XBadges. How soft skills are boosted by video games

Improving persistence, risk taking & spatial reasoning with Flappy Bird, Pacman & Tetris

Sergio Alloza Castillo
of gamification research department
at GECON.es
Barcelona, Spain
salloza@gecon.es

Flavio Escribano
Head of gamification research department
at GECON.es
Barcelona, Spain
fescribano@gecon.es

Abstract—XBadges is a research project based on the hypothesis that commercial video games (nonserious games) can train and promote certain soft skills. With a pre and post test model, Persistence, Spatial Reasoning and Risk Taking were measured before and after in the subjects (15 divided into groups of 5) trained for 2 hours with the video games: Flappy Bird, Tetris and Pacman (respectively). In addition, the emotions of the players were captured too (thanks to a facial and emotional recognition system), to determine if emotions also have something to do with the skills acquisition. The results show how Tetris significantly improves spatial reasoning and how Pacman significantly improves prudence in certain areas of behavior. As for emotions, they reveal that being concentrated helps improve performance and skills acquisition but frustration is also shown as a key element in the process. With the results obtained we can begin to glimpse multiple applications in the education sector and Skilling Development area.

Keywords—Video games; soft skills; training; skilling development; emotions; cognitive abilities; Flappy Bird; Pacman; Tetris.

I. INTRODUCTION

First of all, we need to define what a soft skill is. A lot of authors theorized about that, since there is not an absolute description or classification. Here there are some of the best definitions to get a better understanding about what are the soft skills we are trying to improve through video games:

- Soft skills can be defined as interpersonal, human, people or behavioral skills necessary for applying technical skills and knowledge in the workplace [1].

- Soft skills are a new way to describe a set of abilities or talents that an individual can bring to the workplace [2].

- [3] categorize soft skills as: (1) intrapersonal and interpersonal skills; (2) personal and social skills; and (3) cognitive skills.

- Even there are some researchers that have classified different levels and types of soft skills [4, 5].

But, why we are trying to boost them? Why soft skills?

Why now?

Said in [6]: ‘As organizations become increasingly diverse, the ability to exhibit some soft skills like think critically or decision making with confidence can provide greater opportunities for employment [7, 8]. Communication and soft skills are noted by employers as important skills in the workforce, yet there are highly lacked by recent graduates applying for employment [9, 10, 11].

Employers are looking for methods to reduce the costs of identifying the soft skills through behavioral interviews, but procedures are still expensive and limited and they cannot be used to filter in contexts in which an initial triage of massively received CVs has to be carried out [12].

In the world of formal education the landscape of identification and evaluation of transversal competences is perhaps even more complex than in the labor market since these skills are hardly being evaluated and trained. In most cases, there is no standard and quantitative system available for a teacher that hardly gets to evaluate and train the curricular competencies, it would be very difficult to overload a teaching agenda with an addition in soft skills without increasing too much the teaching load with what that would suppose in costs as much economic as of time.

This is where video games enter into the equation as an identification and training tool. Video games are an element that does not go unnoticed, in fact they have become the most used artifact in the entertainment industry. In addition their presence begins to influence other aspects or areas a priori non ludic.

This study was born from a research project called XBadges, funded by the Ministry of Industry, Energy and Tourism, Government of Spain, 8th in the AEESD call, forming a consortium led by the company COMPARTIA, which subcontracted both GECON.es foundation (gamification experts) and the University of Barcelona (for the creation of an artificial vision system).
The objective of XBadges is to facilitate, with a software, the training and evaluation of soft skills of the users through the use of videogames and to grant certifications. Thus, with XBagdes both the business and education sector would have a tool to meet their current and future needs related to the identification and training of soft skills. Much of the current literature is already investigating the effects of video games on human cognition [13, 14, 15], but very few studies [16, 17, 18, 6] relate these to soft skills.

Specifically, and following an initial review of video games in open source and preliminary soft skills (commented below), the following games and soft skills were chosen hypotheses for research: Pacman (Risk Taking), Tetris (Spatial reasoning) and Flappy Bird (Persistence). These soft skills were chosen based on a review of the competencies most valued by various organizations and institutions [19, 20].

