University of Reading

'It's like having a test but in a fun way': young learners' perceptions of a digital game-based assessment of early language learning

Article

Accepted Version

Courtney, L. and Graham, S. ORCID: https://orcid.org/0000-0002-7743-3977 (2019) 'It's like having a test but in a fun way': young learners' perceptions of a digital game-based assessment of early language learning. Language Teaching for Young Learners, 1 (2). pp. 161-186. ISSN 2589-207X doi: 10.1075/ltyl.18009.cou Available at https://centaur.reading.ac.uk/82230/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1075/ltyl.18009.cou

Publisher: John Benjamins

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

‘It’s like having a test but in a fun way’: Young learners’ perceptions of a digital game-based assessment of early language learning.

Louise Courtney, University of Reading

Suzanne Graham, University of Reading

Bios

Louise Courtney taught French in UK primary schools before completing a PhD in Early Language Learning. She has contributed to large-scale studies evaluating the teaching and learning of languages with younger learners, focusing on primary to secondary school transition, second language literacy development and language assessment. Her research has been published in leading international journals.

Suzanne Graham has researched and published widely in second language learning motivation, vocabulary development and reading and listening comprehension. She led a recent longitudinal study of young learners of French in England, exploring the impact of teacher and teaching factors on motivation and learning outcomes.

Abstract

Assessment is a central challenge within classroom-based early language learning, where there is a need to employ assessment methods which, as well as being valid and reliable for a range of learners, protect rather than diminish motivation. The motivational properties of digital or serious games within language learning are increasingly recognised in the literature, yet the value of digital game-based assessment (DGBA) remains underexplored. This study used a questionnaire to examine how 3437 young language learners of English, Spanish, German, Italian and French perceived a DGBA tool and the extent to which their perceptions were modulated by age, gender and performance levels on the game.

Learners perceived the DGBA tool as fun to play, worth playing again, helpful for telling them about their progress and of moderate difficulty level. Girls were more positive than boys about the game but for all learners levels of positivity were not related to age. There was a significant but weak relationship between positivity and game scores, suggesting that

learners liked the game regardless of their attainment levels. The study's findings are discussed in relation to theories of motivation associated with digital game-based tools and their practical implications for the teaching of early language learners.

Key words

digital game-based assessment; motivation; gender; age; early language learning.

It has become increasingly widespread to introduce a second or foreign language into the elementary or primary school curriculum (Copland, Garton & Burns, 2014; Murphy, 2014; Nikolov, 2016). This is often done with the intention of developing positive attitudes towards future learning at secondary school, as is the stated rationale in England, for example (Department for Education (DfE), 2013). It also reflects a belief that young learners learn a new language quickly, and are necessarily motivated language learners. Increasingly, however, research evidence highlights that **instructed** early language learning is in fact a slow process (Nikolov, 2016) and not always a positive one for all young learners (Courtney, Graham, Tonkyn, & Marinis, 2015), with classroom experiences influencing both learning outcomes and motivation (Graham, Courtney, Marinis, & Tonkyn, 2017; Graham, Courtney, Tonkyn, & Marinis, 2016; Johnstone, 2009). Likewise, a large-scale questionnaire survey of teachers in over 142 countries (Copland et al. 2014) cited the issue of motivating young learners of English as one of the main difficulties they faced. Indeed, Nikolov (2016) comments that an earlier start to language learning may, perhaps unexpectedly, make issues of maintaining motivation more rather than less of a challenge, in so far as "the earlier L2 learning is introduced, the sooner typical classroom activities and topics become boring for young learners" (p.4).

An important element of learners' early language classroom experiences relates to the assessment practices that they are faced with, which can be both positive and negative. On the one hand, assessment has the potential to provide learners with the sense of progress that is important for maintaining motivation (Graham et al., 2016; McKay, 2006). On the other hand, assessment poses challenges, particularly when the focus of early language learning is less on "fun and ease" and more on "standards-based measurement of the outcomes in the target language" as part of "an international trend in educational assessment for

accountability in public educational policies in all subjects and competences" (Nikolov, 2016, p.4). Finding assessment methods which protect rather than diminish motivation, while also providing valid and reliable measurements of learners' attainment, is therefore an important goal for practitioners and designers of curricula alike.

Literature Review

The goal of finding appropriate methods of assessment for young learners is particularly challenging for a number of reasons. The first arises from the characteristics of young language learners, whose attention span is usually more limited than learners at high school or above, and whose capacity for the more abstract, decontextualized thought often needed in assessment tasks does not usually develop until high school or secondary school. McKay (2006) emphasises the importance of assessment tasks that are concrete rather than abstract, that use similar approaches to those used in the teaching of learners (i.e. drawing heavily on game-based approaches), related to learners' realm of experience and using the genres with which they are familiar, and which help them to maintain concentration and motivation to complete the tasks. This implies that assessment through intrinsically motivating content is important. The fear and dislike of assessments reported in a number of contexts such as Hungary (Nikolov, 2003; 2008) and England (Graham et al., 2016) may arise from the fact that the self-esteem of young learners is vulnerable (McKay, 2006), making it important that assessment tasks are "psychologically safe" (p. 10). This may be especially true for learners of lower attainment levels. Early language learning often takes place in mixed-attainment classes which poses an additional challenge. Copland et al. (2014) found that teachers viewed responding to the different levels of learners and their differing learning needs in mixed-proficiency classes as a particular difficulty. A lack of differentiated provision also emerged as a problem in Graham et al. (2016), where learners appeared to be demotivated by

work that they saw as either too easy or too difficult. Therefore, assessments need to allow learners of all attainment levels to demonstrate what they are capable of and in a format that is accessible to all.

Furthermore, the reliability and validity of any assessments are threatened if the anxiety levels of young learners taking them are high (Johnstone, Cavani, Low, & McPake, 2000). As Szpotowicz and Campfield (2016, p. 111) argue, learners are likely to feel more motivated by, and less anxious about, assessment formats that are similar to "regular daily school activities", with positive implications for the validity of the test data they produce as well as for learners' longer- term language learning motivation. Nikolov's (2017) study conducted with 2173 young learners of English in Hungary (grades 1 and 6) found significant, medium to strong correlations between how easy and familiar an assessment task was perceived to be and how much learners reported liking it. By contrast, there were only weak (but significant) correlations between actual test score and reported likeability. These findings suggest that learners respond better to assessment tasks based on familiar task-types, and that assessment tasks can be seen as inherently enjoyable (if well designed), despite learners achieving lower scores.

