A narrative literature review of games, animations and simulations to teach research methods and statistics

Elizabeth A. Boyle [a,*], Ewan W. MacArthur [b], Thomas M. Connolly [b], Thomas Hainey [b], Madalina Manea [c], Anne Kärki [d], Peter van Rosmalen [e]

[a] School of Social Science, Faculty of Health, Education and Social Science, University of the West of Scotland, Paisley Campus, High St., Paisley, Scotland, PA1 2BE, UK
[b] School of Computing, Faculty of Science and Technology, University of the West of Scotland, Paisley Campus, Paisley, Scotland, PA1 2BE, UK
[c] University of Medicine and Pharmacy of Craiova, Craiova, Romania
[d] Satakunta University of Applied Sciences, Tiedepuisto 3, Pori, 28600, Finland
[e] Welten Institute, Research Center for Learning, Teaching and Technology, Faculty of Psychology and Educational Sciences, Open University of the Netherlands, The Netherlands

ARTICLE INFO

ABSTRACT

Basic competence in research methods and statistics is core for many undergraduates but many students experience difficulties in acquiring knowledge and skills in this area. Interest has recently turned to serious games as providing engaging ways of learning. The CHERMUG project was developed against this background to develop games to support students in learning about research methods and statistics. As a first step in designing the CHERMUG games a narrative literature review was carried out to establish whether similar games, animations and simulations already existed. Search terms used in the literature review included varied terms for digital games, simulations and animations, terms relevant to the twin goals of learning and engagement in games and terms for research methods and statistics. Application of the inclusion criteria led to 26 papers which were considered relevant. Synthesis of the papers suggested that there is reason to be optimistic that a game-based approach might be effective in learning in this area.

1. Introduction

There is an increasing recognition of the need for students across many disciplines to acquire competence in research methods and statistics. This is often referred to as statistical literacy, defined by Gal (2005, p. 70) as "the ability to interpret, critically evaluate, and communicate about statistical information and messages". Nikiforidou, Lekka, and Pange (2010, p. 798) also acknowledged the importance of statistical literacy as requiring consideration of the "the synergy of content, pedagogy and technology".

1.1. Difficulties and misconceptions

However, acquiring methodological and statistical expertise poses significant challenges for many students (Tishkovskaya & Lancaster, 2010). The material is challenging because it is highly abstract and requires the consideration of inter-related logical reasoning, critical thinking, data analysis and interpretation and evaluation skills. Students perceive statistics as difficult and boring and this leads to anxiety and a lack of self-efficacy. Castro Sotos, Vanhoof, Van den Noortgate, and Onghena's (2007) systematic review of students' statistical misconceptions found that students have many misconceptions surrounding statistical constructs such as sampling distributions, hypothesis testing and confidence intervals.

---

* Corresponding author. Tel.: +44 1418483783.
E-mail addresses: liz.boyle@uws.ac.uk (E.A. Boyle), ewan.macarthur@uws.ac.uk (E.W. MacArthur), thomas.connolly@uws.ac.uk (T.M. Connolly), thomas.hainey@uws.ac.uk (T. Hainey), madalinamanea2005@yahoo.com (M. Manea), anne.karki@samk.fi (A. Kärki), Peter.vanRosmalen@ou.nl (P. van Rosmalen).

1.2. New ways of teaching research methods and statistics

In recent years there have been criticisms that traditional methods of teaching research methods and statistics do not inspire a passion for research in students and do not allow students to apply their knowledge. Consequently teachers are continually looking for new ways to make research methods and statistics more appealing to their students. The extensive research of Garfield, Ben-Zvi, Chance, Medina, and Roseth (2008) on the characteristics of Statistical Reasoning Learning Environments identified a number of principles which can help students to learn about statistics, such as using real datasets in classroom activities, providing opportunities to explore and analyse data, developing statistical reasoning and using appropriate technological tools to allow students to test hypotheses. Pfannkuch and Wild (2004) argued that students find it difficult to apply their statistical knowledge and to overcome this they argued that statistics teaching should focus more on covering key statistical concepts (especially variation), develop the ability to explore and learn from data, develop statistical argumentation, use formative assessment and try to understand students' reasoning. Tishkovskaya and Lancaster also proposed a number of strategies to help students learn statistics more effectively, such as shifting the focus of statistics curricula to practical tasks, using problem-based learning and real life examples in project work, developing statistical literacy by supporting critical thinking skills and communication about statistics, using technology and on-line resources and targeting misconceptions through discussion and assessment. Many of these suggested principles and strategies seem to align very well with the functionality of serious games.

1.3. Games for learning

Serious games and games for learning have recently been suggested as an engaging way of helping students to learn (Joint Information Systems Committee, JISC, 2007). Interest in serious games emerged initially from speculation that games could provide highly engaging activities which could be utilised in learning (Boyle, Connolly, & Hainey, 2011). More importantly however games offer methods of learning which are highly consistent with modern theories of effective learning which propose that learning activities should be active, situated, problem-based, interactive and socially mediated (Boyle et al.). Nadolski et al. (2008) suggest that serious games can be a useful tool for Higher Education Institutions to develop and deploy, to enhance the student experience and to assist them in achieving the intended learning outcomes.

The CHERMUG quantitative and qualitative games (www.chermug.eu) are digital games designed to support students as they learn about research methods and statistics. The games were conceived against the background of finding a more engaging, activity-based approach to learning in this area and the recognition that games might help to provide this. The CHERMUG games were initially targeted at nursing and social science students for whom methodological and statistical expertise is an area of core competence. In nursing education there is a strong focus on evidence-based practice and critical evaluation modules which emphasise the need for nurses to acquire useable knowledge which makes links between research findings and practice (Barker, 2010; Emanuel, Day, Diegnan, & Pryce-Miller, 2011; Newell & Burnard, 2011; Winters & Echeverri, 2012).

1.4. Aim of the current paper

The aim of the current paper is to report on a narrative literature review which was carried out to establish to what extent a games-based approach is currently being used to teach research methods and statistics. A meta-analysis carried out by Larwin and Larwin (2011) showed the advantages of using Computer-Assisted learning in postsecondary statistics education with an overall effect size of 0.566 from 70 studies. However CAL has been around for much longer than games and there has been time to build up an evidence base for its effectiveness. Since games-based learning is still a relatively new area of research, it was predicted that well-designed studies of games in this subject domain may be scarce at this stage. A previous review of the use of digital games in learning (Connolly, Boyle, Hainey, MacArthur, & Boyle, 2012) confirmed this and indicated that the research designs of papers looking at games are varied and rarely report Randomised Control Trials (RCTs), the design of choice for evaluating educational interventions. For these reasons the focus in the current review was not so much on establishing effect sizes for studies using games for teaching research methods and statistics, but on establishing the state of knowledge in the area. Consequently a narrative review was considered to be most appropriate. While the review was important groundwork for the design of the CHERMUG games, it is also relevant to the games community where there is interest in developing games to support the teaching of core 21st century skills (Dondlinger, 2007).

2. Method

2.1. Search terms

The search terms for the literature review were derived from a previous search carried out on the evaluation of computer games which addressed the variety of digital games that might be played (Connolly et al., 2012). While we were primarily interested in games, search terms for simulations and animations were also included as these might include elements of gamification. e-Learning and Computer Assisted Learning (CAL) were excluded as search terms as these were too wide-ranging for the current review. Terms used for games were:

("computer games" OR "video games" OR "serious games" OR simulation OR "simulation games" OR "games-based learning" OR "MMOG" OR "MMORPG" OR "online games" OR animation)

To narrow down the search to focus on games relevant to the twin goals of games for learning, i.e. learning and engagement, the following terms were also used:

AND (learning OR education OR skills OR engagement OR motivation)

and to narrow down the search to papers which were relevant to games about research methods and statistics the following terms were also included:

AND (statistics OR "research methods")

Reflecting the relatively recent interest in serious games, the time period selected for the search was the ten year span from January 2004 to December 2013. The date of the latest search was on 21st December 2013. Given the very general nature of the search terms for "research methods" and "statistics" and their widespread use in reporting research in journal papers, it was very difficult to provide more focused search terms and it was predicted that a relatively high proportion of papers identified by the search terms might not be relevant.