As for the video games, open source video games were needed to embed internal indicators within the code and measure the skills with our telemetry algorithm (actions done by the player). That telemetry was created thanks to a literature review specific to each skill as we show next. These review was also another reason why we chose the mentioned games, since they have the necessary elements to stimulate the skills but they are still simpler than most current video games (so we can link a specific behavior into the game with the skill acquisition).

For tracking persistence with Flappy Bird, as [21, 22] said: "A subject is persistent when, faced with a situation in which it has to emit responses to reach a given solution (reach a score of 20 in Flappy Bird, for example), it maintains a high response rate (the user keeps trying) despite the low frequency of reinforcement (the user keeps dying)". So, in this case, the telemetry was tracking the tries over the time and how far they reach.

In Pacman the telemetry tracks the behaviors of the player that are risky, like being near a ghost, eliminate them when they are vulnerable but the player knows there is only a little time of invulnerability left, etc. So following [23, 24] we can infer that the behavior behind these tracked actions is risk taking (derived from decision making).

Finally in the case of Tetris, we decided to replicate [25] since there are already a lot of research about what Tetris can do to our minds [26, 27, 28] and that study specifically measure spatial reasoning. The telemetry in this game tracked the complementation of Tetris lines and the time between pieces placement, based on the premise that repeated exposure to changing visual patterns in a 2D virtual space (manipulable under rotation and translation) with progressive difficulty increase curve will decrease the time required for information processing, with processing speed being faster and rotation and translation more effective (set of skills that combine spatial reasoning) as the previous authors argued.

The software also captured the emotions thanks to the system of artificial vision. Once a face is detected, emotion recognition is performed in the corresponding bounding box/area of interest. For the recognition of emotions in images, we based on deep learning, in particular we took benefit of the pre-trained VGG convolutional neural network to be fine-tuned on emotions considering annotating public emotion datasets. As a result, a deep learning model was trained, able to recognize face textures representative of the presence of a particular emotion. Thus allow us to see the effect of the emotions of the users in the data and in the competences acquisition.

The emotions we capture are of Joy, Frustration, Concentration and Boredom. The selection of these emotions has been made taking into account the most studied emotions and the research behind the recognition of emotions [29, 30].

The objective of the research is to contrast the following hypotheses: (-1-) -Commercial video games as a pedagogical differentiator element (nonserious games), improve soft skills, (-2-) -The percentage of emotions generated at a general level correlates with the percentages of improvement of users and (-3-) -Emotions generated by users at specific times vary according to the scores obtained from some indicators (for example: completing a line in Tetris and eliminating ghosts in vulnerability mode A and B in Pacman).

II. METHODS

A. Participants

The sample consists of 15 subjects (12 Males and 3 Females), randomly divided into 3 groups of 5 people each (group 1 with FlappyBird, group 2 with Tetris and group 3 with Pacman). We have chosen 5 as minimum number of people for statistical analyzes to be reliable and valid, as indicated in [31] and taking as reference other studies that also have a reduced sample size [32, 33]. The inclusion criteria applied in the sampling is:

- Age between 18 and 50 years.
- Not accustomed to playing the video games of the research or similar (minimum 1 year without previous experience).

The participants were searched through the social networks, the contacts of the researchers and with the collaboration of Yuzz (center of innovation and Coworking).

B. Materials

On one hand, one of the materials we have is the XBadges software. This software has integrated the three video games previously commented with the added telemetry. It also has the aforementioned system of facial and emotional recognition that allows the recording of the emotions of the players while playing.

On the other hand, we list below the standardized tests that
were used as a reference measure to test whether or not there is a real acquisition of the soft skills and thus to verify the validity of the indicators as measuring instruments:

- **Risk Taking - Domain-Specific Risk-Taking (Pacman)**

  Domain-Specific Risk-Taking (DOSPERT) [34] is a psychometric scale that assesses risk taking in 5 different domains: financial decisions, health/safety, recreation, ethics, and social decisions. The subjects rate the likelihood of specific risk activities for each domain. A second and third part of the questionnaire assesses the perception of the risk magnitude of the expected benefits of the activities of the 1st part. The reduced and revised spanish version of 30 items [35] is used.

