

BMJ Open Effectiveness of serious games and impact of design elements on engagement and educational outcomes in healthcare professionals and students: a systematic review and meta-analysis protocol

Marc-André Maheu-Cadotte,^{1,2,3} Sylvie Cossette,¹ Véronique Dubé,^{1,2} Guillaume Fontaine,^{1,2} Tanya Mailhot,^{1,4} Patrick Lavoie,⁵ Alexis Cournoyer,^{1,6,7} Fabio Balli,⁸ Gabrielle Mathieu-Dupuis⁹

To cite: Maheu-Cadotte M-A, Cossette S, Dubé V, *et al.* Effectiveness of serious games and impact of design elements on engagement and educational outcomes in healthcare professionals and students: a systematic review and meta-analysis protocol. *BMJ Open* 2018;**8**:e019871. doi:10.1136/bmjopen-2017-019871

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-019871>).

Received 3 October 2017
Revised 17 January 2018
Accepted 16 February 2018



For numbered affiliations see end of article.

Correspondence to
Marc-André Maheu-Cadotte;
marc-andre.maheu-cadotte@umontreal.ca

ABSTRACT

Introduction Serious games (SGs) are interactive and entertaining digital software with an educational purpose. They engage the learner by proposing challenges and through various design elements (DEs; eg, points, difficulty adaptation, story). Recent reviews suggest the effectiveness of SGs in healthcare professionals' and students' education is mixed. This could be explained by the variability in their DEs, which has been shown to be highly variable across studies. The aim of this systematic review is to identify, appraise and synthesise the best available evidence regarding the effectiveness of SGs and the impact of DEs on engagement and educational outcomes of healthcare professionals and students.

Methods and analysis A systematic search of the literature will be conducted using a combination of medical subject headings terms and keywords in Cumulative Index of Nursing and Allied Health, Embase, Education Resources Information Center, PsycINFO, PubMed and Web of Science. Studies assessing SGs on engagement and educational outcomes will be included. Two independent reviewers will conduct the screening as well as the data extraction process. The risk of bias of included studies will also be assessed by two reviewers using the Effective Practice and Organisation of Care criteria. Data regarding DEs in SGs will first be synthesised qualitatively. A meta-analysis will then be performed, if the data allow it. Finally, the quality of the evidence regarding the effectiveness of SGs on each outcome will be assessed using the Grading of Recommendations Assessment, Development and Evaluation approach.

Ethics and dissemination As this systematic review only uses already collected data, no Institutional Review Board approval is required. Its results will be submitted in a peer-reviewed journal by the end of 2018.

PROSPERO registration number CRD42017077424.

INTRODUCTION

Education of healthcare professionals and students is of the utmost importance to

Strengths and limitations of this study

- A comprehensive and prospectively registered (International Prospective Register of Systematic Reviews #CRD42017077424) systematic literature review protocol reported according to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols statement.
- An assessment of the overall quality of the evidence regarding effectiveness of serious games (SGs) on each outcome, using the Grading of Recommendations Assessment, Development and Evaluation approach
- The definition of 'SG' in this systematic review is limited by the fact that no consensus prevails in the literature.

promote the adoption of best practices and to improve patient safety. In 2013, the WHO underlined a need to train healthcare professionals to face population health needs and health service delivery adapted to epidemiological and demographic realities.¹ Engagement, defined as a learner's involvement and interest towards an educational intervention, has been positively correlated with educational outcomes such as learning and behaviour change.^{2 3} Designing educational interventions to sustain the engagement of healthcare professionals and students is therefore critical to their effectiveness. As such, the last decade has seen a growing interest from researchers about the use of serious games (SGs) in healthcare professionals' and students' education.⁴

SGs are designed as entertaining and active learning environments which can be

made available on any digital platform (eg, smartphone, computer). Learning in SGs typically occurs through a gameplay that engages the learner in challenges adapted to his in-game skills. Challenges are defined as subjective experiences that solicit the learners' skills.⁵ For example, challenges can require the learner to explore, experiment, compete or cooperate with other learners.^{6,7} Gameplay is defined as a combination of challenges and design elements (DEs).⁸

