Computers & Education 113 (2017) 86–97

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Video games can develop graduate skills in higher education students: A randomised trial

Matthew Barr

School of Humanities, University of Glasgow, Glasgow G12 8QQ, Scotland, UK

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 14 May 2017
Accepted 25 May 2017
Available online 26 May 2017

Keywords:
Adult learning
Cooperative/collaborative learning
Interactive learning environments
Media in education
Post-secondary education

ABSTRACT

This study measured the effects of playing commercial video games on the development of the desirable skills and competences sometimes referred to as 'graduate attributes'. Undergraduate students in the Arts and Humanities were randomly assigned to either an intervention or a control group. Previously validated, self-report instruments to measure adaptability, resourcefulness and communication skill were administered to both groups. The intervention group played specified video games under controlled conditions over an eight week period. A large effect size was observed with mean score change 1.1, 1.15, and 0.9 standard deviations more positive in the intervention group than the control on the communication, adaptability, and resourcefulness scales respectively ($p = 0.004$, $p = 0.002$, and $p = 0.013$ for differences in groups by unpaired t-test). The large effect size and statistical significance of these results support the hypothesis that playing video games can improve self-reported graduate skills. The findings suggest that such game-based learning interventions have a role to play in higher education.
Crown Copyright © 2017 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Graduate attributes are those generic skills such as problem solving, communication, resourcefulness or adaptability which are considered desirable in graduates, particularly where employability is concerned (Barrie, 2006, 2007; Hughes & Barrie, 2010). However, it may be argued that most higher education courses are not explicitly designed to teach or develop graduate attributes. Many commercial video games, on the other hand, require players to exercise a range of such skills and competences in order to progress (Barr, 2013). Advocates of the learning experience afforded by modern video games include Thomas and Brown (2011), who describe a new "disposition" towards learning that games exemplify, requiring players to be both adaptable ("thrive on change") and resourceful ("marshal all their available resources and experiment with them to find multiple ways of accomplishing a task"). Gee (2007) identifies a range of learning principles present in video game design that encourage, for example, critical thinking and reflective learning, which are also commonly cited as desirable attributes in graduates. However, while there are indications that playing video games may be associated with positive cognitive and social effects (Granic, Lobel, & Engels, 2014), there is little empirical evidence for the efficacy of using commercial video games to develop graduate skills. This work, therefore, was intended to measure the effects of playing commercial video games on the attainment of certain graduate attributes, testing the hypothesis that playing selected games can improve student scores on measures of graduate skills. This general hypothesis may be subdivided into hypotheses relating to each of the areas under investigation: communication, resourcefulness and adaptability. These hypotheses are evaluated by

E-mail address: Matthew.Barr@glasgow.ac.uk.

http://dx.doi.org/10.1016/j.compedu.2017.05.016
0360-1315/Crown Copyright © 2017 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

comparing control and intervention group pre- and post-test scores on self-report measures. If these hypotheses are supported, then new opportunities for institutional skills development provision may be revealed, and there are important implications for how video games are perceived by the wider public.

The GAGA Project ("Using games technology to develop graduate attributes") saw the development of a game to help prepare international students for study at the University of Abertay, by introducing that institution's graduate attributes (Lloyd, 2011). However, no empirical evidence of the game's efficacy has been published. Furthermore, this work involved the development of a game for the specific purpose of introducing certain graduate attributes: the study reported here is concerned with the use of existing commercial video games. Adachi and Willoughby (2013), however, demonstrated by means of a four-year longitudinal study that playing strategy and role-playing games predicted self-reported problem solving skills among a sample of 1492 high school-aged participants. Adachi & Willoughby also noted that the empirical evidence for the relationship between playing video games and the development of problem solving skills was limited. Subsequently, Shute, Ventura, and Ke (2015) have shown statistically significant gains in problem solving, spatial skills and persistence in a group of participants asked to play *Portal 2* – one of the titles used in this study – for 8 h, compared with a group asked to play a suite of 2D puzzle games purported to improve such skills.

The Scottish institution at which the study was carried out identifies ten graduate attributes: Investigative, Effective Communicators, Independent and Critical Thinkers, Adaptable, Resourceful and Responsible, Confident, Experienced Collaborators, Subject Specialists, Reflective Learners, Ethically and Socially Aware. As noted by Nicol (2010), the development of graduate attributes in Scotland has drawn heavily on work carried out by Barrie (2006; 2007) in Sydney and Melbourne and, as a result, the attributes extolled by this institution are broadly comparable to those identified by other universities, internationally. Furthermore, it is evident that skills relating to communication, adaptability, and resourcefulness are sought by employers the world over, regardless of the umbrella terms used to describe them.

Since the host institution already purports to develop this list of skills and competencies in its students, this study was designed to determine whether video games offer any advantages for attribute development over-and-above existing university provision. A pilot project indicated that of the attributes listed, effective communication, adaptability and resourcefulness were the most promising candidates for further study, and identified suitable instruments for their measurement.

2. Games

The games used in the pilot project were selected with assistance from colleagues in industry and academia, who were asked to suggest titles that might relate to the list of graduate attributes provided. These suggestions were necessarily mediated by the financial and technical limitations of the study, which dictated that games must run on very modestly-specified Windows PCs and be relatively inexpensive. An additional consideration was the quality of the games. A poor quality game is of little utility here: well-received titles are more likely to be representative of those which players would choose to play on their own time, and a particularly poor game is likely to impact negatively on the participants' willingness to engage in the study. While game quality is somewhat subjective, review scores aggregated by sites such as Metacritic (Metacritic - Movie Reviews, TV Reviews, Game Reviews, and Music Reviews, 2016) are used by industry and consumers alike to determine a game's excellence (Graft, 2011). Metacritic scores are recognized as imperfect (Dring, 2010) but they undoubtedly provide an easily quantifiable means of determining the relative merits of a game. For the purposes of this study, no game with a Metacritic score of less than 80 out of 100 was considered, with scores ranging from 82 (*Lara Croft and the Guardian of Light*) to 95 (*Portal 2*). All of the games, with the exception of *Gone Home* and *Papers, Please*, include a substantial multiplayer component.

The games used in the study were all commercial titles, designed for entertainment purposes rather than with the intention of developing particular skills in players. The brief descriptions below describe the games in general terms while highlighting some of the features most relevant to this study.