- **Persistence - Big Five subscale (Flappy Bird)**

  As a personality test, BFQ allows us to observe patterns and profiles of behavior in users. It has multiple questions grouped in different dimensions. Specifically, the dimension "Conscientiousness", which bifurcates in two sub-dimensions: "Scrupulousness" and "Perseverance". Given the purpose of the experiment, we are interested only in the subscale that measures Perseverance (Persistence). Again a spanish version is used [36].

- **Spatial Reasoning – Fibonicci’s Test (Tetris)**

  We have used the web test [37] that was used in the study [25] as a replica, to measure spatial reasoning ability. This test consists of a series of items showing a series of 3D figures displayed and the subject has to choose the option (between 4) of the same figure, but folded.

**C. Procedure**

In the first place, the subjects completed the questionnaires corresponding to each video game (as a pre-test phase), explained in the previous section. Then each participant played the video game that corresponds to his group for 3 sessions of 40 minutes each, sessions distributed at the convenience of the subjects in a maximum period of 1 week, with no possibility of doing two or three sessions in one day. The design of temporality is based on similar study [28]. In the Flappy Bird group the subjects were able to stop playing whenever they wanted after the 20th minute. When measuring Persistence, we had to leave a margin of time in which the subject decided to play or not, since otherwise we would have been skewing the persistence scores. Finally, at the end of the 3rd session, the subjects had to re-complete the questionnaires (as a post-test phase) so we were able to compare the questionnaire results before and after the game training.

**III. RESULTS**

After the analysis of the data obtained, the following results are presented, grouped by hypotheses:

**D. 1º Hypothesis: Video games & soft skills.**

The data are shown in table form, assembled by indicators/telemetry and standardized tests, by videogame:

- **Flappy Bird**

  Sessions data have been modified by removing the high end values (40 minutes) for the "ceiling effect". Some of the data (about 30%) have been provoked by the end of session time, so we cannot infer that these are the times users would adjust if they had more time to play. The final table after clustering and descriptive analysis of the data is as follows "Table 1":

<table>
<thead>
<tr>
<th>Table 1. Total time indicator data of Flappy Bird</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average</strong></td>
</tr>
<tr>
<td><strong>50’</strong></td>
</tr>
<tr>
<td><strong>60’</strong></td>
</tr>
<tr>
<td><strong>70’</strong></td>
</tr>
<tr>
<td><strong>80’</strong></td>
</tr>
<tr>
<td><strong>90’</strong></td>
</tr>
<tr>
<td><strong>100’</strong></td>
</tr>
</tbody>
</table>

Statistical significance of repeated measures ANOVA accepted with epsilon GG adjustment (Epsilon GG= 0.17, F= 10,003, p= 0.025). The results of the standardized test Persistence Big Five subscale "Table 3" are presented, with statistical significance of Student's t-test for paired samples not accepted (t= -1,176, p= 0,152):

<table>
<thead>
<tr>
<th>Table 3. Data of Perseverance subscale of Big Five</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average</strong></td>
</tr>
<tr>
<td>47</td>
</tr>
</tbody>
</table>

- **Tetris**

Data obtained from the Tetris telemetry "Table 4":

<table>
<thead>
<tr>
<th>Table 4. Telemetry data of Tetris</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average</strong></td>
</tr>
<tr>
<td>681,200</td>
</tr>
<tr>
<td>122,486</td>
</tr>
<tr>
<td>755,369</td>
</tr>
<tr>
<td>90,746</td>
</tr>
</tbody>
</table>

Statistical significance of repeated measures ANOVA accepted with epsilon GG adjustment (Epsilon GG= 0.35, F= 10,176, p= 0.025).
6,614, p= 0.020). The results of the standardized test that measures spatial reasoning are shown below "Table 5" with statistical significance of Student's t-test for paired samples accepted (t= -2.449, p= 0.035):