DEs, which can be seen as building blocks or shared features of games (eg, points, difficulty adaptation, story), are suggested to be instrumental in improving the learner's engagement in SGs.^{4,9,10} Therefore, a SG may be composed of several DEs and these DEs may be similar across SGs.¹¹ A list of DEs, based on a literature review by the authors, is presented in table 1.^{10,12–15} Based on Csikszentmihalyi's flow theory and theoretical propositions of other authors in game design, Pavlas suggests that DEs operate by influencing antecedents of engagement such as: the learner's ability to concentrate on task, his sense of control, the feedback he receives and a deep but effortless involvement.¹⁶ Lameris *et al* underline the difficulty in linking DEs and engagement to specific learning processes in SGs. Nonetheless, it is believed that higher engagement will lead the learner to become deeply involved and to repeatedly take on the challenges offered to improve his in-game performance and, consequently, his educational outcomes.¹⁷ However, while supporting engagement, there are also concerns that DEs can become a source of distraction from the learning content and negatively affect educational outcomes.^{16,18,19}

Previous reviews have highlighted the mixed effectiveness of SGs when compared with other educational interventions (eg, classroom learning) in improving healthcare professionals' and students' engagement or educational outcomes.^{4,20,21} As previous authors pointed out, the integration of an educational purpose to the gameplay is a delicate process that should be considered throughout SGs development, right up to the initial designing phase of the SGs.^{22,23} As such, the mixed effectiveness of SGs in supporting engagement and improving educational outcomes could be explained by the highly variable DEs found in SGs.²⁴ Moreover, the optimal integration of DEs in SGs remains to be discovered.^{22,25} Findings from a recent meta-analysis investigating the effectiveness of SGs on healthy lifestyle promotion underlined the necessity of strong theoretical foundations in designing SGs and the need to further explore which DEs are the most impactful.²⁶ As the development of SGs can be quite expensive and time consuming, describing which DEs have been integrated in the development of SGs for healthcare professionals and students, as well as their impact on engagement and educational outcomes, could help in the efficient development of future SGs.^{20,27} However, a systematic review of the impact of DEs of SGs on engagement and educational outcomes has yet to be published.

Table 1 Design elements in serious games to be assessed

Avatars	The learner chooses or creates a graphical representation of himself in the serious game.
Badges for achievement	A visual representation that serves as a symbol for the learner's achievements
Content unlocking	Access to new aspects of the serious game (eg, higher levels) when certain tasks have been accomplished
Difficulty adaptation	Levels of difficulty of the challenges can be adjusted by the learner (eg, easy, medium and hard) or automatically adjusted to the learner's performance.
Hints	A suggestion or an indication given by the serious game to help the learner achieve a challenge. A learner may choose to receive a hint or the serious game can give hints based on his performance.
Leaderboard	A table or a graph that ranks the learners according to their success based on specific criteria
Performance tables or graphs	A table or a graph that provides information to the learner about the progression of his performance over time
Plot	A narrative discourse that serves to organise the events of a story in a logical or temporal order
Points	Points are awarded or subtracted depending on the learner's performance and serve to numerically represent the progression. Points can therefore serve as a reward and as an immediate feedback tool.
Teams	Learners who work together to achieve a common goal. Teams will be broadly regarded here to encompass either cooperative or versus playing.
Time pressure	A time limit that is allowed for the learner to achieve a specific challenge. A time limit can be illustrated by a countdown timer that indicates to the learner the time remaining.
Virtual goods	Game assets that have a certain in-game value. Virtual goods can sometimes be bought or exchanged using real-world currency.

Therefore, our objectives are twofold: (1) to systematically identify, appraise and synthesise the best available evidence regarding the effectiveness of SGs on engagement and educational outcomes of healthcare professionals and students and (2) to describe which DEs have been integrated in the development of SGs for healthcare

professionals and students and their impact on engagement and educational outcomes.