2.2. Databases searched

The databases searched included those identified as relevant to education, information technology and social science and nursing:

Science Direct; ERIC; EBSCO (consisting of Psychology and Behavioural Science, PsycINFO, SocINDEX, Library, Information Science and Technology Abstracts, CINAHL); ASSIA.

Since more included papers were identified through the ERIC database, the search strategy for that database is described. A connection with the ERIC database was established via the UWS library website. The "collection" tab was selected and the "peer reviewed only" tab. Overall 4772 items were found with 2740 in the 10 year time period. The latest search was carried out on 18/12/2103.

2.3. Grey literature

The European Conferences on Game Based Learning (ECGBL) from 2009 to 2013 and the International Conference on Teaching Statistics (ICOTS) were also searched for relevant papers. Another source of evidence about the relevant literature was the pooled knowledge and expertise of the CHERMUG project partners concerning game design and research methods and statistics.

2.4. Selection criteria for inclusion of papers in the current review

This study aimed to find previous research which used computer games or gamification techniques, animations or simulations to improve or support students' skill and knowledge acquisition or engagement with research methods and statistics.

A number of further criteria were specified to select appropriate studies for inclusion in the review. To be included in the review papers had to (a) describe a digital game, animation or simulation which aimed to teach research methods and statistics, (b) include some element of quantitative or qualitative empirical evaluation of the application, (c) date from January 2004 to December 2013 and (d) include an abstract.

3. Results

3.1. Papers identified

Table 1 shows the number of papers which were identified from each database along with the number of papers which were included in the review.

3.2. Papers excluded

The titles or titles and abstracts of the 4040 papers identified were read to select papers which were possibly relevant. As predicted, due to the very general nature of the search terms for research methods and statistics and their widespread use in reporting the methods and analysis sections of empirical papers, many papers which emerged from the search were not relevant to the review. Papers were excluded for the following reasons:

Table 1
The number of papers identified from each database and included in the review.

	Identified from database	Included in the review
SCIENCE DIRECT	360 identified using abstract title keywords from 2004 to present	6
ERIC	2740 from 2004 to present; peer reviewed only	10
EBSCO (psychology social sciences, nursing and health professions)	480	2
ASSIA	12	0
ICOTS-08	127	1
ECGBL 2009–2013	315	1
Previously known by authors	6	6
Total	4040	26

(a) A number of papers which reported the use of games or game-like activities in teaching research methods and statistics were of interest in the development of the CHERMUG game, but were excluded from this review if they did not include any qualitative or quantitative evaluation of the effectiveness of the intervention which they described (Burguillo, 2010; Nordmore, 2004; Quinn, 2003).

(b) Similarly, several papers identified by the search terms described the wide variety of animations and simulations which is available to teach statistical constructs, but these were excluded from the review if they did not include any empirical evaluation of the effectiveness of the tool in learning (Bowman, 2010; Bulmer, 2004; Bush, Menzies, & Thorp, 2009; Dominguez-Dominguez & Dominguez-Lopez, 2010; Forbes, 2010; Gordon & Gordon, 2009; Lane & Peres, 2006; Stirling, 2010).

(c) Many papers identified by the search terms described digital games, but if the game did not specifically support skills which are required at some stage in the research methods cycle, the paper was excluded. For example several papers describing games to support mathematics were found, but the content was not judged to be sufficiently relevant (Chang, Wu, Weng, & Sung, 2012; Chang et al., 2007; Kebritchi, Hirumi, & Bai, 2012). Similarly a number of papers describing serious games for health and nursing education were identified but these were not relevant to the current review if they did not address research methods and statistics skills (Kanthan & Senger, 2011). More difficult to decide were papers about games for science education which on the face of it could be relevant to methodology. Examples included Magnussen, Hansen, Planke, and Sherson (2013). However if the paper did not discuss methodology and statistics it was excluded.

(d) Many papers discussed aspects of learning with games generally. Examples include Denner, Werner, Campe, and Ortiz (2013), Gros (2007), Mayer (2012) and Wouters and van Oostendorp (2013). However, since the focus of the current review was specifically on games which support research methods and statistics, such papers were excluded from the current review.

(e) Many papers identified by the search terms described modern ideas about effective teaching of research methods and statistics or examined active approaches to learning about research methods and statistics. A sample of these is Diekema, Holliday, and Leary (2011), Gal (2005), Garfield and Ben-Zvi (2009), Garfield et al. (2008), Leppink (2010), Pfannkuch and Wild (2004) and Tishkovskaya and Lancaster (2010). While these papers provided useful guidance about how research methods and statistics should be taught, they were excluded from the review if they did not describe evaluations of game-like activities, animations or simulations to teach these skills.

3.3. Papers selected

Using the inclusion criteria described in 2.4, 26 papers met the inclusion criteria and were identified as relevant to the current review, that is they described and evaluated a game, animation, simulation or other more general e-learning application that in some way supported learning in research methods and statistics.

3.4. Categorisation of papers

As with previous literature reviews in the area of digital games (Connolly et al., 2012), papers identified in the review were diverse with respect to the topics which they addressed. It was useful to organise the 26 papers into categories depending on the kind of e-learning application that the paper described: entertainment game, serious game, animation or simulation, more general e-learning application or web-based tool, and modelling tool. These categories are shown in Table 2, along with the number of papers covering each topic. There were approximately equal numbers of papers in each category. Also shown are the numbers of papers which adopted a quantitative, qualitative or both approaches in the evaluation. All 26 papers are summarised in the table in Appendix A with respect to the authors; the data-base from which the papers came; the name or kind of game, animation, simulation, tool or model described; the research approach and design and aim of the study, the research methods or statistics topic which was addressed and the results of the study.

3.4.1. Papers on entertainment games

Traditional card, dice and board games such as poker, ludo and snakes and ladders were not designed to teach players statistical concepts and principles, but over time players become familiar with the probabilistic characteristics inherent in these games of throwing dice or playing cards and teachers have traditionally used these characteristics to illustrate basic concepts of randomness and probability in teaching introductory statistics (Hogben, 1950).

More recently, the structure and formats of digital entertainment games and television quiz shows have also been used to illustrate statistical concepts. For example Quinn (2003) described how probabilistic features implicit in the well-known "Who Wants to be a Millionaire" game can be used to motivate and interest students in learning concepts of expected value and probability, such as the mathematical probability of answering all 15 questions correctly.

Few papers describing entertainment games reported any evaluation of the value of the games in learning. An exception was Chow, Woodford, and Maes (2010) who described an evaluation of a digital version of the popular game "Deal or No Deal" in helping students

Table 2
Number of papers covering each topic and research approach.

Kind of game or e-learning application	Number of papers	Quantitative	Qualitative	Both
Entertainment game	5	2	2	1
Serious game	5	3	1	1
Animation or simulation	5	3	1	1
General e-learning application or web based tool	7	5	2	0
Modelling tool	4	1	1	2
Total	26	14	7	5

to understand the concept of expected value. In the game, players make a series of choices from 26 briefcases which are identical except that each contains a different monetary value ranging from one cent to one million dollars. Towards the end of the game contestants have to decide whether to accept the dealer's offer of a certain sum of money or not (Deal or No Deal). The offer should be accepted when it is equal to or greater than the contestant's expected value. Getting students to make a prediction which is either subsequently fulfilled or not can help them to understand the simple statistical concept of expected value. Chow et al. tested the students' retention of the concept one week after the intervention and found that 95% of students who played the game remembered the construct compared with only 59% of students who had not played the game.

Bottino, Ferlino, Ott, and Tavella (2007) were also interested in whether playing games with specific characteristics can help to support logical thinking. They carried out a 3 year longitudinal study looking at how playing digitised versions of classic puzzle-type entertainment games can foster the development of logical and scientific reasoning and problem solving abilities in primary school students. Games included PappaLOTTO, a version of the classical Mastermind involving parrots and Hexip, a game which is similar to Battleship but with different rules. The skills required in playing these games include logical reasoning and testing of variables skills which underlie methodological thinking. The authors compared performance on national tests of mathematics for groups who had played the games with those who had not and found that the mean scores for the two experimental groups were higher than those for the two control groups, although the paper did not report whether this difference was significantly different.

