



Simulation and education

Game-based assessment of first aid and resuscitation skills[☆]

Nathalie Charlier*

Faculty of Pharmaceutical Sciences, Katholieke Universiteit Leuven, Campus Gasthuisberg O&N2, Herestraat 49, Box 522, 3000 Leuven, Belgium

ARTICLE INFO

Article history:

Received 8 July 2010

Received in revised form 5 November 2010

Accepted 7 December 2010

Keywords:

Peer-assessment

Training

Game-based

Basic life support (BLS)

Cardiopulmonary resuscitation (CPR)

ABSTRACT

Aim: Effective assessment of motor skills in large-size classes is a challenge in medical education. This case-study investigates whether a game can be considered a valid tool for the summative assessment of first aid and basic life support skills.

Methods: Using a traditional exam as bench-mark, a board game format was experimentally trialed to assess students' competency after taking a first aid course. Fifty-five students were randomly assigned to two groups. Two assessments, a game-based assessment and a traditional test, consisting of a paper-and-pencil test in combination with a skill assessment, were applied to both groups in opposite order. In both formats students acted as judges of other students' efforts. In the game, the student's outcome was equal to the number of cards collected by answering questions correctly as deemed by peers. Similarities between both assessment types included individual testing, type of assessor (peers), content, type of questions and demonstrations, and the use of checklists for skill assessment. The assessment methods differed in format (written or oral test, both in combination with skill assessment) and feedback availability.

Results: Both groups performed equally in the game-based assessment as well as in the traditional test, in spite of the opposite order of the assessments. No significant difference was found between the mean scores on the game-based assessment and the traditional test.

Conclusions: These data suggest that use of a game format for assessment purposes may provide an effective means of assessing students' competence at the end of a practical course.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Most cardiac arrests occur outside the hospital setting where, in most cases, the general public will be responsible for providing initial basic life support (BLS). Incorporating BLS training in the school curriculum would maximize the number of potential cardiopulmonary resuscitation (CPR) providers in the community.^{1–5} Studies have shown that teachers are able to train pupils successfully in CPR.⁶ With this in mind, training the whole school population can be considered a feasible option for disseminating the importance of acquiring CPR skills.

In designing a first aid training, one major challenge is the assessment of competency (knowledge, skills and attitudes (KSA)). Testing technical skills requires methods and assessment instruments that are somewhat different than those used for cognitive skills.^{7,8} Regarding the assessment of technical skills, simulations – manikin simulations,⁹ standardized patient simulations^{10,11} and computer-based simulations¹² – are increasingly being used in medical education to ensure that examinees can demonstrate inte-

gration of prerequisite KSA in a realistic setting.¹³ However, a major drawback of computer-based simulations is that in most cases they do not provide the opportunity to conduct a real physical examination or demonstrate motor skills, such as first aid.^{14,15} Complex manikins are used to realistically simulate clinical cases, but here users are restricted to conducting physical examinations other than those for which the manikins are designed for.¹⁶ Specially trained actors – referred to as “Standardized Patient” (SP) – portray patients with particular health concerns and are able to answer the full spectrum of questions about their condition.¹⁷ However, because of the high costs for training, students are not exposed to a large number of cases and the encounters are often only used for summative assessment and not as formative learning activities.¹⁸ As a consequence, alternative training and assessment methods that are economically and logistically feasible need to be explored.

1.1. Peer assessment and feedback

Peer assessment is a process wherein peers evaluate each other's work, usually along with, or in place of, an expert marker.¹⁹ Peer assessment can be used as a tool for social control, for assessment, for active participation, as learning aid for how to assess.²⁰

Using peer assessment has potential benefits for the assessee and the assessor, for instance sharing the observation or reading

[☆] A Spanish translated version of the summary of this article appears as Appendix in the final online version at doi:10.1016/j.resuscitation.2010.12.003.

* Tel.: +32 16 32 34 53.

