

Motivating children to learn effectively: exploring the value of intrinsic integration in educational games

HABGOOD, MP Jacob <<http://orcid.org/0000-0003-4531-0507>> and
AINSWORTH, Shaaron E

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/3556/>

This document is the Accepted Version [AM]

Citation:

HABGOOD, MP Jacob and AINSWORTH, Shaaron E (2011). Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, 20 (2), 169-206. [Article]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games.

M.P. Jacob Habgood^{1,2} and Shaaron E. Ainsworth¹

Abstract

The concept of *intrinsic motivation* has been considered to lie at the heart of the user engagement created by digital games. Yet despite this, educational software has traditionally attempted to harness games as extrinsic motivation by using them as a sugar-coating for learning content. This paper tests the concept of *intrinsic integration* as a way of creating a more productive relationship between educational games and their learning content. Two studies assessed this approach by designing and evaluating an educational game for teaching mathematics to seven to eleven year olds called *Zombie Division*. Study 1 examined learning gains of 58 children who played either the intrinsic, extrinsic or control variants of *Zombie Division* for two hours, supported by their classroom teacher. Study 2 compared time-on-task for intrinsic and extrinsic variants of the game when 16 children had free choice between them. The results of these studies showed that children learned more from the intrinsic version of the game under fixed time limits and spent seven times longer playing it in free time situations. Together they offer evidence for the genuine value of an intrinsic approach for creating effective educational games. The theoretical and commercial implications of these findings are discussed.

1) School of Psychology and Learning Sciences
Research Institute

University of Nottingham,

University Park, Nottingham, NG7 2RD, UK

Email: Shaaron.Ainsworth@nottingham.ac.uk

Phone +44 115 9515314

2) Department of Computing

Sheffield Hallam University,

Furnival Building, Sheffield, S1 1WB

Email: J.Habgood@shu.ac.uk

Phone +44 114 225 6709

The use of computer games and simulations in education dates back to the 1950's (Cullingford, Mawdesley, & Davies, 1979) when computing was still in its infancy and the commercial videogame industry had yet to emerge. Nonetheless it was the raw engagement power of 80's videogames like Pac-Man that inspired a new generation of educationalists to consider the learning potential of this exciting new medium (Bowman, 1982). These early protagonists were quick to identify the motivational power of videogames as their key asset (e.g. Lepper & Malone, 1987; Loftus & Loftus, 1983) and were able to apply a range of existing motivational (e.g. Csikszentmihalyi, 1975; Deci, 1975; Lepper & Greene, 1975) and behavioral (Ferster & Skinner, 1957) theories to their rationales. However, despite this promising start, the resulting generations of 'edutainment' products have been widely recognized as failing to effectively harness the engagement power of digital games (e.g. Hogle, 1996; Kerawalla & Crook, 2005; Papert, 1998; Trushell, Burrell, & Maitland, 2001). So while the mainstream games industry boomed throughout the 1990s, the educational sector was left behind in terms of technology, revenues and commercial interest. However, the turn of the millennium has seen a rejuvenation of interest in game-based learning with a number of texts extolling the potential of games (e.g. Aldrich, 2004; Gee, 2003; Shaffer, Squire, Halverson & Gee, 2006), paralleled by commercial success of 'self-improvement' titles such as 'Brain Training' and 'Big-Brain Academy' (Nintendo).

This paper offers empirical evidence for the value of a design approach which may help to explain the failure of edutainment to fulfill its educational promise. This approach hinges upon the ability of learning games to effectively harness the *intrinsic motivation* (Deci, 1975) of a game for educational goals by creating an *intrinsic integration* (Kafai, 1996) between a game and its learning content. Furthermore, we suggest that such an integration is created through an intrinsic link between a game's *core mechanics* (Lundgren & Björk, 2003) and its learning content.

Zombie Division is a computer game specifically created to empirically examine the concept of intrinsic integration. The game integrates mathematics into the core-mechanic of a 3D adventure through a combat system in which opponents are mathematically divided in order to defeat them. Three variations of this game were created for evaluation: an *intrinsic* version which integrated mathematics into combat, an *extrinsic* version which had non-mathematical combat and placed identical mathematical multiple choice questions between levels instead, and a *control* version which contained no mathematics at all. The first study compared learning gains between all three versions as a measure of the relative educational effectiveness of the intrinsic approach. The second study compared time-on-task between the intrinsic and extrinsic versions of the game as a measure of the relative motivational appeal of the intrinsic approach.

Defining Intrinsic Integration

The concept of intrinsic integration in educational games is rooted in the more familiar concept of ‘intrinsic motivation’. It is commonly surmised that a person is intrinsically motivated to perform an activity when he receives no apparent rewards except the activity itself (Deci, 1975). Although modern videogames *can* provide external rewards (such as those produced by farming virtual game resources: see Steinkuehler, 2006) they are largely autonomous pursuits which create their own internal motivations for continuing the activity. Game designers can create these internal motivations through the inclusion of aspects such as *challenge, control, fantasy* and *curiosity*, while inter-personal motivations can be added through factors such as *competition, co-operation* and *recognition* (Malone & Lepper, 1987). The inclusion of *challenge* in this taxonomy is derived from the work of Csikszentmihalyi (1988) into flow theory and optimal experience. This proposes that clear goals, achievable challenges and accurate feedback are all required to achieve a state of flow in an activity which requires “*a balance between the challenges perceived in a given situation and the skills a person brings to it*”, suggesting that “*no activity can sustain it for long unless both the challenges and the skills become more complex*” (p.30). There are clear parallels between this and the way that game designers carefully structure the difficulty curves of their games to provide the optimal level of challenge as a player’s skills develop (Habgood & Overmars, 2006, p.158). It is perhaps unsurprising then that feelings of total concentration, distorted sense of time, and extension of self are as common experiences to game players as they are to Csikszentmihalyi’s (1988) rock climbers and surgeons. There is also emerging evidence that, when measured correctly, flow is predictive of learning (e.g. Engeser & Rheinberg, 2008)

The gaming literature provides an overwhelming number of different approaches to defining the essence of a game (Caillois, 1961; Crawford, 1982; Huizinga, 1950; Juul, 2005; Koster, 2005; Salen & Zimmerman, 2004). Yet these differences only serve to highlight Wittgenstein’s (1953) observation on games that “*you will not see something that is common to all, but similarities, relationships, and a whole series of them at that*” (aphorism 66). Therefore we in the interests of practicality we use a definition of a game which seeks to highlight the main differences between games and other forms of entertainment, rather than all the similarities between things we might refer to as a game. This pragmatic definition defines a game as simply an “interactive challenge”, suggesting that games contain an interactive element that distinguishes them from films, and prescribed challenges that distinguish them from toys (Habgood & Overmars, 2006, p87). We therefore see games as something which encompasses a wide spectrum of digital and non-digital applications – including many simulations – and we hope this research could potentially have relevance to all of them. It should also be noted that our definition deliberately avoids assigning a motivational aspect to the definition of a game as the experience of intrinsic motivation is

subjective (does a game stop being a game if it stops being fun?). Nonetheless, the ability of games or simulations to create intrinsic motivation is clearly central to this argument and uninspiring games are not a good model for creating motivating learning games either.

Although digital games may be capable of providing activities which are intrinsically motivating in their own right, it is critical to consider the effect of adding learning content to an intrinsically motivating game. Game designers have come to recognize the role of learning in good game design (e.g. Crawford, 1982; Gee, 2003; Habgood & Overmars, 2006; Koster, 2005). This is not about commercial games containing educational content, but how the enjoyment of games derives from the process of learning itself: i.e. “*the fundamental motivation for all game-playing is to learn*” (Crawford, 1982, p.17). Unfortunately edutainment products have traditionally taken a “chocolate-covered broccoli” (Bruckman, 1999) approach when combining learning content with gameplay. This is where the gaming element of the product is used as a separate reward or sugar-coating for completing the educational content.

It was Malone’s (1980) and Malone and Lepper’s (1987) seminal work on videogames which first considered the problem of creating a more integrated approach to designing educational games. This originally proposed the concept of an *intrinsic fantasy* as providing “*an integral and continuing relationship between the fantasy context and the instructional content being presented*” (1987, p.240). This was contrasted with an *extrinsic fantasy* where “*the fantasy depends on the skill being learned, but not vice versa*” and it was suggested that the learning content of extrinsic fantasies could be easily replaced with something different. Furthermore, it was suggested that “*In general, intrinsic fantasies are both more interesting and more instructional than extrinsic fantasies*” (1980, p.164). We can attempt to clarify this definition by considering contrasting examples. The classic Maths Blaster’s series included a game called “Trash Zapper” which required the player to provide the answer to a simple arithmetic sum ($9+6=?$) by shooting a moving item of rubbish that has the answer written on it. However, you could replace arithmetic sums with a spelling tasks (e.g. ELE?HANT) and attach letters to the rubbish without having a significant impact on the fantasy context of the game. Therefore this game could be considered to be an extrinsic fantasy. Conversely, the same definition might consider football management games to provide intrinsic fantasies for mathematical learning content. This kind of game allows control over team budgets and team statistics and clearly it would completely break the fantasy context of the game to try and balance a payroll, or a team's abilities by spelling words correctly.

However, we have argued that this focus on the intrinsic nature of fantasy is misplaced, suggesting that such fantasy contexts are often purely arbitrary and could be swapped for another so long as the basic mechanics of the game are not altered (Habgood, Ainsworth, & Benford, 2005a). The football example may be intrinsic according to the definition above, but the fantasy-context of football could be

replaced by Smurf volleyball so long as the rule-systems and player interactions with the budgets and player statistics remained the same. The game may not appeal to the same audience, but it would be equally as ‘instructional’ as the original. It is therefore the core-mechanics which embody the rule-systems and player interactions that are intrinsic to the educational value of football management games and not the fantasy context at all.

Consequently we consider the term “intrinsic integration” to be a more appropriate way to describe a situation in which “*a designer integrates the subject matter with the game idea*” (Kafai, 1996 p82). This is not to suggest that the theoretical argument for the role of core-mechanics has been ‘won’ and many researchers still prefer to describe the concept of effective integration with reference to fantasy (Asgari & Kaufman, 2004; Gunter, Kenny, & Vick, 2008). Others acknowledge the wider debate between researchers in attributing the intrinsic motivation produced by games to the role of narrative context (fantasy) or intrinsic goals and rewards (core-mechanics) (Dondlinger, 2007). Nonetheless, it is clear that when some researchers refer to the intrinsic role of fantasy they mean this to include the role of both fantasy and core-mechanics in the intrinsic relationship (e.g. Paras, 2005).

The definition of intrinsic integration evaluated in this paper was first developed in Habgood et al (2005a) and has two central components. Intrinsically integrated games:

1. deliver learning material through the parts of the game that are the most fun to play, riding on the back of the flow experience produced by the game, and not interrupting or diminishing its impact and;
2. embody the learning material within the structure of the gaming world and the player’s interactions with it, providing an external representation of the learning content that is explored through the core mechanics of the gameplay.