The second challenge arises from the difficulties language assessment poses to teachers working with young learners, many of whom have had little or no training or confidence in conducting it, and may lack the curriculum time to do so. These issues have been reported in contexts ranging from Germany and the Netherlands (Edelenbos & Kubanek-German, 2004) and Hungary (Nikolov 2017) to Asia (Butler, 2009a, b). Teachers are not helped by a lack of clarity regarding expectations for young language learners. In most countries in Europe the first level of The Common European

Framework of Reference for Languages (CEFR) A1 is set as a standard for primary age learners. Arguably however it is not sufficiently fine-grained to capture the very early stages of language learning and the small amounts of progress learners might make, especially in contexts where limited time is devoted in the curriculum to language study. Even worse, while in England language learning from age 7 has been compulsory since 2013, there is no national system of assessing achievement below the level of the national examination taken at age 16. Curriculum guidelines merely state that by the end of primary school learners should "make substantial progress in one language" (DFE, 2013, p. 2). Furthermore, a lack of reliable assessment data poses problems for effective transition to secondary education, which is crucial to the success of early language learning. Fundamental to effective transition is the communication of accurate assessment data which enables practitioners to build on learners' current level of proficiency. Without this information secondary schools often adopt a "fresh start" approach to language teaching, thereby demotivating learners and potentially hindering progress (Graham et al., 2017). What is therefore required are assessment tools that can be used repeatedly to span learners both in the primary and early secondary sectors. These tools should assess a range of language skills, demonstrate progression over time and contain content that is appropriate both linguistically and motivationally for learners across a range of ages.

Overall, the characteristics of young language learners require that assessment tools are concrete rather than abstract, use similar approaches to those used in everyday teaching, are related to learners' realm of experience, and use the genres with which they are familiar, in order to maintain their concentration and motivation to complete the tasks. All of these factors suggest an important role for games in assessment. Furthermore, in order to facilitate whole class assessment by teachers who may lack assessment knowledge and expertise, as well as lesson time in which to conduct assessment (Graham et al., 2017), the use of games in

computer format offers promising possibilities. While Digital Game-Based Learning (DGBL) or "serious games" to use an alternative term (Sørensen & Meyer, 2007) is described by Alyaz, Spaniel-Weise and Gursoy (2017) as a "current trend" (p. 250), Digital Game-Based Assessment (DGBA) in language learning is less common and research into its impact even less so (Lay, Patton, & Chalhoub-Deville, 2017).

Research into the impact of serious games on attainment in learning in general and in language learning in particular has produced inconclusive and contradictory findings (Calvo-Ferrer, 2017; Chiu, Kao, & Reynolds, 2012; Girard, Ecalle & Magnan, 2013). Nevertheless, evidence for their capacity to motivate learners across a range of curriculum areas is more widespread (Papastergiou, 2009), with positive reactions to digital learning from older as well as younger learners (Alyaz et al., 2107). The positive motivation of learners towards such games reported by the majority of studies springs perhaps from their widespread use in life outside the classroom (Annetta, Minogue, Holmes, & Cheng, 2009: McFarlane, Sparrowhawk, & Heald, 2002; Papastergiou, 2009, with reference to learning in general, and Sundqvist & Sylven (2014) in relation to English language learning). There are however also sound, theoretical reasons why serious games are motivating. A number of authors highlight their motivational benefits with reference to mainstream motivational constructs/theories, particularly "flow", intrinsic motivation and Self-Determination Theory (see for example, Boyle, Connolly, & Hainey, 2011; Malone & Lepper, 1987; Wrzesien & Raya, 2010). In a 30 year old but highly cited review, Malone and Lepper (1987) identify four main areas – challenge, curiosity, control and fantasy - that contribute to the intrinsic motivation stimulated by digital games. Challenge, or more specifically optimal challenge, is fundamental to intrinsic motivation according to a number of motivational theories (e.g. Bandura's (2001) social cognitive theory), but most especially to Csikszentmihalyi's (2000/1975) concept of *flow (*Sitzmann, 2011). Flow refers to the state of almost total absorption in and concentration on

an activity and persistence with it to the end, in which awareness of time passing seems to disappear (and which digital game players often experience), and which is argued to be heightened when the activity is just beyond an individual's current level of competence, i.e. at an optimal level of challenge (Kiili, 2005). In digital games, the use of different levels of tasks and specific, not too distant goals, but which still involve a degree of uncertainty as to their attainability, encourages learners to persevere to the next stage of the game and has been found to be beneficial for motivation, as Lister (2015) reports in a meta-analysis of studies exploring the motivational effect of DGBL. Continuous, constructive and encouraging feedback is also important for optimal challenge, with care needed with the design of so-called 'failure states', whereby the game player learns that they have performed a task incorrectly. In a study of adult use of DGBL for language learning, Cornillie, Clarebout and Desmet (2012) note that failure states must be "interesting and enjoyable, and that the player can repeatedly fail without compromising the motivation necessary for successfully completing an action or task" (p. 262) in order to protect their sense of flow. This is likely to be even more important for young learners and for those whose sense of competence is fragile. Typically, digital games offer feedback in the form of status bars that show the learners' progress through the game, which Shute (2011) sees as advantageous in that it can enhance learners' "metacognitive processes" (p.519) of reflection and self-knowledge.

Such feedback is also important for instilling a sense of competence, a factor which alongside autonomy and control is central to the Self-Determination motivation framework of Ryan and Deci (2000). These motivational factors are inherent features of most digital games (Boyle et al, 2011; Przybylski, Ryan, & Rigby, 2009). Competence and sense of control or power can be enhanced through the level of fantasy and cognitive curiosity (e.g. desire to solve a problem) typically part of serious games, with fantasy providing vicarious experiences of success or power (Malone & Lepper, 1987). Emotional fantasy, meeting emotional needs, can

be increased if the game player can identify with imaginary characters in the game; for example, if the game allows them to choose their own avatar and to give it their name (Sykes, Oskoz, & Thorne, 2008). A compelling narrative is also important for fantasy and engagement (Stott & Neustaedter, 2013).