In teaching research methods and statistics, many teachers use examples where data has been generated from real-life activities. Two papers were identified which used digital entertainment games incidentally to generate data and to teach aspects of experimental design. Stansbury and Munro (2013) used the Wii dance game, Dance Dance Revolution (DDR), to teach factorial design. Students were randomly assigned to one of four different conditions in a factorial design where choice of song (Lady Gaga's Bad Romance vs. Sean Paul's So Fine) and presence or absence of music were the independent variables with students' dancing scores derived from the game as the dependent variable. Students analysed the data generated by this factorial design. In evaluating the success of the game in helping students to understand factorial design, pre and post-tests on competence on factorial design were administered to the students in the experimental group, a lecture only group and a no content group. The increase in performance from pre- to post-test was bigger for those in the game group than in the two control groups, showing that the use of the game was more effective than either of the controls in improving performance.

Ramler and Chapman (2011) provided a detailed description, although with limited evaluation, of how they used the entertainment game, Guitar Hero, to help students understand the different stages in the design and execution of a research project. Students were encouraged to propose and test a hypothesis concerning the data which they generated as they played the game. An example hypothesis about the distribution of missing notes in playing was: "Notes are missed completely at random". This paper was interesting in showing how games can be used to generate data to test hypotheses, although the evaluation was limited to students' qualitative comments about the project. These were generally favourable, such as: the project helped students "tie together" the material in the course, the project had "an element of fun" and "This course was the first stat course that let me think like a statistician".

A rather different set of research skills demonstrated while playing an entertainment game was reported by Steinkuehler and Duncan who predicted that playing massively multiplayer online games (MMOGs) helps players to develop higher level reasoning and argumentation skills. Steinkuehler and Duncan (2008) looked at players' contributions to online discussion fora in World of Warcraft. They analysed a random sample of 1984 players' posts to these fora using established national benchmarks for scientific literacy. Steinkuehler and Duncan found that players demonstrate higher order scientific reasoning skills, such as using data and argument, building on others' ideas and using system-based and model based reasoning. Such findings suggest that virtual worlds can provide platforms which encourage players to display logical approaches to solving complex problems.

3.4.2. Papers on serious games

In contrast to entertainment games, serious games are games which have been intentionally designed to support learning. The game which was most similar to the proposed CHERMUG game was Operation ARA (Acquiring Research Acumen) described by Halpern et al. (2012). This game aimed to teach players to adopt a more rigorous and critical approach to seeking evidence and developing arguments in carrying out research. Operation ARA was an adventure/mystery game which aimed to teach 17 methodological and statistical concepts. Players are provided with an engaging narrative about helping the Federal Bureau of Science to identify extra-terrestrials who are carrying out poor research. This provided the rationale for developing a more rigorous approach to research, helping players to adopt an evaluative approach in determining whether particular research studies are reliable or flawed. During the game the players are tutored by avatars which are sensitive to players' previous responses. An early evaluation of the game suggested that students who played Operation ARA learnt significantly more than those who did not play the game. There was no difference between different tutoring groups which varied in degree of active engagement with the material in immediate post-test, but at delayed post-test the group which had more active engagement performed significantly better than the group which had less engagement.

Like Steinkuehler and Duncan, Asbell-Clarke et al. (2012) were influenced by their observations that players of MMOGs are actively engaged in activities which have close parallels with the problem solving and reasoning activities of professional scientists in tracking down and analysing data. They developed a mystery game, Martian Boneyards, which aimed to help players develop skills of systematic scientific enquiry in the area of paleontology. Players played the role of scientists, tackling problems collaboratively with the help of a range of tools to assist them in data collection, exploration, data-gathering, analysis, theory-building and collaborating by communicating their ideas to other players. The results showed that players spend a lot of time gathering data, but also have opportunities to analyse data and look at how the evidence that they collected supported specific theories. This study showed that a game-based approach could be useful at the early (inductive and exploratory) data collection and hypothesis development stages of research.

Ancker, Weber, and Kukafka (2011) used an interactive, game-like graphical presentation to convey information about health risks which are difficult to understand when conveyed verbally. This allowed players to experience rather than simply read about the frequency of occurrence of an event to see whether this improved participants' perceptions of risk. Contrary to predictions, Ancker et al. did not find a significant main effect of the different graphical representations on risk feelings or risk estimates for either low-risk or high-risk stories, although they did find that the interactive graphics reduced risk estimates for low numeracy participants. Order effects were also found, suggesting that it is important to consider carefully how statistical information about risk is presented.

Nte and Stephen (2008) described the development and use of a game which illustrated the principles of the normal distribution. The game allowed students to see the construction of a normal distribution of data derived from asking each character emerging from a virtual pub about how many drinks they had consumed. There was no objective evaluation of the impact of the game on students' performance, but questionnaire responses indicated that students agreed that the game was a useful method for helping them learn about statistics, and did not provoke anxiety, although players agreed that they normally did feel anxiety about statistics.

Gresalfi and Barab (2011) described how the use of a video game, Ander city, engaged young students and helped them to learn about statistical constructs, such as central tendency and range, by providing opportunities to take on new roles and experience the consequences of their decisions in thinking about dilemmas that the residents of the city are facing. Students were given data, representations of the data and tools to manipulate the data and they had to use these tools to construct graphs, compare distributions of data, consider how representations and arguments relate to each other and make decisions and develop arguments grounded in the data. The authors provided examples of the dialogue generated by pupils and teachers which illustrated how the game supports deep learning by engaging pupils procedurally, conceptually, consequentially and critically.

3.4.3. Papers on animations and simulations

There is a long tradition of using visual representations, such as graphs, histograms, bar charts and pie charts, to support students' understanding of statistical constructs. Visual representations were first introduced in the early nineteenth century to help to display data in a way which is easier to understand (Lewandowsky & Spence, 1989). Animations go beyond these static representations to provide more powerful graphical representations which display sequences of images which create an illusion of movement. This can facilitate our understanding of two and three dimensional properties of distributions and assist our understanding of changes across time or space. Xie (2013, p. 1) summarised the benefits of animations: "Animations can add insight and interest to traditional static approaches to teaching statistics and reporting, making statistics a more interesting and appealing subject". Simulation is similar to animation but typically allows users to change variables systematically and explore the results of these changes on the visual representation.

As discussed in 3.2, the inclusion of the search terms "simulation" and "animation" in the review revealed the widespread availability of these applications for teaching statistics. However, only a few papers provided empirical evidence of the effectiveness of animations and simulations in supporting learning. Mills (2004) compared students using an Excel simulation to learn abstract statistical concepts related to the sampling distribution and the Central Limit Theorem (CLT) with a control group who learnt in the traditional way. He found no difference in performance between the two groups on the pre-test but students in the experimental group performed significantly better post-test and had more positive attitudes to the instruction than those in the control group.

Wang, Vaughn, and Liu (2011) examined the impact of game-like animation interactivity on novices' learning of the traditionally difficult topic of hypothesis testing (including p-values and type I and type II errors). Wang randomly assigned 123 college students to four different levels of support and increasing animation and interaction from (1) static to (2) simple animation to (3) input to (4) practice. The results showed that the simple animation, input and practice groups were significantly better than the static group in understanding, but the later three groups did not differ. There was no effect of group on student confidence and program perception.

A major problem for students in learning statistics is the misconceptions they can experience. Liu, Lin, and Kinshuk (2010) examined whether students who used Simulation Assisted Learning (SALS) to correct misconceptions were subsequently less prone to misconceptions in understanding correlation. SALS used a cognitive conflict model and dynamically linked multiple representations, which allow the actions performed on one representation to be automatically shown in other representations. 72 Grade-12 students were randomly assigned to either the experimental group who used the SALS or a control group. The results confirmed that the experimental (SALS) group was significantly more effective post intervention than the lecture-based learning group both in terms of reducing students' misconceptions and improving their understanding of correlation.

Neumann, Hood, and Neumann (2012) reported a qualitative analysis of students' perceptions of simulation software. Coding of students' responses to interviews indicated that students perceived that the simulation tool helped their learning by providing immediate feedback, practice and a useful visual aid to learning, as well as providing motivational benefits such as creating interest and engagement in the topic.