We can now consider the two earlier examples of Maths Blaster and football management games with respect to this new definition. "Trash Zapper" certainly delivers learning material through the parts of the game that are the most fun to play and rides on the back of the flow experience. As a result it does not diminish the impact of the game, and the additional cognitive demands of the arithmetic potentially add to the challenge of an otherwise trivial gaming exercise. However, the relationship with the core mechanics and structure of the gaming world is not embodied. Although the mathematical content is attached to the core-mechanic, the two are not actually integrated at all – which is why the learning content could so easily be replaced with spelling. This thin integration means that mathematical representations are not part of the structure of the gaming world so it does not provide scope for the kind of constructivist interactions associated with microworlds and simulations (see below). Interestingly

football management games arguably fail the first part of our definition for intrinsic games, as they do not deliver the learning content through the part of the game that is most fun to play! The model for this genre of game often involves completing the 'numerical chores' between matches in order to reach the reward of watching your side play based on your team selection and substitutions. However, some players choose to skip the opportunity to watch their team play altogether and get straight back to the 'chores', which suggests that administration between games provides a flow experience in its own right. Nonetheless football management games do seem to embody some learning material within the structure of the gaming world and the player's interactions in a way that leads players to engage cognitively with the learning content in the game.

Potential Advantages and Disadvantages of Intrinsic Integration

Having described our approach to intrinsic integration we can now turn to considering what advantages or indeed disadvantages it might bring for learning. The central claimed benefit of educational games and intrinsic integration lies in the potential to more effectively motivate and engage the player in the learning content of a game. (e.g. Garris, Ahlers, & Driskell, 2002; Rieber, 1991). As such flow is often considered to be critical in creating and maintaining this motivational appeal. Integrating learning content into the very parts of the gameplay which give rise to the flow experience should ensure that the benefits of the flow are directed towards educational goals. Conversely, edutainment or extrinsic games which provide gameplay as a reward for learning content are more likely to disrupt flow if players are asked to regularly switch their to another non-flow inducing activity. Moreover the flow state in extrinsic games is therefore experienced in the service of game but not educational goals. However, whilst intrinsic integration in educational games may increase motivation and flow, it is not completely clear how this translates into increased learning (Pintrich, 2003). Mechanisms that have been postulated include persistence, more focused attention, increased arousal, increased affect and alternative strategies (Garris et al., 2002; Martens, Gulikers, & Bastiaens, 2004; Parkin, Lewinsohn, & Folkard, 1982; Pintrich, 2003; Vollmeyer & Rheinberg, 2000).

The second central benefit of intrinsic integration comes from embodying the learning content (the tasks learners must address, the actions they perform to do so and the feedback they receive as a consequence) within the core representational structure of a gaming world. There is a vast literature on the importance of external representations in learning, with much evidence that using appropriate representations (or combinations of representations) can enhance learning outcomes (Ainsworth, 2006; Scaife & Roger, 1996; Winn, 1987 all provide reviews). The value of interactive representations has long been recognised in educational simulations (de Jong, & van Joolingen, 1998) and microworlds (Papert, 1980), where structured learning environments attempt to embody a particular learning domain by

providing interactive representations within a self-contained, rule-governed world (Edwards, 1998) and the synergies between microworlds and digital games are evident (Rieber, 1996). However, while microworlds provide a carefully structured learning environment, they do not generally attempt to structure the motivational environment and manage the flow experience in the same way as digital games. So we suggest that learning should be enhanced if the representational structure embodies the core gameplay mechanisms which give rise to the central flow experience of the game.

However, it is also possible that intrinsic integration may be disadvantageous. It is well known that children find it difficult to apply mathematical knowledge acquired in one context to a different one, even if the mathematical principles are the same (e.g. Nunes, Schliemann & Caraher, 1993). So one concern is that intrinsic integration could create knowledge that is highly specialized to the specific condition of application so that learning in the game ‘stays in the game’ rather than transferring to school mathematics. Children may also apply their learning from extrinsic games more effectively as the format is often much closer to the abstract form of a school context, and situations that are similar tend to promote both enhanced recall (e.g. Tulving & Thompson 1973) and transfer (e.g. Gentner, 1989).

Our approach to intrinsic integration raises these concerns about transfer for two main reasons. Firstly, intrinsic integration typically involves concrete representations rather than abstract ones. Some (although by no means all) existing research suggests that children (in particular) can find it difficult to transfer their understanding from concrete representations to alternative representations (e.g. DeLoache 1991; Kaminski, Sloutsky, & Heckler 2009; Goldstone & Son, 2005). Interacting and playing with concrete representations can make this situation worse (e.g. Uttal, O’Doherty, Newland, Hand & DeLoache (2009)). Secondly, intrinsically integrating learning content within frantic action-based games could make it harder to learn the educational content as the learner must cope with two forms of competing demands simultaneously (the educational and game play elements). This is likely to be true of action-led games such as *Zombie Division* rather than simulations or epistemic games (Shaffer, 2004). There may be a concern that it may also inhibit contemplative reflective activity and so hinder the development of appropriate strategies and as a consequence there may be less transfer from the game (e.g. Berry, 1983).

Given that equally compelling arguments could be made for the advantages and disadvantages of intrinsic integration, we decided to test these arguments by developing and then evaluating an intrinsic and extrinsic version of the same educational game.

Zombie Division: The Educational Game

The learning content of the game is based upon the United Kingdom's National Curriculum targets for Key Stage 2 (7-11 year olds) focusing on number patterns and sequences:

- Recognise and describe number patterns, including two- and three-digit multiples of 2, 5 and 10, recognizing their patterns and using them to make predictions; recognise prime numbers up to 20 and square numbers up to 10×10 ; find factor pairs and all the prime factors of any two digit integer.

The game itself is a 3D adventure game, based around sword-fighting in which the player (acting as the hero Matrices) must use different attacks to mathematically divide opponents according to the numbers displayed on their chests (see Figure 1). The core mechanic could be described as “defeating enemies in combat by attacking each enemy with a divisor that divides their dividend into whole parts”. Each of the player's attacks has a different animation which embodies that divisor and reinforces the association between the divisor and attack. Archaic combat weapons illustrate these relationships (e.g. Divide by 2 – a single swipe of a sword; Divide by 3 – a barge with a triangular shield; Divide by 5 – a punch with a (five-fingered) gauntlet, Divide by 10 – a single swipe of a sword and a punch with a gauntlet, etc). Thus, the structure of these attacks embodies additional mathematical relationships in the way that weapons combine. In this way the learning content is integrated within the core mechanic of the game-play. The game also includes secondary game mechanics that revolve around exploration and collection (exploring a non-linear 3D dungeon and collecting keys), but these are not integrated with the mathematical content of the game. An arbitrary fantasy context for the game is provided (linking the numbers on skeletons' chests to cursed Olympic athletes), but we argue that this fantasy is extrinsic and could very easily be replaced with an entirely different fantasy context (e.g. evil robot enemies with monitors on their chests) without changing the intrinsic relationship of the learning content with the game's core-mechanic. You could also replace the secondary game mechanics that revolve around exploration and collection with others (e.g. tower-defense: defending a static object from attack) provided the combat mechanic remained the same.

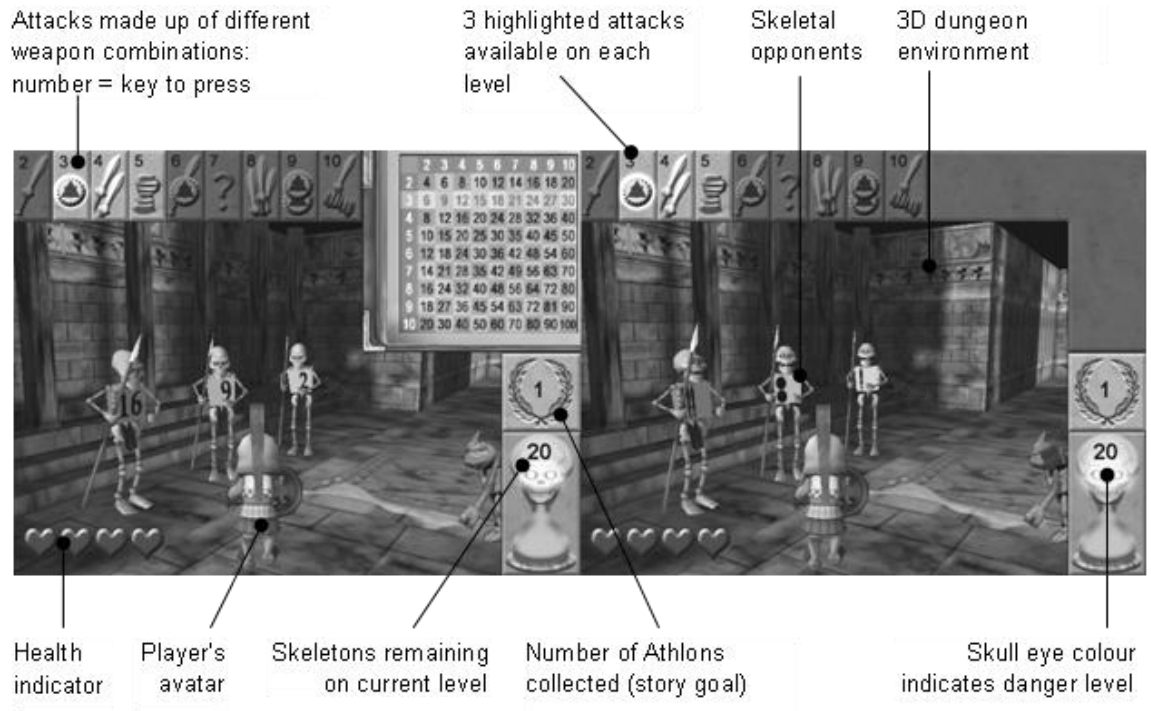


Figure 1. Screenshots of an identical gameplay situation in *Zombie Division* (intrinsic (left) and extrinsic (right) versions of the game. All labeled features are common to both versions).

Each game level contains about twenty enemies (zombie skeletons). When divided with an appropriate attack, the ‘spirit’ of a defeated skeleton rises and then splits into equal sized portions (depending on the divisor), which grow into small ghosts bearing the quotient (these normally disperse but see below). Using an inappropriate attack against an enemy results in the skeleton fighting back and the player losing health, which when exhausted forces the player to start the level again. In this way the player is not just asked to choose between three divisors for each opponent, but must consider whether opponents are dividable at all using their current attacks.