METHODS

The protocol for this systematic review was developed according to the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P; see online supplementary file 1) and follows the guidance provided by the Cochrane Handbook for Systematic Reviews of Interventions.^{28 29} The protocol was also registered prospectively on the International Prospective Register of Systematic Reviews (# CRD42017077424).

Eligibility criteria

The eligibility of studies will be assessed based on study characteristics, participants, interventions assessed, comparators and the outcome measures for inclusion in the systematic review and the meta-analysis.

Study characteristics

Primary studies published in English or in French from 1 January 2005 to 31 December 2017 will be considered. Previous reviews on SGs have identified a growth in published literature starting mid-2000s,^{24 30} and no experimental study assessing a SG for healthcare professionals' and students' education was found before that date.⁴ As suggested, Effective Practice and Organisation of Care (EPOC) Cochrane Review Group, randomised controlled trials and cluster randomised controlled trials will be included in the current review.³¹

Types of participants

Studies with healthcare professionals and healthcare students from undergraduate to postgraduate education, either in a clinical or an academic setting will be considered. Clinical settings include all environments in which healthcare is provided. Studies conducted exclusively among patients or students receiving non-healthcare-related education (eg, high school students) are beyond the scope of this review as the objectives of the games evaluated in these populations (eg, illness self-management in patients, the learning of mathematical concepts in students) differ from those developed for healthcare professionals and healthcare students.

Types of interventions

Studies assessing a SG as a stand-alone intervention or as part of multicomponent intervention (eg, combined with workshop, classroom or digital-based learning) will be considered. For this systematic review, we define SGs as interactive and entertaining digital software with a primary educational purpose that engage the learner through various challenges.^{5 8 32 33}

Types of comparators

We will consider studies where the comparator is any type of educational intervention.

Types of outcome measures

Studies reporting at least one measure of engagement or educational outcomes will be considered.

We will retain the definition and the indicators of engagement reported by Perski *et al*, who defined engagement as: (1) the extent of the learner's involvement and (2) a subjective experience characterised by affect, attention and interest.³ These two dimensions will be considered individually. Regarding involvement, we will look individually at the duration and the frequency of SGs usage, either self-reported or electronically measured by the SGs. Regarding subjective experience, we will consider individually all self-reported measures of the learner's experience while using the SG.

We will define the educational outcomes, which are learning and behaviour change, according to the levels of evaluation proposed by Kirkpatrick.³⁴ Learning represents the knowledge that was acquired (ie, factual knowledge; eg, knowledge about the physiopathology of a specific disease), the skills that were improved (ie, how to perform a certain procedure; eg, ultrasound needle placement) or the attitudes that were changed (ie, how worthwhile the learner believes it is to apply the learning in his day-to-day role; eg, attitudes towards pain management) after playing the SG. These three dimensions will be considered separately (ie, knowledge, skills and attitude). Behaviour change represents a change in the day-to-day role of the healthcare professionals or the students after playing the SG (eg, coronary heart disease patients' referral to cardiac rehabilitation by healthcare professionals). For this review, we will consider studies reporting both subjective (eg, self-reported evaluation) and objective measures (eg, quiz) of learning and behaviour change.

Information sources

Bibliographical databases

Eligible primary studies will be identified through a comprehensive search of six bibliographical databases: Cumulative Index of Nursing and Allied Health (EBSCO), Embase (Ovid), Education Resources Information Center (ProQuest), PsycINFO (APA PsycNET), PubMed National Center for Biotechnology Information and Web of Science—Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) (Institute for Scientific Information (ISI)—Thomson Scientific).

Handsearching

Relevant journals will be handsearched for additional articles. Such journals include, but are not limited to: *Games for Health Journal*, *Games*, *G|A|M|E* *The Italian Journal of Game Studies*, *International Journal of Computer Games Technology*, *International Journal of Serious Games* and *JMIR Serious Games*. We will also handsearch for additional articles the reference lists of identified studies and previous systematic reviews related to the use of serious games in healthcare care professionals and healthcare students.