Borderlands 2 (Borderlands 2-Gearbox Software, 2016) is a co-operative role-playing first-person shooter game developed by Gearbox Software. Players work together to obtain loot and weaponry while defeating a range of foes against a colourful, if violent, cartoonish backdrop and attendant story. A variety of play styles are supported through the choice of character classes presented to the player, ranging from a tank-like "Gunzerker" to a stealthier assassin. The game allows for local area network (LAN) multiplayer, meaning the co-operative elements functioned within the university infrastructure and did not require an internet connection. *Borderlands 2* also permits players to drop in and drop out as required – a participant who arrived after others had already embarked on a mission could straightforwardly join the team without being forced to wait for the beginning of the next mission, or requiring the others to start again from the beginning.

Minecraft (Minecraft, 2016) is a procedurally-generated sandbox game with construction, exploration and survival elements. In single player mode, players are free to explore the world and collect ('mine') resources such as stone, wood and metal to create ('craft') a range of buildings, tools and weapons. Multiplayer mode is similarly non-prescriptive in terms of what it permits (or requires) players to do: the main difference is that the world is shared, so players may choose to work together.

Valve's *Portal 2* (Official Portal 2 Website, 2016) is described by the developer as "a hilariously mind-bending adventure that challenges you to use wits over weaponry in a funhouse of diabolical science". The game features a robust co-operative mode, and the co-operative portion of the game allows for split-screen play, meaning two people can play together on the same machine. Participants were asked to play *Portal 2* in pairs and provided with printed instructions on how to host or

connect to a co-operative game. The configuration of the game allowed players who had already completed a number of the co-operative levels to advance to subsequent stages and thus minimise repetitious play.

Lara Croft and the Guardian of Light (Lara Croft and the Guardian of Light, 2016) was included in the study as a result of the game's emphasis on co-operation to solve puzzles and progress. A fixed isometric view of the action is presented and the game is intended to be played with a collaborator. One player assumes the role of the gun-toting Lara while the other plays as Totec, a Mayan warrior who comes equipped with a spear that may be used to create impromptu ladders and bridges. Co-operative players share the same screen (although online co-op is an option in most versions of the game) and for this study both players were provided with a games controller similar in design to that used with the Xbox 360 games console.

The rationale for the inclusion of *Warcraft III: Reign of Chaos* (Blizzard Entertainment, 2016) was based on its strategic multiplayer mode, which may be played over a local network without an internet connection. The game is played on a three dimensional map with up to four races (Orcs, Humans, Night Elves and Undead) vying for domination. Each player controls one of these races and must collect resources (gold and lumber) to develop and construct buildings, units and weaponry with the ultimate aim of obliterating their opponents from the map. The multiplayer mode of the game supports team play, meaning that participants in the study could work together (even as different races) to defeat a computer-controlled adversary. Unlike *Lara Croft*, many different multiplayer configurations are supported, from the previously described two-versus-one scenario through to any combination of human and computer teams. The computer-controlled adversary may also be handicapped somewhat to accommodate inexperienced human players.

Valve's *Team Fortress 2* (Team Fortress 2, 2016) is a multiplayer-only game that may be hosted on a local server, again avoiding the need for an internet connection to facilitate matchmaking. The game sees the competing teams thrown into conflict on a time-limited or objective-based map. When a team meets the victory conditions – or time runs out – the next map is loaded and a new objective pursued. Each map operates in a pre-determined game mode, such as Capture the Flag or King of the Hill, with the objective of each mode explained by means of a short video shown at the beginning of play. In Capture the Flag mode, for example, both teams are tasked with stealing a briefcase of intelligence from the opposing team's base and transporting it back to their own, with the briefcase standing in for the more traditional flag. The winning team is that which captures the enemy intelligence three times. Players must therefore decide how much emphasis to place on defence of their own intelligence versus making an offensive move to capture the enemy's briefcase, with different team members assuming different roles as agreed.

The Fullbright Company's *Gone Home* ('About Gone Home', 2012) might be described as a first-person interactive story or adventure (the designers term it a "story exploration video game") wherein the player, assuming the role of a young woman returning to her family home after a year-long absence, explores the apparently abandoned house. In doing so, the player may uncover a number of storylines, the most significant of which relates to the protagonist's younger sister. There are no explicit goals and interaction is relatively limited, with plot developments uncovered by reading discarded letters and examining ephemera such as concert ticket stubs and television viewing guides.

Papers, Please (Papers, Please, 2016) is a puzzle/simulation game in which the player is cast as an immigration officer, deciding whom to admit and whom to turn away from the border of the fictional former communist state of Arstotzka. The player performs this role by critically assessing the documentation presented by each potential immigrant in light of the ever-changing rules and regulations imposed by the state. As well as exercising critical judgement and adapting to change, the player is presented with an opportunity to reflect on the ethical and social consequences of their in-game actions, not only in terms of the lives of the fictional immigrants and existing citizens of Arstotzka (terrorist attacks are a distinct possibility, should the 'wrong' person be permitted access to the country) but also in terms of the personal price to be paid by the family of the player's character if quotas are missed.

3. Methods

Duran's Communicative Adaptability Scale (CAS) (1983; 1992) is a self-report measure of communication ability, framed in terms of communicative adaptability, which Duran defines as "the ability to perceive socio-interpersonal relationships and adapt one's interaction goals and behaviours accordingly." The scale has been used in a number of studies and produced Cronbach's alpha coefficients between 0.74 and 0.84, indicating good reliability (Duran, 1992). While CAS is here used as a measure of communication skill, a salient feature of this scale is its focus on adaptability. 'Adaptable' is another of the university's identified graduate attributes, and Duran (1992) states that "the most basic form of communication competence is fundamental competence", which may be defined as "an individual's ability to adapt effectively to the surrounding environment over time" (Spitzberg & Cupach, 1984).

The adaptability implicit in Duran's Communicative Adaptability Scale depends upon both "cognitive (ability to perceive) and behavioural (ability to adapt) skills". The scale comprises 30 statements (e.g. "I feel nervous in social situations") that relate to the six dimensions (Social Composure, Social Confirmation, Social Experience, Appropriate Disclosure, Articulation and Wit) of the "social communication repertoire". Respondents are asked to indicate the degree to which each statement applies to them, on a scale from 1 ("never true of me") to 5 ("always true of me") and the responses summed for each dimension. Certain responses (e.g. "I sometimes use words incorrectly") are reversed before summing, e.g. a 2 becomes a 4. The scale was developed and validated by means of two studies: one involving adult teachers, and the other involving university students. It is therefore thought to be highly applicable for use in this study, which involves adult university students.