E-mail address: Nathalie.Charlier@pharm.kuleuven.be

burden among multiple assessors since students are considered to be 'surrogate' or 'assistant' teachers.²¹ Concerns however exist on the validity and reliability of peer-generated grades. Over 70% of the studies find the reliability and the validity of the peer assessment adequate as compared to expert graders,^{22,23} whereas a minority find them variable.^{19,24} Nevertheless, several studies conclude that concerns about peer evaluation reliability and validity should not be a barrier to implementing peer evaluations, provided that appropriate scaffolds are applied, e.g. providing guidelines, checklists, or other tangible scaffolding to students, introducing decision making by teams instead of individual peers, etc.^{24–27}

Peer assessment can also prove to be helpful in terms of providing feedback. Peer feedback can confirm existing information, add new information, identify errors, correct errors, improve conditional application of information, and aid the wider restructuring of theoretical schemata.²⁸ Although peers are not experts in the domain, their feedback can be a trade-off against expertise in terms of being understandable, timely, frequent, extended, individualized and reassuring.²⁹

1.2. Game-based assessment

In general, summative assessment induces a high degree of stress/anxiety, which may have debilitating health, emotional and educational effects.³⁰ To overcome these issues game approaches may be used. Games are interactive, motivating and challenging.^{26,31} They provide an opportunity to integrate and demonstrate knowledge and skills, and they give direct feedback. To date, researchers tend to concentrate on the use of games to enhance student learning and delivery of courses.^{32–34} Studies have been focused on using games as formative assessment instruments.³⁵ Little research exists on the use of games as a summative assessment tool.

The research question in this study was: Can a board game serve as a valid alternative tool for the summative assessment of first aid competence?

2. Methods

2.1. Participants

The sample consisted of students taking the one-year preservice teacher training programme during or after their Master's degree. The course, a first aid training module (including BLS), is part of a compulsory, one-semester course (4 ECTS). The module is organized in four sessions of 4 h during four consecutive weeks. Summative assessment – the focus of this study – took place during the last session. All students freely signed a consent form agreeing on using their results for research purposes.

2.2. Procedure

In this study, the validity of a game as alternative assessment tool has been investigated. Therefore, a traditional test consisting of a written test and a demonstration of CPR and two first aid skills was used as a bench-mark. As a consequence all participants were assessed twice, by both a traditional test and a game-based assessment. Since both scores were included in the final marks, a possible impact of one assessment on the outcome of the other assessment need to be investigated. We randomly assigned participants to two groups, receiving the assessments in opposite order. Group A ($N=28$) started with the game, group B ($N=$ initially 27) with the traditional test. Supervision of each group was performed by an assistant who was not involved in the course itself. Both assistants were familiar with the game and had expert knowledge on the content of the course, but were instructed only to supervise and mark

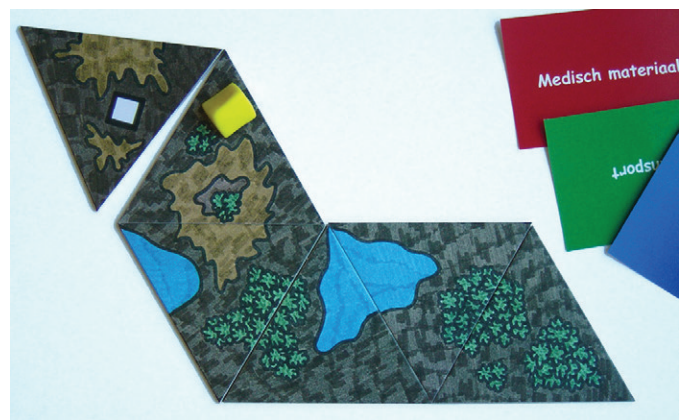


Fig. 1. (Top view) Example of the game showing five tiles with the following features: swamps (pale brown), fields (dark brown), seas (blue), bushes (green) and a building lot (white square). Another white square is covered by a first aid post of one of the players (yellow). On the top right corner some question cards are shown.

irregular observations. At the start of the game-based assessment, the rules of the game were orally explained and participants were randomly assigned to four-member groups (Group B contained 1 three-member group). Participants were instructed to perform the demonstrations during the game in a central area, named the 'skill zone'. A video camera was set up to record all demonstrations and served as a control instrument for the assistant in case of irregularities (e.g. discussions, cheating). Since the first aid game was new to all players, a test round of 15 min was played to master the game rules. Subsequently, groups were given 90 min to play the game. Players were instructed to finish the last round to assure an equal number of answered questions or demonstrations by each player within their own group.