Two additional mechanics help children develop their mathematical competency. As players progress through a game, skeletons gain weapons with which to parry attacks. For example, an opponent with a sword could parry attacks that divide them by 2, 4, 6, 8 and 10. In these cases the player would be forced to consider other (potentially less obvious) divisible factors of the opponent’s dividend in order to defeat them. This mechanism, for example, stops children dividing all even numbers by two. The second mechanic is the inclusion of giant skeletons which can only be defeated outright by dividing by a large number such that the quotients of the resultant parts are smaller than 10. Quotients larger than 10 rise again as new skeletons in their own right and continue to attack the player with increased ferocity. Again this is included to encourage children to use a fuller repertoire of divisors.

Only three different attacks (divisors) are available to the player at any point in the game in order to ensure that there are always some enemies that cannot be defeated with the available choices. They begin with the divisors two, five and ten and gradually gain access to different attacks as the game progresses. The skeletons themselves gain more ferocity which encourages faster responses from children (in the initial levels skeletons stand around rather passively, but as levels increase they move around, block areas of the dungeon or pursue Matrices). More docile skeletons are generally found in levels with new divisors and the hostility of the skeletons on each level is indicated through the color of the skeleton's eyes on the player's user interface (ranging from green to glowing red).

Children received two main sources of in game help; Gargle (an animated pedagogical agent) and a magical book of times tables (the multiplication grid). Gargle accompanies Matrices throughout the game providing just in time instructions and help. He is first seen in an initial training level where the players learn to navigate and fight passive 'clockwork' skeletons before meeting them for the first time in the game. He also helps children understand how the multiplication grid can be used to help decide whether a number is divisible by a particular divisor (which is provided in a small tutoring component of the game based around sharing bones). As the game progresses, Gargle can provide oral instructions as to how to play the game, including task direction, but is typically silent unless children are experiencing problems.

The game was created specifically for this study (and different versions created to test our approach to intrinsic integration, see below). The design and development of *Zombie Division* was undertaken as an iterative process with regular input from both girls and boys from the target audience. The instigation of the project followed an intensive period working with children of the same age, teaching them how to make their own educational games (Habgood, Ainsworth, & Benford, 2005b). The initial design for *Zombie Division* was created as a cardboard prototype and piloted with boys and girls from a range of different mathematical abilities. Despite our own initial reservations about the gender-neutrality of the fantasy context, it appeared to have cross-gender appeal for the target age group when presented within a school setting. It is also worth noting that the game has a number of similar game-mechanics to the "Zelda" series of games which have become notable as a franchise which appeals to both male and female audiences. Regular piloting also allowed us to address gameplay issues (such as navigation and combat systems) which may otherwise have favored success by the more game-literate children (usually boys).

Zombie Division was designed to be consistent with a 7+ rating under the PEGI European game age-rating system (e.g. includes occasional violence to non-realistic fantasy characters). Its theatrical title seemed to provide instant appeal to its target audience, but the game itself did not contain any of the

gruesome or gory content that the word zombie might evoke. The skeletons were rather comical and children were more frightened of the doors (which slammed behind them in a spooky way) than the skeletons. Nonetheless, even the youngest children enjoyed the creepy atmosphere, and it seemed to add to their engagement with the game rather than putting them off.

Study 1: The impact of intrinsic integration on learning

This study was designed to test the effectiveness of intrinsic integration in educational games as used within traditional classroom settings by assessing learning outcomes. Three versions of the game were created which differed only in how the math content was delivered (intrinsic, via in game action; extrinsic, end of level quiz or control, none at all). The classroom setting imposed a number of additional constraints on the study. The overall time children spent learning was fixed in terms of the time per session (15-20 minutes) and strictly controlled in terms of the total playing time (2¼ hours). This amount of time was chosen to be realistic within the framework of the UK Numeracy Strategy which would typically devote this amount of time to teaching multiplication and division every day for a week and then revisit the topic frequently under a cyclic curriculum. Finally, we also wanted their teachers to play an active role in supporting children's learning with the game.

Post-play debriefing is considered by many to be critical to the effective application of educational simulations and games (e.g. Squire, 2004; Garris, Ahlers & Driskell, 2002; Lederman, 1992). Sanford & Williamson (2005) go further to suggest that "*The outcomes of any lesson-based computer activity will depend on the introduction of the task, the interventions made during the activity and the way that the activity is set in the context of students' wider educational experience*" (p.11). Even outside of games, the role of the teacher is often seen as an important facet of computer-based learning (e.g. Tabak, 1997). Therefore the way in which Zombie Division is framed and supported within the children's wider educational context is likely to have a large effect on the learning outcomes of the game. The potential learning gains in both intrinsic and extrinsic conditions would have almost certainly been maximized by the addition of supporting activities pre, post and in-parallel with the gaming interventions. However, the purpose of this study was not to maximize overall learning gains, but compare learning gains between intrinsic and extrinsic approaches over limited amount of time on task. We did not want to bias the learning gains produced by just over two hours of gaming with another two hours of teaching.

We therefore decided to include a controlled teacher-led reflection session delivered after the children had had a chance to become familiar with the game, but before they had got too far. This would aim to get children to reflect on the mathematical context of the game (in the intrinsic and extrinsic versions) and scaffold the conceptual process of making the link between solving division problems using multiplication facts. The reflection session would be structured to include identical mathematical content

in all three conditions, but made relevant to the context of each version of the game (except for the control group where the learning content has no relevance to their version of the game and so would be taught in the abstract case). We believed this would provide the right balance between the requirement for some form of debriefing to make effective use of the game and the need not to make this a study about teaching mathematics around the instructional anchor (Brandsford et al, 1990) of a videogame.

Method

Participants

All children attended a primary school in a low-income area on the outskirts of a large city in the north of England. The attainment of the school's intake was below the national average and the percentage of final year students achieving expected levels in mathematics was below national averages for the preceding year. The 30 girls and 28 boys were aged between 7 years 1 month and 8 years 10 months (mean 8 years 0 months). All participants had prior experience of using the computers in the school's ICT suite.

Design

This study used a two-factor [3 by 3] mixed design. The first factor, 'game' was between-groups with three levels (intrinsic, extrinsic or control) which determined which version of the game children played. The second factor, 'time' was within-groups and also had three levels (pre, post and delayed-tests). The fifty-eight children were randomly assigned to one of the three conditions such that pre-test scores, gender and age within each condition did not differ. Consequently, 20 children were assigned to the intrinsic, 20 to extrinsic and 18 to the control condition.

Materials

Facilities

Children sat tests and played with *Zombie Division* within the school's ICT suite which contained twenty PCs running Windows 2000 with accelerated 3D graphics support, and audio output through stereo headphones. The teacher-led reflection sessions were carried out in a standard classroom, using an interactive whiteboard running PowerPoint to present the teaching material.

Game Versions

Three versions of the game were created for this study. The base version was the intrinsic version (as described above) and the extrinsic and control versions were based upon it. The key practical and theoretical difference was that dividing skeletons no longer involved any mathematical content. In the extrinsic version the mathematical content was provided as an end-of-level quiz and in the control version

it was excluded all together. In the intrinsic version, the dividend displayed on each skeleton's chest provided the player with a way of determining its vulnerability to different attacks. The same result was achieved in the extrinsic and control versions by replacing the dividend with a symbolic representation of which attacks can divide that skeleton (e.g. displayed by combinations of sword, gauntlets and shield). Thus exactly the same dividend and divisors were present in the extrinsic version but the mathematical relationship is hidden because the numbers were no longer displayed. For example, the number 16 can be divided by 8, 4 and 2, but the symbol for a divisor of 8 (three swords) naturally includes symbols for a divisor of four (two swords) and a divisor of two (one sword) as well. This had the additional bonus of making the symbols require a level of logical interpretation, keeping the challenge of defeating skeletons at a more comparable level to the intrinsic version. It also meant that for dividends within the range of 1-99 divided by divisors in the range of 1-10, only the numbers 60 and 90 need to be represented by more than three symbols.

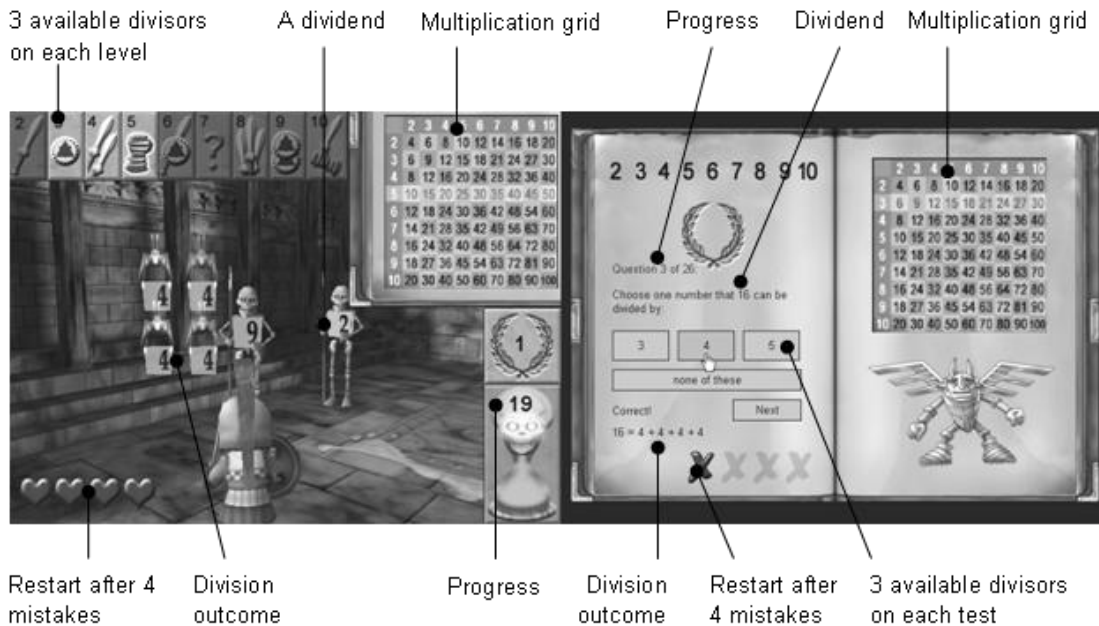


Figure 2. Explanations of mathematical learning content in the intrinsic (left) and extrinsic (right) versions of *Zombie Division*. Comparable features are labeled.

In the extrinsic version, the mathematical content was now reintroduced at the end of each level in the form of a multiple-choice quiz. This quiz required the player to divide the same dividends as found on the skeletons in the intrinsic game, using exactly the same choice of divisors (weapons) that were available to defeat those skeletons (including “none of these” to be equivalent to leaving a skeleton). The extrinsic version therefore provided identical learning content delivered away from the flow-inducing game-play, and presented as abstract mathematical questions (see Figure 2). The control version simply

omits the mathematical end of level quiz. A comparison of the different versions of the game is presented in Table 1 (stressing the gameplay) and Table 2 (highlighting the mathematical content).