Many of these characteristics of serious games were identified as being the features of an 'ideal' digital language learning game in studies by Butler and colleagues. They firstly determined which game characteristics appealed to 3945 children (aged 4-12) in an online English-learning site in Japan (Butler, Someya, & Fukuhara, 2014), by examining frequency of play of certain language learning serious games. These characteristics included being cognitively demanding (challenge), stimulating curiosity, and giving players control, as well as involving multiple players (even if computer-generated rather than real). A follow-up, qualitative study of 82 learners aged 11-12 added further characteristics to this list, such as clear rules, objectives, obstacles, outcomes and feedback, repetition, sounds and visual effects, stories and fantasy (Butler, 2015).

Gender and age differences were also explored by Butler et al. (2014) regarding learners' preferences. While no real age differences in playing frequency (i.e. likeability) were noted for the easiest level of the game (Bronze), at the linguistically more challenging Silver and Gold levels, frequency of playing decreased with age, especially around ages 8-9 and 10-12. By contrast, significant gender differences were noted on only two sub-sections of the game (a card game and a car-racing game), with no age-gender interactions. There is however some evidence (e.g. McFarlane et al., 2002) that boys like playing educational games more than girls do and partake in computer game playing in English outside the classroom to a greater extent than girls. In addition, a literature review of girls' and boys' preferences regarding computer games suggests that the two sexes enjoy different aspects:

males appear to generally prefer games with an active and competitive element, females those that incorporate logic, puzzles and skill training (Procci, Bohnsack, & Bowers, 2011; Romrell, 2013). Nevertheless, even though boys in a study by Paperaastergiou (2009) reported playing computer games more often outside school than did girls, as in the study by Butler et al. (2014) no gender differences emerged in how motivating they found an educational game in Computer Science. It seems likely that certain design features can be used to maximise a game's appeal to both genders; for example, allowing players to customize the physical appearance of the game avatar is considered to make a game more appealing to females whose interest might otherwise be lower than males' (Procci et al., 2011). As stated earlier, however, compared with their use for learning, the potential of serious games for assessment remains largely unexplored and especially in relation to language learning (Lay et al., 2017), and from the perspective of how they are perceived by learners. This point is no trivial matter, given what we have outlined above regarding the importance of assessment tools.

In sum, assessing young language learners presents challenges for teachers, with evidence suggesting that more traditional formats have negative implications for learners' motivation and anxiety. There is a strong theoretical base and a growing empirical base to support the motivational potential of digital game-based learning, yet research into learners' perceptions of digital-game based assessment is very limited, especially in the area of language learning and in relation to the age, gender and attainment level of the learner. These issues were explored within the context of a digital game for language assessment (DGBLA) used with young learners in four countries, addressing the following research questions:

1. How is a digital game-based assessment of language competence perceived by young learners of English, Spanish, Italian, German and French?
2. To what extent are learners' perceptions modulated by their age, gender and levels of performance on the game?

Methods

Participants and Research Design

Schools known to the research team were invited to participate in the study, using convenience sampling. Overall, 3,437 young learners across four countries (England, Germany, Italy and Spain) played the DGBLA (at one of two levels) and then completed an accompanying questionnaire. The mean age for the participants was 9.3 years with an age range of 6 years to 13.59 years. Boys and girls were fairly evenly represented within the sample. Table 1 displays the descriptive statistics for the participants by country.

Table 1: Participant Details by Country

	England	**Spain**	**Italy**	**Germany**	**Total**
Total no. of responses	1378	799	651	609	3437
Mean age at testing in years (range)	9.25 (7.16-11.82)	9.00 (6-11.79)	8.78 (6.98-10.48)	10.31 (7-13.59)	9.3 (6-13.59)
No. of boys	679	402	306	319	1706
No. of girls	699	397	345	290	1731

Within the sample from England, 438 were learners of Spanish, 622 learners of German, 102 learners of Italian and 216 learners of French. Participants from all other countries were

learning English. Across the countries there was a range of different starting ages for language learning and amount of weekly lesson time for language learning, which for reasons of space are not discussed here. Of the total sample, 2620 learners completed the Level 1 game, and 817 the Level 2 game.

The Game: *The Language Magician*

The game was designed to provide diagnostic information on language proficiency in the areas of vocabulary, listening, reading and writing (speaking was not assessed because of the technical challenges of using voice recognition software reliably with young learners). Two levels were created: Level 1, aimed at learners after 50 to 70 hours of instruction, and Level 2, for learners with 70 to 100 hours of instruction. Versions were created in English, German, Italian, Spanish and French. Tasks were created that assessed the knowledge and skill expected across the countries involved in the project at each respective level, following a model of progression that moved from the more simple (e.g. phoneme, word-level recognition), to sentence and text-level comprehension and production. There are 90 tasks in total at each level of the game, split across five floors corresponding to different rooms in the tower. Table 2 below provides overview details of what is tested at each level, with full details available in the game manual at https://www.thelanguagemagician.net/resources/. Task types are broadly similar at Level 1 and 2 but with more difficult content (for example, at Floor 2, Level 2, identifying the position of objects from sentences including prepositions). Each learner has two attempts at each item; the language content on the second attempt remains the same but the answers are randomly shuffled. The scoring criteria are as follows: three points for correct answer at first attempt, two points for correct answer at second attempt, one point for attempting the item and zero points if the learner does not attempt the answer the item. In an assessment of the game results reported elsewhere

(https://www.thelanguagemagician.net/latest-research-results-from-the-eu-project-the-language-magician/), it was found to be effective in discriminating between different levels of attainment, to be of a "good" level of difficulty (McCowan & McCowan, 1999), that is, appropriate for the learners at whom it was aimed and to correlate strongly with a measure of general linguistic ability, an adapted C-test.

Table 2: Details of Language Tasks per Floor of the Game (number of items, Level 1/2)

	Skill	Sample task
Floor 1	Listening & reading (22/12)	Match word/sentence to picture or written option
Floor 2	Listening (21/19)	Minimal pairs
Floor 3	Reading (17/21)	Colour a picture to match short description
Floor 4	Integrated activities (15/18)	Transcribe short phrase
Floor 5	Writing (15/20)	Complete a short dialogue choosing from options given

The Questionnaire

Learners' perceptions of the game were assessed through a short online questionnaire (see Appendix A). Based on a tool developed for tracking motivation over time among learners aged 10 to 12 in England (Graham, Marinis, Tonkyn, & Courtney, 2014), the questionnaire contained 14 closed items that used a four-point Likert scale and asked questions about language learning in general, feelings of self-efficacy and attitudes to the game. The internal consistency of the whole questionnaire was .822, as measured through Cronbach's Alpha.