2.3. The game

Due to the absence of a first aid board or digital game, we developed our own educational game. Several trials testing the game were conducted beforehand to optimize the game, and the questions and demonstrations. Given the positive effects of peer assessment and peer feedback, we opted to apply these principles in the game.

The game board is a landscape of a developing country built by the players as the game progresses. The game starts with a single terrain triangular tile face up and 69 others shuffled face down for the players to draw from. On each turn a player draws a new terrain tile and places it adjacent to the tiles of the progressing landscape in such a way that it extends features on the tiles it abuts: swamps must connect to swamps, fields to fields, seas to seas and bushes to bushes (Fig. 1). After placing the new tile, that player chooses a blue, red, green or brown card with a true/false, short answer, performance and open essay question respectively. This question is read aloud by a competitive peer player. If the answer or performance is deemed as being correct by the peers, the player keeps the question card. If not, the card is placed aside. As soon as a player has collected a blue card, a red card and a green card, he is allowed to station a first aid post. Building is only allowed on a specific feature on the tile marked by a white square. If this feature is present on the newly placed tile, the player may opt to station a first aid post in exchange for the three collected cards. In a later round the player can transform his first aid post to a hospital in exchange for a brown question card. Players are able to sabotage one another by natural disaster. Six of the 70 tiles display a malaria epidemic, a bush fire or a tsunami. They can be used to destroy a first aid post of a competitive player. Hospitals cannot be destroyed by disasters.

At the end of the game, all first aid posts (1 point) and hospitals (3 points) placed by each player are counted. The player collecting the most points wins the game.

2.4. Research instruments

2.4.1. Group A: game-based assessment

The outcome of the game (number of first aid posts and hospitals) was not used as assessment score, due to a factor of luck (i.e. the chance of drawing a tile with a white square) and sabotage (by natural disaster). Instead a summative individual score was generated by comparing the amount of individually collected question cards (representing the level of acquired knowledge and skills) to the number of play rounds. To overcome concerns about peer evaluation reliability and validity, checklists indicating the order and description of each individual step of the (complex) manipulation were used in the peer skills assessment (Appendix 1 published online only). These checklists have been used extensively during the first three training sessions of the course in order to familiarize students with these instruments. Also, the final decision had to be made via consensus by three peers, increasing the objectivity of the peer-assessment.

2.4.2. Group B: traditional test

Participants had 90 min to complete a paper-and-pencil test and demonstrate three skills. The written test consisted of five true/false, four short answer and three essay questions, assessing their knowledge on first aid. At the end of the exam, all copies were collected and redistributed among the students. Each student assessed the written test of a peer using a mark of 0 or 1 for each item, which was consistent with the scoring procedure in the game.

For the skill assessment, groups of two students were individually and in private assessed on their first aid skills within the timeframe of the written test. Under the supervision of the evaluator each student performed the skill on a peer student who subsequently scored the performance using the checklists mentioned previously (Appendix 2, published online only).

To assure validity, the question items and demonstrations were direct translations from a standard work written by emergency medicine experts. The questions covered all course objectives and comprised the different levels of the cognitive, psychomotor, and affective domains.³⁶ Furthermore, similar, but not identical questions were used in the traditional and game-based assessment.

2.5. Data analysis

All quantitative data were listed using the SPSS 16.0 software for statistical analysis. Descriptive statistics were calculated for the two groups and the two assessments to be compared. Students' scores on the two assessments were compared using a Wilcoxon Signed Ranks test. A Mann–Whitney *U*-test was performed to evaluate the mean scores of both groups on each assessment method separately. A Pearson correlation was measured between the individual students' scores on the game-based assessment and the traditional test.

