Comparing the Potential of Commercial Off-The-Shelf and Educational Video Games for Adult Foreign Language Education: an Experimental Study

Frederik De Grove, Jan Van Looy, Peter Mechant
IBBT-MICT, Ghent University, Belgium
frederik.degrove@ugent.be
j.vanlooy@ugent.be
peter.mechant@ugent.be

Abstract: The goal of this paper is to explore the experiences evoked by playing a commercial and two digital language learning games. More particularly, it deals with the differences in the playing and learning experiences of adult foreign language learners (N=62). While results of the experimental design suggest that the commercial game evokes better playing and learning experiences, these findings are partly neutralized by the attitude of the participants towards learning through video games and by being a gamer or not. This raises questions as to how video games should look to appeal to a public of learners that is not familiar with gaming in general and with digital game-based learning in specific.

Keywords/Key Phrases: CALL, foreign language learning, game-based learning, game experience, COTS, experimental.

1. Introduction

Using video games for educational instruction has gained importance in theory and practice over the past decade. A recurring question in this field of research concerns what kind of games to use: existing commercial games or special-purpose educational ones (see e.g. Van Eck, 2006). The former often boast significantly larger budgets permitting to invest more in the quality of the experience whereas the latter are more directly aimed at compatibility with predefined learning outcomes. While there is merit in discussing possible dangers and opportunities on a theoretical level, insights gained from a user-centric perspective are largely absent. The aim of this study is to measure and compare how adult foreign language learners experience playing three different games. More concretely, an experimental design is used to compare play and learning experiences of two foreign language learning games and one non-educational commercial game in a foreign language. First, we provide a brief overview of existing literature on game and learning experiences and the impact of external stimuli. Finally, we report on the empirical exploration of the game and learning experiences evoked by playing the three games.

2. Theoretical framework

2.1 Game experience

The concept of game experience has become a major topic of interest when studying video games (see e.g. Trepte and Reinecke, 2011, Gajadhar et al., 2010, Klimmt et al., 2009, Nacke et al., 2009, Nacke and Lindley, 2008, De Kort and Ijsselsteijn, 2008, Gajadhar et al., 2008, Ijsselsteijn et al., 2007, Mäyrä, 2007, Vorderer et al., 2004, Fu et al., 2009, Weibel et al., 2008, Sweetser and Wyeth, 2005). Although the complex nature of experiences makes it difficult to formulate a clear definition, game experience is commonly conceptualized as that which evokes or inhibits enjoyment when playing games (Vorderer et al., 2004, Ritterfeld et al., 2009, Cowley et al., 2008). A regularly used concept on what makes an activity enjoyable is that of flow. Flow refers to "an optimal, intrinsically motivating experience induced by an activity in which one is fully absorbed" (Csikszentmihalyi, 1990). Such an activity is characterized by a balance between challenge and skills, the merging of action and awareness, clear goals and feedback, concentration on the task at hand, control, a loss of self-consciousness and the transformation of time. As noted by Sweetser and Wyeth (2005) these elements strikingly fit the activity of playing video games. It is therefore not surprising that flow or elements connected to flow have been used on a regular basis to explore game experience. While flow elements are recurring concepts in experience research, academic inquiry on the topic is not limited to these experiences. In a study Klimmt et al. (2007) explored the effect...
of control and effectance on enjoyment and found effectance to be an important underlying factor in evoking game enjoyment. Combining flow theory with the technology acceptance model, Holsapple and Wu (2006) explored the antecedents and effects of flow in online gaming. Using the Game Experience Questionnaire Nacke and Lindley (2008) measured how adjusting the difficulty of a level in Half-life 2 influenced several experience dimensions and found that challenge and tension differed significantly between sessions. Using the same questionnaire De Grove et al. (2010b) explored how the game experience of a serious game differed during subsequent design stages and found significant differences for challenge and competence. To our knowledge, however, no research has been performed to explore the differences in game experience between commercial and special-purpose educational games when both are used with an educational purpose in mind. Considering the available academic literature, it is difficult to predict whether a commercial game will evoke better game experiences in a learning environment than an educational one. Learners might feel more attracted to an educational game due to the adapted content and related (learning) expectations as opposed to a commercial game with a more attractive story and appealing audiovisual cues but without obvious learning goals.