The four game-related questions asked learners to state how far they agreed that the game was fun to play, that the tasks were easy, that they would like to play the game again, and that the game told them how well they were doing in the target language. A further five open questions then invited learners to write in comments on what they liked the most about the game, whether there was anything they did not like about it, what was the hardest and the easiest aspect of the game, and what they had learnt from playing the game. For the purposes of this paper, only responses to the game-related open and closed questions will be analysed and presented. Findings related to the overall learner motivation for language learning are reported elsewhere (Courtney & Graham, in preparation).

Procedures

Before learners played the game, each class teacher entered demographic information, including the date of birth, gender and hours of instruction for each child on the teachers' dashboard. Learners then played the game using either a tablet or computer with headphones. Immediately after they had finished the game, participants were required to complete an online version of the questionnaire. Game item and questionnaire responses for each individual participant were then sent to the game research backend for extraction and analysis. Consent for the use of learners' data was gained from school head teachers and parents/carers.

Data Analysis

For quantitative analysis of the questionnaire responses, the csv file extracted from the research backend of the game was converted to SPSS format. Mean scores for individual questionnaire items were calculated. The three components of the questionnaire, namely attitudes to language learning, self-efficacy and attitudes to the game, emerged from principal

components analysis with Direct Oblimin Rotation performed on the whole questionnaire (Kaiser-Meyer-Olkin, = .87, Bartlett's Test of Sphericity < .001; variance explained, 30.65%, 12.37%, and 7.94% respectively), followed by assessment of internal consistency as measured by Cronbach's Alpha. Reliability for the attitudes towards the game component was rather low (.544). As it included a small number of items (four), reliability was also assessed by calculating the mean inter-item correlation value (Pallant, 2013), at .25. Hence internal consistency was deemed acceptable (i.e. between .2 and .4 as recommended by Briggs & Cheek (1986), although we acknowledge that a higher level would have been preferable; see Limitations below). As such we calculated a combined score for the four items, which we termed 'positivity', as well as exploring each item individually. Non-parametric statistical tests were applied to the questionnaire data which were not normally distributed and individual items were at ordinal level[1.] Qualitative analysis of the open item responses was conducted using a grounded approach to thematic analysis and data were coded and analysed using Nvivo.

Results

Research Question 1: Learners' Perceptions of the Game

The mean score for the positivity scale (combining all four game-related questions) for the entire cohort was 3.40, indicating that overall learners held very positive attitudes to the game. Since the mean/median of a four-point Likert scale is 2.5 we considered score of 2.5 or higher to indicate a positive response. As well exploring overall positivity, we also analysed responses to each of the four game-related closed questions for all learners, in order to gain more detailed insights. Table 3 indicates that learners considered that playing the game was fun, with few participants rating that statement as 2 or below. The vast majority

(94%) agreed that they would like to play the game again and 88% felt that the game told them how well they were doing at language learning. On the other hand, it appears that the game was challenging for a number of learners, with 40% of participants disagreeing that the game was easy to play. This is to be expected with a mixed ability cohort, from different contexts and with differing amounts of language instruction. We sought further insights into how learners' sense that the game was not easy compared with their views that it was fun and worth playing again. To do so we assessed whether learners' level of agreement varied across each questionnaire item, using within-subject non-parametric tests (Friedman's test followed by post-hoc Wilcoxon signed ranks tests, with Bonferroni adjustment, reducing the alpha level to .008).

Table 3: Mean Scores, Standard Deviations, Range and Percentage of Participants Scoring 2 or Below for each Game-related Closed Item

Questionnaire item	**Mean Score (s.d.)**	**Range**	**% participants scoring 2 or below**
1_0: The game is fun to play	3.69 (.62)	1-4	5
1_1: The tasks in the language game are easy	2.79 (.86)	1-4	40
1_2: I would like to play the game again	3.70 (.63)	1-4	6
1_3: The game tells me how well I'm doing (in the language learned)	3.39 (.78)	1-4	12

Results showed significant differences between perceiving the game as easy on the one hand and on the other considering it fun to play ($z = 36.86$, $p < .001$), worth playing again ($z = 36.04$, $p < .001$), or informative about progress ($z = 26.84$, $p < .001$). This indicates that learners found the game enjoyable and worthwhile even if (or maybe because) they did not find it easy.

Qualitative analysis.

In order to gain further insights into the quantitative findings, responses to the open-ended items for the learners from England only were analysed. This country was chosen because learners there have been found in previous research (European Commission, 2012) to be less positive about language learning and were therefore of particular interest from a motivation perspective. The first open question (Q1_4) asked: What did you like most about the game? Of the 1378 learners from England completing the game, 1163 learners provided a response to this question[1]. An examination of the responses to Question 1_4 identified eleven key themes. Text-based queries and hand-coding were conducted in Nvivo to quantify the frequency of occurrence of codes related to these eleven key themes (Table 4).

Several responses indicated overall High Positivity to the game (amounting to 7.8% of responses containing the word "love" or the statement that the learner liked "everything", or that the game was "awesome", and contrasting with only seven responses that indicated that "nothing" was liked or that it was "not fun"). Fantasy and emotional engagement, identified in the literature review as important motivational aspects of serious games, appeared frequently in learners' response. References to liking the storyline/adventure, relating to, for example, animals, magic and the avatar, comprised 38.3% of the responses to question 1_4. The following examples (with spelling and grammar uncorrected) illustrate

engagement with the storyline and game concept:

> *It was alot of fun to learn more about spanish and I liked that you could save the animals. I really liked it* (boy, 11.15 years)
>
> *That we get to dress up and rescue the poor animals* (girl, 9.24 years)
>
> *I liked that we could design our own character and it was basically really fun i would definitely play this game again* (boy, 9.73 years)

A number of learners (refs = 49) commented positively on the look and feel of the game and the sound effects:

> *the grathics are brilliant* (boy, 10.69 years)
>
> *i liked Winivil's voice and the way the game looks* (girl, 9.68 years)

The responses also indicated that some learners felt they had learned something new from playing the game. Not only did a good proportion (14.4%) state that they considered playing the game was fun, but also a not insignificant number (10.5%) claimed that it helped them to learn new elements of the target language, even though the game was designed as an assessment rather than a learning tool. Below are some examples of the responses in relation to "learning" and feelings of "getting better":