3. Results

3.1. Population demographics

Fifty-five students (42 women and 13 men) participated in this study. The sexes were equally distributed among the two groups. During the review of the peer-grading of the traditional test, a total score for two students was lacking. Therefore, we excluded the 2 students (female) from the analysis. The following data apply to the remaining 53 students.

The mean age was 23.08 (± 1.89) years old. At the beginning of the course 38 participants (71.70%) held a master's degree, 15 participants were master students. Educational background of the participants varied from 'science, engineering & technology' (24.53%), 'humanities & social sciences' (32.07%) to 'sports science' (43.40%). None of the students had participated in a first aid course prior to the study. All sport science students and graduates received an introduction on CPR during their master education. However, no significant difference was found between the scores of the sport science group and the non-sport science group on each of the assessments ($p_{\text{game}} = .267$; $p_{\text{trad}} = .646$).

3.2. Game versus traditional test outcomes

In group B one outlier value was identified for the traditional test outcome. However, excluding the outlier case did not result in a change of statistical significance in the analysis and was therefore included in the analysis.

Students played an average of 12 rounds (median = 11, range = 6), answering or demonstrating an equal amount of questions and demonstrations within their group.

The mean scores of group A (game-based assessment first) were higher than those of group B (traditional test first) on each of the two assessments, but the difference was not statistically significant as revealed by the Mann–Whitney *U*-test on either the game-based assessment ($p = .442$) or the traditional peer-graded test ($p = .057$). These results show no significant impact of the order of the assessments on the outcomes.

Both groups scored higher on the game-based assessment than on the traditional test. A Wilcoxon Signed Ranks test revealed however no significant difference in mean scores between the two assessments within each group ($p_{\text{group A}} = .084$ and $p_{\text{group B}} = .109$), indicating no significant advantage of a previous assessment on the outcomes of the next.

A significant correlation between the individual outcomes of the game-based assessment and the written test is observed (Pearson correlation = .305, $p = .026$).

Detailed results are shown in Table 1 and Appendix 3 (published online only).

4. Discussion

Mean scores of the game-based assessment did not significantly differ from those of the traditional test, suggesting that the board game can serve as a valid alternative assessment instrument. In this study, several factors may favor the choice of the board game over the traditional exam.

While traditional tests are most likely to induce a high degree of stress/anxiety resulting in a poor performance, games can be fun, motivating, and challenging and therefore able to dispel some

Table 1

Assessment score percentages of both groups on the traditional test (peer-graded) and the game-based assessment.

	Score traditional test	Score game-based assessment	
Group A (N = 28) game – traditional test	M = 89.17 SD = 5.16 Min = 77.10; Max = 95.85	M = 92.54 SD = 7.55 Min = 77.80; Max = 100.00	$p = .084^b$
Group B (N = 25) traditional test – game	M = 85.46 SD = 8.57 Min = 55.20; Max = 95.85 $p = .057^a$	M = 89.14 SD = 12.09 Min = 63.65; Max = 100.00 $p = .442^a$	$p = .109^b$

^a *p*-Value of the Mann–Whitney *U*-test.

^b *p*-Value of the Wilcoxon Signed Ranks test.

fear of examinations.^{26,31} In this study, winning the game is not necessarily equivalent to an excellent performance in the assessment. Players who do not like playing games and who have no insight or strategy could have performed poorly in building their first aid posts and hospitals, resulting in a low game result. To avoid this we opted to disconnect the results of the game play from the assessment score.

Secondly, feedback has proven to induce a positive effect on learning.²⁸ Both groups obtained a higher mean score on the game-based assessment in comparison to the traditional test, suggesting a positive effect of feedback. In addition, we would expect that students of group A (game-based assessment first) would benefit from the feedback in the subsequent traditional test. In contrast, a decrease in mean scores (however not significant) on the traditional test was shown. These observations however need to be interpreted with care. First, the quality of peer feedback has not been measured. The results could be attributed to poor (in quantity and quality) feedback. Simply telling a student that his answer is wrong does not help him perform very much better on a second attempt.³⁷ Related to this, questions in the traditional test, although covering the same content, did not reappear in the game-based assessment. As a result, a learning effect might have been present although not observed, due to the absence of second attempts.