RQ1: To what extent do game experiences differ between commercial and special-purpose educational games when both are used in an educational language learning context?

2.2 Learning experience

Educational games are regularly defined as those games of which the primary goal is not enjoyment but education (Susi et al., 2007). This is interesting for two reasons. First, as discussed in the previous paragraph, enjoyment is seen as a core experience of playing video games and, second, academic literature on digital game-based learning in general and on flow in specific conceptualizes learning as an effect of enjoyment or flow (Kiili, 2005, Hoffman and Novak, 2009). As such, enjoyment is conceptualized as the motivational basis for digital game-based learning (see e.g. Garris et al., 2002, O Neil et al., 2005, Squire, 2005, Michael and Chen, 2006). Video games are intrinsically motivating because they are enjoyable. It is this trait that is used as a lever to facilitate learning (Chuang, 2007). Authors like Gee (2003) and Prensky (2003) argue that the motivational nature of video games combined with educational content will make learning more effective. It should, however, be noted that the idea of implicit learning underlies these assumptions. With educational games in general and foreign language learning games in particular, it could be useful to consider the idea of explicit learning. In this vein, it has been argued that a certain amount of conscious attention is needed for successfully taking in and learning formal aspects of a foreign language (Doughty and Long, 2003).

With respect to language learning and video games, DeHaan (2010) found vocabulary recall to be impeded by the extraneous cognitive load evoked by the interactivity of a music game. Hence, when comparing educational games with commercial video games, it is not clear how learning experiences will differ. More concretely, it is unclear whether the investment in the quality of experience of commercial games will lead to better or worse learning experiences compared to video games designed with specified learning outcomes in mind. The former possibly being more enjoyable yet cognitively more demanding and the latter being less enjoyable but offering explicit learning opportunities.

RQ2: To what extent are learning experiences between commercial and special-purpose educational games different when both are used with educational purposes in mind?

2.3 Priming

Little research on digital game-based learning has explored the effect of using an external stimulus on learners during gameplay. Priming is a tool that is widely used to explore the nature of underlying cognitive and linguistic representations (Mayr and Buchner, 2007, Whishaw and Kolb, 1995). The relation between the preceding stimulus (the prime) and the subject’s response enables researchers to make inferences about the nature of the subject’s representations. According to the media priming literature, exposure to media violence can prime subsequent aggressive behaviours, cognitions, or perceptions (see e.g. Roskos-Ewoldsen et al., 2007). Other researchers demonstrated the influence of language priming
In game research, Nelson and Strachan (2009) for example found a priming effect for two different types of video games. They concluded that playing an action video game results in faster reaction times and lower accuracy, while playing a puzzle game results in slower reaction times and higher accuracy. Playing video games can be conceptualized as an automotivational activity (cf. supra). On the other hand, educational settings commonly make use of external stimuli such as grades. Therefore, we explore how learners react to an external stimulus when playing an intrinsically motivating video game.

**RQ3:** To what extent does priming influence the game and learning experience?

### 3. Method

#### 3.1 Design

The study consists of a 3x2 mixed model experimental design (N=62) whereby each participant plays three games (within-subjects factor). Half have been primed with the message that they will have to take a language proficiency test afterwards (between-subjects factor). The games are each played for twenty minutes in random order and before the testing session each participant takes a language test. All participants are adults who are learning German. Game- and learning-related variables are measured by administering a short questionnaire after playing each game.

#### 3.2 Subjects

Participants were recruited through schools offering adult language learning courses for German and among first-year university students with German as a major. To minimize pre-existing attitudes towards video games, the experiment was described as participating in a research project on foreign language learning software. An incentive of 20 euro was foreseen for each participant. Interested learners were asked to fill out their contact details and were invited by e-mail to come to the IBBT Game Lab. In total 62 participants took part of which 42 were female and 20 were male. There were no significant differences between the priming conditions on account of gender ($\chi^2=0.03; df=1; p=.86$). The mean age was 21.45 (SD=4.82) years.

#### 3.3 Games

Three different games were used, one commercial and two foreign language learning games.