3.4.4. Papers on general e-learning applications and web based tools

Search terms in the current review focused on games, animation and simulations but did not include more general e-learning applications web-based tools. Nonetheless a number of papers describing such tools for supporting research methods and statistics were identified in the review. Bush, Menzies and Thorp described the huge array of web-based tools and teaching resources, ranging from class surveys to individual simulation experiments, which are available for teaching statistics but they did not evaluate their impact in learning. Archee and Gurney (2006) described the use of interactive online Flash video tutorials to teach students SPSS. This study included only very minimal formative evaluation of staff and student feedback about the intervention but this indicated that this blended learning approach provided personalized learning which allowed students to proceed at their own pace.

A few papers reported between group comparisons of e-learning applications with more traditional methods of learning statistics. Burruss and Furlow (2007) compared the performance of a group of students using Microsoft Excel spread-sheets and Macromedia Flash movies to demonstrate key statistical concepts with a control group taught the traditional way. There was no significant difference between groups on a chi-square skills test following the intervention, but the experimental group performed better than the control group on conceptual definitions of chi square. In addition the control group showed improvement in the affect, cognitive competence and difficulty scales of the attitudes towards statistics measure, while the treatment group only showed more positive attitudes from the pre-test to the post-test. The results suggest that introductory material on statistics should include visual cues and conceptual questions to assist learning.

Dinov, Sánchez, and Christou (2008) looked at how the availability of the Statistics Online Computational Resource (SOCR) helped undergraduates learn. The evaluation involved three different undergraduate classes, which all included a treatment group which used the SOCR resources and a control group which received classical instructional methods. Pooling the results across all courses the impact of the SOCR resources was large. There was also less variance in scores for the treatment groups suggesting that many students benefited from the resources. The SOCR e-learning resources were varied but Dinov et al. did not report which were more useful than others.

Given their prevalence in teaching research methods, it is frequently assumed that using data analysis packages, such as SPSS, helps students to acquire a better understanding of statistics. Basturk (2005) evaluated this assumption, comparing students in a lecture only group with a lecture plus SPSS group. Results showed that those in the lecture plus SPSS group gained higher averages on the midterm and final exams than the lecture only group and this difference was even larger for inferential statistics than descriptive statistics. This confirms the value of using the SPSS data analysis package to assist students in understanding statistics.

Rather than comparing an experimental and control group, González, Jover, Cobo, and Muñoz (2010) adopted an interesting design in evaluating the impact of a web-based tool, e-status. González et al. felt that ethically all students should be given the opportunity to use the application and consequently the course content was divided into Parts A and B, where the difficulty and length of course content in each part were reasonably equal. Approximately half the students had access to e-status on Part A and half on Part B. Results indicated that performance on the problems where students used e-status was significantly better than that on problems where e-status was not used by 4.8%. For the 94 students who actually employed e-status, the effect size was 0.63. González et al. concluded that e-status had a measurable and positive effect in improving student performance.

Callaghan, Lea, Mutton, and Whittlesea's (2011) qualitative study used a staff focus group to identify the areas where students had most difficulties in learning research methods. The areas identified were research design, the language of research, interpreting data, reliability, validity, significance, descriptive statistics, qualitative and quantitative methods, sampling and population. Video resources were developed to demonstrate the teaching of these constructs and focus groups with students indicating that they found these videos to be extremely useful in supporting their learning.

Krause, Stark, and Mandl (2009) examined two features of how the Koralle e-learning environment might lead to better performance in learning correlations and scatterplots: they compared students learning in dyads with students learning alone and also compared students who got feedback about their performance with those who did not. Contrary to expectations, Krause et al. found that working cooperatively in dyads did not lead to improved performance, but that feedback did. In spite of this, Krause et al. argued that the evidence for the benefits of working in groups in other areas is strong and it would be useful to pursue the idea of collaborative learning in research methods.

3.4.5. Papers on modelling

The majority of the papers describing animations, simulations and e-learning tools provide support to students in understanding the properties of statistical distributions. However, there are many other important research skills that students need to acquire, especially at the early stages of the research cycle, such as experimental design, developing and testing research questions and hypotheses, and identifying relevant variables. A number of papers were identified in the review describing modelling applications which deal with these earlier stages in the research methods cycle.

Lehti and Lehtinen (2005) described the ALEL (artificial laboratory for exploratory learning) program which supports students in creating their own experimental designs by defining sequences of actions based on real examples. In a before and after design, performance on a group using ALEL was compared with two control groups. Results indicated that, while all groups improved from pre-test to post-test, only students in the ALEL group performed significantly better in the post-test than the pre-test.

Wu, Hsu, and Hwang (2010) described the development of a technology enhanced modelling tool (Air Pollution Modelling Tool) to support students as they engaged in modelling of air pollution. The tool makes the research process easier by providing scaffolding which decomposes the task into smaller sub-tasks, such as identifying multiple variables and providing multiple representations, to help students visualize data and relationships. Wu et al. did not include an evaluation of this tool and this paper was not selected in the review, but Wu (2010) reported both qualitative and quantitative evaluation of the APoME tool. Qualitative analysis of process videos, class videos, and interview transcripts revealed four different processes in modelling: making a plan, identifying and connecting variables, designing and examining a model and generating and applying findings. In a pre-test, post-test design Wu found that students' understanding about air quality was significantly improved after they engaged in the APoME activities. Students performed better on modelling abilities such as planning, identifying variables, and testing models.

Valanides and Angeli (2008) described the use of a computer modelling tool (Model-It) to help pre-service primary school teachers to think about the early stages of research design, such as developing and refining a research question and operationalizing variables. The tool allowed the teachers to build qualitative models of a domain by creating objects that correspond to observable entities in the real world, such as sun, soil, plants, air etc. The teacher then represented variables associated with each object such as light, water, growth, carbon dioxide etc. and then identified these as causal or not. Relationships between variables (increases, decreases etc.) can also be modelled and represented graphically. The results of the study showed that the teachers managed to build models but they found the task complex and their models tended to be very simple.

Garfield, del Mas, and Zieffler (2010) described the use of model-eliciting activities (MEAs) in teaching methodological and statistical concepts. MEAs are "open-ended problems that encourage students to build mathematical models in order to solve complex problems". Garfield et al. argued that this modelling approach engages students in statistical thinking by allowing them to "experience the statistical enquiry cycle". The paper included only minimal formative evaluation in the form of authors' comments about the value of the model.

4. Discussion

This narrative review focused on identifying papers which provide empirical evidence about the value of digital games, animations and simulations in learning about research methods and statistics. Although many irrelevant papers were identified, application of our inclusion criteria identified 26 relevant papers. Taken together these papers were generally positive about the impact of such applications in leading to more successful learning in research methods and statistics.

With respect to the research approach used in the studies, 14 papers adopted a quantitative approach, 7 a qualitative approach with 5 reporting both quantitative and qualitative data (see Table 2). Several quantitative papers described between group comparisons which showed better performance on the target skill by the group using the electronic support compared with a control group. This was true for

studies in all the different categories: entertainment games (Chow, Woodford, & Maes), serious games (Halpern et al.,), animations (Wender & Muehlboeck), simulations (Liu, Lin, & Kinshuk), elearning tools (Dinov) and models (Lehti & Lehtinen). Consistent with Connolly et al. (2012) there was an absence of RCTs, although some higher quality studies including Stansbury and Munro, Liu, Lin and Kinshuk, Lehti and Lehtinen, Halpern et al. and Wu carried out pre-tests to confirm that there were no differences in performance on the target skills prior to the intervention. Few studies reported the random allocation of participants to experimental conditions. While this is a requirement of an RCT, in educational interventions participants are more frequently allocated to conditions at the class level for pragmatic reasons. Allocation of participants to conditions by class clearly suffers from the criticism that the classes may differ in some salient respects.

