An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games
Boyle, Elizabeth; Hainey, Thomas; Connolly, Thomas M.; Gray, Grant; Earp, Jeffery; Ott, Michela; Lim, Theodore; Ninaus, Manuel; Pereira, João; Riberio, Claudia

Published in:
Computers and Education

DOI:
10.1016/j.compedu.2015.11.003

Publication date:
2016

Document Version
Peer reviewed version

Link to publication in Heriot-Watt University Research Gateway

Citation for published version (APA):
Accepted Manuscript

An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games

Elizabeth A. Boyle, Thomas Hainey, Thomas M. Connolly, Grant Gray, Jeffrey Earp, Michela Ott, Theodore Lim, Manuel Ninaus, João Pereira, Claudia Riberio

PII: S0360-1315(15)30075-0
DOI: 10.1016/j.compedu.2015.11.003
Reference: CAE 2941

To appear in: Computers & Education

Received Date: 6 July 2015
Revised Date: 29 October 2015
Accepted Date: 4 November 2015


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games

Elizabeth A. Boyle\textsuperscript{1}, Thomas Hainey\textsuperscript{2}, Thomas M. Connolly\textsuperscript{2}, Grant Gray\textsuperscript{1}, Jeffrey Earp\textsuperscript{3}, Michela Ott\textsuperscript{3}, Theodore Lim\textsuperscript{4}, Manuel Ninaus\textsuperscript{5}, João Pereira\textsuperscript{6} and Claudia Riberio\textsuperscript{6}

\textsuperscript{1}School of Media, Culture and Society, University of the West of Scotland, High St, Paisley PA1 2BE, Scotland; email: liz.boyle@uws.ac.uk; grant.gray@uws.ac.uk
\textsuperscript{2}School of Engineering and Computing, University of the West of Scotland, High St, Paisley PA1 2BE, Scotland; email: thomas.connolly@uws.ac.uk; thomas.hainey@uws.ac.uk
\textsuperscript{3}Istituto Tecnologie Didattiche (ITD), Consiglio Nazionale delle Ricerche (CNR), Via De Marini, 6 - 16149 Genova; email: ott@itd.cnr.it; jeff@itd.cnr.it
\textsuperscript{4}Herriot Watt University, Institute of Mechanical, Process and Energy Engineering, Riccarton, Edinburgh, EH14 4AS, Scotland; email: T.Lim@hw.ac.uk
\textsuperscript{5}Department of Psychology, University of Graz, Graz, Austria; email: manuel.ninaus@uni-graz.at
\textsuperscript{6}INESC-ID, Lisboa, Portugal; email: jap@inesc-id.pt; claudia.sofia.ribeiro@inesc-id.pt

Correspondence to: Elizabeth A. Boyle, School of Media, Culture and Society, University of the West of Scotland, High St, Paisley PA1 2BE, Scotland; Tel: +44141 848 3783; email: liz.boyle@uws.ac.uk

This work has been co-funded by the EU under the FP7, in the Games and Learning Alliance (GaLA) Network of Excellence, Grant Agreement nr. 258169.
An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games

Abstract

Continuing interest in digital games indicated that it would be useful to update [Authors’, 2012] systematic literature review of empirical evidence about the positive impacts and outcomes of games. Since a large number of papers was identified in the period from 2009 to 2014, the current review focused on 143 papers that provided higher quality evidence about the positive outcomes of games. [Authors’] multidimensional analysis of games and their outcomes provided a useful framework for organising the varied research in this area. The most frequently occurring outcome reported for games for learning was knowledge acquisition, while entertainment games addressed a broader range of affective, behaviour change, perceptual and cognitive and physiological outcomes. Games for learning were found across varied topics with STEM subjects and health the most popular. Future research on digital games would benefit from a systematic programme of experimental work, examining in detail which game features are most effective in promoting engagement and supporting learning.

Keywords: Computer games, serious games, entertainment, engagement, learning, systematic literature review
1. Introduction

In recent years there has been an upsurge of interest in the use of digital games for learning and behaviour change with an increasing number of conferences, journals, projects and societies devoted to the topic. Much of this interest has been speculative, discussing the potential of games to provide new methods for supporting learning. To address concerns about the lack of empirical evidence about the effectiveness of games, [Authors] carried out a literature review that aimed to identify research evidence about the positive impacts of games. The review confirmed that research on games was very diverse with respect to the focus of the studies, the outcomes reported, underlying theoretical models and methodological approaches. [Authors] advocated a narrative review to capture the diversity of the selected studies and they developed a multi-component analysis of games and their outcomes which provided a useful framework for organising the research along key variables, making links between the different ways of categorizing games and outcomes, and identifying emerging trends, hot topics and gaps in the literature.

The effort and timescale involved in carrying out and publishing literature reviews means that they are out of date fairly quickly. [Author’s] review was published in 2012, but examined papers published between January 2004 and February 2009. Reflecting continuing interest in digital games, the aim of the current study was to update [Authors], looking at empirical papers published in the five year period since the end of that review, i.e. between March 2009 and February 2014. Like the previous review, the current review recognises the heterogeneity of the research in this relatively new field and aims to provide some organisation and structure to this area.