Search strategy for bibliographical databases

The search strategy was collaboratively and iteratively developed with the assistance of a librarian. We initially developed the search strategy for PubMed using a combination of medical subject headings and keyword related to the following key concepts: serious games, healthcare professionals/healthcare students and effect on educational outcomes. The search strategy was then translated for the other bibliographical databases (see online supplementary file 2).

Study records

Data management

All collected references will be imported and managed in EndNote V.X8 (Clarivate Analytics). Based on the eligibility criteria, we will manage and categorise references in specific files and subfiles inside the software. Full texts will be uploaded in the software at the corresponding stage of the selection process.

Selection process

Three reviewers will independently perform the selection process of the collected references. The titles and the abstracts will be screened by applying the previously stated eligibility criteria. A full-text assessment of the remaining references will then be performed. We will resolve all disagreements through discussion and consensus. Reasons for excluding references will be documented and reported in a PRISMA flow diagram at the full-text assessment stage of the selection process.³⁵

Data extraction process

Two independent reviewers will independently perform the data extraction process using an adapted electronic format of the EPOC data extraction template. The original template was developed by the EPOC editorial team to serve as a guide in extracting data from primary studies.³⁶ The reviewers will pilot the form by individually extracting data from a study and by comparing completed forms. Precisions will be added, and wording will be revised, if needed. Notwithstanding data regarding the studies eligibility, we will extract data regarding the impact of the DEs of the SGs on the previously stated outcomes. Data related to the DEs will be extracted during actual gameplay if the SG is publicly available. When it is not possible, data related to the DEs will be extracted based on the information provided in the article. A Kappa statistic will serve to illustrate agreement between the two independent reviewers on the extraction of DEs due to the challenge that could represent this step. Data needed to assess the risk of bias (eg, allocation concealment, blinding) will also be extracted. All data items to be collected are listed below. When data adjusted for baseline differences between groups are available, we will use them to compute effect sizes. When adjusted data are not available in the article, we will use the unadjusted data. All corresponding authors will be contacted in the case of unclear or missing data in the published articles. We

will enter all data in RevMan V.5.3 (The Cochrane Collaboration), a software which allows data management and analysis during the process of a systematic review and a meta-analysis.³⁷

Data items

To guide the extraction of data items related to DEs, a list of DEs was identified through a literature search by the authors (see [table 1](#)).^{10 12–14} Accordingly, we will extract the following data items from the included studies:

- ▶ To a descriptive purpose: study setting; participants' inclusion and exclusion criteria; study aim; design employed; unit of allocation (eg, individual, group); start and end dates of the study; duration of participation; unit of measurement; time points measured; tool used to measure the outcome; validity of the tools; name of the SG evaluated; theoretical framework used for the SG development; cost and duration of the SG development; clinical topic addressed in the SG; method of delivery of the comparator intervention (eg, e-learning, face to face), duration and frequency of use of the comparator intervention, clinical topic addressed in the comparator intervention;
- ▶ To a meta-analytic purpose: study population; sample size; statistical differences at baseline between groups; DEs in the SG evaluated (see [table 1](#)); duration of SG usage; frequency of SG usage; outcomes related to the learner's subjective experience while using the SG (eg, interest, flow), knowledge acquisition, skill improvement and attitude and behaviour change; risk of bias data (as outlined below).

Assessment of risk of bias

Two authors will independently assess the risk of bias of included studies using the EPOC criteria.³⁸ Nine criteria are used and, for each one of them, the study can be judged at 'low risk', 'high risk' or 'unclear risk' of bias. A high risk of bias diminishes the reliability of the study's results. The criteria account for aspects regarding the allocation sequence and concealment, measurements and characteristics of baseline outcome, incomplete outcome data, selective outcome reporting, contamination, blinding, and other risk of bias. Specific guidance to evaluate the risk of bias individually for each criterion is provided by EPOC. Each criterion will be considered independently and no attempt at assigning an overall score at each study will be made, as suggested by EPOC.³⁸ We will resolve all disagreements about the risk of bias assessment of the two authors through discussion and, if needed, with the help of a third review author.