Building on the work of Pulakos, Arad, Donovan, and Plamondon (2000; 2002), Ployhart & Bliese (2006) suggest a self-report measure of adaptability based on their own Individual ADAPTability (I-ADAPT) theory, in which they define adaptability as an "individual difference construct that influences how a person interprets and responds to different situations". The I-ADAPT-M measure was developed with the practicalities of its use in mind, while addressing all eight of the dimensions of adaptability identified by Pulakos et al. It is, therefore, relatively short (taking around 10 min to complete) and its self-report nature is intended to "simplify administration and scoring, and to enhance applicability to multiple contexts". The measure has been tested and refined in collaboration with subject matter experts and validated by means of a confirmatory factor analysis and a construct validity study of the original 40-item measure "found strong support for convergent and discriminant validity" (Ployhart & Bliese, 2006). Intended for use with adults across multiple contexts, the I-ADAPT-M measure is thought to be highly applicable here.

Ployhart & Bliese present their I-ADAPT theory as a nomological network of knowledge, skill, ability and other characteristics (KSAOs), performances and situations. The I-ADAPT-M measure is a 55-item survey that asks respondents to indicate how well each of the items describes their preferences, styles and habits at work, with each item aligned with a particular dimension of adaptability. For example, the item "I work well with diverse others" relates to cultural aspects of adaptability. The instrument is freely available for research purposes.

While no obvious measures exist for the responsibility aspect of the 'Resourceful and Responsible' attribute, Zauszniewski, Lai, and Tithiphontumrong (2006) offer a 28-item Resourcefulness Scale, developed and validated in a two-phase study with chronically ill elderly patients. The authors found that the scale had acceptable internal consistency (Cronbach's alpha = 0.85) and state that it may be used with younger and middle-aged adults as well as the elderly, making it applicable for use with the young adults sampled here.

The 28-item Resourcefulness Scale measures resourcefulness along two dimensions: personal ("the ability to independently perform daily tasks") and social ("to seek help from others when unable to function independently"). 16 of the 28 items relate to personal resourcefulness (e.g. "When faced with a difficult problem, I try to approach its solution in a systematic way"), with the remaining 12 designed to measure social resourcefulness (e.g. "When I am feeling sad, it helps to talk to other people"). Items are presented on a 6-point scale ranging from 0 ("not at all like me") to 5 ("very much like me") and respondents asked to indicate how descriptive each of the items is of them.

Student participants were recruited to the study by means of an email invitation with a link to an online form. The email was targeted at year one and year two students in the College of Arts and explained that the study may involve playing video games and completing surveys. The email indicated that participants who completed all assigned tasks would be entered into a prize draw for an Amazon voucher at the end of the semester. Potential participants were not given any advance indication of what the tests might be intended to measure. Participants were randomly assigned to one of two groups: a control group, asked to complete the attribute-measuring tests at the beginning and the end of the study; and, an intervention group that would be asked to play selected video games under controlled conditions and to complete the tests.

Basic demographic information was collected for all participants, including age, gender and subjects studied. In addition, participants were asked to estimate the frequency with which they played video games. A total of 100 students were recruited (mean age 20.45, 52.2% female). Of the 50 students randomly assigned to each group, 36 completed the initial round of testing per group. 20 (40%) of those assigned to the control group completed both pre- and post-tests, while 16 (32%) of the intervention group saw the study through to completion. Groups were broadly comparable in terms of gender, age, year of study and time spent playing video games per week (Table 1).

Table 1
Summary of week 1 test scores and demographic information by control/intervention group.

		Control	Intervention	*p*
N		36	36	
Measures				
Communicative Adaptability Scale (mean (SD))		100.14 (8.92)	99.06 (17.88)	0.746
I-ADAPT-M (mean (SD))		202.69 (19.70)	200.36 (37.65)	0.743
Resourcefulness Scale (mean (SD))		82.75 (19.75)	81.44 (23.33)	0.798
Demographic information				
Note that one participant in both groups did not complete the demographic survey, so N = 35 for these data.				
Year of study (%)	Year 1	22 (62.9)	24 (68.6)	0.801
	Year 2	13 (37.1)	11 (31.4)	
Age (mean (SD))		19.80 (3.41)	21.09 (5.95)	0.271
Gender (%)	Female	18 (51.4)	20 (57.1)	0.346
	Male	14 (40.0)	15 (42.9)	
	Other	3 (8.6)	0 (0.0)	
Hours spent playing video games per week (%)	0	10 (28.6)	9 (25.7)	0.973
	1-4	12 (34.3)	14 (40.0)	
	4-8	6 (17.1)	6 (17.1)	
	>8	7 (20.0)	6 (17.1)	
Retention (%)	Completed	20 (55.6)	16 (44.4)	0.48
	Lost to follow up	16 (44.4)	20 (55.6)	

The game labs operated on a drop-in basis, open for students to come and play the specified games between the hours of 9am-5pm every Tuesday, Wednesday and Friday. The flexible drop-in structure partially addresses a common criticism of laboratory-based video game studies, where an arbitrarily defined time limit on play does not mirror the circumstances under which players normally play games (Egenfeldt-Nielsen, Smith, & Tosca, 2008). While participants were asked to log 120 min of play on most games (60 min each on *Gone Home* and *Papers, Please*; 14 h total game play across eight games), time management was the students' responsibility, meaning they could choose to play for "just 5 min more" or leave when they had a class to attend. Participants did occasionally opt to play for longer than the prescribed period, either because they were simply immersed in an enjoyable experience, or because it is more natural to stop playing at a suitable juncture in the game, for example, at the completion of a level or mission. In this sense, the lab was arranged to provide better ecological validity than would have been afforded by imposing a rigid temporal structure on proceedings. Under normal conditions, no player would choose to stop when they are in the middle of a game they are enjoying, and many players – particularly those less accustomed to lengthy sessions of video game play – might find being asked to endure 2 h of an unfamiliar game in a single sitting prohibitively tiring.

The control and intervention groups were assessed for similarity at baseline (taken to mean the point at which the first battery of tests was completed, following randomisation) by comparing demographic factors and baseline test scores by attribute (Table 1). Due to concern about retention bias, these baseline features were also compared between those with and without week 8 (end of study) scores (Table 2). All comparisons were by Fisher's exact test for categorical variables, and by unequal variances *t*-test assuming unequal variance between groups for continuous variables (Welch's *t*-test). Total scores for each attribute were calculated, in accordance with the published scoring mechanisms, for each participant at each time point they completed testing.