2. Method

2.1 Databases searched

Since initial searches had suggested that a very large number of papers would be found, the electronic databases searched in this review were a subset of those used in the previous review [Authors]: Science Direct, Applied Social Sciences Index and Abstracts (ASSIA), BioMed Central, Education
Resources Information Center (ERIC), Ingentaconnect and Institute of Electrical and Electronics
Engineers (IEEE). Limiting the number of databases selected provided a manageable number of papers
with ACM, EBSCO, Extended Academic ASAP, Emerald, Index to Theses and the ECGBL conferences
excluded from the current review.

2.2 Search terms

The search terms were also derived from the previous review and addressed the variety of games
that might be played:

(“computer games” OR “video games” OR “serious games” OR “simulation games” OR “games-
based learning” OR MMOG OR MMORPG OR MUD OR “online games”) as well as terms for the possible outcomes or impacts of playing games:

AND (evaluation OR impacts OR outcomes OR effects OR learning OR education OR skills OR
behaviour OR attitude OR engagement OR motivation OR affect)

2.3 Technology-assisted search

To assist in searching the electronic databases an automated retrieval system was used to identify articles
according to the search terms specified by the researcher. The software returned a large number of papers
(see Table 1, column 2). However many of these were irrelevant and to focus on a more manageable set
of papers, this set of papers was further screened by searching all titles and abstracts of papers for the
terms “game”, “gaming”, “MMOG” and “MMORG”. Papers selected on this basis were collated and
duplicates removed (see Table 1, column 3).

2.4 Selection of papers for inclusion in the review

To select appropriate studies for inclusion in the review, abstracts of these papers were read and a
number of further criteria were applied based on those used in the original review. Selected papers had to
(a) include empirical evidence relating to the impacts and outcomes of playing games, (b) date from
March 2009 to February 2014, (c) include an abstract and (d) include participants over the age of 14
years.

2.5 The multi-component coding of selected papers, games and learning outcomes
The papers selected on the basis of the inclusion criteria were coded using the data extraction pro-forma that was developed by Authors (2012) and further details can be found there. To summarise, papers were coded with respect to:

- **Variables related to the study**: the research design used in the study, categorized as a randomized control trial (RCT); quasi-experimental; survey; correlational or qualitative design
- **Game variables**: whether the game was a digital or non-digital game; whether the primary purpose of the game was as an entertainment game, a game for learning or a serious game; the game genre; the platform for the game; the subject discipline addressed in the game
- **Game outcome variables**: general or specific: whether the games described were games generally or a specific game; learning outcomes: knowledge acquisition, affective and motivational, perceptual and cognitive, behaviour change, physiological and social/soft skills outcomes, although slight revisions were introduced whereby motor outcomes were classified under physiological outcomes and a new category of skills outcomes was introduced

Two additional study-related variables were added in the current study:
- the country and continent where the first author of the paper was based
- the general aim of the study

2.6 Quality of the studies

Papers were quality assessed according to the 5 criteria used by Authors (2012): (1) quality of research design for answering the question: high, e.g. RCT; medium, e.g. quasi-experimental controlled study or low, e.g. pre-test/post-test design; (2) appropriate methods & analysis; (3) generalizability of findings; (4) relevance of the focus of the study (including conceptual focus, context, sample and measures) for addressing the research question and (5) trustworthiness of findings. Each paper was scored between 1 and 3 on these five dimensions and total scores for each paper varied between 5 and 15.

2.7 Consistency and reliability of coding

Seven researchers took part in the identification, coding and quality coding of papers but, for consistency, one experienced researcher oversaw all the coding. A sub-sample of 21 of the 143 papers
(15%) was coded independently by two of the coders. The inter-rater reliability for the total scores was .87, showing good agreement between the two coders concerning the quality of the papers.

3. Results

3.1. Papers identified by search terms

Table 1, column 2, shows that a very large number of papers (54,580) published in the time period March 2009-February 2014 was identified either manually or using the automated search. As discussed in section 2.1.3, this set of papers was further screened to focus on a still large, but more manageable, set of 7,117 relevant papers (see Table 1, column 3).

3.2. Papers selected using our inclusion criteria

Applying the four inclusion criteria to these papers, 512 papers were identified (see Table 1, column 4) and, to provide a more manageable task, the current review focused on the 143 papers that were quality rated as 9 or over, since these papers provide the strongest evidence about the impact of games. Appendix A provides summaries of all of these 143 papers and all are referenced in the coded papers reference section of this paper. The largest number of high quality papers was found in Sciencedirect, followed by ERIC, Ingenta, ASSIA, BioMed Central and IEEE.

Table 1

Total number of papers identified from each database, identified following screening, selected as meeting inclusion criteria and quality coded over 9

<table>
<thead>
<tr>
<th>Database</th>
<th>Number of papers found by database search</th>
<th>Number of papers identified following screening for games terms</th>
<th>Number of papers selected as meeting inclusion criteria</th>
<th>Number of papers rated over 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioMed Central</td>
<td>758</td>
<td>59</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>ASSIA</td>
<td>138</td>
<td>138</td>
<td>47</td>
<td>13</td>
</tr>
<tr>
<td>ERIC</td>
<td>1078</td>
<td>992</td>
<td>161</td>
<td>31</td>
</tr>
</tbody>
</table>
3.3 Analysis of game variables

This section describes the results of the analyses of variables related to the kind of game used.

3.3.1 Digital or non-digital games

All papers scoring over 9 addressed digital games, apart from one board game, RESORTES, described by Speelman, García-Barrios, Groot, and Tittonell (2014) which was a game for smallholder participation in the design of more sustainable agricultural landscapes.