Assessment of selective reporting of outcomes

We will assess the selective reporting of outcomes by comparing the outcomes reported in the primary study with the ones stated in the published research protocol. If no published research protocol is available, we will check if the study was prospectively registered and compare the outcomes reported in the primary study with the ones in

the registration form. If the study was not prospectively registered, we will compare the outcomes presented in the Methods section with the ones reported in the results section. In all cases, the corresponding authors of studies for which there are discrepancies in the outcomes reported will be contacted by email to obtain relevant unreported data.

Assessment of publication bias

If more than 10 studies are included in the meta-analysis, we will assess publication bias by constructing a funnel plot with RevMan V.5.3 (The Cochrane Collaboration).³⁷ The interpretation of the funnel plot will follow the guidance included in Cochrane Handbook for Systematic Reviews of Interventions.²⁸ If not, the assessment of publication bias will be done qualitatively.

Data synthesis

Quantitative data synthesis

The meta-analysis will evaluate the effectiveness of SGs on healthcare professionals' and students' engagement (ie, duration and frequency of SGs usage; outcomes related to the learner's subjective experience while using the SG) and educational outcomes (ie, knowledge acquisition, skill improvement and attitude and behaviour change). The meta-analysis will only include low risk of bias studies based on the assessment using the EPOC criteria.³⁸ In the case of serious reporting or publication bias, no meta-analysis will be performed. This decision will be made by consensus between review authors while also considering clinical and methodological diversity.²⁸

When multiple trial arms are reported in a single trial, we will include only the relevant arms according to the intervention received and the comparison made (ie, SG vs any type of educational intervention). If multiple independent comparisons are relevant in a single study (ie, no common group between comparisons), we'll include individually all comparisons in the meta-analysis. If one or more groups are shared between comparisons, we'll first try to combine several groups to create a single pair-wise comparison. If it's not possible, we'll split the participants in the 'shared group' into multiple subgroups, with smaller sample size, to allow pair-wise comparisons. No minimal number of participants per arm will be required for a study to be initially included in the meta-analysis.

Random-effects models will be used to pool weighted effect sizes. It is assumed, by using random-effects models, that intervention effect across studies follow a distribution and are not all giving an estimate of the same intervention effect.²⁸ The decision to use random-effect models was made due to the expected variability between SGs (notably on DEs) and in study design. Pooled weighted effect sizes will be calculated with 95% CI for each outcome. Continuous outcomes that were obtained by using different measures will be grouped by using standardised mean differences. The significance of the effect sizes will be determined using Cohen's classification (<0.2=negligible; 0.2–0.49=small; 0.5–0.8=moderate;>0.8

=large).³⁹ We will define a statistically significant result by a two-sided alpha of 0.05. All quantitative data synthesis will be realised using RevMan V.5.3 (The Cochrane Collaboration).³⁷

Assessment of heterogeneity

Statistical heterogeneity represents the inconsistency of the studies' results and the percentage of variation across studies that is not due to chance.²⁸ We will assess statistical heterogeneity using the χ^2 test and the I^2 statistic within RevMan V.5.3 (The Cochrane Collaboration).³⁷ An I^2 value superior to 50% will be considered as a high level of heterogeneity.

Subgroup analysis

We will try to explain the source of statistical heterogeneity by exploring clinical and methodological diversity through subgroup analyses. Following the guidance of the Cochrane Handbook for Systematic Reviews of Interventions, caution is required in comparing the magnitude of effect between subgroups as these results remain observational. Therefore, these subgroup analyses will serve primarily to investigate heterogeneity and will be conducted when at least two studies can be included in an individual subgroup.