> *it was fun and helps you get better at French* (girl, 10.51 years).
>
> *i liked it because i learnt alot of german and it is fun* (girl, 10.44 years)
>
> *It helped me lern more spanish and i liked it* (boy, 8.4 years)

Challenge is another aspect identified in the literature as key for serious games if they are to be motivating. The analysis of the closed questionnaire items had already indicated that some learners found the game difficult. This was confirmed in the open responses, where only 15 learners used the word 'easy' to describe the game. A much larger number of learners (n=65) mentioned difficulty or challenge, but as a positive aspect of the game, as exemplified in the following:

> *I liked that it was challenging and hard because the questions were in German* (girl, 11.06)

> *What i liked most about it is that it's challenging and makes me like Italian more and it tought me some new words* (boy, 10.88)

Only one learner held a negative view of the difficulty of the game in relation to this question, although they did still think the game was fun:

> *I didn't really like anything, it was fun but too hard* (boy, 10.43)

A small number of learners (n=10) indicated that playing the game was motivating, both while playing the game, and in terms of learning the language more broadly, as is exemplified by the responses below:

> *It made German a very fun and inthusiastic* (girl, 8.5 years)

> *that its a fun game to make children play more and find out more spanish* (girl, 8.65 years)

Even though some learners were aware they were being assessed, their responses indicated that they saw the game as a fun and non-threatening way of assessing language. Furthermore, a small number of learners (n=23) liked the fact that tool enabled them to test their knowledge of the language they were learning:

> *its more fun than a normal french assessment* (boy, 10.98 years)

> *What i liked about the game is that it tells me how much i know Spanish* (boy, 8.75 years)

> *That it gives you a chance to see how well you are doing in the spanish lessons* (girl, 10.92)

Table 4: Frequencies for Response Codes for Open Question 1_4 *(like about the game)*

Theme	**Code**	**No. of references**	**% of total references**
Storyline/adventure	Animals	218	14.4
	Spells	72	4.8
	Avatar	133	8.8
	Storyline	66	4.4
	Magic/Magician/Magical	37	2.4
	Stars	37	2.4
	Collecting Accessories	16	1.1
	Total	**579**	**38.3**
Fun	Fun	**218**	**14.4**
Learning	Learning	150	10

	Getting Better	7	0.5
	Total	**157**	**10.5**
High positivity	Everything	60	4
	Love	46	3
	Misc. High Positivity	12	0.8
	Total	**118**	**7.8**
Motivating	Motivating	**10**	**0.6**
Perceptions of ease/difficulty	Hard/challenging	65	4.3
	Easy	15	1
	Total	**80**	**5.3**
Game content	Tasks	**203**	**13.4**
Game design features	Graphics, illustrations	27	1.8
	Sound	22	1.5
	Total	**49**	**3.3**
Testing	Testing in a fun way	8	0.5
	Testing knowledge	23	1.5
	Total	**31**	**2**
Negativity	Nothing	6	0.4
	Not fun	1	0.1
	Total	**7**	**0.5**
Miscellaneous	Miscellaneous	**61**	**4**
	Overall Total	**1513**	**100**

The overwhelmingly positive responses to Q1_4 were also echoed in the responses to Q1_5, which asked learners what they disliked about the game: almost 60% of learners of the 1057[1] learners who gave a response to this question said *'nothing'.* However, in contrast to the learners who liked the game because it was challenging (see Table 4), a good proportion considered the difficulty of the game as something that they disliked (13.2%, see Table 5). It should be noted, however, that of the 141 learners commenting that the game was difficult, only seven had a mean positivity score of 2 or below on the closed questionnaire items, indicating that the vast majority of those who thought the game was hard were still positive about it overall. Some of the reasons cited were just that it was "*too hard"* (boy, 10.43 years), and in some cases learners complained that they had yet to learn the content; "*sometimes it was hard and i havent leant most of it"* (girl; 10.90). Also, activities such as "working at sentence level" and the "writing/spelling" tasks were quoted by a small number of learners as items they disliked. Related to the perceived difficulty of the tasks was the reference by many learners (n=92) to needing more time to complete the game and sensing frustration from the imposed time limit; "*i dident like you coudent take your time. It was hard"* (girl, 8.68 years).

Table 5: Frequencies for Response Codes for Open Question 1_5 (*dislike about the game*)

Theme	Code	No. of references	% of total references
High Positivity	**Nothing**	**607**	**56.7**
Perceptions of ease/difficulty	**Hard/difficult**	**141**	**13.2**
Game Design Features	**Time Limit**	92	8.6
	Look and Feel	80	7.5
	More Avatar Options	20	1.9
	Less Speaking Time	13	1.2
	Total	**205**	**19.2**
Storyline/adventure	**Magician**	33	3.1
	Boring	6	0.6
	Total	**39**	**3.7**
Miscellaneous	**Miscellaneous**	**59**	**5.5**
Game content	**Sentence Level Tasks**	11	1
	Spelling/Writing Tasks	9	0.8
	Total	**2**	**1.8**
	Total	**1071**	**100**

The other less frequently given responses mainly related to the storyline/adventure including the magician (who was considered by some as scary) and to the game design features,

comprising references to: the look and feel of the game, the avatar (not enough options for clothing); the graphics and sound effects; and the length of the introduction to the game where the animals speak to the player to set the scene (coded as *less speaking time*).

The third open item asked learners what part of the game they found easiest. As might be expected for a game designed to become progressively more challenging, 44% of the 1025[1] responses coded referred to the first floor (or beginning) of the game that focused on listening and reading (see Table 6). The responses also highlighted wide variation in how easy/difficult learners found the game. While 31 learners claimed everything was easy, a further 37 stated that nothing was easy, and 15 learners were not sure. Although 61 learners made reference to specific content or tasks that they found easy, a large number did not mention the task content at all, suggesting that they were more focused on the game as a game and less on its language assessment aspects. They instead emphasised features of the in-game activities such as casting spells (n=105), feeding the animals (n=54), customising the avatar (n=50), collecting stars (n=47), saving the animals (n=30), listening to the story (n=8) and collecting accessories (n=3).