Table 1. Comparison of the key gameplay differences between the intrinsic extrinsic and control versions of the game

Feature	Intrinsic	Extrinsic	Control
Skeleton	Numbers on chests	Weapon symbols on chests	As Ext.
Attacks	Number relates to divisor & function key	Number relates to function key	As Ext.
Successful attack	Skeleton splits into a proportional number of ghosts bearing the quotient	No ghosts.	As Ext.
New weapon tutorial	Introduced to the numbers that they divide	Introduced to the symbols that they divide	As Ext.

Table 2. Comparison of the key mathematical features of the intrinsic, extrinsic and control versions of the game

Feature	Intrinsic	Extrinsic	Control
Division Context	Dividing skeletons	Multiple-choice questions at end of level	None
Division Control	Function keys	Controlled using mouse	None
Multiplication grid	Appears during gameplay	Appears with questions	None
Feedback on Incorrect choice	Player knocked back and loses health. Sent to restart position if health reaches zero.	Told answer is incorrect and asked to choose again. Sent back to start of test if chances run out.	None
Feedback on Correct choice	Splits into a proportional number of ghosts bearing the quotient	Told answer is correct and dividend, divisor and quotient are displayed	None

Test Materials

Outcome test

The time-limited computer-based test consisted of 63 multiple choice questions with four options in each case (one correct + three distractors). The first three questions were interface practice questions (e.g. "Select the number of legs that a dog has: 4, 5, 6 or 7"). Of the remaining 60 questions, 40 were division questions equally comprised of two formats: a) Dividend-based, where the child was asked to select the divisor that divides a given dividend (e.g. "Select one number that 45 can be divided by: 4, 6, 9 or none of these"); and b) Divisor-based, where the task was to select the dividend that can be divided by a given divisor (e.g. "Select one number that can be divided by 5: 35, 13, 29 or 41"). There were five questions on each divisor from 2 to 10 (excluding the divisor 7 as it was not included as part of the games' learning content). In addition five dividend-based questions were included where the answer was 'none of these'. The remaining 15 questions were more conceptual in nature: three tested knowledge of the heuristic patterns associated with numbers that divide by 2, 5 and 10; and twelve tested an understanding of relationships between divisors (e.g. that all numbers which divide by 9 can also be divided by 3) or for applying rules outside of normal limits (i.e. dividends greater than 100).

The order of the questions was initially randomized, but remained consistent between subjects and between tests. The software timed 15 minutes from the start of Question 4 (the end of the practice questions) and automatically stopped the test at the end of this time period. However, the time was not displayed on the screen and no feedback was provided on the choices made. A multiplication grid was provided in the corner of the screen similar to the one found in the game

Challenge Level

Given the concerns raised about the potential for intrinsic integration to reduce transfer, we created 'challenge' levels to allow us to compare questions in the abstract pre/post tests and the same questions contextualised within the challenge levels of the game. Thus, the challenge level acted as a game-based test and consisted of two specially constructed levels of the game that directly replicated a portion of the outcome test's division problems within the game environment. All three groups played the gaming elements of these 'challenge levels' with the learning content embedded (or omitted) appropriately for their group's condition.

In the extrinsic version these questions were asked in the normal way at the end of each level (e.g. "Select one number that can be divided by 5: 35, 13, 29 or 41"). In the intrinsic version, each question was posed within the context of a separate room within the challenge levels. The weapons (and therefore divisors) available to the player changed to match each question as they entered its associated room.

Divisor-based questions were posed in terms of offering the player a choice of three weapons with which to divide a single skeleton. An exit to the room was also included to provide an option equivalent to 'none of these'. A correct answer was only recorded if the skeleton was defeated with the correct attack on the player's first attempt. Dividend-based questions were posed in terms of a choice of four skeletons to divide with a single weapon and again a correct answer was only recorded if the right skeleton was chosen on the player's first attempt. In addition, the gameplay demands were reduced within the challenge levels as the dungeons were linear, the skeletons were immobile and keys were provided when required.

Teacher Led Reflection

The reflection sessions were included to help children to reflect on the mathematical content of the game and scaffold the conceptual process of making the link between solving division problems using multiplication facts. Children were taught in three separate groups according to their experimental condition. All the teaching materials were created by the researchers and tailored to the context of each group's game, but contained identical learning content (including the numerical examples) and followed the same structure. Each session lasted for half an hour consisting of 15 mins of direct instruction followed by ten minutes of collaborative exercises. The instruction addressed three issues:

- Division as sharing: Children were shown a number of objects (bones or balls) and asked how they would work out if they could be divided into two equal-sized sets. A volunteer was then asked to come and draw circles around the sets. The class then confirmed this by counting the number of objects in each set.
- Tables and rules: The class was asked to suggest other techniques they could use to work out whether a number of objects can be divided equally into a number of sets. This continued until the class offered 'using times-tables' as a solution or the teacher eventually intervened with this suggestion. The class was also reminded of the numeric patterns for the 2, 5 and 10 times table, if they had not already been discussed.
- The multiplication grid: The class was asked how they could solve division problems for times tables they didn't know, without counting objects. They were presented with the multiplication grid and shown how it can be used to answer division problems. They then worked through four example questions, checking if a specific dividend could be divided by a specific divisor.

This was followed by 10 minutes of exercises carried out in pairs or groups of three. This contained twelve divisor-based division problems with an option of three divisors to divide a given

dividend. A multiplication grid was available. Worksheets also contained three blank questions for the children to create their own questions for their partners at the end. During this period the teacher provided individual support to any child that needed it.

Procedure

Figure 3 shows the schematic for the study. In total children spent four hours in the study which was spread over 34 days.



Figure 3. Schematic of Study.

Stage 1. Pre-test

The pre-tests were carried out in three half-hour sessions ten days before the main body of the study. Groups of up to twenty children were selected at random to complete the 15 minute timed pre-test in the ICT suite. The task was explained to children with the aid of a demo which emphasized the presence of the multiplication grid to help them with the test. They were informed of the 15-minute time limit, but told that they were not expected to finish all of the questions and encouraged not to treat it as a race. Each child was then allocated a PC and allowed to begin the test in their own time. Children that finished before the end of their time limit were asked to sit quietly until the entire group had finished.

Stage 2. The Game

The children first played *Zombie Division* ten days after the pre-test. Each group (intrinsic, extrinsic and control) played their version of the software without children from the other group present. Each playing session lasted for approximately twenty minutes, with a half hour turnaround on successive groups. The order of groups was rotated on each day of the study, with the first group beginning at 10:00 and the last group finishing at 11:30. Each child's position in the game was saved at the end of each playing session and the game resumed from precisely the same point at the start of the next one. Each child played the game twice in this stage.

Stage 3. The Teacher Led Reflection

All the reflective sessions were delivered by the same practicing teacher who taught older children within the school. None of the children had been formally taught by this teacher before, although some level of familiarity through every day school life can be assumed. These took place immediately preceding the children's third game session when all groups had played the game for an average of forty minutes.

All sessions were observed by the researcher and there was no diversion from the teaching materials provided.

Stage 4. Further Game Sessions

The children played the game on two more days until they had accumulated a total of one hundred minutes playing time. At this point the software automatically stopped the game and the child was sent back to their class. A number of catch-up sessions were run for absentees to ensure that all children had played for their allotted time before taking the post-test.

Stage 5. Post-Test

The post-tests were carried out on the day after the children had completed one hundred minutes of playing time with the game. Children were divided into three new groups containing an equal number of children from each condition. These groups were tested in three consecutive sessions in an identical way to the pre-test but with the addition of the challenge level. In order to prevent any distraction, children were not allowed to begin the game-based test until the entire group had finished their outcome tests.

Stage 6. Final Game Session

Two weeks after the post-test, children had a final opportunity to play their version of the game. This brought their total playing time up to exactly 135 minutes.

Stage 7. Delayed Test

The delayed-tests were carried out on the same day as the final playing session. The children were divided into mixed condition groups and tested in three consecutive sessions as before. The challenge levels were taken in the same way, two days later.

Results

Of the 54 children who completed all stages of the study, three have been excluded from the analysis. One child was identified as having special educational needs in mathematics and two demonstrated significantly better mathematical knowledge than the other children before the study started as they scored 81.5% and 71.8% at pre-test (3.6 and 2.9 SDs above the mean respectively) Analyses were conducted on the remaining 51 children's data to explore the impact of the game condition on both the process and outcomes of learning.

Learning Outcomes

Learning was measured by examining the percentage of correct answers that children gave on the tests (correct answers / total answers x 100). As the tests were timed, this measure was chosen to make sure

that strategies that may take longer to perform but are more accurate as a result (such as using the multiplication grid) were not penalized.

A two-way mixed measures ANOVA with three levels of the within-subjects factor ‘time’ (pre, post and delayed test) and three levels of the between-subjects factor ‘game’ (intrinsic, extrinsic and control) was performed (see Figure 4). This revealed a significant main effect of time ($F(2,96) = 34.86$, $MSE = 4006.41$, $p < 0.001$, $\eta^2 = 0.42$). Post-hoc comparisons using the Bonferroni adjustment for multiple comparisons showed students’ post-test scores were higher than pre-test scores (11.69, $p < 0.001$) as were delayed tests (14.46, $p < 0.001$) but there were no significant difference between the post-test and delayed test.

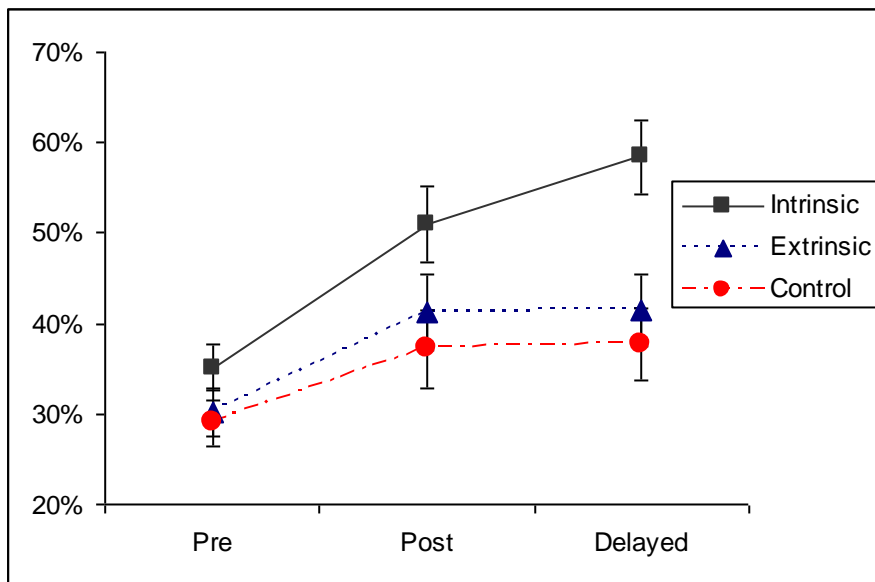


Figure 4. Mean Percentage Score by Time and Game Condition.