Finally, using a game format creates the opportunity of assessing large-size classes simultaneously. Peer assessment might be an alternative to overcome costly and time-consuming assessment methods. To overcome concerns of reliability, observation matrices, extensively used during training, served as skill assessment instruments. Also a final decision was made by at least two peers to increase accuracy, as was suggested in the literature.²⁶

There are some potential limitations to this study. First, the number of participants in this study is relatively low, due to its implementation in an existing course. Some of the *p*-values lean to a near-significance level which might be caused by a low statistical power. The focus point of this case-study was the implementation of an alternative assessment instrument in a first aid course for students in the teacher training. All students taking the course participated, resulting in an ecological valid study. Secondly, the target group addressed is a limitation for the generalization of findings in this case-study. Future studies using various settings such as secondary school students, emergency medicine personnel, first aid training for laypersons will provide more insight in the applicability and suitability of this method. Alternative settings can involve larger populations which creates opportunities to also study specific student characteristics (e.g. gender). Thirdly, this case-study was implemented in an existing program of 16 h. Hence, a pre-test measuring prior knowledge could not be taken from the participants. Instead a questionnaire investigating participants' experience of BLS and first aid was handed out. To measure gains in knowledge and performance, a before–after intervention-format should be conducted in future studies. Fourthly, the measurement of first aid competencies in a simulated situation such as a game-based assessment may not necessarily reflect performance in a real-life experience. In addition, the effect of a summative assessment format on the presence of first aid competencies at a later time point remains to be determined. Finally, the results of this study are produced using a particular board game. Prudence is in order when extrapolating these results to other game-based assessments, since every game has its own specific characteristics, such as presence of feedback.

5. Conclusions

No significant difference was found between the game-based assessment and the traditional test outcomes, indicating that the game format applied in this study might be a valid alternative

assessment instrument and might provide an effective means of assessing student competence at the end of a practical course.

Further research is needed on the use of games as summative assessment instruments. While many studies are available on the quality assessment of games, the use of games as learning and teaching tools and the use of games in formative assessment, there is a lack of literature on hands-on experience and lessons learned from research implementing games as summative assessment tools.

Conflict of interest statement

None to declare.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.resuscitation.2010.12.003.