Table 6: Frequencies for Response Codes for Open Question 1_6 (*found easiest*)

Theme	Code	No. of references	% of total references
Floors	Floor 1	454	44.3
	Floor 2	39	3.8
	Floor 3	22	2.1
	Floor 4	6	0.6

	Floor 5	7	0.7
	Total	**528**	**51.5**
In-game activities	Casting spells	105	10.2
	Feeding animals	54	5.3
	Customising avatar	50	4.9
	Collecting stars	47	4.6
	Saving animals	30	2.9
	Listening to the story	8	0.8
	Collecting accessories	3	0.3
	Total	**297**	**29**
Game content	Particular content	20	2
	Making potion	15	1.5
	Yes/No questions	11	1.1
	Colouring	9	0.9
	Spelling	3	0.3
	Making	3	0.3

	sentences		
	Total	**61**	**6**
Perceptions of ease/difficulty	Easy	31	3
	No-nothing	37	3.6
	Don't know	15	1.5
	Total	**83**	8.1
Miscellaneous	**Miscellaneous**	**56**	**5.5**
	Total	**1025**	**100**

In contrast, when asked about the most difficult elements of the game, the responses of the 1090[1] learners who responded to this question were largely focused on language assessment content with the upper floors being considered as the most difficult (Table 7), with a gradual decrease towards the lower floors. The upper floors (4 and 5) included writing (n=79) and sentence/text level tasks (n=64), and learners' comments highlighted difficulties with more literacy-oriented activities:

> *the spelling because i knew what it was but i didnt know how to spell it* (girl, 10.41 years)

> *The hardest part of the game was a sentance written in German and you had to choose what the sentance was about* (boy, 10.96 years)

The learners who suggested that everything or nothing was difficult may perhaps have felt that no floor was more difficult than another.

Table 7: Frequencies for Response Codes for Open Question 1_7 (*found most difficult*)

Theme	Code	No. of references	% of total references
Floors	Floor 5	266	23.5
	Floor 4	95	8.4
	Floor 3	82	7.3
	Floor 2	48	4.2
	First Floor	23	2
	Total	**514**	**45**
Game content	Writing tasks	79	7
	Sentence Level Tasks	64	5.7
	Colouring	44	3.9
	Making potion	34	3
	Particular Content	26	2.3
	Total	**247**	**21.8**
Perceptions of ease/difficulty	Everything	121	10.7
	Nothing	68	6
	Hard/difficult	16	1.4
	Easy	2	0.2
	I don't know	15	1.3
	Total	**222**	**19.6**
In-game activities	**In-game Activities**	**56**	**5**
Miscellaneous	**Miscellaneous**	**92**	**8.1**

	Total	**1131**	**100**

Research Question 2: Learners' Perceptions in Relation to Age, Gender and Levels of Performance on the Game

The results of this research question are based on analysis of data from all 3437 participants across the four countries. Given the variation in hours of instruction and differing starting ages for language teaching across the participant countries, and the fact that the game was designed for learners who had received a specific number of hours of instruction rather than being of a certain age, it was important that the assessment game was appropriate and appealing to learners from across the age range of 7-13 years. This was investigated through a Pearson correlation, to determine whether age at testing was related to how learners perceived the game. The result shows that there was no significant correlation between age at testing and game positivity score ($r=-.020$, $p=.268$), indicating that the game appealed to learners from across the age range.

Turning to gender, the results displayed in Table 8 show that ratings for individual questionnaire items for both boys and girls were high, suggesting that the game appealed to both groups of learners. However, the girls had a significantly higher positivity score overall and they scored higher for each individual question. There were significant differences in responses to questions 1_0 and 1_2 showing that, contrary to some previous literature on digital game-based learning (e.g. McFarlane et al., 2002), the girls found the game more fun to play and were more willing to play the game again than boys were.

Table 8: Means (standard deviations) and Results of Between-groups, Mann–Whitney U Tests, Boys and Girls

	Girls	**Boys**	**Mann-Whitney U**
1_0: The game is fun to play	3.74 (.58)	3.64 (.66)	*u* = 1199952.50 *p* <.001
1_1: The tasks in the language game are easy	2.81 (.83)	2.77 (.90)	*u* = 1147068.50 *p* =.207
1_2: I would like to play the game again	3.77 (.53)	3.62 (.72)	*u* = 1117258.00 *p* <.001
1_3: The game tells me how well I'm doing	3.42 (.74)	3.36 (.81)	*u* = 1214531.00 *p* =.172
Overall game positivity score	3.43 (.47)	3.34 (.55)	*u* = 1283846.50 *p* <.001

In relation to test scores and attitudes to the game, the results in Table 9 show that there was a low albeit significant correlation between individual items and overall game positivity score on the one hand and test scores on the other. This suggests that learners' positive perceptions of the game were related to their actual performance on the test, but only weakly. In other words, learners enjoyed playing the game and were willing to play it again, regardless of how they performed on it.

Table 9: Spearman Correlations for Level 1 and Level 2 Game Scores and Questionnaire Responses

	Level 1 Total score	**Level 2 Total score**
1_0: The game is fun to play	.126**	.060
1_1: The tasks in the language game are easy	.085**	.169**
1_2: I would like to play the game again	.154**	.117**
1_3: The game tells me how well I'm doing	.098**	.090*
Overall game positivity score	.154**	.175**

Note: ** significant at .01 level, * significant at .05 level

Discussion

Our findings can be summarised as follows. 3,437 young learners across four countries (England, Germany, Italy and Spain) perceived the digital game-based assessment tool used in our study as fun to play, worth playing again, helpful for telling them about their progress in language learning and of moderate difficulty level. That they were more likely to judge the game as fun, worth playing again and informative about progress than to judge it as easy

indicates that digital games can be useful for assessing learners on challenging language content without having a negative impact on their sense of enjoyment. Some learners did mention perceived difficulty of the game as an aspect they disliked, but this did not seem to prevent them from being positive about the game overall. It might however indicate that they had been put forward by their teacher to take the test before they had completed the suggested number of hours of instruction, highlighting the need to make clear in the game guidance documents for teachers how the assessment tool should be used.

As discussed in the Literature Review, assessment formats that are perceived positively by young learners are likely to have better reliability and validity than those that they view more negatively. Learners' responses to the open questionnaire items indicated that the elements of fantasy, storyline, emotional engagement, challenge and sense of competence contributed to their positivity towards the game, in line with motivational theories such as flow (Csikszentmihalyi, 2000/1975), Self-Determination Theory (Ryan & Deci, 2000) and a growing body of empirical research.