The key issue in designing and developing effective serious games is to have a clear understanding of what the desired learning outcomes are and to establish a transparent mapping between the affordances of the game and desired learning outcomes (Connolly et al., 2012). The knowledge and skills required in understanding research methods and statistics are broad-ranging, complex and interconnected and include logical and scientific reasoning, critical thinking, use of evidence and argument, understanding different representations of data, data analysis, interpretation of results and evaluation skills. Given the breadth of these skills, it is necessary to specify clearly which skills are being targeted in developing a specific game. The review found that games, animations and simulations supported players in acquiring a wide range of methodological and statistical skills, including factorial design (Stansbury & Munro), the sampling distribution and the Central Limit Theorem (Mills), experimental design and statistical inference (Lehti & Lehtinen), expected value (Chow, Woodford, & Maes), hypothesis testing (Wang, Vaughn, & Liu), scientific reasoning and argumentation skills (Steinkuehler & Duncan) and understanding the properties of the normal curve (Nte and Stephen).

In addition to supporting these varied skills, several papers reported that players found games, animations and simulations interesting and engaging (Neumann, Hood, & Neumann) and enjoyable (Ramler & Chapman; Asbell-Clarke et al.). Engagement was largely reported through qualitative analysis of players' comments rather than more rigorous quantitative approaches. Gresalfi and Barab found that pupils' and teachers' dialogues demonstrated procedural, conceptual, consequential and critical engagement in the Anders city game. Butler, Forsyth, Halpern, Graesser, and Millis (2011) claimed that games provide "hard fun" meaning that players experience great satisfaction after completing tasks which are very challenging.

Games, animations and simulations contribute to more effective learning in a number of ways. Games and e-learning applications seem to provide activities which mirror those identified by Garfield et al. (2008) and others in their research on the characteristics of successful and effective "Statistical Reasoning Learning Environments". This suggested that effective teaching should provide students with active, problem based methods of learning, where they can apply their knowledge to real world datasets, take part in practical tasks, think about different representations of data, develop critical thinking and statistical argumentation skills and have opportunities to communicate about statistics.

Animations and simulations work by extending the visual power of static graphical representations to help participants comprehend the properties of statistical distributions more easily. In addition animations and simulations are visually engaging and this seems to help dispel some of the anxieties that many students experience in this difficult content area (Nte and Stephens). The most effective animations/ simulations are also interactive and allow students to change variables and experience the effects of these changes (Wang et al.). Modelling might be of particular interest in developing a game to support the early stages of thinking about a research question, as it provides the kind of generative and experimental thinking which might be useful in a games-based approach where players are required to make decisions and can experience the effects of these decisions.

The most obvious way in which traditional entertainment games support understanding of methodological and statistical principles is where the logical structure or probabilistic nature inherent to the games illustrates these principles. This is true even with the simplest dice and coin tossing games (Dunn, 2005). Nordmore (2004) provides a tongue-in-cheek example of how playing the simple game, Poohsticks, helped Winnie the Pooh to develop an awareness of basic elements of scientific reasoning and hypothesis testing. In the game Pooh had to make predictions about which of two cones thrown into a river from a bridge will emerge first from the other side of the bridge based on their size.

In the current review, Chow et al. (2010) described how the intrinsic structure of the digitised version of the "Deal or No Deal" game helped students to understand the concept of expected value in probability theory, while Bottino et al. found provisional evidence that the logical and scientific thinking required in playing mind games transfers to higher scores on National Maths tests. However the use of Commercial Off-The-Shelf games to learn methodological and statistical concepts is constrained by the nature of the logical and probablistic structure of the activities built into the games.

There was evidence too that players of MMOGs engage in the kind of higher level discussion, argument and critical thinking required in scientific thinking (Steinkuehler & Duncan). There is much interest in how this informal learning can be honed in the more formal learning environments of serious games.

Did the study reveal features of serious games which supported students' learning about research methods and statistics? The serious game which was conceptually closest to the proposed CHERMUG game was Halpern et al.'s Operation ARA which aimed to teach scientific method and experimental design via a mystery/role playing game with an "aliens" narrative. This game covered a broad range of methodological concepts from the different phases of the research cycle. Both Operation ARA and the Martian Boneyards game (Asbell-Clarke et al.) used a mystery game genre where players take on the role of a scientist, formulating hypotheses and searching for information or evidence to test these hypotheses. The mystery game seems a promising format for supporting sustained skills of scientific enquiry since the game activities reflect the nature of scientific reasoning, where players solve different kinds of puzzles and challenges at different stages and where certain facts are known and others have to be worked out according to the narrative.

4.1. Limitations

The content areas used in the different games/electronic activities were very varied. For example Gresalfi's hypothesis was about which brand of bike is safer to offer for rental in the park, Ramler and Chapman's related to the distribution of missing notes in playing music,

Ancker's game related to health decision making, while Wang et al.'s animation looked at which bag a specific voucher comes from. Methodological and statistical skills are highly generic in the sense that once they have been acquired they should be applicable across a range of content. However the transferability of these skills to different kinds of problem would need to be tested.

Papers which looked at general issues to do with developing serious games were excluded from consideration in the current review since they have been dealt with elsewhere, but also because there was an assumption that there are game features which are specifically important for teaching research methods and statistics and which are different to those found in other kinds of games, animations and simulations. Could this claim be supported? Many of the feature discussed above: the inherent logical or probablistic structure of entertainment games, the dynamic visual properties of animations and simulations, the use of the mystery game genre for scientific enquiry, the use of different representational formats, do seem to be, if not unique to methods and statistics, at least to be relatively more important in teaching these complex skills. We should not forget however that the characteristics, functionality and mechanics of games generally, such as narrative, challenge, goals, dialogue, feedback, rewards, cooperation and competition, will also be useful for games for learning about research methods and statistics.

While this review does provide preliminary evidence that games, animation and simulations can assist our understanding of research methods and statistics, Härdle, Klinke, and Ziegenhagen (2007) offer a more pessimistic appraisal of the value of e-learning applications to support learning in statistics. They reviewed a variety of e-learning applications ranging from simple Excel spread-sheets to integrated learning environments on the basis of HTML and Java tools. While acknowledging that e-learning applications can help students they concluded that "the utility is marginal in comparison with the effort invested".

4.2. Conclusion

The current narrative review revealed that there are currently few rigorous evaluations of game-based approaches in this area, although the general thrust of the evidence was positive and there is reason to be optimistic about the potential of a games-based approach. The activities provided by games, animations and simulations provide active approaches to learning which are consistent with modern ideas about how best to teach research methods and statistics. Progress will be made by looking in more detail at how characteristics of games, both at the higher level of game genre but probably more usefully at the lower level of game mechanics, lead to more effective learning. To make the most effective use of electronic resources such as games in teaching, they should be incorporated within the context of well-developed and coherent curricula for teaching research methods.

The CHERMUG games were initially targeted at nursing and social science students. However with the focus on key competences required by students and workers in 2020, research methods and statistics skills are clearly very important for other disciplines. The literature reviewed indicates that the potential benefits of such games could be on a very large scale.

Acknowledgments

Appendix A. Summaries of the 26 selected papers.

	Research approach, study design and aim	Topic	Authors	Game	Results
Entertainment games (*N* = 5)					
1	**Quantitative**: A mixed-design ANOVA looking at pre- to post-test performance of a group of students who used the game (*N* = 26) to generate data compared with a lecture-only group (*N* = 23) and a no-content control group (*N* = 16).	Factorial design	Stansbury and Munro (2013) **ERIC**	Dance Dance Revolution (DDR) on Wii	There was a significantly greater increase from pre-test to post-test in content knowledge (understanding of factorial design) when using the game (Wii DDR) compared to the lecture-only and no-content control conditions.
2	**Qualitative**: Very limited qualitative evaluation in the form of students' comments about the project.	Statistical thinking about data collection, hypothesis development, data analysis.	Ramler and Chapman (2011) **ERIC**	Guitar Hero	Detailed description of how Guitar Hero was used to support statistical thinking in the module; students' comments about the project were generally favourable.
3	**Qualitative**: Analysis of 1984 players' contributions to online discussion fora.	Scientific reasoning and argumentation skills	Steinkuehler and Duncan (2008) **ERIC**	Massive multiplayer online game (MMOG) World of Warcraft (WOW)	Coding of the posts on the WOW fora showed that the players demonstrated high level scientific reasoning skills: using data and arguments, system-based reasoning.