References

- Lester CA, Weston CF, Donnelly PD, Assar D, Morgan MJ. The need for wider dissemination of CPR skills: are schools the answer. *Resuscitation* 1994;28:233–7.
- Lester C, Donnelly P, Weston C, Morgan M. Teaching schoolchildren cardiopulmonary resuscitation. *Resuscitation* 1996;31:33–8.
- Lubrano S, Romero S, Scoppi P, et al. How to become an under 11 rescuer: a practical method to teach first aid to primary schoolchildren. *Resuscitation* 2005;64:303–7.
- Miro O, Jimenez-Fabrega X, Espigol G, et al. Teaching basic life support to 12–16 year olds in Barcelona schools: views of head teachers. *Resuscitation* 2006;70:107–16.
- Connolly M, Toner P, Connolly D, McCluskey DR. The ABC for life' programme—teaching basic life support in schools. *Resuscitation* 2007;72:270–9.
- Toner P, Connolly M, Laverty L, McGrath P, Connolly D, McCluskey DR. Teaching basic life support to school children using medical students and teachers in a "peer-training" model. Results of the "ABC for life" programme. *Resuscitation* 2007;75:169–75.
- Vargas AL, Boulet JR, Errichetti A, Van Zanten M, Lopez MJ, Reta AM. Developing performance-based medical school assessment programs in resource-limited environments. *Med Teach* 2007;29:192–8.
- Norcini JJ, McKinley DW. Assessment methods in medical education. *Teach Teach Educ* 2007;23:239–50.
- Perkins GD. Simulation in resuscitation training. *Resuscitation* 2007;73:202–11.
- Reznick RK, Blackmore D, Dauphinee WD, Rothman AI, Smee S. Large-scale high-stakes testing with an OSCE: report from the Medical Council of Canada. *Acad Med* 1996;71:S19–21.
- Whelan GP. Educational commission for foreign medical graduates: clinical skills assessment prototype. *Med Teach* 1999;21:156–60.
- Kneebone R. Simulation in surgical training: educational issues and practical implications. *Med Educ* 2003;37:267–77.
- Tekian A. Assessing communication, technical and affective responses: Can they relate like a professional? In: Tekian A, McGuire CH, McGaghie WC, editors. *Innovative simulations for assessing clinical competence*. Chicago: Department of Medical Education, University of Illinois; 1999. p. 105–12.
- Garret BM, Callear D. The value of intelligent multimedia simulation for teaching clinical decision-making skills. *Nurse Educ Today* 2001;21:382–90.
- Bergin R, Youngblood P, Ayers MK, et al. Interactive simulated patient: experiences with collaborative e-learning in medicine. *J Educ Comput Res* 2003;29:387–400.
- Gordon JA, Wilkerson WM, Shaffer DW, Armstrong EG. Practicing medicine without risk: students' and educators' responses to high-fidelity patient simulation. *Acad Med* 2001;76:469.
- Eagles JM, Calder SA, Nicoll KS, Sclare PD. Using simulated patients in education about alcohol misuse. *Acad Med* 2001;76:395.
- Hubal RC, Kizakevich PN, Guinn CI, Merino KD, West SL. The virtual standardized patient. *Stud Health Technol Inform* 2000;70:133–8.
- Topping K. Peer assessment between students in colleges and universities. *Rev Educ Res* 1998;68:249–76.
- Gielen S. Peer assessment as a tool for learning. Ph.D. Dissertation. Belgium: Katholieke Universiteit Leuven; 2007.
- Forbes D, Spence J. An experiment in assessment for a large class. In: Smith R, editor. *Innovations in engineering education*. London: Ellis Horwood; 1991. p. 97–101.
- Sadler PM, Good E. The impact of self and peer-grading on student learning. *Educ Assess* 2006;11:1–31.
- Pare DE, Joordens S. Peering into large lectures: examining peer and expert mark agreement using peerScholar, an online peer assessment tool. *J Comput Assist Learn* 2008;24:526–40.

24. Falchikov N, Goldfinch J. Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Rev Educ Res* 2000;70:287–322.
25. Cho K, Schunn CD, Wilson RW. Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *J Educ Psychol* 2006;9:891–901.
26. Desrochers MN, Pusateri Jr MJ, Fink HC. Game assessment: fun as well as effective. *Assess Eval High Educ* 2007;32:527–53.
27. Topping K. Peer assessment. *Theor Pract* 2009;48:20–7.
28. Butler DL, Winne PH. Feedback and self-regulated learning: a theoretical synthesis. *Rev Educ Res* 1995;65:245–81.
29. Cole D. Change in self-perceived competence as a function of peer and teacher evaluation. *Dev Psychol* 1991;27:682–8.
30. Cassady JC, Gridley BE. The effects of online formative and summative assessment on test anxiety and performance. *J Tech Learn Assess* 2005;4.
31. Patel J. Using game format in small group classes for pharmacotherapeutics case studies. *Am J Pharm Educ* 2008;72 (article 21).
32. Klassen KJ, Willoughby KA. In-class simulation games: assessing student learning. *J Inform Technol Educ* 2003;2:1–14.
33. McDonald KK, Hannafin RD. Using web-based computer games to meet the demands of today's high-stakes testing: a mixed method inquiry. *J Res Teach Educ* 2003;35:459–72.
34. Bochennek K, Wittekindt B, Zimmermann SY, Klingebiel T. More than mere games: a review of card and board games for medical education. *Med Teach* 2007;29:941–8.
35. Wang TH. Web-based quiz-game-like formative assessment: development and evaluation. *Comput Educ* 2008;51:1247–63.
36. Bloom BS. Taxonomy of educational objectives, handbook I: the cognitive domain. New York: David McKay Co Inc.; 1956.
37. Lee Y-J, Palazzo DJ, Warnakulasooriya R, Pritchard DE. Measuring student learning with item response theory. *Phys Rev Spec Topic: Phys Educ Res* 2008;4, 010102–1–6.