Analysis also revealed a time by game interaction ($F(4,96) = 5.86$, $MSE = 1025.794$, $p < 0.006$, $\eta^2 = 0.11$). Simple main effects analysis showed that all groups improved over the tests (intrinsic ($F(2,47) = 24.89$, $p < 0.001$, $\eta^2 = 0.51$); extrinsic ($F(2,47) = 6.78$, $p < 0.003$, $\eta^2 = 0.22$) and control ($F(2,47) = 3.97$, $p < 0.025$, $\eta^2 = 0.15$)). Multiple comparisons (Bonferroni corrected) showed that children who had played the intrinsic game scored significantly more at post-test than they had at pre-test (mean 15.91, $p < 0.001$) and gained still more from post-test to delayed test (mean 7.04, $p < 0.04$). The extrinsic group improved from pre to post-test (mean 10.97, $p < 0.006$) but made no further improvement, whereas for the control group the delayed score test was significantly higher than their pre-test (mean 8.78, $p < 0.03$). The only test which showed any differences between the groups was the delayed test ($F(2,48) = 7.49$, $p < 0.001$, $\eta^2 = 0.24$). Post-hoc comparisons using the Bonferroni adjustment for multiple

comparisons showed that children who had played the intrinsic game scored significantly higher than children in either the extrinsic group (mean 16.94, $p < 0.14$) or the control (mean 20.66, $p < 0.002$). In summary, children in all conditions learned from the experience (game plus teacher led reflection) but children in the intrinsic condition learned the most.

This analysis was repeated with the additional between group-factor of gender. There were no significant differences between boys and girls ($F(1,45) = 1.16$) and no interaction between gender and any other factor including game condition.

Time to answer questions

As there were differences between conditions in accuracy of response to tests it is also informative to consider if there were differences in time taken to answer questions (see Table 3). This measure was examined using a [3 by 3] mixed ANOVA with three levels of the within-subjects factor ‘time’, and three levels of the factor ‘group’ (intrinsic, extrinsic and control). This revealed a significant main effect of time ($F(2,96) = 5.09$, $MSE = 145.3$, $p < 0.008$, $\eta^2 = 0.10$). Pairwise comparisons (Bonferroni corrected) showed that children took longer to answer the questions at the delayed test than pre-test (mean 2.95, $p < 0.02$). A main effect of group ($F(2,48) = 3.67$, $MSE = 423.6$, $p < 0.04$, $\eta^2 = 0.13$) found that children in the control condition took longer to answer questions than children in either of the other two conditions (intrinsic mean 1.19, $p < 0.06$, extrinsic mean 2.95, $p < 0.02$). However, overall children who took more time to answer questions were more accurate (pre-test ($r = .35$, $p < 0.02$), post-test ($r = .5$, $p < 0.001$) and delayed test (.41, $p < 0.003$)).

Table 3. Time per question in seconds by condition and time table

	Intrinsic	Extrinsic	Control
Time	(n = 17)	(n = 17)	(n = 17)
Pre-test	16.8 (7.2)	12.7 (5.3)	14.1 (4.6)
Post-test	18.0 (9.8)	12.5 (5.6)	13.0 (5.7)
Delayed-test	21.6 (10.8)	15.6 (10.8)	15.1 (4.7)

Challenge Levels

The challenge levels provide a direct comparison between questions in the abstract pre/post tests and the same questions contextualized within the context of the game as a means of exploring transfer. Twenty of

the questions from the assessment were repeated (in appropriate format) in the challenge levels of the intrinsic and extrinsic games (there is no math content in the control condition). Children played these levels on two separate occasions: once following the post-tests and once following the delayed tests. Table 4 shows the mean percentage scores for the challenge levels and for the equivalent 20 items in the outcome tests (note one child in the intrinsic condition missed the test).

Table 4. Outcomes scores by environment, time and game condition

	Intrinsic (n = 16)		Extrinsic (n = 17)	
	Challenge	Outcome	Challenge	Outcome
Post	59.1% (23.0)	62.1% (18.8)	45.3 % (16.2)	47.9% (16.4)
Delayed	75.6% (16.2)	57.9% (20.0)	57.9% (20.8)	49.0% (10.8)

Analysis by 2 (post, delayed) by 2 (challenge, outcome) by 2 (intrinsic, extrinsic) ANOVA showed a main effect of time ($F(1,31) = 32.3, MSE = 2598.5, p < 0.001, \eta^2 = 0.51$) with children scoring higher on the delayed test and a main effect of condition ($F(1,31) = 6.41, MSE = 8400.9, p < 0.02, \eta^2 = 0.17$) with children in the intrinsic condition scoring higher than children in the extrinsic condition. Overall children's performance on the different types of test did not differ but there was a test by time interaction ($F(1,31) = 7.92, MSE = 1081.4, p < 0.008, \eta^2 = 0.20$). There were no differences between the challenge and outcome test at the post-test ($F(1,31) = 0.39$) but children scored higher on the challenge level at delayed test ($F(1,31) = 9.15, p < 0.005$). Children also significantly increased their scores from post to delayed test but only on the challenge level ($F(1,31) = 32.3, p < 0.001$). There was no test type by condition interaction ($F(1,31) = 0.1$) so children in the extrinsic condition whose tests had similar formats to the quiz section of their game did not perform better than children in the intrinsic condition whose tests were dissimilar.

Game Performance

The process logs produced by the game provide a valuable source of data in the form of a time-stamped commentary on the game as it is being played. For this study, over two and a half thousand log files from the extrinsic and intrinsic condition were mined for the purposes of post-hoc analysis.

Before performance on the intrinsic and extrinsic versions can be compared, it is necessary to operationalise measures of performance which are as equivalent as possible. Some variables are the same

in both games (e.g. level refers to the amount of the game explored), but others vary depending upon the game version. A math task is always presented in the form of a skeleton in the intrinsic version, and a quiz question in the extrinsic version. In both cases a math task is a dividend-based question with a choice of up to three divisors to divide a given dividend. These divisors are provided in the form of different weapons in the intrinsic version, and multiple-choice answers in the extrinsic version. The player also has the option of rejecting all the divisors provided if none of them would divide the dividend. In the intrinsic version this involves maneuvering to avoid combat with the skeleton, while in the extrinsic version a player selects an alternative answer marked 'none of these'. Furthermore, for the purposes of this analysis a math task is assessed in terms of the first attempt made upon the dividend. This is said to begin when the player enters the same room as the skeleton and end if the player attacks the skeleton or leaves the room again without dividing it. This means there can be one of five outcomes depending on whether the dividend is a target or distracter:

Correct Outcomes

1. Target Answered Correctly (TAC) – the player correctly divides the dividend by one of the available divisors on the first attempt (extrinsic chooses the right number for math quiz item/intrinsic chooses correct weapon for skeleton).
2. Distracter Left Correctly (DLC) – the player correctly rejects all of the available dividends for an indivisible dividend on the first attempt (extrinsic chooses "none of these" /intrinsic does not fight skeleton).

Incorrect Outcomes

3. Target Answered Incorrectly (TAIN) – the player incorrectly attempts to divide the dividend by one of the available divisors on the first attempt (extrinsic chooses incorrect number for math quiz item/ intrinsic chooses incorrect weapon for skeleton).
4. Target Left Incorrectly (TLIN) – the player rejects all of the available divisors for a divisible dividend on the first attempt (extrinsic chooses "none of these" when a correct response exists /intrinsic chooses not to fight a skeleton who could be defeated).
5. Distracter Answered Incorrectly (DAIN) – incorrectly attempts to divide an indivisible dividend by one of the available divisors on the first attempt (extrinsic chooses a number for math quiz item when correct response is "none of these"/intrinsic chooses to fight a skeleton who is not defeatable).

A total of 2.46% of the data were not able to be analyzed due to errors such as the software

crashing, an error in data-logging or restarting level.

Table 5. Game measures by condition

	Intrinsic (n = 17)	Extrinsic (n = 17)
Level Reached	7.59 (3.59)	6.88 (3.24)
Unique Maths Tasks	156.8 (84.5)	120.9 (77.0)
Attempts at all Maths tasks	382.2(180.5)	198.2 (75.0)
Accuracy	70.1% (7.9)	79.1% (10.5)

Analysis by one way MANOVA on the variables shown in Table 5 revealed a multivariate effect of condition ($F(4,29)=7.12$, $p<0.001$, $\eta^2 = 0.50$). Univariate analyses showed that children reached the same level regardless of condition ($F(1,32) = 0.36$) and performed the same number of unique math tasks (quiz question or skeleton dividing) ($F(1,32) = 1.00$). However, children in the extrinsic condition performed significantly less math tasks overall ($F(1,32) = 15.07$, $MSE = 23796$ $p< 0.001$, $\eta^2 = 0.32$) and were more accurate ($F(1,32) = 7.86$, $MSE = 683.65$ $p< 0.01$, $\eta^2 = 0.20$).

It might be expected that children's performance in the game would relate both to their initial mathematical understanding and what they learned from their experiences. Although children's pre-test scores were significantly related to all game measures except accuracy, there was no relation between game performance and what children learned (Table 6)

Table 6. Correlations between indices of game performance, pre-test and learning outcome

	2	3	4	5	6
(1) Level Reached	.99***	.49**	.41*	.43*	.18
(2) Unique Maths Tasks		.55**	.38*	.44*	.20
(3) All Maths Tasks			-.23	.37*	.25
(4) Accuracy				.06	-.14
(5) Pre-test					-.01
(6) Gain					

Note. * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$ (two-tailed test of significance).

Analysis of the result of a math task (percentage of a particular result given the total number of unique math tasks) is displayed in Table 7. This analysis helps reveal which particular aspects of mathematic and/or game-play children found difficult in the different conditions.

Table 7. Results of math task by condition

	Intrinsic		Extrinsic	
	(n = 17)		(n = 17)	
	M	SD	M	SD
Target Answered Correctly	56.85%	7.13	66.65%	10.98
Distracter Left Correctly	13.27%	3.00	12.86%	3.39
Target Answered Incorrectly	4.75%	2.81	3.07%	2.19
Target Left Incorrectly	20.56%	6.40	7.78%	5.92
Distracter Answered Incorrectly	4.24%	2.48	5.70%	2.95

Analysis by MANOVA on the variables shown in Table 7 showed a multivariate effect of condition ($F(5,28)=22.83$, $p<0.001$, $\eta^2 = 0.80$). Univariate analysis showed that the children in the extrinsic condition were significantly better at answering a target divisor question which had a valid dividend ($F(1,32) = 9.53$, $MSE = 0.82$, $p< 0.004$, $\eta^2 = 0.24$) but not at correctly leaving a distracter question which did not have a valid dividend ($F(1,32) = .57$). In terms of incorrect responses, children in the extrinsic condition were much less likely to leave a target incorrectly ($F(1,32) = 36.57$, $MSE = .14$, $p< 0.001$, $\eta^2 = 0.53$).