**Ausflug am Wochenende nach München** (Ausflug) is an interactive text-based game produced by the Utah State University. It is specifically developed to learn German and was released in 2007. Learners play the role of Karin Moller, a student making a trip to Munchen. The game does not offer graphics or sounds. The game requires input of the player in the form of typed text.

**Who is Oscar Lake?** (Who is) is a point-and-click graphical language learning game released in 1995. It was developed with the aim of improving second language acquisition and is available in several languages such as Spanish, German and Italian. In the game, the player needs to solve a diamond theft. The game makes use of sound (spoken language) and graphics.

**Geheimakte Tunguska** (Geheimakte) is a commercial point-and-click adventure game developed by Animation Arts and released in Germany in 2006. It provides graphics and sound (spoken language). While open for discussion, gameplay, graphics and sound can be considered to be more engaging than that offered by **Who is** or **Ausflug**.
3.4 Measures

3.4.1 Game experience
To measure game experience dimensions, the core game experience questionnaire (GEQ) developed during the FUGA project was used (Poels et al., n.d.). As each participant needed to fill out this questionnaire three times, each construct was measured using only two items rated on 5-point likert scales (totally agree to totally disagree). The selection of these items was based on previous research using the GEQ (De Grove et al., 2010a). Experience dimensions included in the study were positive affect (Chronbach’s $\alpha$ = .85 to .87), negative affect (Chronbach’s $\alpha$ = .85 to .88), immersion (Chronbach’s $\alpha$ = .69 to .83), challenge (Chronbach’s $\alpha$ = .51 to .78), skill (Chronbach’s $\alpha$ = .63 to .84) and frustration (Chronbach’s $\alpha$ = .70 to .83). Sample items were “I felt skillful” and “I felt bored”.

3.4.2 Perceived learning
To account for learning effects, two ad hoc measures were used: active (Chronbach’s $\alpha$ = .69 to .80) and passive learning (Chronbach’s $\alpha$ = .84 to .90). Both constructs were measured using 6 items rated on 5-point likert scales (totally agree to totally disagree). Active learning asks whether respondents have the impression to have learned something on account of speaking or writing German. Passive learning refers to a perceived improvement in understanding and reading a language. Sample items were “playing this game has improved my reading skills” and “my writing skills have improved by playing this game”.

3.4.3 Attitude towards learning games
As it can be expected that the attitude towards the learning tool would influence learning experiences (see e.g. Kirkpatrick, 1998), an ad hoc measure was created to probe the participant’s attitude towards learning a foreign language through video games (Chronbach’s $\alpha$ = .72). The scale consists of three items on a 5-point likert scale (totally agree to totally disagree). Sample items were “I don’t think that video games are suited to learn a foreign language” and “learning a foreign language through video games is pleasant”.

3.4.4 Gaming Frequency
Respondents were asked how frequently they played video games. Those indicating to play less than once a month are considered as non-gamers (N=48). Others are considered as gamers (N=11). A Chi$^2$ test showed that there were no significant differences in distribution of gamers and non-gamers for the between-subjects condition (Chi$^2$ =1.93; df=1; p=.290).

3.4.5 German proficiency
Before playing the games, each respondent filled out a German proficiency test consisting of 40 multiple choice questions ($M_{\text{score}}$=29.8; SD=5.93). This test was provided by a professional language training institute in Flanders. An independent t-test showed that there were no differences between the priming conditions on account of test score ($F=.28; t=-.014; p=.989$).

4. Results

4.1 Game experiences
A repeated measures mixed model (GLM) was used to explore the effect of type of game and priming on each experience dimension. Significant differences between video games were found for negative affect ($F(2,59)=11.30; p<.001$; Partial $\eta^2=.16$), immersion ($F(2,59)=13.16; p<.001$; Partial $\eta^2=.18$) and challenge ($F(2,59)=6.01; p<.005$; Partial $\eta^2=.09$). No significant differences were found for skill ($F(2,59)=.87; p=.42$;
Partial $\eta^2=0.01$, positive affect (F(2,59)=.75; p=.47; Partial $\eta^2=0.01$) and frustration (F(2,59)=.13; p=.88; Partial $\eta^2=0.002$). Moreover, no significant effect of priming was found for any experience dimension. Pairwise comparisons show that the difference for negative affect lies between both educational games on the one hand and the commercial game on the other. For immersion, all three games differ significantly from each other and for challenge, only *Who is* and *Geheimakte* differ significantly. Table 1 gives an overview of mean scores of each game on the different experience dimensions.