3.3.2 Primary purpose of the game

Of the 143 high quality papers, 38 described games for learning while 34 described serious games. Since further analysis suggested that there were few differences between games for learning and serious games and the terms seem to be used more or less synonymously in the literature, for the purposes of this review, it was thought to be more informative to collapse games for learning and serious games into one category and to contrast these with entertainment games. There were similar numbers of papers looking at entertainment games (71) and those looking at games for learning (72).

3.3.3 Game genre

Game genre was very varied, with simulations (14), simulation games (10) and role-playing games (12) the most popular, followed by drill and practice games (9), MMORPGs (8), strategy games (8), puzzles (6) and adventure games (4). 7 studies, all looking at games for learning, discussed the use of game features rather than game genre and 30 studies, mostly concerning entertainment games, examined a variety of games. 23 of the 24 simulations or simulation games were games for learning. 7 of the 12 studies of role playing games, 5 of the 8 studies of MMORPGs and 6 of the 9 studies of drill and practice games concerned entertainment games.
3.3.4 Subject area

Games are being used across many subject disciplines, with games for Science, Technology, Engineering and Maths (STEM) subjects the most popular (24), closely followed by health (23). Of the STEM games there were 12 in science (e.g. Forsyth et al., 2012), 7 in computing (e.g. Coller & Scott, 2009), 2 in engineering (e.g. Joiner et al., 2011), 2 in Maths and 1 categorised generally as STEM. Games for health addressed topics such as exercise (Pichierri, Murer, & Bruin, 2012), rehabilitation (Li, Sevcenco, & Yan, 2013; Hurkmans, Ribbers, Streur-Kranenburg, Stam, & Berg-Emons, 2011), quality of life (Chen, Hsieh, Wei, & Kao, 2012) and advanced life support training (Cook, McAlloon, ONeill, & Beggs, 2011). Games for business (8) (e.g. Tan, Tse, & Chung, 2012), languages (6) (e.g. Chen & Yang, 2013) and economics (3) (e.g. Huang, Johnson, & Han, 2013) were also found, with a variety of games in other subjects such as planning, civic learning, relationships and problem solving. Games used in STEM subjects, business and economics tended to be games for learning, while games used in health were equally divided between entertainment games, such as Wii fit, and games for learning. 13 out of 15 papers categorised under social studies involved entertainment games. Games were also used in informal, extracurricular learning contexts, such as substance abuse (Verduin, LaRowe, Myrick, Cannon-Bowers, & Bowers, 2013) and healthy eating (Peng, 2009). As might be expected, the majority of the games where subject area was categorised as “not relevant” were entertainment games. Few generalisations could be made about the kind of game genre used in specific subject areas, although a variety of game genres was used in health and STEM subjects, while 6 of the 8 games for business and economics were simulations.

3.3.5 Specific or generic game outcomes

The majority of studies, 78%, examined the impact of playing a specific game with 22% looking at the impact of games generally. Chi-square analysis looking at the relationship between the purpose of the game with respect to specific or generic outcomes of playing tended to significance, ($\chi^2(1) = 2.023$, p=0.099), indicating that studies looking at the impact of playing a specific game were marginally more likely to concern games for learning, while studies looking at the impact of playing games generally were marginally more likely to involve entertainment games.

3.3.6 Platform/delivery
The video console was the most popular platform for delivering games, especially entertainment games, with 44% of studies reporting this method of game delivery. Games delivered via PC or laptop were also popular, with 38% of studies, mostly involving games for learning, using this platform. 18 papers described online games both for entertainment (9) and learning (9), suggesting that these have become more popular since the previous review in 2012, when only 9 were found. 2 games were mobile games.

3.4 Analysis of variables related to the study

This section reports the analysis of variables related to the study.

3.4.1 Number of papers with respect to geographical location

The coding of papers with respect to the country and continent where the first author of the paper was based, showed that research on games has an international dimension, with the most prolific regions being North America (62), especially the USA (53), and Europe (45) where Britain (11), Germany (6), Belgium (6), the Netherlands (4) and Spain (3) were most involved in games research. A sizeable amount of research was also carried out in Asia (26) especially in Taiwan (11) and China (7) with fewer studies reported in South America (5) and Australasia (5).

3.4.2 Study design

The current study adopted standards of rigour that were consistent with a quantitative approach. The randomised control trial (RCT) is the design of choice in evaluating educational interventions (Woolfson, 2012) and 18 RCTs were found in this review (see Table 2). The most popular study design, used both in studies of entertainment games and games for learning, was the quasi-experimental design (72). Although less rigorous than RCTs, this design can nevertheless provide valuable evidence about why games might work and which game features might support learning. A number of studies were like RCTs but, for pragmatic reasons, participants could not be randomly allocated to groups, since they were tested in the context of the whole class (McKenzie, 2013). Similarly it is not always possible or meaningful to carry out a pre-test to establish baseline similarity of the two groups as required in an RCT. For example many studies comparing visual-spatial performance of game
players and non-game players capitalise on the fact that the game players have been playing games for a long time and consequently the target skills were not originally pre-tested. Similarly, where the purpose of a game is to teach new skills, as in Knight et al.’s (2010) study of triage skills, it is not meaningful to pre-test performance since participants would have no prior knowledge of the target skills.

