of “Emily.”

Zakaryan et al. (14) have made a good suggestion, namely, that graduate or senior students be used as game moderators to enrich their own knowledge and develop their teaching skills in active learning. At PMS, students have successfully been used to write the answers to formative assessment items (2), and facilitating interactive learning sessions as a game moderator is another way in which students can potentially contribute to assessment and peer learning as well as learning themselves.

As many of the articles referred to in this paper have shown, games can make a valuable contribution to learning and assessment, and further use of this or other innovative game techniques seems warranted to improve and maximize learning. We highlight the need to carefully prepare the learning setting so games can facilitate, and not detract from, the achievement of desired learning outcomes. Care is also need to prepare an assessment that has content validity for vocational education in medicine. As Howarth-Hockey and Stride (5) remind us, when students interact with and learn from each other, they are trained to view peers as potential resources for future learning experiences. This is a desirable outcome for medical education, and, potentially, formative quizzes and games could also be used to promote interprofessional learning among health professionals.

ACKNOWLEDGMENTS

The authors thank Julie Belka and Deb Kirvell, in Exeter and Plymouth, respectively, for entering in to the spirit of the quiz and obtaining prizes to motivate and make it fun for students and to the first cohort at the Peninsula Medical School for completing the evaluation.

While J. N. Hudson is now an academic at the University of Wollongong, the formative assessment activity took place at the Peninsula Medical School in 2003, when she was a foundation staff member of this new medical school.

REFERENCES