1. Population: whether current healthcare professionals or students are the focus of the study.
2. Intervention: whether the serious game was delivered as a stand-alone intervention or as part of a multicomponent intervention.
3. Publication year: whether the study was published before or after 2014. As SG development is intrinsically linked to the technological state at a given moment, it is suggested that the overall quality of SGs should have improved in the last couple of years, and similarly their effectiveness in supporting engagement and improving educational outcomes.³⁰ In parallel, the New Media Consortium declared in 2014 that the use of SGs was to be significantly experimented by educational institutions in the next 2 to 3 years.⁴⁰

Sensitivity analyses

Sensitivity analyses will be performed to evaluate if the exclusion of small studies would affect the findings of the meta-analysis and the statistical heterogeneity for each previously stated outcome. If, after exclusion of small studies (if appropriate), at least three studies included in the meta-analysis evaluated SGs containing a specific DE (see table 1), we will perform sensitivity analyses to evaluate the impact of this DE on engagement and educational outcomes. This minimal number of studies (three) is necessary to minimise, during analysis, the potential risk of homogeneity in the DEs integrated to the SGs.

Assessment of the overall quality of evidence

The quality of the evidence regarding the overall effectiveness of SGs and each outcome will be assessed by using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.⁴¹ GRADE

formalises the process of evaluating the overall quality of evidence and formulating recommendations. For each individual outcome, there are four levels of quality of evidence (very low, low, moderate, high) which represent our confidence in the estimate of effect. Quality of evidence depends on factors such as risk of bias, unexplained heterogeneity and indirectness of the evidence. Two authors will independently assess and justify the quality of evidence regarding the overall effectiveness of SGs on each selected outcome. All disagreements will be resolved through discussion and consensus.

Qualitative data synthesis

A qualitative data synthesis will be presented whether or not a meta-analysis will have been performed. Descriptive data (eg, year published, country of origin, population, sample size, study design, name of the SG assessed, clinical topic addressed, outcomes measured) of each included study will be presented in a tabular form. A summary table containing an overview of all DEs included in the SGs assessed will also be presented. The results according to the overall effectiveness of SGs and the impact of DEs on engagement and educational outcomes will be presented narratively.

DISCUSSION

This systematic review will fulfil important needs in the designing of SGs for healthcare professionals' and students' education. As previous systematic reviews focused on a narrative evaluation of the overall effectiveness of SGs in the healthcare professions,^{4 21 42 43} to our knowledge, this will be the first systematic review to describe and evaluate of the impact of the prespecified DEs on engagement and educational outcomes. Previous systematic reviews were able to quantify the impact of several DEs in simulation (eg, range of task difficulty, deliberate practice)^{44 45} and internet-based education (eg, integration of an online discussion forum and audio files)⁴⁶ by adopting a similar approach to the one we propose. Depending on the quantity of data available, planned sensitivity analyses will allow the quantitative evaluation of the impact of the DEs on selected outcomes. However, we acknowledge that possible homogeneity in DEs integrated to SGs could represent a limit to planned sensitivity analysis.

The protocol of this systematic review is prospectively registered (PROSPERO #CRD42017077424) and is reported according to the PRISMA-P guidelines.²⁹ Our research strategy was collaboratively designed with the assistance of a librarian and will extensively cover several relevant databases. However, as clear criteria have yet to be established in the literature regarding what can be considered a SG, we will adopt an inclusive definition that does not discriminate based on the SG genre (eg, quiz, adventure).^{5 8 32 33} In the case of ambiguity regarding the inclusion of specific studies due to the nature of the of intervention, review authors will assess their eligibility through discussion and consensus. Nonetheless, this

limitation will be addressed in the reporting of the full systematic review and considered in the formulation of recommendations. Moreover, as the integration of the selected DEs in SGs developed for healthcare professionals and students is currently unknown, it is currently planned to describe the integration of these DEs in developed SGs and their alignment with theoretical foundations, as stated by the authors.