<table>
<thead>
<tr>
<th>Pre phase</th>
<th>Post phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>12.4</td>
</tr>
<tr>
<td>Standard error</td>
<td>1.631</td>
</tr>
</tbody>
</table>

- Pacman

Data obtained from the Pacman telemetry "Table 6 ":

<table>
<thead>
<tr>
<th>10'</th>
<th>20'</th>
<th>30'</th>
<th>50'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>48,930</td>
<td>49,008</td>
<td>54,733</td>
</tr>
<tr>
<td>Standard error</td>
<td>9,350</td>
<td>10,094</td>
<td>9,551</td>
</tr>
<tr>
<td>90'</td>
<td>110'</td>
<td>120'</td>
<td>130'</td>
</tr>
<tr>
<td>755,369</td>
<td>773,241</td>
<td>779,983</td>
<td>792,213</td>
</tr>
<tr>
<td>90,746</td>
<td>107,316</td>
<td>115,917</td>
<td>113,825</td>
</tr>
</tbody>
</table>

Statistical significance test of repeated measures ANOVA not accepted with epsilon GG adjustment (Epsilon GG= 0.22; F= 2,499, p= 0.170).

Regarding the results of DOSPERT (test that measures Risk Taking), no statistically significant differences were found in general or in any of the subscales except for Safety and Health (only in the part of the test that measures probability of behavior), where a significant difference of means was found through the Student's t-test for paired samples (mean pre= 22.6 & mean post= 19; t= 2.882, p= 0.022).

E. 2° Hypothesis: Emotions & improvement percentage.

The correlations between the four emotions (J= Joy, C= Concentration, F= Frustration and B= Boredom) and the percentage of improvement of the three video game indicators (FB= Flappy Bird, T= Tetris and P= Pacman) are shown "Table 7":

<table>
<thead>
<tr>
<th>J</th>
<th>C</th>
<th>F</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>r= -0.86</td>
<td>r= -0.85</td>
<td>rho= 0</td>
</tr>
<tr>
<td>T</td>
<td>rho= 0.3</td>
<td>r= 0.66</td>
<td>r= -0.53</td>
</tr>
<tr>
<td>P</td>
<td>r= 0.16</td>
<td>r= 0.93*</td>
<td>r= -0.77</td>
</tr>
</tbody>
</table>

F. 3° Hypothesis: Emotions & video game indicators.

Next, the averages of the percentages of emotions present in the moments in which the indicated criteria were fulfilled are presented "Table 8 ", following the abbreviations of the previous hypothesis. The indicators are: completing a line in Tetris (reflected in the table as Tetris) and eliminating a ghost in vulnerability mode A and B in Pacman (reflected as Pacman 1 and Pacman 2 in the table):

<table>
<thead>
<tr>
<th>Indicators</th>
<th>J %</th>
<th>C %</th>
<th>F %</th>
<th>B %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetris</td>
<td>7.56</td>
<td>50.32</td>
<td>34.21</td>
<td>7.85</td>
</tr>
<tr>
<td>Pacman 1</td>
<td>5.65</td>
<td>25.36</td>
<td>63.61</td>
<td>5.37</td>
</tr>
<tr>
<td>Pacman 2</td>
<td>4.29</td>
<td>25.23</td>
<td>65.46</td>
<td>5.07</td>
</tr>
</tbody>
</table>

IV. DISCUSSION

Interpreting the results, we can affirm the following premises, again by hypothesis:

G. 1° Hypothesis: Video games & soft skills.

- Flappy Bird

Significant differences were detected in the data obtained through the Flappy Bird indicators, with a total time of 130 minutes of training (F= 10,003; p= 0.025). However, we cannot say that these changes reflect an improvement in persistence capacity outside Flappy Bird, given the nonsignificance changes in the Big Five subscale measures (t= -1,1766; p= 0.152).