Table 2
Number of papers addressing the different game purposes by study design

<table>
<thead>
<tr>
<th>Purpose of game</th>
<th>Entertainment</th>
<th>Game for Learning</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlational</td>
<td>14</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>Qualitative study</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Quasi-experimental</td>
<td>39</td>
<td>33</td>
<td>72</td>
</tr>
<tr>
<td>RCT</td>
<td>7</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Survey</td>
<td>10</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>72</td>
<td>143</td>
</tr>
</tbody>
</table>

Correlational designs were typically used to examine the impact of variables related to the game, the players or the context of play on the outcome, while surveys were commonly carried out in the earlier stages of game evaluation to assess acceptance of games. These study designs were less likely to make it over the quality threshold, several correlational studies reported more complex statistical analysis, such as structural equation modelling (SEM). A handful of well executed qualitative studies (5) also made it over the quality threshold to provide higher quality evidence that was included in the review.

3.4.3 Study design and purpose of game

Table 2 shows that quasi-experimental designs were prevalent both in studies of games for learning and entertainment game; RCTs, surveys and qualitative designs were more likely to be used in studies of games for learning, while correlational and quasi-experimental designs were more popular in studies of entertainment games. Chi-square analysis confirmed that the relationship between the purpose of the game and study design was significant ($\chi^2(4) = 11.95, p=.018$).

3.4.4 General aim of the studies
Given the heterogeneous nature of the empirical research identified in this review, the studies were also categorised in terms of the general aim of each study. The most frequently occurring studies were those comparing a game group with a control group with respect to a specified outcome (47) and those looking at how characteristics of players impact on outcomes (47). 25 studies examined how specific features of a game impact on play, while 24 looked at aspect of the context of play. Table 3 also shows that studies of entertainment games and games for learning had different aims, and chi square analysis confirmed this significant relationship between the aim of the study and the purpose of the game ($\chi^2(3) = 18.164, p<0.01$). Studies that compared a game group with a control group, studies looking at the features of games and studies looking at the context of play were more likely to be about games for learning, while studies looking at player characteristics were carried out mainly with entertainment games.

Table 3
Number of papers addressing the different game purposes by the general aim of the study

<table>
<thead>
<tr>
<th>Aim of the study</th>
<th>Entertainment</th>
<th>Game for Learning</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study comparing a game group with a control group</td>
<td>21</td>
<td>26</td>
<td>47</td>
</tr>
<tr>
<td>Study looking at the characteristics of players</td>
<td>34</td>
<td>13</td>
<td>47</td>
</tr>
<tr>
<td>Study looking at the context of play</td>
<td>6</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>Study looking at the features of a game</td>
<td>10</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>72</td>
<td>143</td>
</tr>
</tbody>
</table>

3.5 Analysis of game outcomes

This section reports the analysis of variables related to the game outcomes.

3.5.1 Learning and behavioural outcomes of games

The learning and behavioural outcomes of games are an important focus of this review and table 4 shows that the most frequently occurring outcome was knowledge acquisition, followed by perceptual and cognitive, affective and behaviour change, with fewer papers reporting physiological, skills and soft
and social skills outcomes. However entertainment games and games for learning address different kinds of outcomes with knowledge and skill acquisition almost exclusively studied in games for learning, with affective, behaviour change and physiological outcomes more likely to be studied with entertainment games.

Table 4

Number of papers addressing different purposes by the learning and behavioural outcomes* of the game

<table>
<thead>
<tr>
<th>Outcomes of playing game</th>
<th>Entertainment game</th>
<th>Game for learning</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective</td>
<td>15</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Behaviour change</td>
<td>15</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>Perceptual and Cognitive</td>
<td>12</td>
<td>9</td>
<td>21</td>
</tr>
<tr>
<td>Knowledge acquisition</td>
<td>5</td>
<td>42</td>
<td>47</td>
</tr>
<tr>
<td>Physiological</td>
<td>16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Skill</td>
<td>2</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Soft and social skills</td>
<td>6</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>71</strong></td>
<td><strong>72</strong></td>
<td><strong>143</strong></td>
</tr>
</tbody>
</table>

*The majority of papers reported one main outcome and these are reported in the table.

3.5.2 Learning and behavioural outcomes by study design

Table 5 shows that different study designs were used to address different outcomes. RCTs were used most frequently in studies of knowledge acquisition and perceptual and cognitive outcomes; quasi-experimental designs were used to study all the different learning outcomes, but especially knowledge acquisition, physiological and perceptual and cognitive outcomes; correlational designs were used most frequently with affective, social and soft skills and behaviour change, while survey designs were used most with knowledge acquisition, affective and behaviour change outcomes.

Table 5

Number of papers using each study design by learning and behavioural outcomes
3.6 Discussion of specific learning and behavioural outcomes

In this section of the review, a selection of papers that illustrates the varied learning and behavioural outcomes of playing games is discussed in more detail, considering especially the purpose of the game, subject discipline and study design.