Author affiliations

¹Faculty of Nursing, Université de Montréal, Montreal, Quebec, Canada

²Montreal Heart Institute Research Center, Montreal, Canada

³Centre de recherche du Centre hospitalier de l'Université de Montréal, Montreal, Canada

⁴Department of Anesthesiology, Faculty of Medicine, Université de Montréal, Montreal, Canada

⁵William F Connell School of Nursing, Boston College, Boston, Massachusetts, USA

⁶Resident in Medicine, Clinician-Scientist Program, Université de Montréal, Montréal, Montreal, Canada

⁷Hôpital du Sacré-Cœur de Montréal, Montreal, Canada

⁸Concordia University, Montreal, Canada

⁹School of Librarianship and Information Science, Université de Montréal, Montreal, Canada

Acknowledgements This research was carried out as part of the doctoral studies of the first author and for which he received scholarships from the following: Montreal Heart Institute Foundation, Ministère de l'Éducation et de l'Enseignement supérieur (MEES), Réseau de recherche en interventions en sciences infirmières du Québec (RRISQ), Faculty of Nursing of the Université de Montréal, Fondation J.A. DeSève, Fondation Alice-Girard, and Faculté des études supérieures et postdoctorales of the Université de Montréal (FESP). We also wish to thank Kate Johnson for her careful manuscript editing.

Contributors All authors contributed to the conception of the protocol. MAMC wrote the initial draft of the protocol. SC, VD, GF, TM, PL, AC, FB and GMD critically revised the initial draft for important intellectual content. All authors gave final approval of the manuscript submitted. All authors agreed to be accountable for all aspects related to this work.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Disclaimer The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Competing interests None declared.

Patient consent Detail has been removed from this case description/these case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. World Health Organization. *Transforming and scaling up health professionals' education and training: World Health Organization Guidelines 2013*: World Health Organization, 2013.