To assess the primary research question, a summary measure of "score change" was calculated for each attribute by subtracting week 1 score from week 8 score for each participant with available data. Thus, each participant is assigned a score change for each attribute, which is negative if their score worsened, and positive if their score improved. The distribution of score changes was assessed in both groups (control and intervention) for each attribute:

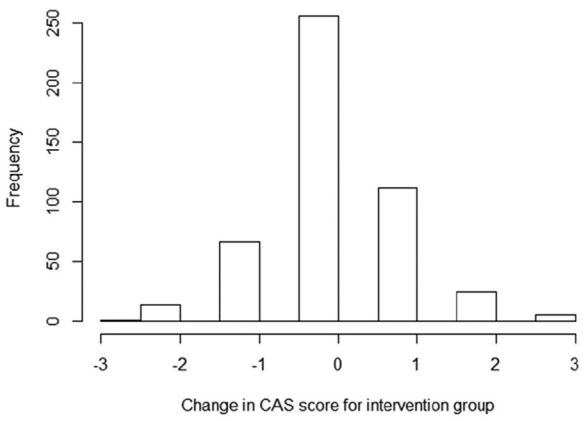
- Score changes were assessed for normality graphically by histogram (Fig. 1);
- Each participant's week 1 score was plotted against week 8 score in a scatter plot, such that participants with positive score changes lay above the diagonal, and negative score changes below the diagonal (Fig. 2B);
- Differences in score change between the groups were formally assessed by calculating a Cohen's *d* for difference in the means for the groups and tested using *t*-test assuming unequal variance (Table 3).

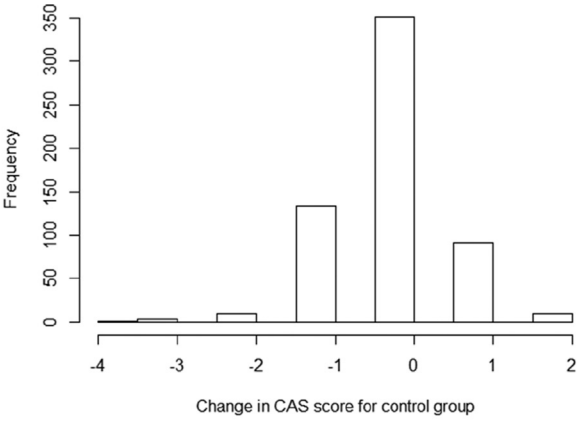
4. Results

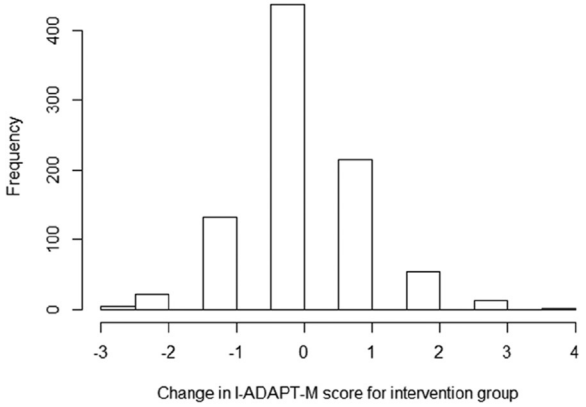
Test scores for communication, adaptability and resourcefulness all showed significant increases for the game-playing invention group over the control group, and 95% confidence intervals calculated for the difference between mean scores for the control and intervention groups did not cross zero (see Fig. 3, Tables 4,5 and 6). All three measures produced approximately normal distributions of scores. Cronbach's alpha coefficients calculated for baseline data indicated very good

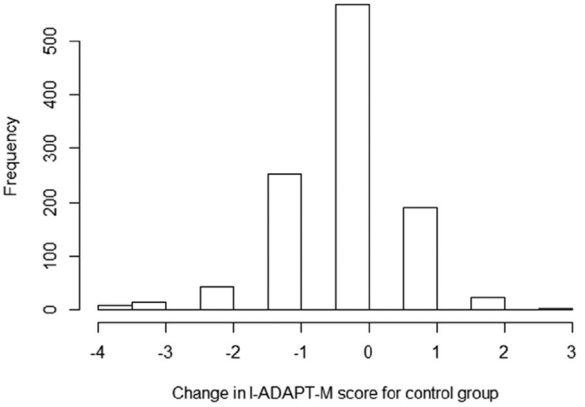
Table 2
Summary of week 1 test scores and demographic information by completed/lost to follow-up.

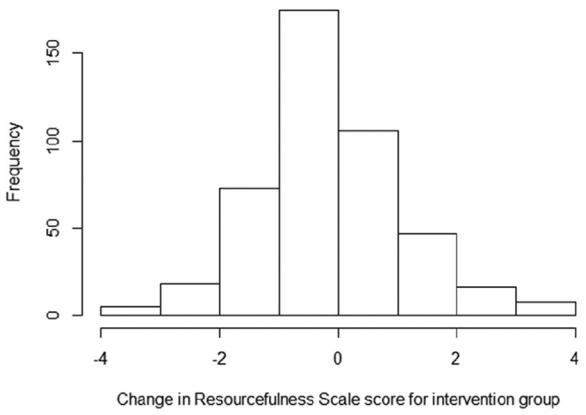
		Completed	Lost to follow-up	*p*
N		36	36	
Group (%)	Control	20 (55.6)	16 (44.4)	0.48
	Intervention	16 (44.4)	20 (55.6)	
Measures				
Communicative Adaptability Scale (mean (SD))		97.72 (8.41)	101.47 (17.94)	0.26
I-ADAPT-M (mean (SD))		200.22 (19.06)	202.83 (37.97)	0.713
Resourcefulness Scale (mean (SD))		80.94 (18.09)	83.25 (24.60)	0.652
Demographic information				
Note that one participant in both groups did not complete the demographic survey, so N = 35 for these data.				
Year of study (%)	Year 1	19 (54.3)	27 (77.1)	0.078
	Year 2	16 (45.7)	8 (22.9)	
Age (mean (SD))		21.06 (4.28)	19.83 (5.36)	0.293
Gender (%)	Female	20 (57.1)	18 (51.4)	0.714
	Male	13 (37.1)	16 (45.7)	
	Other	2 (5.7)	1 (2.9)	
Hours spent playing video games per week (%)	0	7 (20.0)	12 (34.3)	0.32
	1–4	14 (40.0)	12 (34.3)	
	4–8	5 (14.3)	7 (20.0)	
	>8	9 (25.7)	4 (11.4)	











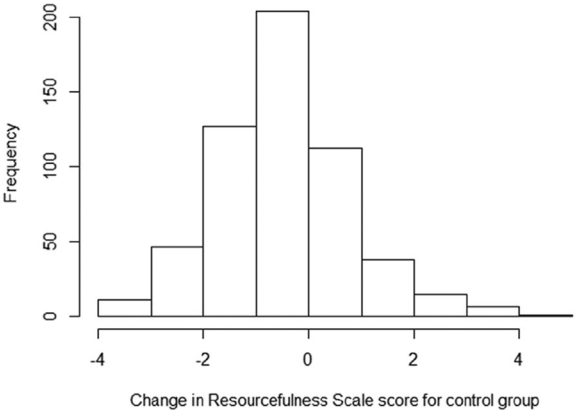


Fig. 1. Histograms showing distributions of score change between week 1 and week 8 for each measure (CAS, I-ADAPT-M and Resourcefulness Scale), for both control and intervention groups.