(continued)

	Research approach, study design and aim	Topic	Authors	Game	Results
4	**Quantitative**: Between groups comparison of retention of the concept of expected value (illustrated in the game) of 32 students who played the online game ($N = 32$) with students who did not play the game ($N = 29$).	Expected value	Chow et al. (2010) **SCIENCE DIRECT**	Online version of "Deal or no deal" game	95% of students (30/32) using the online game correctly answered questions about expected value one week after the intervention, but only 59% of students (12/29) who had not played the game.
5	**Both**: To identify the design and interface characteristics of digital mind games that can foster the development of reasoning and problem solving abilities in primary school students. Scores on national mathematics tests of 2 classes which played the games were compared with 2 classes which did not play game.	Reasoning and problem solving abilities	Bottino et al. (2007) **SCIENCE DIRECT**	Digital mind games (brainteasers and puzzlers) such as PappaLOTTO and Hexip	The characteristic of mind games that can foster the development of reasoning and problem solving abilities includes (a) adjustable level of difficulty, (b) availability of hints and ability to sustain and orient cognitive activities. The mean scores of the two experimental groups were higher than scores for the two control groups, although this difference was not tested statistically.
Serious games ($N = 5$)					
6	**Quantitative**: An RCT compared the game group ($N = 58$) with a no-game control group ($N = 78$) in a pre-test, post-test design. Students were tested on 17 scientific reasoning concepts.	Critical thinking and scientific reasoning skills	Halpern et al. (2012) **SCIENCE DIRECT**	Operation ARA (Acquiring Research Acumen) – an adventure/mystery game which provides a narrative about alien invaders.	Students who played operation ARA had higher proportional learning gains compared with students in the control group who did not play the game.
7	**Both**: Descriptive data and qualitative analysis of the behaviours and communications of 613 players of the game to establish how the game develops skills of systematic scientific enquiry in the area of paleontology; experts rated the quality of the evidence that the players produced.	Systematic scientific enquiry in the area of paleontology	Asbell-Clarke et al. (2012) **Previously known from IGI GLOBAL.com**	Martian Boneyards, a Mystery game	Players spent on average 30 h in the game (with some playing up to 200 h) over a four month period suggesting that players enjoyed playing the game. Players spent most of their time gathering data with 30% spent in in-game interactions in the analysis and theory building phases of the game. Analysis of the dialogue showed that players asked questions, made claims and substantiated claims with evidence with commendable accuracy and depth.
8	**Quantitative**: A questionnaire study which compared the effects of interactive, game-like graphics with static graphics on risk perception and health decision making. ($N = 165$; 100 online and 65 from clinic.)	Risk perception and health decision making (risk feelings, quantitative risk estimates, and intentions to take protective action).	Ancker et al. (2011) **Previously known**	A game-like interactive graphic presentation of statistical information	The different graphics did not effect on risk feelings or estimates, although the difference between low and high numeracy respondents was reduced with the interactive search graphic. Order effects suggested that the way in which information is presented can effect risk feelings and perceptions in health decisions.
9	**Quantitative**: Survey of players' perceptions ($N = 55$); of usefulness of the game for learning about the properties of the normal curve.	Properties of the normal curve	Nte and Stephen (2008) **ecgbl**	Player is a virtual scientist in a game environment	Students agreed that the game helped them understand about the normal distribution and caused no anxiety.
10	**Qualitative**: Study of pupils' and teachers' dialogues. To illustrate how Ander city can encourage children (aged 9–16) to take a more active role as statisticians by providing engaging and realistic decision making activities.	Decision making activities	Gresalfi and Barab (2011) **ERIC**	Ander city, the statistics unit from the online multiplayer video game, Quest Atlantis	Examples of pupils' and teachers' dialogues illustrated procedural, conceptual, consequential and critical engagement and how these support more effective learning.

(continued)

	Research approach, study design and aim	Topic	Authors	Game	Results
Animations and simulations (*N* = 5)					
11	**Quantitative**: Between groups comparison of performance on abstract statistical concepts. Students were randomly assigned to an experimental group (*N* = 14) (who used the excel simulation) with students who were assigned to a control group (*N* = 17) who learnt in the traditional way.	Abstract statistical concepts related to the sampling distribution and the Central Limit Theorem	Mills (2004) **ERIC**	Simulation using excel	Students in the experimental group performed significantly better post-test and had more positive attitudes to the instruction than those in the control group. Students with high spatial ability performed better than those with low spatial ability, but there was no interaction between spatial ability and graphics condition.
12	**Quantitative**: Compared the effects of computer-animated graphics with static graphics in teaching abstract statistical concepts (the multiplication of two matrices, the covariance of two random variables, the method of least squares in linear regression, *a* error, *b* error, and strength of effect). (*N* = 112; students were randomly assigned to an animation or static group.)	Abstract statistical concepts	Wender and Muehlboeck (2003) **Previously known**	Animated diagrams	Animated graphics led to significantly better learning on a 10-item comprehension test for these statistical concepts presented immediately after the training. The factor *animation* showed a significant effect [$F(1,111) = 9.09$, $p < 0.003$] accounting for 44% of the variance ($h^2 = 0.44$).
13	**Quantitative**: Compared the effects of 4 different kinds of interactive animation (static group, simple animation group, input group and practice group) on understanding of hypothesis testing (*N* = 123; students were randomly assigned to 4 groups).	Hypothesis testing	Wang et al. (2011) **SCIENCE DIRECT** **ebsco**	Interactive animation programme by Adobe Flash to facilitate hypothesis testing	Animation interactivity impacted on students' improvement in understanding but did not have a significant impact on confidence or program perception. The simple animation, input and practice groups were significantly better that the static group in understanding.
14	**Both**: Compared performance on misconceptions in understanding correlation. 72 Grade-12 students were randomly assigned to either the experimental group, who used the SALS simulation, or a control group, who received lecture only material	Misconceptions in understanding correlation	Liu et al. (2010) **ebsco**	SALS simulation	McNemar tests showed that students in the SALS group showed improvements in all 10 misconceptions, while students in the control group showing improvements in only 3 misconceptions. ANCOVA of post-test scores with pre-tests scores and learning time as covariates showed higher scores for the SALS group than the control group.
15	**Qualitative**: Students' comments in interviews of the interactive assessment approach were conceptually coded (*N* = 38).	Perceived advantages of the interactive assessment approach	Neumann et al. (2012) **ERIC**	Computer-based interactive simulations including JAVA applets and FLASH-based application in the assessment of statistical concepts	Qualitative coding of students interviews found that the interactive assessment approach benefitted learning by providing immediate feedback, a visual aid to learning and practice. Motivational benefits were also found such as creating interest and engagement.
Web based tools and e-learning applications (*N* = 7)					
16	**Quantitative**: Compared performance on computational skills and conceptual knowledge about chi-square for the treatment group (who used Excel and Flash movies) (*N* = 38) with a control group who learnt traditionally (*N* = 32).	Computational skills and conceptual knowledge about chi square	Burruss and Furlow (2007)	Microsoft Excel spread-sheet and Macromedia Flash movies	No significant difference was found between the 2 groups on a chi-square skills test following the intervention, but the experimental group performed better than the control group on conceptual definitions of chi square.