In spite of this we can establish new lines of investigation guiding the video game Flappy Bird as a measure of persistence more sensitive than the standardized test itself, since although the change in the questionnaire is not significant, the average of the scores of the same one rises (47 vs 48.8). Under this line we would face a nonsignificance caused by a small sample size, memory bias when repeating the same test in just 1 week, little training time or any other variable outside the game. Supporting this new hypothesis, one of the indicators that theoretically relates more to Persistence, “playing time, number of tries” (response frequency), did show significant changes (t= -2.818; p= 0.033) indicating that players spent more time in the game the longer they played.

- Pacman

As can be seen previously, significant changes have been detected in one of the DOSPERT subscales. In particular, against the approach of the initial hypothesis, there is a decrease in the probability of risky behavior in the area of Health and Safety (t= 2.882; p= 0.022), so we can say that playing Pacman with a training time of at least 90 minutes, increases the prudence in the mentioned area.

- Tetris

Tetris training with a minimum of 70 minutes of play has been shown to significantly improve spatial reasoning ability (F= 6,61449348; p= 0.02). These results fit the replica of the study [25] where they also relate the same video game and
spatial reasoning, obtaining similar results. In addition this research also specifies the improvement effect of Tetris since the sessions have been carried out with a lower sample size in regard to the original study.

We also emphasize in a general way, that not having measured other soft skills, we are leaving aside relations that can be significant. A good way to evolve the research would be to expand the range of capabilities to measure and relate them to different (or the same) video games.

We also discuss the limitations of the memory effect in the tests complementation of the post phase (pre-post test design), the small sample size per group and the short temporal design of experimental sessions, so that the results obtained could be underestimated (statistical error Type II).

\[ H. 2^\text{nd} \text{Hypothesis}: \text{Emotions} \& \text{improvement percentage.} \]

The results obtained regarding the emotions related to the percentage of improvement of the indicators, do not follow the initial approach. In fact, only one correlation of the 12 (4 emotions * 3 games), Concentration & Pacman, is significant (r = 0.93, p = 0.02), showing that the more concentrated Pacman is played, more is improved in the video game.

The value of the correlation is very sensitive to the number of data available to analyze, so if the study had been carried out with a large number of people and therefore, there would be many more data to analyze, the value of the correlations would oscillate as well as their statistical significance, confirming perhaps the initial hypotheses that to more presence of boredom less percentage of improvement, or greater the presence of joy is, greater the percentage of improvement, for example.

\[ I. 3^\text{rd} \text{Hypothesis}: \text{Emotions} \& \text{video game indicators.} \]

To our surprise, joy was not one of the most prevalent emotions when these indicators were met. In particular, in Tetris, when completing lines during the games, the subjects showed high concentration percentages (50,30%) while the other emotions did not have as much presence. An example of what is commented "Fig. 1". In Pacman, in reaching and eliminating the ghosts in vulnerability mode A and B, the prevailing emotion was in both cases frustration (with more than 63% in both cases).

Contrary to expectations, concentration and frustration are present in moments where the user is positively reinforced by the video game. Perhaps we are facing here an implicit relationship between these emotions and the acquisition of skills.

The analysis and global interpretation of the results suggest that video games can be useful tools to enhance or boost certain soft skills, as well as the presence of emotions is closely linked to the motivation of the players and their evolution in the game. With this findings the video games win value as a training tool for soft skills, offering a new tool to the market with possibility of application in multiple sectors.

In the business world, on one hand, providing employees with a tool for identifying and training soft skills required in certain jobs, and on the other, to employers, offering a more automated CV screening tool.

As for the academic world, showing the use of video games as a methodology to enhance soft skills that go unrecognized in most academic curriculum and thus better prepare students to adapt to the context that awaits them.

Another sector where XBadges idea could be applied is eHealth. There are many diseases or pathologies that impair certain skills. Although, in particular, more research is needed in this field, alleviating certain symptoms or improving dysfunctional skills with video games could prove to be an effective method in addition to engage to the patient.

And above all, regardless of the sector of application, XBadges offers information to the population about the positive influence of their play habits on their minds and behavior, since players will continue to play the same, but knowing that they are boosting their abilities.

Acknowledgments. We thank Mercè Muntada (mmuntada@compartia.net) for her support as the leader of Compartia & leader of the project consortium.

REFERENCES