**Table 1**: Means and SD for each game on game experience dimensions (letters in superscript indicate pairwise comparisons results).

<table>
<thead>
<tr>
<th></th>
<th>Skill</th>
<th>Challenge (***</th>
<th>Negative Affect (***</th>
<th>Positive Affect</th>
<th>Immersion (***</th>
<th>Frustration</th>
</tr>
</thead>
</table>
| *Ausflug* | 2.72 (.12)
|           | 3.26 (.12)$^{ab}$ | 2.85 (.12)$^{a}$ | 3.01 (.12)$^{a}$ | 2.38 (.10)$^{a}$ | 2.48 (.12)$^{a}$ |
| *Who is*  | 2.54 (.11)$^{a}$ | 3.00 (.12)$^{a}$ | 2.76 (.11)$^{a}$ | 3.06 (.11)$^{a}$ | 2.69 (.11)$^{b}$ | 2.55 (.11)$^{a}$ |
| *Geheimakte* | 2.58 (.10)$^{a}$ | 3.45 (.09)$^{b}$ | 2.19 (.09)$^{b}$ | 3.18 (.10)$^{a}$ | 3.06 (.12)$^{c}$ | 2.54 (.12)$^{a}$ |

When taking into account whether the respondent is a gamer, the effect of type of game disappears for negative affect and immersion. For challenge, a significant effect is found (F(2,59)=6.93; p<.05; Partial $\eta^2=0.10$). This effect also differs between games (F(2,59)=3.61; p<.05; Partial $\eta^2=0.06$). The effect of being a gamer on challenge is stronger for *Ausflug* ($B_{NG}=-0.9; t=-3.46; p<0.001$) than for *Geheimakte* ($B_{NG}=-.48; t=-2.29; p<.05$) and is not significant for *Who is* ($B_{NG}=-.09; t=-.31; p=.75$) (Figure 1). Thus, non-gamers (NG) felt less challenged when playing *Ausflug* and *Geheimakte* than gamers. Moreover, while there are no significant differences for skill or positive affect between games, these dimensions differ between gamers and non-gamers (F_{skill}(1,60)=14.71; p<0.001; Partial $\eta^2=.20$; F_{positive affect}(1,60)=7.58; p<.05; Partial $\eta^2=.11$).
4.2 Learning experiences

Results for learning experiences show significant differences for active \((F(2,59)=15.50; \ p<.001; \ \text{Partial } \eta^2=.21)\) and passive \((F(2,59)=7.98; \ p=.001; \ \text{Partial } \eta^2=.12)\) learning. No significant effects were found for priming. Pairwise comparisons show that, for passive learning, differences lie between the language learning games and the commercial game while for active learning, these differences lie between \textit{Ausflug} on the one hand and \textit{Geheimakte} and \textit{Who is} on the other. Table 2 gives an overview of mean scores for each game on both learning dimension.

\textbf{Table 2:} Means and standard deviations for each game on learning dimensions (letters in superscript indicate pairwise comparisons results).

|               | Active learning (***
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Ausflug}</td>
<td>2.58 (.083)\textsuperscript{a}</td>
</tr>
<tr>
<td>\textit{Who is}</td>
<td>2.97 (.092)\textsuperscript{b}</td>
</tr>
<tr>
<td>\textit{Geheimakte}</td>
<td>3.15 (.085)\textsuperscript{b}</td>
</tr>
</tbody>
</table>

When controlling for possible influences of attitude towards learning games on active learning \((F(1,59)=16.98; \ p<.001; \ \text{Partial } \eta^2=.22)\), previous differences disappear \((F(2,118)=2.48; \ p=.088; \ \text{Partial } \eta^2=.040)\). Whereas this effect is significant for passive learning too \((F(1,59)=18.99; \ p<.001; \ \text{Partial } \eta^2=.24)\), it does not account for all variation between games \((F(2,118)=4.38; \ p<.005; \ \text{Partial } \eta^2=.069)\).
When controlling for the effect of German proficiency, no significant effects are found for active (F(1,60)=.77; p=.38; Partial $\eta^2=.013$) or passive learning (F(1,60)=.23; p=.63; Partial $\eta^2=.004$).