(continued)

	Research approach, study design and aim	Topic	Authors	Game	Results
	Students were randomly assigned to treatment or control group.				
17	**Quantitative**: A between groups comparison of students using the SOCR treatment group ($N = 88$) with students in the control group ($N = 83$) which used a classic instructional method. The design was quasi-experimental as students were not randomly assigned to groups.	General statistical abilities and skills	Dinov et al. (2008) **NIH PUBLIC ACCESS**	The Statistics Online Computational Resource (SOCR), a large online resource for teaching probability and statistics which includes instructional materials, statistical calculators, interactive tools for data analysis and visualisations, computational and simulation applets etc.	The impact of SOCR was large and many students benefited from the online learning. There was also less variance in the results for the SOCR group showing more consistency in performance.
18	**Quantitative**: A between groups comparison of students of students' understanding of statistics in a group who used the SPSS data analysis programme as well as lectures ($N = 65$) with a group who only received the lectures ($N = 140$).	Statistics	Basturk (2005) **ERIC**	Computer Assisted Instruction (CAI) for SPSS	Students in the lecture plus SPSS group gained higher averages on the mid-term and final exams than those in the lecture only group.
19	**Quantitative**: A within groups comparison of the effects of e-status of statistical abilities. Half the students had access to e-status on Part A ($N = 61$) and half on Part B ($N = 60$) of the course.	General statistical abilities and skills	González et al. (2010) **ERIC**	e-Status, a web-based e-learning tool with statistical exercises for students and immediate feedback as they tackled these exercises.	Performance on the problems where students used e-status was significantly better than that on problems where e-status was not used.
20	**Quantitative**: A pre-/post-test design looking at the effect of working alone or in groups and feedback intervention (feedback available or not) on performance on statistical examples concerning correlations and scatterplots. ($N = 137$).	Correlations and scatterplots	Krause et al. (2009) **SCIENCE DIRECT**	e-Learning environment: Koralle to support learning of correlations and scatterplots	Feedback led to improved performance but working cooperatively in dyads did not.
21	**Qualitative**: 15 staff members participated in 3 focus groups to establish areas where students find most difficulties in learning research methods; 27 focus groups of students evaluated the video resources.	Students' difficulties in learning research methods	Callaghan et al. (2011) **SCIENCE DIRECT**	Web-based video resource	Analysis of students' focus group responses showed that they found the web-based video resources useful in learning about the topics identified by the staff and compared favourably with current resources; students liked the combination of audio and visual material on the animated PowerPoint presentations.
22	**Qualitative**: To examine whether a blended learning approach could provide support for learning SPSS. Formative evaluation based on observation and feedback from staff and students.	SPSS skills	Archee and Gurney (2006)	Online Flash video tutorials (created with Captivate) used in teaching SPSS	The main benefit was that the Flash tutorials helped students to work independently at their own pace.
Modelling tools ($N = 4$)					
23	**Quantitative**: A between groups comparison of performance of the treatment group (who used the ALEL learning environment) with 2 control groups (an article group, which had to analyse an article, and a statistics	Experimental design and statistical inference	Lehti and Lehtinen (2005) **eric**	The ALEL (artificial laboratory for exploratory learning) program: presents information on real examples, requires students to carry out multiphase activities typical of professional practice to obtain	There was an interaction between treatment group and pre- and post-test scores with students in the ALEL group performing better than the two control groups on the post-test.

(continued)

	Research approach, study design and aim	Topic	Authors	Game	Results
	group, which focused on experimental design and statistical thinking). 32 students of educational science took part.			information, provides simulations showing different representations of the information and shows the problem-solving path students have taken.	
24	**Both**: Qualitative analysis of process videos, class videos, and interview transcripts to identify different processes in modelling. Performance scores pre- and post-intervention were measured.	Identifying the multiple variables and providing multiple representations to help students visualize data and relationships	Wu (2010) Mentioned in Wu et al.'s (2010) article identified on **ERIC**	Air Pollution Modelling Tool (APoMT)	Four different processes in modelling were identified: making a plan, identifying and connecting variables, designing and examining a model and generating and applying findings. A paired t-test showed students' conceptual understandings and modelling abilities concerning air quality were improved from pre- to post-test.
25	**Both**: Investigation of the use of Model-it on elementary teachers' ($N = 47$) skills in developing qualitative models in science and designing science lessons.	Skills in designing science lessons	Valanides and Angeli (2008) **SCIENCE DIRECT**	MODEL-IT, a computer-modelling tool	Model-It enabled the majority of teachers (78%) to build models that had a correct structure, but 28% found it difficult to identify appropriate entities and variables. 87% constructed models with 2 or 3 entities.
26	**Qualitative**: Minimal formative evaluation – authors' comments about the value of the model	Statistical thinking such as the recognition of the need for data, transnumeration, variation and the use of models	Garfield et al. (2010) **ICOTS-8**	Model eliciting activities (MEAs)	The authors commented that MEAs engage students in the types of thinking which are fundamental to statistical thinking such as the recognition of the need for data, transnumeration, variation and the use of models.

References

Ancker, J. S., Weber, E. U., & Kukafka, R. (2011). Effects of game-like interactive graphics on risk perceptions and decisions. *Medical Decision Making, 31*(1), 130–142.

Archee, R., & Gurney, M. (2006). Flash me now: using blended learning to teach SPSS to undergrads. *IADIS Virtual Multi Conference on E-learning*.

Asbell-Clarke, J., Edwards, T., Rowe, E., Larsen, J., Sylvan, E., & Hewitt, J. (2012). Martian boneyards: scientific inquiry in an MMO game. *International Journal of Game-based Learning, 2*(1), 52–76.

Barker, J. (2010). *Evidence-based practice for nurses*. London: Sage Publications.

Basturk, R. (2005). The effectiveness of computer-assisted instruction in teaching introductory statistics. *Educational Technology & Society, 8*(2), 170–178.

Bottino, R. M., Ferlino, L., Ott, M., & Tavella, M. (2007). Developing strategic and reasoning abilities with computer games at primary school level. *Computers & Education, 49,* 1272–1286.

Bowman, A. (2010). Statistical cartoons: the role of graphics in understanding statistics, ICOTS-8.

Boyle, E., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing,* 69–74.

Bulmer, M. (June 2004). Virtual worlds for teaching statistics, CAL-laborate.

Burguillo, J. C. (2010). Using game-theory and competition-based learning to stimulate student motivation and performance. *Computers & Education, 55*(2), 566–575.

Burruss, G. W., & Furlow, M. H. (2007). Teaching statistics visually: a quasi-experimental evaluation of teaching chi-square through computer tutorials. *Journal of Criminal Justice Education, 18*(2), 209–230.

Bush, S., Menzies, G., & Thorp, S. (2009). An array of online teaching tools. *Teaching Statistics, 31*(1), 17–20.

Butler, H., Forsyth, C., Halpern, D., Graesser, A. C., & Millis, K. (2011). *Secret agents, alien spies, and a quest to save the world: operation ARIES! Engages students in scientific reasoning and critical thinking*. In R. L. Miller, R. F. Rycek, E. Amsel, B. Kowalski, B. Beins, K. Keith, et al. (Eds.), *Programs, techniques and opportunities* (Vol. 1). Syracuse, NY: Society for the Teaching of Psychology. Available from the STP web site: http://www.teachpsych.org/teachpsych/pnpp/.

Callaghan, L., Lea, S. J., Mutton, L., & Whittlesea, E. (2011). Enhancing health students' understanding of generic research concepts using a web-based video resource. *Nurse Education in Practice, 11,* 375–379.

Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: a review of the empirical evidence from research on statistics education. *Educational Research Review, 2,* 98–113.

Chang, K.-E., Wu, L.-J., Weng, S.-E., & Sung, Y.-T. (2012). Embedding game-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education, 58,* 775–786.

Chang, S. -B., Deng, Y. -C., Cheng, H. N. H., Liao, H. -C., Yu, F. -Y., Chan, T. -K. (2007). Implementation and evaluation of EduBingo for arithmetic drill. The First IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning (DIGITEL'07).

Chow, A. F., Woodford, K. C., & Maes, J. (2010). Deal or no deal: using games to improve students learning, retention and decision making. *International Journal of Mathematical Education in Science and Technology, 42*(2), 259–264.

Connolly, T. M., Boyle, E. A., Hainey, T., MacArthur, E., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59,* 661–686.

Denner, J., Werner, L., Campe, S., & Ortiz, E. (2013). Using game mechanics to measure what students learn. In *Proceedings of the 7th ECGBL* (pp. 344–351). Porto: Instituto Superior de Engenharia do Porto.

Diekema, A. R., Holliday, W., & Leary, H. (2011). Re-framing information literacy: problem-based learning as informed learning. *Library & Information Science Research, 33*(4), 261–268.

Dinov, I. D., Sánchez, J., & Christou, N. (2008). Pedagogical utilization and assessment of the statistic online computational resource in introductory probability and statistics courses. *Computers & Education, 50,* 284–300.