3.6.1 Knowledge acquisition outcomes

RCTs evaluating the effectiveness of serious games for knowledge acquisition were reported across varied subject disciplines and tended to report that playing the game led to better performance than the control condition. Arnab et al. (2013) showed that their serious game about adolescent sexual relationships was more effective than the control condition in helping adolescents understand about sexual coercion. Nishikawa and Jaeger (2011) found that simple computer-based simulations and games in political science were as effective as traditional classroom lectures in the short run and produced better concept retention in the long run. Peng (2009) showed that students using the RightWay Café game increased their knowledge of nutrition, their self-efficacy for healthy eating, the perceptions of the benefits of healthy eating, and showed a greater intention to commit to healthy eating than students in a control group. Suh, Kim, and Kim (2010) found that elementary school students who studied English as a second language with online role playing games, gained higher scores in listening, writing and reading.
than those who studied English with face-to-face classroom instruction. Less encouraging results were reported by Rondon, Sassi, and de Andrade (2013) who found that game-based learning was comparable to traditional learning in producing short-term gains in teaching the anatomy and physiology of the head and neck, but the traditional lecture was more effective in improving students’ short and long-term knowledge retention. McKenzie (2013) found that both game-based and face-to-face teaching resulted in significantly higher knowledge about and confidence in dealing with aggressive situations but, contrary to expectations, face-to-face teaching led to significantly higher knowledge scores compared with the game-based online learning activity.

Surveys of knowledge acquisition outcomes for games were reported by Proctor and Marks (2013), who found that secondary school teachers were less positive in their perceptions of computer-based games and technology for classroom instruction than primary school teachers and Appel (2012), who found that those who are more involved in the social use of PC and Mac computers for gaming and using social media also had better practical and theoretical computer knowledge. Qualitative studies also provided valuable contributions to our understanding of the outcomes of playing games. Turkay and Adinoff (2012) found that gamers identified four main areas where they thought they learned from playing games: game mechanics, game narratives, each other and by becoming interested and studying outside resources. Poplin (2012) identified both advantages and disadvantages of using their online serious game, NextCampus, to support playful public participation in urban planning. Benefits were the potential for joy and playfulness in the game, while the drawbacks concerned the high cost and complexity of developing the game as well as unclear borders between serious games and entertainment games.

Only 5 of the studies looking at knowledge acquisition in games involved entertainment games. deHaan, Reed, and Kuwada (2010) looked at the use of a music video game for teaching English, Minnery and Searle (2014) described the use of SimCity4 in teaching town planning and Charsky and Ressler (2011) looked at the use of Civ III in teaching history. While all were relatively successful, the small number of entertainment games used for knowledge acquisition suggests interest is moving away
from using Commercial Off the Shelf (COTS) games for learning, in favour of designing educational games or other technological solutions that can target the desired learning objectives more precisely.

3.6.2 Skill acquisition outcomes

Studies of skills tended to use higher quality RCT and quasi-experimental designs. Knight et al.’s (2010) game led to improvements in performance on the triage skills, “tagging accuracy” and “step accuracy”, compared with a control group which used a card game. Creutzfeldt, Hedman, and Fellnder-Tsai (2013) explored medical students’ retention of knowledge and skills as well as their proficiency gain after pre-training using a Multiplayer Virtual World with avatars for cardio-pulmonary resuscitation team training, while Delasobera et al. (2010) found that simulator and multimedia educational tools were more effective than reading about Advanced Cardiac Life Support (ACLS) in India.

3.6.3 Perceptual and cognitive outcomes

21 studies in the current review provided additional evidence to support previous research documenting the attentional and visual perceptual benefits of playing action video games (Boot, Kramer, Simons, Fabiani, & Gratton, 2008). These studies also tended to adopt rigorous experimental designs, with 5 RCTs and 12 quasi-experimental studies. Boot et al. (2010) confirmed the advantages of variable priority training, which trains people to flexibly distribute their attention over multiple aspects of a task, compared with full emphasis training, where the emphasis is on the whole task. Deveau, Lovcik, and Seitz (2014) showed that playing a specially designed video game improved visual perceptual abilities across a range of visual characteristics, including central and peripheral acuity and contrast sensitivity, compared with control subjects who showed no change in vision.

A range of effects was reported using a post-test only design. Karle, Watter, and Shedden (2010) found that video game players had an advantage in task switching performance, compared with non-video game players, probably due to their superior ability to control selective attention. Romano Bergstrom, Howard, and Howard (2012) found that both game players and musicians showed significantly greater implicit learning of sequential context than a control group on both speed and accuracy measures. Dye and Bavelier (2011) found that participants who played action games showed enhanced performance on tasks involving the ability to deploy attention over space, time and objects, while Chiappe, Conger, Liao,
Caldwell, and Vu (2013) demonstrated that playing action videogames increases players’ ability to take on additional tasks by increasing attentional capacity.

It has been more difficult to establish firm evidence of the value of playing games in supporting higher level cognitive processing although Yang’s (2012) RCT showed that a game for building virtual cities, New York City tycoon, promoted students’ problem solving skills, in the areas of finding causes, finding solutions and avoiding problems, and increased motivation compared with a control group which showed no improvement. In an RCT, Bozoki, Radovanovic, Winn, Heeter, and Anthony (2012) found a relatively small effect size on a six-week, moderate-intensity computer game-based cognitive intervention with high-functioning seniors, with no transfer effects. Coller and Scott (2009) described how students who created games learned more about numerical concepts about the movement of cars rather than those who played the games.

Two good quality studies provided interesting qualitative analyses of the cognitive processes required in playing games. Martinovic et al. (2014) developed a matrix, where simple games were categorised in terms of the primary and secondary cognitive information processing features that they support, such as visual perception, memory and executive functions. Baniqued et al. (2013) also carried out a rigorous qualitative task analysis of the cognitive abilities which underlie success on a range of casual video games but they also validated their classification, confirming that games categorized as tapping working memory and reasoning were strongly related to performance on an established battery of working memory and fluid intelligence tasks.