5. Discussion

As expected, the commercial game scored higher on several experience dimensions such as immersion and challenge and lower on negative affect. The text-based game, however, scores equally high on challenge as the commercial one, suggesting that a rich audiovisual environment is not a necessary condition to challenge a player. When looking at the effect sizes, the influence of the type of game is small to moderate for immersion (16%), negative affect (18%) and challenge (9%). An important factor to take into account is the influence exercised by being a gamer or not. Not only did this explain the differences between immersion and negative affect for the different games. It also proved to be a significant factor for skill and positive affect. For adult learners, this raises questions as to the applicability of video games for language learning. While no data are available on the distribution of gamers in the population of language learners, our experience is that a substantial number of them do not play games. Therefore, a central question is what a video game should look like to appeal to gamers and non-gamers alike. With challenge and skill being important dimensions of the game experience, the need for adaptive gaming environments in terms of gameplay can be an important issue (RQ1).

As for learning experiences, some surprising results occurred. Whilst the commercial game has been developed with no educational intentions in mind, it scores highest for passive learning and equally high for active learning as Who is. However, in line with the effect of being a gamer, it is remarkable to see how the attitude towards learning through games influences these learning experiences. When accounting for attitude (with an effect size of 22%), the differences for active learning disappear indicating that neither game- nor content-related characteristics accounted for them. Likewise, there is a substantial effect of attitude (24% of variation) for passive learning. As learners with more or less equal language proficiency skills were recruited on purpose, the non-significant effect of the language skill level on learning effects is not surprising. Moreover, the commercial game seems to evoke slightly better game experiences and a better passive learning experience. These differences, however, are rather small (RQ2). A possible explanation is that learning experiences evoked by a commercial game are of a different nature (implicit) than those evoked by an educational game (explicit). Although mean scores on learning experiences differ little, their antecedents might be different. The available data, however, do not allow us to confirm this hypothesis.

Furthermore, priming has no effect on either game or learning experiences (RQ3). The most obvious explanation for this finding is that the context in which participants played the games influenced the possible effect of the stimulus. As real-life consequences were absent if participants failed the language test afterwards, the power of this external motivator may have been insufficient or even non-existent. Related to this weak motivator is the possibility that participants forgot about the test once they started playing the game. Debriefing interviews seem to confirm this train of thought. With the available data, however, it is not possible to explore to what extent this forgetfulness is to be attributed to intrinsic or external stimuli.

6. Conclusion

The main finding of this study is that there is a significant influence of non-game-related characteristics on the game and learning experience. Despite the variation in the type of games that were used during this experiment, a substantial part of the variation of these experiences is explained by the attitude towards learning with games and being a gamer or not. Considering that the mean age is 21 years this cannot be attributed to the fact that participants were not so-called digital natives. As the learning topics that can be integrated into video games are varied, so is the public of learners. While some authors have enthusiastically embraced the idea of digital-based learning, especially for the digital natives (Prensky, 2005, Prensky, 2003) the empirical findings of this study call for some restraint. It appears that a
substantial part of learners regard the idea of language learning through video games negatively which affects their potential playing and learning experiences.

7. Limitations and future research

This study was cross-sectional in design. None of the participants was familiar with the games that were used. As most games have a learning curve, it is possible that experiences would have been different when participants were allowed to play longer. In a similar vein, studying how game and learning experiences change over a longer period of time in a naturalistic setting could provide valuable information. Moreover, the mean scores of the game experience dimensions suggest that none of the games were experienced as highly enjoyable. Using other games could provide different results. Moreover, learning experiences could have been evoked through different mechanisms: implicit or explicit. Future research should consider taking these different sources of learning into account. With a certain extent of explicit learning being favoured for foreign language learning (cf. supra), it might also be useful to explore which effects are found when using other learning content and other learners.

8. Acknowledgments

The work reported in this article was supported by the Interdisciplinary Institute for Broadband Technology (IBBT) and was conducted in the context of the IBBT-ICON LLINGO project. The language proficiency test used during the study was provided by the Business Language and Communication Centre (BLCC).

References