2. Donkin L, Christensen H, Naismith SL, et al. A systematic review of the impact of adherence on the effectiveness of e-therapies. *J Med Internet Res* 2011;13:e52.
3. Perski O, Blandford A, West R, et al. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017;7:254–67.
4. Wang R, DeMaria S, Goldberg A, et al. A systematic review of serious games in training health care professionals. *Simul Healthc* 2016;11:41–51.
5. Hamari J, Shernoff DJ, Rowe E, et al. Challenging games help students learn: an empirical study on engagement, flow and immersion in game-based learning. *Comput Human Behav* 2016;54:170–9.
6. Pfahl D, Koval N, Rube G, 2001. An experiment for evaluating the effectiveness of using a system dynamics simulation model in software project management education. *Software Metrics Symposium, 2001 METRICS 2001 Proceedings Seventh International; 2001. IEEE.*
7. Westera W, Nadolski RJ, Hummel HGK, et al. Serious games for higher education: a framework for reducing design complexity. *J Comput Assist Learn* 2008;24:420–32.
8. Salen K, Zimmerman E. *Rules of play: Game design fundamentals*. Cambridge, MA: MIT Press, 2004.
9. Nevin CR, Westfall AO, Rodriguez JM, et al. Gamification as a tool for enhancing graduate medical education. *Postgrad Med J* 2014;90:685–93.
10. Sailer M, Hense JU, Mayr SK, et al. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. *Comput Human Behav* 2017;69:371–80.
11. Lundgren S, Björk S, 2003. Game mechanics: describing computer-augmented games in terms of interaction. *Proceedings of TIDSE*
12. Werbach K, Hunter D. *The gamification toolkit: dynamics, mechanics, and components for the win*: Wharton Digital Press, 2015.
13. Nagle A, Novak D, Wolf P, et al. 2014 IEEE 3rd International Conference on; 2014. IEEE. *The effect of different difficulty adaptation strategies on enjoyment and performance in a serious game for memory training. Serious Games and Applications for Health (SeGAH).*
14. Yildirim IG. Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems; 2016. ACM. *Time pressure as video game design element and basic need satisfaction.*
15. Lu AS, Thompson D, Baranowski J, et al. Story immersion in a health videogame for childhood obesity prevention. *Games Health J* 2012;1:37–44.
16. Pavlas D. *A model of flow and play in game-based learning: The impact of game characteristics, player traits, and player states*. Florida: University of Central Florida Orlando, 2010.
17. Lameris P, Arnab S, Dunwell I, et al. Essential features of serious games design in higher education: linking learning attributes to game mechanics. *Br J Educ Technol* 2017;48:972–94.
18. Csikszentmihalyi M. *Beyond boredom and anxiety*. Hoboken, New Jersey: Jossey-Bass, 2000.
19. Schrader C, Bastiaens T. Learning in educational computer games for novices: the impact of support provision types on virtual presence, cognitive load, and learning outcomes. *The International Review of Research in Open and Distributed Learning*. 2012;13:206–27.
20. Boyle EA, Hainey T, Connolly TM, et al. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Comput Educ* 2016;94:178–92.
21. Ricciardi F, De Paolis LT. A comprehensive review of serious games in health professions. *Int J Comput Game Technol* 2014;2014:1–11.
22. Arnab S, Lim T, Carvalho MB, et al. Mapping learning and game mechanics for serious games analysis. *Brit J Educ Technol* 2015;46:391–411.
23. Giessen HW. Serious games effects: an overview. *Procedia Soc Behav Sci* 2015;174:2240–4.
24. Girard C, Ecalle J, Magnan A. Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *J Comput Assist Learn* 2013;29:207–19.
25. Olszewski AE, Wolbrink TA. Serious gaming in medical education: a proposed structured framework for game development. *Simul Healthc* 2017;12:240–53.
26. DeSmet A, Van Ryckeghem D, Compennolle S, et al. A meta-analysis of serious digital games for healthy lifestyle promotion. *Prev Med* 2014;69:95–107.
27. Borrajo F, Bueno Y, de Pablo I, et al. SIMBA: a simulator for business education and research. *Decis Support Syst* 2010;48:498–506.
28. Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons: Hoboken NJ, 2011.
29. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
30. Wouters P, van Nimwegen C, van Oostendorp H, et al. A meta-analysis of the cognitive and motivational effects of serious games. *J Educ Psychol* 2013;105:249–65.
31. Cochrane Effective Practice and Organisation of Care Review Group (EPOC). What study designs should be included in an EPOC review and what should they be called? 2014 <https://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/EPOC> (accessed 12 Dec 2017).
32. Bergeron B. *Developing serious games*. Newton Centre, MA: Charles River Media, 2006.
33. Stokes BG. Videogames have changed: time to consider serious games? *Dev Educ J* 2005;11:6–13.
34. Kirkpatrick DL. Luminary perspective: evaluating training programs. Biech E, ed. *ASTD handbook for workplace learning professionals*. Alexandria, VA: ASTD Press, 2008:485–91.
35. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
36. *Cochrane Effective Practice and Organisation of Care Review Group (EPOC). Data collection checklist*. Ottawa, Canada: Institute of Population Health, University of Ottawa, 2002.
37. The Cochrane Collaboration. RevMan5 2017. <http://community.cochrane.org/tools/review-production-tools/revman-5> (accessed 12 Dec 2017).
38. Cochrane Effective Practice and Organisation of Care (EPOC). Suggested risk of bias criteria for EPOC reviews 2017. <http://epoc.cochrane.org/epoc-specific-resources-review-authors> (accessed 12 Dec 2017).
39. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum, 1988.
40. Johnson L, Adams Becker S, Estrada V, et al. *NMC Horizon Report: 2014 Higher Education Edition*. Austin, Texas: The New Media Consortium, 2014.
41. Dijkers M. Introducing GRADE: a systematic approach to rating evidence in systematic reviews and to guideline development. *KT Update* 2013;1:1–9.
42. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. *Br J Surg* 2012;99:1322–30.
43. Castro-Sánchez E, Kyrtasis Y, Iwami M, et al. Serious electronic games as behavioural change interventions in healthcare-associated infections and infection prevention and control: a scoping review of the literature and future directions. *Antimicrob Resist Infect Control* 2016;5:1–7.
44. Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *Med Teach* 2013;35:e867–98.
45. Cheng A, Eppich W, Grant V, et al. Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. *Med Educ* 2014;48:657–66.
46. Cook DA, Levinson AJ, Garside S, et al. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. *Acad Med* 2010;85:909–22.