Dominguez-Dominguez, J., Dominguez-Lopez, J. A. (2010). A visual approach in the teaching of statistics and probability, ICOTS-8 (2010) Invited Paper.

Dondlinger, M. J. (2007). Educational video games design: a review of the literature. *Journal of Applied Educational Technology, 4*(1), 21–31.

Dunn, P. K. (2005). We can still learn about probability by rolling dice and tossing coins. *Teaching Statistics, 27*(2), 37–41.
Emanuel, V., Day, K., Diegnan, L., & Pryce-Miller, M. (2011). Developing evidence-based practice among students. *Nursing Times, 107*(49/50), 21–23.
Forbes, S. (2010). Data visualisation: a new statistical literacy tool for statistical offices. Proceedings of the 7th Australian Conference on Teaching Statistics.
Gal, I. (2005). Statistical literacy: meanings, components, responsibilities, the challenge of developing statistical literacy. *Reasoning and Thinking, 1*, 47–78.
Garfield, J., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: implementing a statistical reasoning learning environment. *Teaching Statistics, 31*(3), 72–77.
Garfield, J. B., Ben-Zvi, D., Chance, B., Medina, E., & Roseth, C. (2008). Creating a statistical reasoning learning environment. *Developing Students' Statistical Reasoning, 1*, 45–63.
Garfield, J., del Mas, R., & Zieffler, A. (2010). Developing tertiary-level students' statistical thinking through the use of model-eliciting activities. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS-8, July, 2010), Ljubljana, Slovenia.* Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE].
González, J. A., Jover, L., Cobo, E., & Muñoz, P. (2010). A web-based learning tool improves student performance in statistics: a randomized masked trial. *Computers & Education, 55*(2), 704–713.
Gordon, S. P., & Gordon, F. (2009). Visualizing and understanding probability and statistics: graphical simulations using Excel. *PRIMUS, 19*(4), 346–369.
Gresalfi, M., & Barab, S. (2011). Learning for a reason: supporting forms of engagement by designing tasks and orchestrating environments. *Theory Into Practice, 50*(4), 300–310.
Gros, B. (2007). Digital games in education: the design of games-based learning environments. *Journal of Research on Technology in Education, 40*(1), 23–38.
Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: a computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity, 7*, 93–100.
Härdle, W. K., Klinke, K., Ziegenhagen, U. (2007). On the utility of e-learning in statistics, SFB 649 discussion paper, No. 2007,050.
Hogben, L. (1950). *Chance and choice by cardpack and chessboard.* In *An introduction to probability in practice by visual aids* (Vol. 22). London: Max Parrish and Co. Limited.
Joint Information Systems Committee JISC. (2007). *Game-based learning: briefing paper (online).* Available: http://www.jisc.ac.uk/media/documents/publications/gamingreportbp.pdf.
Kanthan, R., & Senger, J.-L. (2011). The impact of specially designed digital games-based learning in undergraduate pathology and medical education. *Archives of Pathology and Laboratory Medicine, 135*, 135–142.
Kebritchi, M., Hirumi, A., & Bai, H. (2012). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 58*, 775–786.
Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction, 19*, 158–170.
Lane, D. M., Peres, S. C. (2006). Interactive simulations in the teaching of statistics: promise and pitfalls. ICOTS-7, 1–6.
Larwin, K., & Larwin, D. (2011). A meta-analysis examining the impact of computer-assisted instruction on postsecondary statistics education: 40 years of research. *Journal of Research on Technology in Education, 43*(3), 253–278.
Lehti, S., & Lehtinen, E. (2005). Computer-supported problem-based learning in the research methodology domain. *Scandinavian Journal of Educational Research, 49*(3), 297–323.
Leppink, J. (2010). Adjusting cognitive load to the student's level of expertise for increasing motivation to learn. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia.* Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE].
Lewandowsky, S., & Spence, I. (1989). The perception of statistical graphs. *Sociological Methods and Research, 18*, 200–242.
Liu, T.-C., Lin, Y.-C., & Kinshuk. (2010). The application of Simulation-Assisted Learning Statistics (SALS) for correcting misconceptions and improving understanding of correlation. *Journal of Computer Assisted Learning, 26*(2), 143–158.
Magnussen, R., Hansen, S. D., Plank, T., & Sherson, J. F. (2013). Scientific discovery games for authentic science education. In *Proceedings of the 7th ECGBL* (pp. 344–351). Porto: Instituto Superior de Engenharia do Porto.
Mayer, I. (2012). Towards a comprehensive methodology for the research and evaluation of serious games. *Procedia Computer Science, 15*, 233–247.
Mills, J. D. (2004). Learning abstract statistics concepts using simulation. *Educational Research Quarterly, 28*(4), 18–33.
Nadolski, R. J., Hummel, H. G., van den Brink, H. J., Hoefakker, R., Slootmaker, A., Kurvers, H., et al. (2008). Emergo: a methodology and toolkit for developing serious games in higher education. *Simulation & Gaming, 39*(3), 338–352.
Neumann, D. L., Hood, M., & Neumann, M. (2012). An evaluation of computer-based interactive simulations in the assessment of statistical concepts. *International Journal for Technology in Mathematics Education, 19*(1), 17–23.
Newell, R., & Burnard, P. (2011). *Research for evidence-based practice in healthcare* (2nd ed.). Oxford: Wiley-Blackwell.
Nikiforidou, Z., Lekka, A., & Pange, J. (2010). Statistical literacy at university level: the current trends. *Procedia Social and Behavioral Sciences, 9*, 795–799.
Nordmore, E. D. (2004). Of poohsticks and *p*-values: hypothesis testing in the hundred acre wood. *Teaching Statistics, 26*(2), 56–58.
Nte, S., & Stephen, R. (2008). Videogame aesthetics and e-learning: a retro-looking computer game to explain the normal distribution in statistics teaching. In *Proceedings of the 2nd ECGBL* (pp. 341–347). Barcelona: Universitat Oberta de Catalunya.
Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking, the challenge of developing statistical literacy. *Reasoning and Thinking, Part I,* 17–46.
Quinn, R. J. (2003). Exploring the probabilities of 'Who wants to be a millionaire'? *Teaching Statistics, 25*(3), 81–84.
Ramler, I. P., & Chapman, J. L. (2011). Introducing statistical research to undergraduate mathematical statistics students using the guitar hero video game series. *Journal of Statistics Education, 19*(3), 1–20.
Stansbury, J. A., & Munro, G. D. (2013). Gaming in the classroom: an innovative way to teach factorial designs. *Teaching of Psychology, 40*(2), 148–152.
Steinkuehler, C., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology, 17*, 530–543.
Stirling, D. (2010). Improving lectures with cast applets. ICOTS-8 (2010) Invited Paper Refereed.
Tishkovskaya, S., Lancaster, G. A. (2010). Teaching strategies to promote statistical literacy: Review and implementation. ICOTS-8 Contributed Paper.
Valanides, N., & Angeli, C. (2008). Learning and teaching about scientific models with a computer-modeling tool. *Computers in Human Behavior, 24*, 220–233.
Wang, P.-Y., Vaughn, B. K., & Lee, M. (2011). The impact of animation interactivity on novices' learning of introductory statistics. *Computers & Education, 56*(1), 300–311.
Wender, K. F., & Muehlboeck, J.-S. (2003). Animated diagrams in teaching statistics. *Behavior Research Methods, Instruments, Computers, 35*(2), 255–258.
Winters, C. A., & Echeverri, R. (2012). Teaching strategies to support evidence-based practice. *Critical Care Nurse, 32*(3), 49–54.
Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education, 60*(1), 412–425.
Wu, H.-K. (2010). Modelling a complex system: using novice-expert analysis for developing an effective technology-enhanced learning environment. *International Journal of Science Education, 32*(2), 195–219.
Wu, H.-K., Hsu, Y.-S., & Hwang, F.-K. (2010). Designing a technology-enhanced learning environment to support scientific modelling. *The Turkish Online Journal of Educational Technology, 9*(4), 58–65.
Xie, Y. (2013). Animation: an R package for creating animations and demonstrating statistical methods. *Journal of Statistical Software, 53*(1), 1–27.