Study 2: The impact of intrinsic integration on choice of game

Study one showed that the intrinsic version of *Zombie Division* produced greater learning gains than either an extrinsic version or a control version without learning content. This study was conducted under a strict *in vivo* experimental regime whereby time on task was completely controlled. However, whilst educational games are used in schools they are also increasingly available as leisure activities. In these situations, it is important that intrinsic games are not just effective but that children choose to use them. It could be that the central benefit of intrinsic integration will be to produce greater time-on-task.

Consequently, this study compared the time children chose to spend playing the different versions of the

game when they were provided with a free choice between them.

There is a long history of measuring motivation as a function of time-on-task, including some of the early seminal research on intrinsic motivation (Daniel & Esser, 1980; Deci, 1971). Furthermore, task persistence is considered one of the key mechanisms by which motivation can influence learning (e.g. Dweck, 1986; Vollmeyer & Rheinberg, 2000). For example, Vollmeyer & Rheinberg, (2000) found that learners with high motivation showed increased task persistence and that when this was combined with initially lower levels of knowledge resulted in increased learning. Therefore time-on-task was chosen as the primary dependant variable with which to compare the relative motivational appeal of intrinsic and extrinsic approaches for *Zombie Division*.

Method

Participants

All children attended the same primary school as in study one, but were taken from a different (older) year group. The 5 girls and 11 boys were between the ages of 9 years 10 months and 11 years 2 months (mean 10 years 4 months). All participants had prior experience of using the computers in the school's ICT suite and were members of an after-school computer club.

Design

This study used a single repeated measure design, game, with two levels (intrinsic, or extrinsic) which reflected the amount of time children played each game.

Materials

Facilities

The intervention was carried out within the ICT suite at the school, using the normal facilities used for the after-school club. The suite contained twenty, relatively new PCs running Windows 2000 with accelerated 3D graphics support, and audio output through stereo headphones. Children could switch between different versions of the game using a menu that appeared each time the game was launched. This allowed children to choose between the intrinsic and extrinsic versions of the game, and provided them with a visual reminder of the differences between the two versions. The order that the options appeared in the menu was randomized each time so that either version would appear on the left or right with an equal probability. Quitting the game would return the player to this menu, where they could switch versions again. When switching versions their exact position was resumed with intrinsic skeletons becoming extrinsic or visa versa. In this way it was ensured that switching between versions neither

provided a gameplay penalty, nor an advantage, so children could not use it as a way to ‘game the system’ (e.g. Baker et al., 2004).

Procedure

Introduction

Children were introduced to *Zombie Division* as a group by demonstrating the two different versions running side by side on two separate PCs. Both games were saved at identical positions within the same game level so that the differences between intrinsic and extrinsic versions were apparent. Children were shown how combat worked in both versions, emphasizing the mathematical content of the intrinsic version, alongside the quiz that appeared at the end of each level in the extrinsic version. They were introduced to the ‘intrinsic’ and ‘extrinsic’ terminology and shown how the version switching menu worked. Children were told that their game position would not be lost by switching versions and that they were expected to try both versions. This introduction took ten minutes.

Game Intervention

All the children played *Zombie Division* on their own PCs for the remainder of the first club session. In subsequent sessions each child could choose to continue playing the game or return to their normal club pursuits (and freely switch between the two). Each club session lasted for approximately one hour, with around 45-50 minutes of playing time. This continued for two more club sessions after the first, providing a maximum of around 135 minutes (2¼ hours) playing time for each child. The children’s positions in the game were saved at the end of each playing session and the game resumed from precisely the same point at the start of the next session. The group was reminded several times over the course of the sessions that they were expected to try playing both versions of the game.

Group Interview

In the fourth and final club session children took part in a group interview with the two different versions running side by side. They were asked to summarize the differences between the two versions and which they preferred. Each child was given the opportunity to explain why they preferred the option they did, and the group was encouraged to discuss which version was the most fun to play and which was the most educational.

Results

The mean number of minutes children spent playing the two different versions of the game was analyzed using a paired sample t-test. Participants spent over seven times longer playing the intrinsic version at 75.7 minutes (sd = 35.5) compared to the extrinsic (10.28 minutes; sd = 10.28) ($t = 7.38$, $df = 15$, $p < 0.001$,

$r = .89$). Analysis of the influence of gender showed that overall girls (median 114) choose to spend longer playing Zombie Division than boys (median 73) ($U=10$, $p<0.052$, $r = .5$), however they did not differ in the proportion of time they spent on the intrinsic version ($U=23$). Table 8 shows the results of the group interview concerning children's responses to different versions of the game. Children not only had more positive perceptions of the intrinsic version of the game, but they also had quite a sophisticated appreciation of the mechanisms of intrinsic and extrinsic integration too. These data suggest that children's decision to spend more time playing the intrinsic version was clearly deliberate and resulted from their positive perceptions of the intrinsic game mechanics.

Table 8. Children's responses to the different game versions

	Intrinsic	Extrinsic
Likes	it's not as hard – it's quick and easy" "it's easier to learn division [...] instead of having to figure out what the symbols are you just have to figure out what to divide by" "it's easier [...] because you get to learn division" "it's better to learn doing it by intrinsic, because it's quicker" "it's easier to learn your times tables" "it's fun" "you don't have to do a test at the end" "more fun because it's like subliminal advertising with maths" "it's like mixing paint [...] the maths in the game with the fun [...] you don't really think you're doing that much"	"it can help you like learning your times tables and doing your SATs"
Dislikes	"it's not faster because on the beginning of every level there's [...] a help thing" [teachers would think it's] "too much fun – and hasn't got a test"	"because you need to do all the maths at the end and that's what you might lose interest in" "the version with the math test at the end wasted time [...] you don't get as far as you do in the other one" "it just tells you what to use" [fighting skeletons] "it's not a challenge" [fighting skeletons] "the maths test at the end was just got boring." "very slow and boring" "you think: oh I've had the fun part, now I have to do a test – I'm just going to turn it off and not bother"

Discussion

Learning Outcomes

The results of these studies provide a strong argument in favor of the intrinsic integration of a game with its learning content (Kafai, 1996; Malone, 1980; Rieber, 1996), in contrast to extrinsic environments

which provide a separate extrinsic motivation or reward for completing learning content. At the beginning of Study 1, children had limited familiarity with the mathematics addressed in *Zombie Division* as they scored an average of 31% at pre-test. Over the course of the study children in all conditions improved their understanding to end with a delayed test score of 46%. However, although all children learned, children in the intrinsic condition improved the most. At delayed test, these children (with a mean score of 58%) significantly outperformed children in both control (38%) and extrinsic (41%) conditions (Figure 4). This improvement was the same for both boys (12.84%) and girls (16.25%).

The intrinsic result might also initially seem surprising given that the children's performance during the game had not favored the intrinsic condition. There was no difference between conditions in the level reached or the number of unique math tasks but children in the extrinsic condition were more accurate (the measure most closely related to the tests). We must acknowledge that may be partly an artifact of the way that accuracy is operationalised. In this extrinsic condition, a child must make a conscious decision to choose "none of these" when a correct response exists, but in the intrinsic condition it is impossible to know if a child chose not to fight a skeleton, or just exited the room without seeing it. Thus, our data mining used a 'best guess' heuristic, which is likely to overestimate the number of conscious rejections of skeletons. According to this heuristic, around 21% of the responses children made in intrinsic condition were of this form compared to only 8% in the extrinsic condition.

However, this over-estimation of rejections is unlikely to be the sole explanation for the difference between the intrinsic and extrinsic conditions. When performing maths tasks in the extrinsic condition, children only need to manage the demands of the division problems. In contrast, during the intrinsic game they must navigate dungeons, seek out keys, and respond to increasingly assertive skeletal attacks as the levels advance. It is therefore unsurprising that children in the extrinsic condition (67%) were more accurate at answering a target question than children in the intrinsic condition (57%). Arguably, children in the intrinsic condition may even have better mathematical skills within the game as well but that these are being masked by the difficulties of the gaming aspects of *Zombie Division*. Evidence for this proposal can be seen from the challenge level data. In these levels, the mathematics remains the same but the game play demands are significantly reduced in the intrinsic condition (skeletons don't attack; less need to navigate, etc). Children in the intrinsic condition scored higher than those in the extrinsic condition on these levels at both post and delayed test (scoring an impressive 76% in the delayed challenge levels).

One of the reservations held about the use of intrinsic games is that they will encourage the development of overspecialized knowledge that does not transfer to everyday school mathematics, but we found no evidence to support this concern. The learning outcomes tests presented the math problems in an

abstract quiz form which was much closer to the form of the extrinsic condition and yet results still favored the intrinsic condition. We also developed challenge levels within the game to directly test this concern as they contained a subset of the outcome tests with 20 questions presented in appropriate game format (skeleton or quiz). Consequently a significant decrease in performance between the challenge level and the test would have indicated children were failing to transfer their mathematical skills from the computer game to maths tasks more generally. Overall, children performed equally well on the test and the challenge but at delayed test they did score significantly better on the challenge level than the test itself. Thus, there is some evidence that children were more engaged with the math content in the game than they were when the same content is presented as a test. However, this difference between challenge level and test performance was the same in both conditions, so there is no evidence to suggest that learning content transfers less effectively from the intrinsic version of the game (where practice and challenge tasks were on skeletons and test tasks were on quiz items) than the extrinsic (where all tasks were in quiz format).

A second concern might be that intrinsic games with their increased time pressure, affect and arousal might encourage children to respond too rapidly to questions thereby promoting speed at the expense of accuracy. This had been a concern in the design of the game and led to the introduction of 'slow' levels with passive skeletons as well as penalties for guessing (health reduction) to encourage mathematical thought and reflection rather than stabbing at keyboards! Happily nothing in this data supports the worry that children had learn to guess rather than work out the answers. Children took longer to answers questions on the delayed test than they did at the pre-test and children who took longer to respond were more accurate. Again, there were no differences between the intrinsic and extrinsic conditions on this measure suggesting that intrinsic players who had been under more pressure to respond faster during learning had not come to rely on a guessing strategy.

Why did intrinsic integration foster learning?

We argued that intrinsic integration could support learning by ensuring that the activities which give rise to the central flow experience of the game involve appropriate external representations that players interact with using core gameplay mechanisms. As such our definition integrates motivational and cognitive processes to explain how intrinsic games can help learning.