3.6.4 Physiological outcomes

All 16 studies reporting physiological outcomes involved entertainment games, with 14 of these adopting high quality RCT or quasi-experimental designs. These studies typically examined objective measures of player behaviour or user experience in game playing, or changes in physical fitness linked to exer-gaming. Granek, Gorbet, and Sergio (2010) found that the basic cortical network for processing complex visually guided reaching is altered by extensive video-game play. Bailey and West (2013) found contrasting effects of playing an action compared with a non-action video game on both neural activity related to target processing and the perception of emotion in facial expression.
Mathiak et al. (2011) were interested in neural correlates of reward in playing games and they studied players’ affective reactions and their neural responses, when a player failed in a fight during a first-person shooter game, while expecting to win. They found that participants who reported greater negative affect after playing the game had less right temporal pole (rTP) activity associated with failure, suggesting that the rTP may be involved in evaluating failure events in a social context.

In an RCT, Vernadakis, Gioftsidou, Antoniou, Ioannidis, and Giannousi (2010) compared balance and stability in a group using the Nintendo Wii with those on a traditional balance training program. They found an improvement in balance for both groups but no differences between the groups. Several studies reported the benefits of exer-games on physical fitness in special populations, including older, inactive adults (Maillot, Perrot, & Hartley, 2012; Pichierri et al., 2012), stroke patients (Hurkmans et al.) and those with depression (Rosenberg et al., 2010).

3.6.5 Affective outcomes

Of the 22 papers addressing affective outcomes, a range of study designs was used with twice as many looking at entertainment games (15) as games for learning (7). Several studies reported experimental accounts of players’ subjective experiences while playing games. Keller, Ringelhan, and Blomann (2011) found support for the balance hypothesis, traditionally viewed as the central tenet of flow theory, which proposes that the balance between player skills and the challenge provided by an activity leads to feelings of flow that make the activity intrinsically rewarding. Smets, Abbing, Neerinckx, Lindenberg, and van Oostendorp (2010) found that participants in a game-based group showed higher arousal levels, a more intense feeling of presence, better situation awareness and faster performance when needed than a group which used a “classical” storyboard-based evaluation.

Correlational designs provided useful evidence about how player variables influence aspects of engagement. Relatively few studies in the current review examined motives for playing but, using the ARCS model of motivational processing, Huang, Huang, and Tschopp (2010) found that the motivational processing components, attention, relevance, and confidence, were significant predictors of satisfaction with the game. Using Keller’s (2008) Integrative Theory of Motivation, Huang, Johnson and Han (2013) found that game appeal, game involvement and game structure predicted motivational processing.
cognitive processing and satisfaction with a game. Several studies adopted the Technology Acceptance Model (TAM) which identifies perceived ease of use and usefulness as key factors in explaining users’ intentions to use specific technologies (Bourgonjon, Valcke, Soetaert, & Schellens, 2010; Bourgonjon et al., 2013; Park, Baek, Ohm, & Joon Chang (2014).

Shaker, Asteriadis, Yannakakis, and Karpouzis (2012) used features extracted from player gameplay and visual behaviours to model user experience in games. Participants were asked to play two Super Mario games and then report their preferred game for the 3 variables: engagement, challenge and frustration. This preference data was linked back to players’ behavioural data to allow the researchers to build a corpus of knowledge to maximise engagement, enjoyment and learning experience.

Time spent playing and game playing habits provide behavioural indicators of engagement in games. Coller and Scott (2009) found that students who learned by creating games rather than playing them spent twice as much time on their coursework outside of class on a game-based numerical methods course and they used this as an index of engagement on the course.

A number of studies aimed to clarify how subjective experience in games is measured by developing or refining existing questionnaires. For example Procci, Singer, Levy, and Bowers (2012) used confirmatory and exploratory factor analysis to investigate the psychometric properties of the Dispositional Flow scale, which was originally developed by Jackson and Eklund (2004). Procci et al. found that this measure was not a good fit for measuring the flow experience and urged that there is a need for further research to pin down the idea of flow in games more precisely and to develop better measures of the construct. In an exploratory factor analysis of the User Engagement Scale (UES), Wiebe, Lamb, Hardy and Sharek (2014) found that a four factor solution (focused attention, perceived usability, aesthetics, and satisfaction) provided a better fit than the six factors identified in the original UES (O’Brien & Toms, 2008). Their revised scale (UESz) was more predictive of game performance than the Flow State Scale which measured the nine dimensions of flow originally identified by Csikszentmihalyi (1990). Christou (2011) examined links between appeal, measured by Hassenzahl and Monk’s (2010) AttrakDiff questionnaire, and immersion, measured by Jennett et al.’s (2008) immersion questionnaire, in two games with experienced and inexperienced players. While appeal and immersion were correlated,
they were both considered to be useful constituents of player experience. Immersion was unaffected by level of player experience, while appeal was affected both by the game played and level of player experience.

3.6.6 Behaviour change outcomes

Behaviour change outcomes were studied most frequently in entertainment games using quasi-experimental, surveys and correlational designs. Just as Anderson & Bushman (2001) and Anderson (2004) found that playing violent entertainment games leads to increased aggression, 3 papers in the current study reported that playing video games with prosocial content led to increases in prosocial behaviours (Greitemeyer, 2010; Greitemeyer & Osswald, 2010; Gentile et al., 2009).