Contrary to Butler et al. (2014) we found no effect of age on how positive learners were towards the game, with positive implications for its use across the transition point between primary and secondary education. A number of studies (for example, Butler et al., 2014; Paperaastergiou, 2009) have found that serious games are as popular with girls as with boys. In our study, girls were more positive about the game than boys were, perhaps because the language tasks incorporated in it involved the elements of logic, puzzles and skill development identified by Procci et al. (2013) as appealing to females. We found a significant but weak relationship between individual aspects of positivity and overall positivity on the one hand and test scores on the other, suggesting that learners liked the game regardless of their levels of attainment. DGBA may thus offer an important solution to the challenge identified by teachers in the study by Copland et al. (2014), that of meeting the

needs of a range of young learners in mixed-attainment classes. Furthermore, while learners' responses to the open items regarding what they found easy suggested that they were focused on the game as a game rather than as a test, so that it became a form of "stealth assessment" (Shute, 2011), those responses regarding what was difficult did indicate that they realised it was a test. This did not however seem to have a negative impact on their positivity towards the game, as summed up by one ten year old boy. He claimed there was nothing he disliked about it *"because it is like having a test but in a fun way"*.

Conclusions, Limitations and Future Research Directions

At the start of this article we outlined the importance of developing assessment methods which, as well as being easy to use by teachers and at whole-class level, protect rather than diminish motivation, while also providing valid and reliable measurements of learners' attainment. Our findings regarding learners' positivity towards the game irrespective of their levels of attainment suggest that digital game-based assessment offers an important way forward in the field of early language learning, although the small number of items used to assess learners' attitudes towards the game is acknowledged as a limitation of our study, as is the cross-sectional nature of our data. The extent to which the use of DGBA has a positive impact on learners' motivation for language study in the long-term is an interesting question for future research.

NOTE 1: For all open questions, not all learners answered each question, and some gave a response that produced more than one code. Hence the number of coding references may not match the number of respondents.

REFERENCES

Alyaz, Y., Spaniel-Weise, D., & Gursoy, E. (2017). A study on using serious games in teaching German as a foreign language. *Journal of Education and Learning, 6,* 250-264. doi:10.5539/jel.v6n3p250

Annetta L.A., Minogue J., Holmes S.Y. & Cheng M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education 53*, 74–85. doi: http://dx.doi.org/10.1016/j.compedu.2008.12.020

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52,* 1-26. doi: https://doi.org/10.1146/annurev.psych.52.1.1

Boyle, E., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing, 2,* 69-74. doi: https://doi.org/10.1016/j.entcom.2010.12.002

Briggs, S.R., & Cheek, J.M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality, 54,* 106-148. doi: https://doi.org/10.1111/j.1467-6494.1986.tb00391.x

Butler, Y. G. (2009a). How do teachers observe and evaluate elementary school students' foreign language performance? A case study from South Korea. *TESOL Quarterly, 43,* 417-444. doi: https://doi-org//10.1002/j.1545-7249.2009.tb00243.x

Butler, Y. G. (2009b). Issues in the assessment and evaluation of English language education at the elementary school level: Implications for policies in South Korea, Taiwan, and Japan. *The Journal of Asia TEFL, 6*, 1-31.

Butler, Y.G. (2015). The use of computer games as foreign language learning tasks for digital natives. *System, 54*, 91-10. doi: http://dx.doi.org/10.1016/j.system.2014.10.010 0346-251X

Butler, Y. G., Someya, Y., & Fukuhara, E. (2014). Online games for young learners' foreign language learning. *ELT Journal, 68,* 265-275. doi: https://doi.org/10.1093/elt/ccu008

Calvo-Ferrer, J.R. (2017). Educational games as stand-alone learning tools and their motivational effect on L2 vocabulary acquisition and perceived learning gains. *British Journal of Educational Technology, 48*, 401-402. doi: https://doi.org/10.1111/bjet.12387

Chiu, Y. H., Kao, C. W., & Reynolds, B. L. (2012). The relative effectiveness of digital games-based learning types in English as a foreign language setting: A meta-analysis. *British Journal of Educational Technology, 43*, 104–107. doi: https://doi.org/10.1111/j.1467-8535.2012.01295.x

Copland, F., Garton, S. & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly, 48,* 38–762. doi: 10.1002/tesq.148

Cornillie, F., Clarebout, G., & Desmet, P. (2012). Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. *ReCALL, 24*, 257– 278. doi: https://doi.org/10.1017/S0958344012000146

Courtney, L., Graham, S., Tonkyn, A., & Marinis, T. (2015). Individual differences in early language learning: A study of English learners of French. *Applied Linguistics, 6,* 824–847. doi: https://doi.org/10.1093/applin/amv071

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass. (Original work published 1975)

Department for Education (2013). *Languages programmes of study: Key stage 2. National Curriculum in England*. Available at: https://www.gov.uk/government/publications/national-curriculum-in-england-languages-progammes-of-study

Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of diagnostic competence. *Language Testing, 21*, 259-283.

European Commission (2012). *First European survey on language competences: Executive summary.*

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta- analysis of recent studies. *Journal of Computer Assisted Learning, 29*, 207–219. https://doi.org/10.1111/j.1365- 2729.2012.00489.x

Graham, S., Courtney, L., Marinis, T., & Tonkyn, A. (2017). Early language learning: The impact of teaching and teacher factors. *Language Learning,* 67, 922–958. doi: https://doi.org/10.1111/lang.12251

Graham, S., Courtney, L., Tonkyn, A., Marinis, T. (2016). Motivational trajectories for early

language learning across the primary-secondary school transition. *British Educational Research Journal, 42/4,* 682–702. doi:10.1002/berj.3230

Graham, S., Marinis, T., Tonkyn, A., & Courtney, L. (2014). *Primary Modern Languages: The impact of teaching approaches on attainment and preparedness for secondary school language learning*. Final Report to the Nuffield Foundation

Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In J. Enever, J. Moon, & U. Raman (Eds.), *Young learner English language policy and implementation: International perspectives* (pp. 31–42). Reading: Garnet Education Publishing Ltd.

Johnstone, R., Cavani, J., Low, L., & McPake, J. (2000). *Assessing modern languages achievement: A Scottish pilot study of late primary and early secondary pupils*. Available at: https://www.scilt.org.uk/Portals/24/Library/publications/SCILT_2000_AssessingMLAchievement.pdf

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and Higher Education, 8,* 13–24.