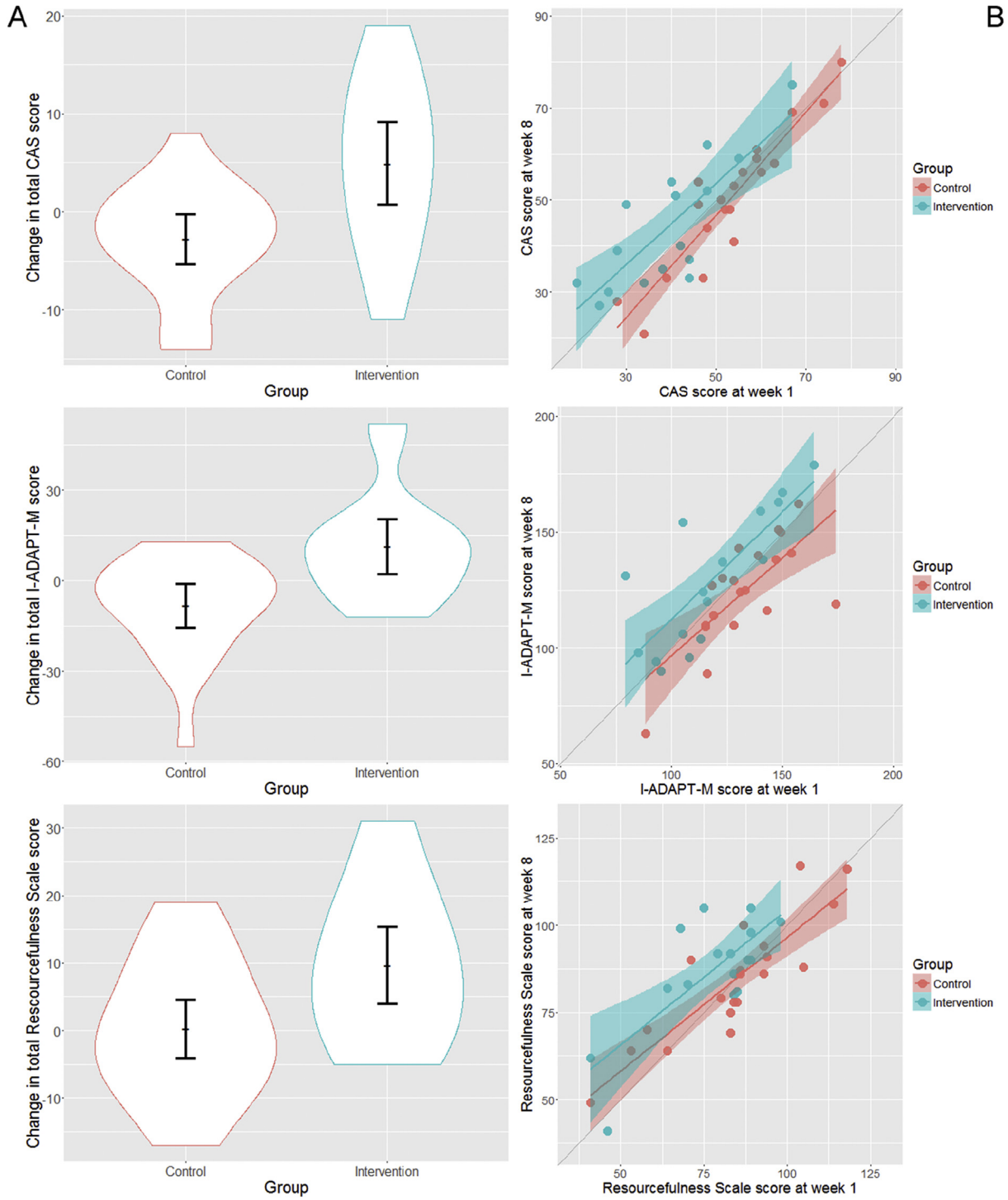


Fig. 2. Left panel (A) shows distributions of total Communicative Adaptability Scale (CAS), I-ADAPT-M and Resourcefulness Scale score change from week 1 to week 8 by control and intervention groups. Plots are kernel density ('violin') plots. Error bars are 2 standard errors of the mean. Right panel (B) shows scatterplots of scores for the same measures for week 1 and week 8 for both groups. Line of equality for week 1 and week 8 scores is shown (diagonal), and a line of best fit (least squares method) is plotted for each group.