A couple of serious games for behaviour change were found including Verduin (2013) who compared relapse prevention skills in outpatients on a substance abuse treatment program for those who used the computer simulation game, Guardian Angel, and those who viewed educational slides about treatment for alcohol use disorders. Verduin found lower ratings on the Obsessive Compulsive Drinking Scale for those using the game, suggesting they had fewer obsessive thoughts about alcohol, which is generally linked to more positive outcomes with respect to relapse. Schönbrodt and Asendorpf (2011) described an innovative game looking at the effects of players’ behaviour towards a virtual spouse on their real world relationship satisfaction and interpersonal motives. They found links between behaviours in the virtual world and intimacy motives and satisfaction with the real-world relationship, suggesting that some sort of transference between the real world and the virtual world took place.

3.6.7 Social and soft skills outcomes

As with affective and behaviour change outcomes, the majority of papers addressing social outcomes involved entertainment games and used mainly correlational or survey designs. Herodotou, Kambouri and Winters (2014) tested the claim that game players play primarily to mitigate their dissatisfaction with real life. They found that players’ basic psychological needs for autonomy, competence, and relatedness, described in self-determination theory, were not good predictors of their motivation for digital game-play, suggesting that gamers don’t play games as a means of satisfying their basic needs and improving their well-being. In addition, contrary to media stereotype, gamers were found
to be highly social individuals. Kowert and Oldmeadow (2013) found positive links between amount of video game involvement and emotional expressivity and emotional control, suggesting that links between online video game involvement and social skills are not as negative as has commonly been supposed.

Social outcomes were also reported for games for learning. González-González, Toledo-Delgado, Collazos-Ordoñez, and González-Sánchez’s (2014) social videogame helped hospitalised children communicate and express emotions. Bachen, Hernández-Ramos, and Raphael (2012) confirmed that a simulation game (REAL LIVES) was an effective way to promote empathy and interest in other cultures, while Haferkamp, Kraemer, Linehan, and Schembri (2011) found that their virtual environment (DREAD-ED) provided an innovative methodology for training soft skills in the area of disaster communication.

4. Discussion

Using the same search terms on a smaller number of databases, the current update to the systematic review of Authors found many more papers reporting empirical evidence of the positive outcomes of playing games (512) than the previous review (129). This illustrates the increased interest in the positive impacts of digital games during the five year period from 2009 to 2014 compared to the previous five year period. Given the very large number of relevant papers identified, the current review focused on summarising the 143 papers that were rated as higher quality using our quality criteria (see 2.6). The papers were of similar quality with 18 RCTs (12.5%) and 72 quasi-experimental designs (50.3%) of the 143 selected papers, compared with 9 (12.8%) and 34 (48.6%) out of 70 in the previous study.

Once again the multidimensional analysis of games and their outcomes provided an informative framework for organising the diverse research on games and examining links between the categorising variables, helping to summarise trends in the research. The design adopted in the studies was related to the learning outcomes and subject discipline, with higher quality designs being used in studies of knowledge acquisition, perceptual and cognitive and physiological outcomes and games for health and science. These more rigorous designs are the prevalent research approach favoured in these “harder” disciplines, where there is a long standing recognition of the need for evidence from well-controlled studies, and where outcomes are relatively straightforward to measure. In contrast, social and
behavioural outcomes in “softer” disciplines such as business and social science tend to be more complex and more difficult to operationalise and measure, and consequently studies of games looking at social and behavioural outcomes and in business and social science were more likely to report correlational, survey and qualitative designs.

Extending the multidimensional classification to consider the geographical location of the first author confirmed that research on games is being carried out across the globe, especially in North America and Europe. Coding the papers in terms of the general aim of each paper confirmed that, researchers recognise the importance of studying player characteristics, game features and the context of play as variables that influence game outcomes. This analysis is consistent with models and theories of game design, such as de Freitas’ Four-Dimensional Framework (de Freitas, Rebolledo-Mendez, Liarokapis, Magoulas, & Poulovassilis, 2010) and activity theory (Peachey, 2010), that emphasise the need to consider the coordination of a range of components in designing and evaluating games.

More high quality papers using the term “serious games” were identified in the current review (34) than in the previous review (8) suggesting that the term has entered the mainstream in recent years. However the term seems to be used interchangeably with games-based learning and these terms were collapsed for the purposes of this review. As in the previous review, similar numbers of papers about entertainment games and games for learning were found, but these studies differed in their aims, the platform used for delivery and the learning and behavioural outcomes they addressed.

Since games are frequently championed as a novel, engaging and active new method for supporting 21st century skills and behaviour change (Dondlinger, 2007), it was disappointing that games for learning are still used most frequently to support knowledge acquisition. This seems to be a rather pedestrian use of games compared with the speculation about their potential. There was however some encouraging evidence that games were successfully used for skill acquisition in health and town planning and for behaviour change in substance abuse and satisfaction with marital relationships, as well as supporting collaborative interactions, soft skills and empathy. While games for learning were found that successfully addressed topics as varied as healthy eating, triage training, operations management and advanced life support retraining, they were used most widely in health and the STEM subjects, with
fewer games in business and language. These topics reflect the needs and curricular interests of individuals in late secondary school and tertiary education, the age range of this review.