Researchers have proposed a number of mechanisms by which motivation and flow could enhance learning, these include persistence, more focused attention, increased arousal, increased affect and alternative strategies (Dweck, 1986; Garris et al., 2002; Martens et al., 2004; Parkin et al., 1982; Pintrich, 2003; Vollmeyer & Rheinberg, 2000). Study 2 was a test of the assertion that increased motivational appeal should lead to increased task persistence and the results strongly supported this claim.

However, task persistence cannot provide an explanation for the increased learning outcomes of Study 1 as time on task was strictly controlled in this study. Consequently, other explanations are required to explain increased learning.

Zombie Division was created with the explicit intention of increasing learners' attention, arousal and affect – all components of the flow experience. However, most of the research on flow does not describe how flow enhances learning (beyond increasing persistence). Therefore, all we can do is postulate some possible mechanisms by which this may have increased learning outcomes for the intrinsic game. Firstly, it is well established that direct attention during encoding of material enhances its recall (e.g. Murdock, 1965; Baddeley, Lewis, Eldridge, & Thomson, 1984). Arousal has a more complicated relationship to performance and learning, albeit one that has been understood for a considerable time. Yerkes & Dodson (1908) famously described a U shaped curve whereby arousal increases performance up to a level before decreasing again. One can therefore speculate that optimizing tasks so that they are challenging but achievable (as in the way that flow theory predicts) should aim to keep learners in an optimal state of arousal. The intrinsic gameplay naturally provides the player with some level of control over their own state of arousal during the mathematics, as they can decide how quickly to seek out and engage with skeletons. Finally, compared to learning mathematics in the extrinsic condition, learning in intrinsic condition is more emotionally charged (although presumably the learner's emotions would change as they win or lose to skeletons). Research on the relationship between affect and learning is still relatively young (e.g. Pintrich, 2003) but, for example, there is reason to believe that increased affect during encoding (e.g. gameplay) can enhance retrieval at a delay (test) (e.g. Parkin et al., 1982).

The intrinsic version of Zombie Division may also have encouraged children to use better strategies to learn the mathematics. Partly this was designed into the structure of the game, as new game mechanics were introduced as the levels advanced to encourage exploration of different strategies (e.g. parrying and 'giant' skeletons). However, few children actually progressed to the levels inhabited by giants in the studies we conducted. Nevertheless, it could be that children in the intrinsic condition were spontaneously employing more effective strategies to combat the skeletons (and hence solve the maths problems). Evidence for this comes indirectly from other research which has shown that games that include more fantasy can encourage children to use more complex mathematical operations (Cordova & Lepper, 1996), that it can encourage use of more systematic strategies (Vollmeyer & Rheinberg, 2006) and more exploratory behaviour (Martens et al, 2004). Unfortunately, the design of studies does not allow us to explore this directly but we can report our observation of children pausing the intrinsic game to work out their approaches to skeletons lurking in the next room. This also serves to highlight that while fantasy may not be the correct focus for intrinsic integration, it should not detract from its value within

the overall game concept.

In contrast to the motivational aspects of intrinsic integration which had predicted almost uniformly that it should lead to better learning, the review of the more cognitively oriented literature concerning the role of interacting with concrete fast-paced external representations had revealed a more mixed picture. For example, there were concerns that making the mathematical symbols more concrete in the intrinsic game would lead to decrease learning and transfer (e.g. DeLoache 1991; Goldstone & Son, 2005; Uttal et al, 2009). Moreover, intrinsic integration breaks many of the ‘rules’ for using concrete representations (e.g. Brown, McNeil & Glenberg, 2009). It encourages learners to play and interact with representations (Uttal et al, 2009) so potentially making it harder for learners to see them as representations rather than objects of interest in their own right (e.g. DeLoache 1991) and certainly gives significance to features of the environment which are not relevant for learning (zombie skeletons!). However, the representations in *Zombie Division* are different to concrete representations that have typically been researched. Firstly, children do not have extensive experience of zombie skeletons in everyday life which might encourage them to ground their understanding in inappropriate ways. Secondly, the skeletons are a concrete context for presenting abstract symbols rather than an alternative representational system such as Cuisenaire or Dienes blocks. Moreover, because children must engage mathematically with the skeletons in order to solve the problems there is no sense in which the environment is ‘doing the work’ for them (e.g. Martin, 2009).

Why did children prefer intrinsically integrated games?

The children in Study 2 demonstrated an overwhelming difference in preference for the intrinsic and extrinsic versions of *Zombie Division*. They spent on average over seven times longer playing the intrinsic version of the game than the extrinsic. This provides clear support for the hypothesis that intrinsic integration increases motivation. There were only a small number of girls in this study (5 girls compared to 11 boys) probably because of its setting within a computer club. However, analysis of any gender differences suggested that whilst the girls played *Zombie Division* more than the boys, there were no gender differences in preference for intrinsic or extrinsic games.

The interview data (see Table 8) reveals why the children preferred playing *Zombie Division*. They tended to see the intrinsic game as easier and quicker. Unsurprisingly it was also seen as more enjoyable. Only two explanations were provided for disliking the intrinsic version – one concerning the enforced in-game tutorial and one because the participant saw the game as not fitting into the school context, speculating that teachers would not approve of it. The children’s perceptions of school requirements was echoed in the only justification provided for liking the extrinsic version when a child commented it would help them on tests. They were correspondingly able to explain their dislike of the

extrinsic version in many ways seeing it as slower and less fun. But perhaps surprisingly they also saw it as too easy both in terms of the math and the game content. Thus, ease was seen as a positive attribute in the intrinsic version and a negative in the extrinsic providing insight into the subtle and important nature of perceived challenge in educational games (e.g. Malone, 1981). Intriguingly, two children articulated principles of intrinsic integration with one stating that it is “*more fun because it’s like subliminal advertising with maths*”; and another “*it’s like mixing paint [...] the maths in the game with the fun [...] you don’t really think you’re doing that much*”.

Is game-based learning for all?

One concern that might arise when considering using intrinsically integrated games in classrooms is whether in so doing a particular subset of the population may be disadvantaged, for example, non-gamers or girls. The results of these studies did not find *Zombie Division* disadvantaged these groups. In Study 1, girls improved their scores by an average of 16% compared to 13% for the boys; a non significant difference. Just like the boys they learned more from the intrinsic game than the extrinsic game (26% in intrinsic and 14% in extrinsic for girls compared to 20% and 9% for boys). There were also no differences in in-game performance, for example on average girls progressed to level 7 and boys to level 8 and their in-game accuracy was 74% and 75% respectively. Study 2 included only a few girls and was conducted in an after school computer club, which are not ideal conditions to explore gender differences. However, the girls spent considerably longer playing the intrinsic game than boys did (as boys spent less time playing *Zombie Division* and more time on other club activities). Consequently, there is no evidence that the central game mechanic – attacking skeletons – caused the girls in these studies any anxiety. However, it should be remembered that *Zombie Division* had been iteratively developed with both girls and boys and some issues that had been observed to have potential concerns for girls addressed (e.g. boys were found to more easily understand the parrying mechanic and so more explanation was provided in the final version). Contrary to media concerns but in line with academic discussion and research (e.g. Kafai, 2008), it seems that there is no simple relationship between gender and games and that an ideal for developing games should involve early participation by both boys and girls as developers to create intrinsic games with a wide range of core mechanics.

We also have no evidence that children’s gaming skills influenced what they could learn from *Zombie Division*. Unfortunately, we do not have demographic data on children’s use of digital technologies and games outside the school that would have allowed us to test this relationship explicitly. However, we can look at whether their performance in the game influenced what they learned. Firstly, we find that prior mathematical knowledge (pre-test scores) does predict game performance (e.g. accuracy in encounters, level reached) suggesting that math knowledge is important in progressing through the game.

Secondly, no measure of game performance (or pre-test score) predicts learning gains suggesting children with all levels of mathematical and game-play skills can learn successfully from *Zombie Division*. Again one reason for this successful result may be the iterative development of the prototype whereby lack of game-play skills that had been observed to cause children problems (such as navigation issues and problems in withdrawing from skeletons) were ameliorated by changes to the design.

Classroom Implications

Although Study 1 showed that children who played the intrinsic version of the game learned the most, children in the control group also made significant improvements during the course of the study (and indeed were not reliably different to children in the extrinsic condition). This illustrates the power of debriefing in combination with the motivational appeal of games. We had expected some improvement in the control group as a result of the teacher-led reflection session but are surprised by the degree of improvement as this lesson had no relevance to their version of the game, and so wasn't reinforced in any way before the post-tests the following week. However, the teacher running these sessions reported that all three groups were unusually enthusiastic and attentive to her lesson – and she attributed this to the children's excitement about their involvement with the game. So it appears that the children's motivation for the game may have transferred to their learning in the classroom context as well.

Therefore this study left the strong impression that there is also significant potential for using games like *Zombie Division* as motivational anchor (Bransford et al, 1990) for classroom learning. There is certainly potential for creating a whole range of supporting materials based around the content and characters in *Zombie Division*. Furthermore, we believe that the intrinsic nature of the game naturally lends itself to the creation of intrinsic supporting materials, which go beyond simply including visual images from the game. It is easy to conceive other characters that would add, subtract or multiply the values of skeletons as well as a whole range of different mathematically based foes – all of which could be cheaply and easily included in paper-based classroom resources.

So in line with other research into game based learning (e.g. Squire, 2004), our experiences with *Zombie Division* seem to support the idea that teachers have a critical role to play in maximizing the educational potential of intrinsic games. While this is not something which our research has explicitly shown (as we controlled the teacher's contribution so as not to bias the results), we do not believe that games should – or could – replace traditional methods of education, but should simply form another part of the toolkit available to teachers in creating engaging and effective learning experiences for their pupils.

Design Implications

Given that the outcomes of intrinsic integration are desirable, both our research and practical experience

would suggest that there is a logical hierarchy to designing an intrinsically integrated game. This prioritizes the learning content, followed by the game mechanics and then finally the fantasy context (in line with our theorizing). Fertile learning content for creating intrinsically integrated games will include concepts that can exist and interact within a common world, rather than separate unrelated content. These links can then be used to create layers of game mechanics, which interact and create emergent game play strategies that reinforce the learning goals. Subordinate to this, the fantasy can then be worked around the game mechanics to bring them together into a coherent and motivating context. This approach is not an attempt to detract from the considerable motivational relevance of fantasy contexts in game design (e.g. Lepper & Cordova, 1996), but an acknowledgement of the primary role of core mechanics in creating an intrinsic relationship between games and their learning content. Furthermore, we would suggest that designers of educational games should give equal consideration to the offline resources available to parents and teachers in order to support learning with their game, as our own findings add to the growing research that the games on their own may be unable to offer a complete learning experience.