Lay, A., Patton, E., & Chalhoub-Deville, M. (2017). A case for the use of the ability-in language user-in context orientation in game-based assessment. *Language Testing in Asia, 7*, 1-17. doi: https://doi.org/10.1186/s40468-017-0045-0

Lister, M. C. (2015). Gamification: The effect on student motivation and performance at the post-secondary level. *Issues and Trends in Educational Technology*. Available at: https://journals.uair.arizona.edu/index.php/itet/article/view/18661/18410

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, Vol. 3: Conative and affective process analysis* (pp. 223–253). Hillsdale, NJ: Lawrence Erlbaum.

McCowan, R.J., & McCowan, S. C. (1999*). Item analysis for criterion-referenced tests*. New York: Buffalo State College

McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games*. Available at: http://www.teem.org.uk/publications/teem_gamesined_full.pdf

McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.

Murphy, V. (2014). *Second language learning in the early school years: Trends and contexts*. Oxford: Oxford University Press.

Nikolov, M. (2016). Trends, issues, and challenges in assessing young language learners. In M. Nikolov (Ed.) Assessing young learners of English: Global and local perspectives (pp. 1-17). *Educational Linguistics*, 25. doi 10.1007/978-3-319-22422-0_1

Nikolov, M. (2017). Students' and teachers' feedback on diagnostic tests for young EFL learners: Implications for classrooms. In M. P. García Mayo (Ed.), *Learning foreign languages in primary school: Research insights* (pp. 249-266). Bristol: Lawrence Erlbaum.

Pallant, J. (2013). *SPSS survival manual : A step by step guide to data analysis using IBM SPSS* (4th ed.). Crows Nest, NSW: Allen & Unwin

Papastergiou, M. (2009). Digital game-based learning in high-school computer science education: Impact on educational effectiveness and student motivation. *Computers and Education, 52*, 1–12. doi: http://dx.doi.org/10.1016/j.compedu.2008.06.004

Procci, K., Bohnsack, J., & Bowers, C. (2011). Patterns of gaming preferences and serious game effectiveness. In Shumaker, R. (Ed.). *Virtual and mixed reality: Systems and applications* (pp. 37-43). Berlin: Springer-Verlag

Przybylski, A.K., Ryan, R.M., & Rigby, C.S. (2009). The motivating role of violence in video games. *Personality and Social Psychology Bulletin, 35*, 243–259. doi: https://doi.org/10.1177/0146167208327216

Romrell, D. (2013). *Gender and gaming: A literature review.* Paper presented at the annual meeting of the AECT International Convention, Hyatt Regency Orange County, Anaheim, CA, Oct 29, 2013 Available at: https://www.researchgate.net/publication/284542074_Gender_and_Gaming_A_Literature_Review [accessed Aug 13 2018].

Ryan, R, M., & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist, 55*, 68–78.

Shute, V. J. (2011). *Stealth assessment in computer-based games to support learning. Computer games and instructio*n. Charlotte, NC: Information Age Publishers. Available at: http://myweb.fsu.edu/vshute/pdf/shute%20pres_h.pdf

Sitzmann T. (2011) A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64,* 489–528. doi: https://doi.org/10.1111/j.1744-6570.2011.01190.x

Sørensen, B.H., & Meyer, B. (2007). Serious games in language learning and teaching – a theoretical perspective. *Situated Play,* Proceedings of DiGRA 2007 Conference, 559-566. Available at: http://www.digra.org/wp-content/uploads/digital-library/07312.23426.pdf

Stott, A., & Neustaedter, C. (2013). *Analysis of gamification in education*. Technical Report 2013-0422-01, Connections Lab, Simon Fraser University, Surrey, BC, Canada, April. Available at: http://clab.iat.sfu.ca/pubs/Stott-Gamification.pdf

Sundqvist, P., & Sylvén, L.K (2014). Language-related computer use: Focus on young L2 English learners in Sweden. *ReCALL, 26*, 3-20. https://doi.org/10.1017/S0958344013000232

Sykes, J. M., Oskoz, A., & Thorne, S. L. (2008). Web 2.0, Synthetic immersive environments, and mobile resources for language education. *CALICO Journal, 25*, 528-546. doi: 10.1558/cj.v25i3.528-546

Szpotowicz, M., & Campfield, D.E. (2016). Developing and piloting proficiency tests for Polish young learners. In M. Nikolov (Ed.) *Assessing young learners of English: Global and local perspectives* (pp. 109-137). *Educational Linguistics, 25*. doi 10.1007/978-3-319-22422-0_1

Wrzesien M., & Raya M.A. (2010). Learning in serious virtual worlds: evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*, 178–187. doi:10.1016/j.compedu.2010.01.003

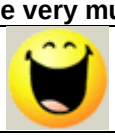
Appendix A: The Questionnaire

Note: Only the section of the questionnaire relevant to this publication is reproduced here for reasons of space. A version was provided in either English, German or Italian for use in each of the countries involved in the study.

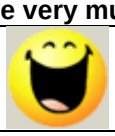
SECTION 1

Please say whether you agree or disagree with these statements. Click a smiley face!

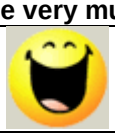
a) *The Language Magician* game is fun to play.

Agree very much	Agree	Disagree	Disagree very much
			

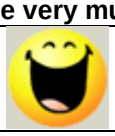
b) The tasks in *The Language Magician* game are easy.

Agree very much	Agree	Disagree	Disagree very much
			

c) I would like to play *The Language Magician* game again.

Agree very much	Agree	Disagree	Disagree very much
			

d) *The Language Magician* game tells me how well I'm doing in (French).

Agree very much	Agree	Disagree	Disagree very much
			

Can you please answer the following questions with a short statement?

1. What did you like most about *The Language Magician* game? Write your answer here.

_____

2. Was there anything that you didn't like about the game? Write your answer here.

_____

3. Which part of the game do you think was easiest?

_____

4. Which part of the game was hardest?

_____

5. Was there anything you have learnt in the game today? If yes, what was it?

_____

Louise Courtney: l.m.courtney@reading.ac.uk

Institute of Education
University of Reading, London Road Campus
4, Redlands Road
Reading, UK
RG1 5EX

Suzanne Graham: s.j.graham@reading.ac.uk (corresponding author)

Institute of Education
University of Reading, London Road Campus
4, Redlands Road
Reading, UK
RG1 5EX