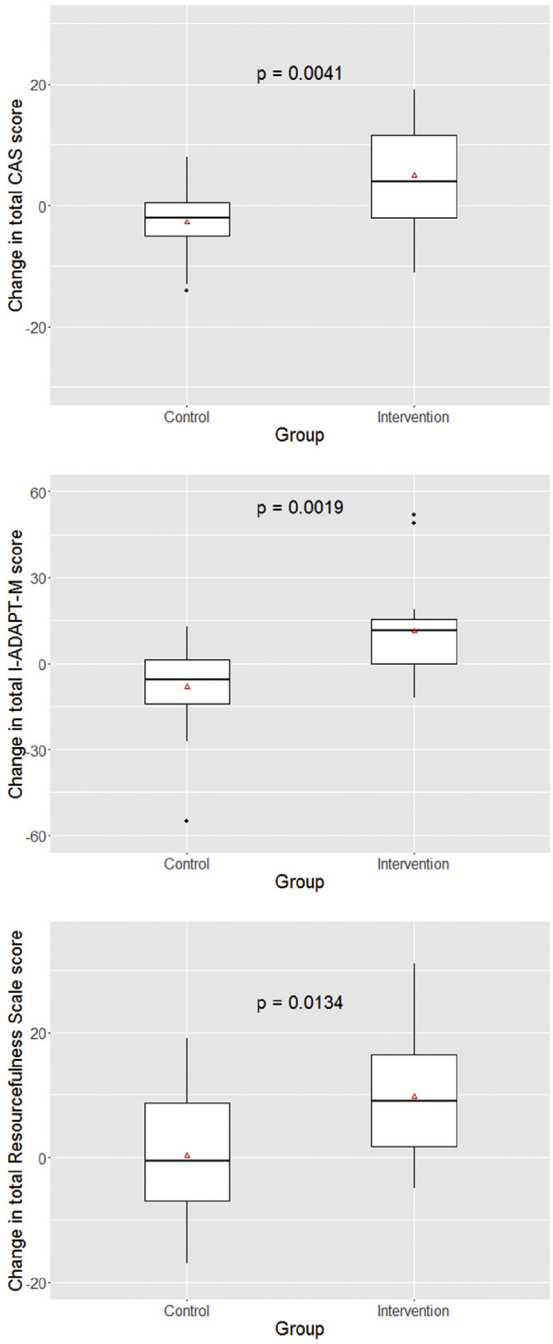


Fig. 3. Box plots comparing distributions of total Communicative Adaptability Scale (CAS), I-ADAPT-M and Resourcefulness Scale score change from week 1 to week 8 between control and intervention groups. The horizontal line is the median score change for the group, the small red triangle represents mean change, the box represents interquartile range, whiskers show the two standard deviation range used to define outlier, and outliers are plotted as dots. Included as an alternative means of visualizing those data shown in Fig. 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

internal consistency for each measure ($\alpha = 0.88$, $\alpha = 0.91$ and $\alpha = 0.84$ for Communicative Adaptability Scale, I-ADAPT-M and Resourcefulness Scale, respectively).

When normalized by standard deviation, the differences in mean score between week 1 and week 8 indicate a large effect size (Cohen's *d* of between 0.9 and 1.15) for all three measures. Changes in score are summarized in Table 3.

The violin plots below (Fig. 2A) describe the distribution of score change for each measure.

Table 3
Summary of score changes from week 1 to week 8.

	Control		Intervention		Difference in means			*p*
	Mean	SD	Mean	SD	Absolute	95% CI	Cohen's *d*	
CAS	−2.8	5.65	4.94	8.41	7.74	12.79 to 2.69	1.1	0.004
I-ADAPT-M	−8.25	15.99	11.31	18.07	19.56	31.32 to 7.8	1.15	0.002
Resourcefulness	0.25	9.71	9.69	11.42	9.44	16.77 to 2.11	0.9	0.013

Table 4
Summary of changes in Communicative Adaptability Scale scores for control and intervention groups.

Change	Group		Row total
	Control	Intervention	
negative (N)	15	5	20 (56%)
negative (N/row total)	0.75	0.25	
negative (N/column total)	0.75	0.31	
positive (N)	5	11	16 (44%)
positive (N/row total)	0.31	0.69	
positive (N/column total)	0.25	0.69	
Column Total	20 (56%)	16 (44%)	36 (100%)

Table 5
Summary of changes in I-ADAPT-M scores for control and intervention groups.

Change	Group		Row total
	Control	Intervention	
negative (N)	12	4	16 (44%)
negative (N/row total)	0.75	0.25	
negative (N/column total)	0.60	0.25	
positive (N)	8	12	20 (56%)
positive (N/row total)	0.40	0.60	
positive (N/column total)	0.40	0.75	
Column Total	20 (56%)	16 (44%)	36 (100%)

Table 6
Summary of changes in Resourcefulness Scale scores for control and intervention groups.

Change	Group		Row total
	Control	Intervention	
negative (N)	12	3	15 (42%)
negative (N/row total)	0.8	0.2	
negative (N/column total)	0.6	0.19	
positive (N)	8	13	21 (58%)
positive (N/row total)	0.38	0.62	
positive (N/column total)	0.4	0.81	
Column Total	20 (56%)	16 (44%)	36 (100%)

On all three scales, students in the game-playing invention group were more likely to show a positive score change than the control group students. 5 of 20 (25%) students in the control group had improved communication scores versus 11 of 16 (69%) in the intervention group ($p = 0.24$, Fisher's exact test). Similar differences were found for adaptability (8 (40%) versus 12 (75%); $p = 0.03$, Fisher's exact test) and resourcefulness (8 (40%) versus 13 (81%); $p = 0.014$, Fisher's exact test). Thus, the effect was not seen in every intervention group participant, but the group effects were highly significant.

Mean score change on the communication scale was −2.8 (SD = 5.65) in the control group and 4.94 (SD = 8.41) in the intervention group (absolute difference in means = 7.74, 95% CI 2.69 to 12.79, Cohen's *d* 1.1). Mean score change on the adaptability scale was −8.25 (SD = 15.99) in the control group and 11.31 (SD = 18.07) in the intervention group (absolute

difference in means = 19.56, 95% CI 7.8 to 31.32, Cohen's *d* 1.15). Mean score change on the resourcefulness scale was 0.25 (SD = 9.71) in the control group and 9.69 (SD = 11.42) in the intervention group (absolute difference in means = 9.44, 95% CI 2.11 to 16.77, Cohen's *d* 0.9). Outliers were observed in both groups, but the measured effect of the intervention was broadly similar across the range of baseline scores (Fig. 2B). As a sensitivity analysis, data was re-analysed with outliers (as defined by > 2 SD from mean) excluded. This made no substantive difference to the results of *t*-tests.

5. Issues and limitations

Certain challenges were encountered during the study, some of which may be relevant if the ideas described in this paper were to be applied in another formal educational setting. These issues, and limitations of the study in terms of its design and implementation, are discussed here.

As noted above, twelve Windows-based PC workstations were obtained and all games were played on this platform. The predominant – and most cost-effective – means of procuring games on PC is Valve's Steam service, which provides digital downloads (replete with a form of Digital Rights Management, or DRM) at competitive prices, especially during their regular sale events. Thus, this was the mechanism by which most games were obtained. It was known from the pilot project that the network ports required for Steam were not open on the university firewall and a support call to have these ports opened was logged with the relevant IT service. Quite understandably, opening ports is a non-trivial undertaking for a large institution's networking team, which also happened to be engaged in rolling out the IT infrastructure for a recently expanded campus at the time. As a result, these ports were not open at the commencement of the study. For the most part, this issue caused little more than inconvenience – the researcher used a mobile internet hotspot and a USB Wi-Fi dongle to install Steam and the required games on each machine, which simply required time and patience. Potentially more problematic was the use of Valve's Steam service to facilitate matchmaking, mediating the connections between players that are required for multiplayer gaming. By default, Steam is used to facilitate the multiplayer component of *Portal 2* and *Team Fortress 2* – both titles produced by Valve. An unforeseen additional complication with Steam matchmaking related to the limitations placed on newly-created Steam accounts. In order to purchase and install multiple copies of each game, a Steam account with a unique associated email address was created for each machine in the lab and games bought – as gift purchases, via the researcher's own Steam account – for each. However, in order to "protect our users from spamming, phishing, and other abuse, Steam prevents some accounts from accessing certain community and social features" (Limited User Accounts - Managing Your Account Features - Knowledge Base - Steam Suppor, 2016). To this end, Steam limits the ability of accounts which have spent less than five US dollars to engage in multiplayer activity, such as sending friend invites. As all of the purchasing was done through a single account, this restriction remained in place for all of the accounts used in the lab.