However, there are also practical and economic factors that present commercial barriers to the application of this research. Intrinsic games may be both more motivating and more effective than their extrinsic equivalents, but they are also more difficult and more expensive to develop which makes it harder to justify a business case. The very nature of extrinsic games means that they are more separate from their learning content, and so can be reapplied more cost-effectively to new educational purposes. Intrinsic games in contrast are far more difficult to apply to new learning content and must be largely redeveloped from scratch in order to address different learning goals. Unfortunately this is an issue that designers will have to wrestle with until the market can demonstrate a financial advantage to creating intrinsic games in addition to any motivational and learning benefits.

Conclusion

Our research acknowledges the motivational significance of fantasy in games, but argues that it is core-mechanics – rather than fantasy – that is critical to creating an intrinsic relationship with the learning content of a game. We have explored the value that this definition of intrinsic integration can bring to educational games and found benefits both in terms of motivation and learning outcomes. There is clearly much more work that could be done to tease apart the components of this intrinsic relationship, or to explore the best way of using intrinsic games within a classroom context. Nonetheless we believe this work goes some way to establishing the value of and relevance of this issue as worthy of future investigation within the field of game-based learning.

Acknowledgements

The authors wish to acknowledge and thank Jenny Habgood, Louise Ash, Lucy Button and Lizzy Evans (Reignhead Primary School), Elaine Cockburn (Gilt Hill Primary School), Sarah Peacock (Sheffield E-Learning Centre West), Vincent Aleven (Carnegie Mellon University) and Steve Benford (University of Nottingham) for their help and support in conducting this research. Our thanks are also extended to the many children who collaborated with us on the design and evaluation of Zombie Division. This research was partially supported by a PhD studentship awarded to the first author from the Learning Sciences Research Institute.

- References
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction* 16(3): 183-198.
- Aldrich, C. (2004) Simulations and the future of learning. San Francisco John: Wiley.
- Asgari, M., & Kaufman, D. (2004). *Relationships among computer games, fantasy, and learning*. Paper presented at the International Conference on Imagination and Education, Vancouver.
- Baddeley, A. D., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, 113(4), 518-540.
- Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the cognitive tutor classroom: when students "game the system". *Proceedings of the Proceedings of ACM CHI 2004: Computer-Human Interaction*, 383-390.
- Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. *The Quarterly Journal of Experimental Psychology Section A*, 35(1), 39-49.
- Bowman, R. F. (1982). A "Pac-Man" theory of motivation: Tactical implications for classroom instruction. *Educational Technology*, 22(9), 14-16.
- Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C. K., & Williams, S. M. (1990). Anchored instruction: Why we need it and how technology can help. In D. Nix & R. Spiro (Eds.), *Cognition, Education and Multimedia* (pp. 115-141). New Jersey: Lawrence Erlbaum Associates.
- Brown, M. C., McNeil, N. M., & Glenberg, A. M. (2009). Using concreteness in education: Real problems, potential solutions. *Child Development Perspectives*, 3(3), 160-164.
- Bruckman, A. (1999). *Can educational be fun?* Paper presented at the Game Developers Conference '99, San Jose, CA.
- Caillois, R. (1961). *Man, Play and Games* (B. Mayer, Trans.). New York: Free Press.
- Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, 88(4), 715-730.
- Crawford, C. (1982). *The art of computer game design*. Retrieved 23rd May, 2007, from <http://www.erasmatazz.com/free/AoCGD.pdf>
- Csikszentmihalyi, M. (1975). Play and intrinsic rewards. *Journal of Humanistic Psychology*, 15(3), 41-63.
- Csikszentmihalyi, M. (1988). The flow experience and human psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *Optimal Experience* (pp. 15-35). Cambridge, UK: Cambridge University Press.
- Cullingford, G., Mawdesley, M. J., & Davies, P. (1979). Some experiences with computer based games in civil engineering teaching. *Computers and Education*, 3, 159-164.
- Daniel, T. L., & Esser, J. K. (1980). Intrinsic motivation as influenced by rewards, task interest, and task structure. *Journal of Applied Psychology*, 65(5), 566-573.
- de Jong, T. & W. R. van Joolingen (1998). "Scientific discovery learning with computer simulations of conceptual domains." *Review of Educational Research* 68(2), 179-201.
- Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum Press.
- DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development*, 62, 736-752.

- Dondlinger, M. J. (2007). Educational video game design: A review of the literature. *Journal of Applied Educational Technology*, 4(1), 21-31.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040-1048.
- Edwards, L. D. (1998). Embodying mathematics and science: Microworlds as representations. *Journal of Mathematical Behaviour*, 17(1), 53-78.
- Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, 32(3), 158-172.
- Ferster, C. B., & Skinner, B. F. (1957). Schedules of reinforcement. *Psychological Reports*, 3(4), 695-695.
- Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, 33(4), 441-467.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity, analogy, and thought* (pp. 199-241). New York: CUP.
- Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences*, 14, 69-110.
- Gunter, G. A., Kenny, R. F., & Vick, E. H. (2008). Taking educational games seriously: using the RETAIN model to design endogenous fantasy into standalone educational games. *Education Tech Research Development* (56), 511-537.
- Habgood, M. P. J., Ainsworth, S., & Benford, S. (2005a). Endogenous fantasy and learning in digital games. *Simulation and Gaming*, 36(4).
- Habgood, M. P. J., Ainsworth, S., & Benford, S. (2005b). *Intrinsic fantasy: motivation and affect in educational games made by children*. Paper presented at the 2005 AIED workshop on Motivation and Affect in Educational Software, Amsterdam.
- Habgood, M. P. J., & Overmars, M. (2006). *The game maker's apprentice: Game development for beginners*. Berkley: CA: APress.
- Hogle, J. G. (1996). *Considering games as cognitive tools: In search of effective "edutainment"*. Retrieved 4th April, 2004, from <http://www.twinpinefarm.com/pdfs/games.pdf>
- Huizinga, J. (1950). *Homo Ludens: A study of the play element in culture* (Unknown, Trans.). Boston, MA: Beacon Press.
- Juul, J. (2005). *Half real: Video games between real rules and fictional worlds*. Cambridge, MA: MIT Press.
- Kafai, Y. B. (1996). Learning design by making games: Children's development of strategies in the creation of a complex computational artifact. In Y. B. Kafai & M. Resnick (Eds.), *Constructionism in practice: Designing, thinking and learning in a digital world* (pp. 71-96). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kafai, Y.B. (2008) Considering gender in digital games: Implications for serious game designs in the learning sciences. Paper presented at ICLS 2008: Utrecht.
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, 320(5875), 454-455.
- Kerawalla, L., & Crook, C. (2005). From promises to practices: The fate of educational software in the home. *Technology, Pedagogy and Education*, 14(1), 107-125.
- Koster, R. (2005). *A theory of fun for game design*. Scottsdale, AZ: Paraglyph Press.
- Lederman, L. C. (1992). Debriefing: Towards a systematic assessment of theory and practice. *Simulation and Gaming*, 23(3), 145-160.
- Lepper, M. R., & Greene, D. (1975). Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. *Personality and Social Psychology*, 31(3), 479-486.
- Lepper, M. R., & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, Learning and Instruction: III. Conative and affective process analyses* (pp. 255-286). Hillsdale, NJ: Erlbaum.
- Loftus, G. R., & Loftus, E. F. (1983). *Mind at play. The psychology of video games*. New York: Basic Books.
- Lundgren, S., & Björk, S. (2003). *Game Mechanics: Describing Computer-Augmented Games in Terms of Interaction*. Paper presented at the 2003 Technologies for Interactive Digital Storytelling and Entertainment conference, Darmstadt, Germany.
- Malone, T. W. (1980). *What makes things fun to learn? Heuristics for designing instructional computer games*. Paper presented at the Association for Computing Machinery Symposium on Small and Personal Computer Systems, Pal Alto, California.

- Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, Learning and Instruction: III. Conative and affective process analyses* (pp. 223-253). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Martens, R. L., Gulikers, J., & Bastiaens, T. (2004). The impact of intrinsic motivation on e-learning in authentic computer tasks. *Journal of Computer Assisted Learning*, 20(5), 368-376.
- Martin, T. (2009). A theory of physically distributed learning: How external environments and internal states interact in mathematics learning. *Child Development Perspectives*, 3(3), 140-144.
- Murdock, B. B. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology*, 56(4), 413-419.
- Nunes, T., Schliemann, A.-L., & Caraher, D. (1993). *Street mathematics and school mathematics*. New York: Cambridge University Press.
- Papert, S. (1980). *Mindstorms: Children, computers and powerful Ideas*. New York: Basic Books.
- Papert, S. (1998). Does easy do it? Children, games and learning. *Game Developer*, 87-88.
- Paras, B. (2005, June 2005). *Game, Motivation, and Effective Learning: An Integrated Model for Educational Game Design*. Paper presented at the Proceedings of DiGRA 2005 Conference: Changing Views - Worlds in Play, Vancouver, BC.
- Parkin, A. J., Lewinsohn, J., & Folkard, S. (1982). The influence of emotion on immediate and delayed retention. *British Journal of Psychology*, 73(3), 389-393.
- Pintrich, P. R. (2003). A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology*, 95(4), 667-686.
- Rieber, L. P. (1991). Animation, incidental learning and continuing motivation. *Journal of Educational Psychology*, 83(3), 318-328.
- Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, 44(2), 43-58.
- Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. Cambridge, MA: The MIT Press.
- Sandford, R., & Williamson, B. (2005). *Handbook on games and learning*. Bristol: Futurelab.
- Scaife, M. & Y. Rogers (1996). "External cognition: How do graphical representations work?" *International Journal of Human-Computer Studies* 45(2): 185-213.
- Shaffer, D. W. (2004). Pedagogical praxis: The professions as models for postindustrial education. *Teachers College Record*, 106(7), 1401-1421.
- Shaffer, D. W., Squire, K. D., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan*, 87(2), 104-111.
- Squire, K. D. (2004). *Replaying history: Learning world history through playing Civilization III*. Indiana University, Bloomington, IN.
- Steinkuehler, C. A. (2006). The mangle of play. *Games and Culture*, 1(3), 199-213.
- Tabak, I., & Reiser, B. J. (1997). *Complementary Roles of Software-based Scaffolding and Teacher-Student Interactions in Inquiry Learning*. Paper presented at the 1997 conference on Computer Support for Collaborative Learning.
- Trushell, J., Burrell, C., & Maitland, A. (2001). Year 5 pupils reading an "interactive storybook" on CD-ROM: Losing the plot? *British Journal of Educational Technology*, 32(4), 389-401.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352-373.
- Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, 3(3), 156-159.
- Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, 10(4), 293-309.
- Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review*, 18(3), 239-253.
- Winn, B. (1987). Charts, graphs and diagrams in educational materials. in D. M. Willows and H. A. Houghton (Eds) *The psychology of illustration: I. Basic research*. New York, Springer: 152-198.
- Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans. Third ed.). Malden, MA: Blackwell Publishing.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of Stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.