Several of the studies demonstrating that playing games can change behaviour in a positive way, involved entertainment games. This unintentional learning by playing games is also seen in the studies demonstrating the visual perceptual benefits associated with playing entertainment games. It seems that this “unintentional” learning found in entertainment games could provide insights into engagement and learning in serious games.

Simulations were by far the most popular game genre both in the current review and original review. Simulations support learning by providing virtual activities and procedures that reflect or replicate those required in the real world, frequently using visually compelling environments. While a variety of other game genre were used, including role playing games, adventure games, strategy games, problem solving and puzzle games, their relative scarcity compared to simulations does point to a possible concern for games for learning. It seems to be much easier to design simulations that model aspects of learning, than to identify and utilise other game features in a way that aligns with and supports learning outcomes. This raises the deceptively tricky issue of the defining characteristics of a game.

Some progress is being made in understanding how specific game features engage players and support learning. Experimental studies examined features such as competition, uncertainty of information and varying training schedules. There was mixed support for competition in making performance more effective. DeLeeuw and Mayer found that adding competitive features to a computer-based circuit game improved memory for what was presented in the game, but did not improve deep learning as measured by transfer. Staiano, Abraham, and Calvert (2012) found that youths in a competitive exer-game significantly improved in executive function skills and lost more weight than those in either a cooperative exer-game or a no-play control group. However Peng and Hsieh (2012) found that games with a cooperative goal structure led to higher motivation than those with a competitive goal structure. Ozcelik, Cagiltay, and Ozcelik (2012) confirmed that introducing an element of uncertainty enhanced learning, while Boot et al. (2010) confirmed the advantages of variable priority training compared with full emphasis training.
Progress has been made too in understanding features that make games more engaging. More precise definitions of subjective experience constructs, such as flow, engagement and appeal have been proposed, and measures which allow greater differentiation between these constructs have been developed. Huang, Johnson, and Han (2013) found that game features such as animation, graphics, fun, rules and goals were important in engaging students and supporting learning in an online instructional game, while King, Delfabbro, and Griffiths (2010) found that rewards such as earning points, finding rare game items and fast loading times were rated as the most enjoyable and important aspects of video game playing. However research on game features is currently quite piecemeal and a more comprehensive and systematic program of experimental studies mapping game features to engagement and learning should be carried out. Arnab et al. (2014) have recently made some proposals about how this might be achieved.

4.1 Limitations

As with all literature reviews, the current review does not claim to be comprehensive, but summarises the research on games based on the search terms used, the databases included and the time period of the review. The level of specificity and distinctiveness of different learning outcome categories is an ongoing discussion in education generally (Cedefop, 2009) and also in analysing serious games (Arnab et al, 2014). The categorisation of learning outcomes used in this review generally worked well but there were still some tricky boundary decisions and ambiguities. For example, categorising behaviour change, skills and social skills was tricky. Is a social skill a skill or a social outcome? Problem solving was categorised as a “skill” or a “perceptual and cognitive outcome” depending on whether a cognitive or educational framework was being adopted in that study. Even the decision about whether an outcome is positive or negative is not straightforward as engagement can turn into addiction and cheating can be beneficial!

Due to the quantitative focus in the current study, only 5 qualitative studies were included and it is possible that some high quality qualitative studies were excluded from consideration. Such studies could provide useful and detailed information about what works in games and what does not. While the current reviews provide a useful overview of games research, the escalating volume of this research suggests that it will be more useful for future literature reviews to focus on more specific issues. The review uncovered a number of more focused reviews that have already been published about the use of games for physical
activity (Peng, Crouse & Lin, 2013), health (Primack, Carroll, McNamara, Klem, King, Rich et al., 2012),
the cognitive and motivational effects of playing games (Wouters, van Nimwegen, van Oostendorp & van
der Spek, 2013) and engagement in games (Boyle, Connolly, Hainey & Boyle, 2012).

4.2 Conclusions

Taken together the original review and the current update illustrate the increased interest in the
positive impacts and outcomes of games, and provide a valuable summary organising the diverse research
that has been carried out on the positive outcomes of playing games since 2004. The term “serious
games” has become mainstream during this time, but is used interchangeably with games for learning.
Games for learning have been used to promote knowledge acquisition across a wide range of topics and
to a lesser degree skill and social skill acquisition and behaviour change. There has been a move away
from using COTS games for learning due to difficulties in integrating them into the curriculum and an
acceptance that it can be more useful to develop games that address specific curricular objectives. Despite
the intense interest in games, it is important to realise that developing games for learning can be very
complex and costly and still provides significant challenges. Future research will benefit from detailed
experimental studies that systematically explore which game features are most effective in promoting
engagement and supporting learning. Analysis of engagement and informal learning in entertainment
games can also provide valuable insights into game mechanisms that can then be applied to games for
learning.

5. Acknowledgements

6. References

Admiraal, W., Huizenga, J., Akkerman, S., & ten Dam, G. (2011). The concept of flow in

Adolescence, 27*, 113–122.

behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: a meta-


7. Coded papers


4. Acknowledgements

This work was co-funded by the EU under the FP7, in the Games and Learning Alliance (GaLA) Network of Excellence, Grant Agreement nr. 258169.
Highlights

- The current systematic literature review updates Author (date)
- The review looks at impacts and outcomes of playing digital games from 2009 to 2014
- Multi-component coding of papers, games and learning outcomes was used
- Many papers were found with 143 papers providing high quality evidence
- Games for entertainment and learning addressed different outcomes