However, solutions were found for both games. As noted above, a local dedicated server was created for *Team Fortress 2*, to which participants' games connected instead of looking to Steam for potential matches. This was straightforward to accomplish as LAN play of this nature is supported by default in the game. The solution for *Portal 2* was slightly less straightforward, as LAN co-op is not an option available to players when they launch the game and, while the functionality does exist in the game's code, it may only be accessed by entering command line instructions via a normally-concealed console. A more user-friendly workaround was found on the Steam Users' Forum which, through the modification of one of the game's configuration files, allowed the option to connect to games on specified lab computers to be added to the game's menu. Participants were then instructed to choose the relevant menu option, depending on whether they were to host or connect to a game. In the latter case, they were instructed to choose the option that would connect them to the machine hosting the game, as identified by its IP address, which was clearly displayed on each machine.

MMORPGs *World of Warcraft* (*WoW*) and *Star Wars: The Old Republic* were also considered for inclusion but these games both require still more network ports to be opened, so this was not a realistic aim. While technical issues precluded any experimentation with these titles, the learning curve and overall complexity of such MMORPGs was also taken into consideration. Certainly, in the 2 h allocated to most of the other games, it seems unlikely that newcomers to the MMORPG genre would have progressed much beyond the most rudimentary stages of the game and thus would not experience the most relevant aspects of the genre, particularly the group-based questing. Further, if existing *WoW* players in the experimental cohort wished to use their own characters – an understandable desire if they have invested many hours into their creation and development – this might also have been problematic, as high level characters are generally prevented from grouping with new players in the first place.

More mundane were the issues relating to the available hardware. As noted above, the machines used in the lab were not remarkably well-specified, and certainly not intended for gaming. They did, however, prove perfectly usable for most games, especially when graphical options were adjusted to reflect the limited capabilities of the machines' graphics cards. The exception to this was *Lara Croft and the Guardian of Light* which, even on the lowest performance settings, was somewhat sluggish and occasionally unresponsive. *Lara Croft* was also one of the few games to rely on the supplied games controllers rather than the keyboard and mouse, and the build quality of the controllers – styled after the very popular Xbox 360 controller – was such that the effects of sluggish controls were amplified by the peripherals' shortcomings. These shortcomings, which mostly related to the left analogue stick typically used for movement, were not immediately obvious. On delivery, the controllers appeared to be of excellent quality and offered a more than acceptable recreation of the experience afforded by the rather more expensive Xbox-branded version of the peripheral. However, after a period of use, the analogue

stick became unreliable, particularly in the diagonal directions – the very directions in which a player of an isometrically-presented game is likely to wish to move.

Such issues were not critical, however. They might have been considered so if they had resulted in a significant number of participants abandoning the study in response to hardware or software problems but, while several participants passed comment on the unsatisfactory nature of their experience with *Lara Croft*, none stated that it was unplayable or that they wished to abandon play. Of the 20 participants who completed the surveys associated with the preceding game (*Portal 2*), 17 completed those associated with *Lara Croft*.

In general, then, technical issues were of relatively little concern, although the nature of the potential problems outlined here should be borne in mind by educators considering similar projects. Other limitations relate to the experimental design of the study.

While the demographic information collected at the beginning of the study was thought sufficient, and is complete enough to account for important variables such as gender and year of study, it might have been useful to collect data on participants' previous study and work experience. For example, a student who has previously attended university or college, or was employed for any substantial period, would presumably possess more finely-honed graduate or work-related skills. The randomised nature of the study addresses concerns about the potential effects that variations in such experience might have, but questions remain about the impact of the game-playing intervention on these participants. Could gains in attribute attainment resulting from the intervention be tempered by previous exposure to opportunities for attribute development, for example?

The experiment was designed with ecological validity in mind, but even the more relaxed, drop-in structure of the lab-based game play is not an accurate reflection of how games are played at home. This is of little concern if this work is viewed as a potential model for introducing selected commercial video games to the higher education experience, as the arrangements in such cases would necessarily be similar to those described here. However, the nature of the intervention means it is difficult to make claims about the effects of playing video games more generally. The lab experience was often quite social in nature, with participants typically surrounded by those with whom they were playing, or who were playing the same game, and thus had a common interest. It is possible that the effects would be different if the games were played online, which might be more typical of the multiplayer games used here.

Participants were also aware that they were being measured by means of the online tests, and observed – however informally – by the researcher in the room. As such, the Hawthorne Effect (French, 1953) must be considered. The Hawthorne Effect, particularly where educational interventions are concerned, has been dismissed by some (see Deslauriers, Schelew, & Wieman, 2011 for a refutation of the phenomenon) but there is little doubt that the atypical circumstances under which the games were played, and their effects measured, could have had some subtle impact. However, given the nature of the study – which was not concerned with how the participants played the games or otherwise behaved under lab conditions – any such effect is considered by be negligible.

The limitations of self-report measures, in general, should also be borne in mind: such measures are susceptible to complications including social desirability bias, response bias, and issues related to participants' understanding of the language used to construct the surveys.

Any future work should also attempt to identify the specific types of game – or game features – that are most effective in raising communication, adaptability and resourcefulness scores. This might be achieved by modifying the experimental design to include multiple intervention groups, each playing a particular game for an extended period of time. However, the titles used here were carefully selected to ensure that there was a focus on multiplayer co-operation and it is likely that this is the most salient feature of the games played, certainly where communication is involved. In this respect, that a game features a co-operative multiplayer mode is probably a more important consideration than the nature of the game play involved.

6. Conclusion

This work demonstrates that playing commercial video games can have a positive effect on communication ability, adaptability and resourcefulness in adult learners, suggesting that video games may have a role to play in higher education. The study also suggests that graduate skills may be improved in a relatively short amount of time, with the gains reported here achieved over a period of eight weeks and representing just 14 h of game play. Certainly, the results of the randomised controlled trial described here suggest that the popular discourse around games' alleged ill-effects should be tempered by considerations of the potential positive outcomes of playing video games. There is a need for further, similarly robust study of these effects, and such work may only be hampered by ill-informed attitudes to the ubiquitous and immensely popular medium of video games.

Funding

The pilot project on which this work is based received seed funding from the RCUK Digital Economy Communities & Culture Network+(EP/K003585/1).

Acknowledgments

I wish to thank Dr. Steve Draper for advice on experiment design, Dr. David Barr for his contribution to statistical analysis and Dr. David Wilson for feedback on this manuscript.

References

About Gone Home. 2012, June 19. Retrieved from https://fullbright.company/gonehome/.
Adachi, P. J. C., & Willoughby, T. (2013). More than just fun and games: The longitudinal relationships between strategic video games, self-reported problem solving skills, and academic grades. *Journal of Youth and Adolescence, 42*(7), 1041–1052. https://doi.org/10.1007/s10964-013-9913-9.
Barr, M. (2013). Can playing video games help develop graduate attributes?. In *Proceedings of the 6th annual university of Glasgow learning and teaching conference. Glasgow, UK*. Retrieved from http://eprints.gla.ac.uk/78489/.
Barrie, S. C. (2006). Understanding what we mean by the generic attributes of graduates. *Higher Education, 51*(2), 215–241. https://doi.org/10.1007/s10734-004-6384-7.
Barrie, S. C. (2007). A conceptual framework for the teaching and learning of generic graduate attributes. *Studies in Higher Education, 32*(4), 439–458. http://doi.org/10.1080/03075070701476100.
Blizzard Entertainment. (2016). *Warcraft III*. Retrieved 25 April 2016, from http://us.blizzard.com/en-us/games/war3/.
Borderlands 2-Gearbox Software. (2016). Retrieved 25 April 2016, from http://www.gearboxsoftware.com/games/borderlands-2.
Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031), 862–864. http://doi.org/10.1126/science.1201783.
Dring, C. (2010, July 20). *EA's Moore: Metacritic mania a 'slippery slope'?*. Retrieved 24 March 2016, from http://www.develop-online.net/news/ea-s-moore-metacritic-mania-a-slippery-slope/0107404.
Duran, R. L. (1983). Communicative adaptability: A measure of social communicative competence. *Communication Quarterly, 31*(4), 320–326. http://doi.org/10.1080/01463378309369521.
Duran, R. L. (1992). Communicative adaptability: A review of conceptualization and measurement. *Communication Quarterly, 40*(3), 253–268. http://doi.org/10.1080/01463379209369840.
Egenfeldt-Nielsen, S., Smith, J. H., & Tosca, S. P. (2008). *Understanding Video Games: The Essential Introduction* (New Ed). Routledge.
French, J. R. P. (1953). Experiments in field settings. In F. J. Gravetter, & L.-A. B. Forzano (Eds.), *Research methods for the behavioral sciences* (pp. 98–135). New York: Holt, Rinehart & Winston.
Gee, J. P. (2007). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.
Graft, K. (2011). Take-Two's zelnick stresses importance of metacritic scores. Retrieved 24 March 2016, from http://www.gamasutra.com/view/news/33092/TakeTwos_Zelnick_Stresses_Importance_Of_Metacritic_Scores.php.
Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist, 69*(1), 66–78.
Hughes, C., & Barrie, S. (2010). Influences on the assessment of graduate attributes in higher education. *Assessment & Evaluation in Higher Education, 35*(3), 325–334. http://doi.org/10.1080/02602930903221485.
Lara Croft and the Guardian of Light. (2016). Retrieved 25 April 2016, from http://www.laracroftandtheguardianoflight.com/.
Limited User Accounts - Managing Your Account Features - Knowledge Base - Steam Support. (2016). Retrieved 8 June 2016, from https://support.steampowered.com/kb_article.php?ref=3330-IAGK-7663.
Lloyd, R. (2011). *The GAGA project: Using games technology to develop graduate attributes*. University of Abertay Dundee. Retrieved from http://www.enhancementthemes.ac.uk/pages/docdetail/docs/case-studies/the-gaga-project-using-games-technology-to-develop-graduate-attributes.
Metacritic - Movie Reviews, TV Reviews, Game Reviews, and Music Reviews. (2016). Retrieved 25 April 2016, from http://www.metacritic.com/.
Minecraft. (2016). Retrieved 25 April 2016, from https://minecraft.net/.
Nicol, D. J. (2010). *The foundation for graduate attributes: Developing self-regulation through self and peer-assessment*. The Quality Assurance Agency for Higher Education. Retrieved from http://qmwww.enhancementthemes.ac.uk/docs/publications/the-foundation-for-graduate-attributes-developing-self-regulation-through-self-and-peer-assessment.pdf.
Official Portal 2 Website. (2016). Retrieved 25 April 2016, from http://www.thinkwithportals.com/.
Papers, Please. (2016). Retrieved 25 April 2016, from http://papersplea.se/.
Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612–624. http://doi.org/10.1037/0021-9010.85.4.612.
Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Borman, W. C., & Hedge, J. W. (2002). Predicting adaptive Performance: Further tests of a model of adaptability. *Human Performance, 15*(4), 299–323. http://doi.org/10.1207/S15327043HUP1504_01.
Ployhart, Robert E., & Bliese, P. D. (2006). Individual adaptability (I-ADAPT) Theory: Conceptualizing the antecedents, consequences, and measurement of individual differences in adaptability. In *Understanding adaptability: A prerequisite for effective performance within complex environments* (Vol. 6, pp. 3–39). Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/doi/full/10.1016/S1479-3601%2805%2906001-7.
Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education, 80*, 58–67. https://doi.org/10.1016/j.compedu.2014.08.013.
Spitzberg, B. H., & Cupach, W. R. (1984). *Interpersonal communication competence*. SAGE Publications.
Team Fortress 2. (2016). Retrieved 25 April 2016, from http://www.teamfortress.com/.
Thomas, D., & Brown, J. S. (2011). *A new culture of learning: Cultivating the imagination for a world of constant change*. Lexington, Ky: CreateSpace.
Zauszniewski, J. A., Lai, C.-Y., & Tithiphontumrong, S. (2006). Development and testing of the resourcefulness scale for older adults. *Journal of Nursing Measurement, 14*(1), 57–68. http://doi.org/10.1891/jnum.14.1.57.