Assessing the impact of the awareness level on a co-operative game

Teruel, Miguel A.; Condori-Fernandez, Nelly; Navarro, Elena; González, Pascual; Lago, Patricia

published in
Information and Software Technology
2018

DOI (link to publisher)
10.1016/j.infsof.2018.02.008

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Information and Software Technology 98 (2018) 89–116

Contents lists available at ScienceDirect

Information and Software Technology

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

Assessing the impact of the awareness level on a co-operative game

Miguel A. Teruel[a], Nelly Condori-Fernandez[b,c], Elena Navarro[a,*], Pascual González[a], Patricia Lago[c]

[a] LoUISE Research Group, Computing Systems Department, University of Castilla- La Mancha, Avda. España s/n, 02071 Albacete, Spain
[b] LBD Research Group, Department of Computer Science, Universidade da Coruña, Campus de Elviña s/n, 15071 A Coruña, Spain
[c] Software and Services Research Group, Department of Computer Science, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081HV Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Co-operative game
Awareness
Gamespace Awareness
Face analysis

ABSTRACT

Context: When playing a co-operative game, being aware of your collaborators (where they are playing, what they are doing, the abilities they have, etc.) is essential for achieving the game's goals. This led to the definition of *Gamespace Awareness* in order to guide in the identification of the awareness needs in the form of a compilation of the awareness elements that a co-operative game should feature.
Objective: Gamespace Awareness does not establish how much awareness information players must be provided with. This constitutes the main motivation for this work: to assess the impact of different levels of Gamespace Awareness elements on a co-operative game.
Method: A multiplayer action game was developed that supports three different awareness configurations, each one featuring different awareness levels (high, medium and low). The impact of these awareness levels was measured as regards game score, time, players' happiness while playing, enjoyment and perceived usefulness. Several techniques such as subjective surveys and facial expression analysis were used to measure these factors.
Results: The analysis of the results shows that the higher the awareness, the better the game score. However, the highest level of player happiness was not achieved with the most awareness-enabled configuration; we found that the players' enjoyment depends not only on their awareness level but also on their expertise level. Finally, the awareness elements related to the present and the future were the most useful, as could be expected in a multiplayer action game.
Conclusions: The results showed that the medium level awareness obtained the best results. We therefore concluded that a certain level of awareness is necessary, but that excessive awareness could negatively affect the game experience.

1. Introduction

The golden age of video games is back. If the original one took place from 1978 (with the release of Space Invaders [1]) to the mid-1990s [2], it can be said that a new boom is here since the popularization of smartphones [3]. Indeed, looking at the figures, the U.S. video game market has grown from sales of $7b in 2003 to $15.4b in 2014, according to the ESA annual report [4]. The reason for this growth is not only the popularization of smartphones (almost everyone now carries around a portable video game platform) but also because the children of that the first golden age, now adults, are still playing and making the sales figures grow. In fact, in 2015 the average game player was 35 years old, so this is no longer just child's play. Furthermore, forecasts suggest that this trend will continue in 2016 with the revival of Virtual Reality [5,6]. It has been estimated that consumer spending on Virtual Reality hardware and software could reach $21.8b by 2020 [7].

Moreover, not only has the sales figures of games grown during the last years, but also the effort devoted to define and apply new Software Engineering techniques that help to manage the increasing complexity of their development during their whole lifecycle. Ampatzoglou and Stamelos stated this need clearly “software engineering techniques are needed for game development in order to achieve greater flexibility and maintainability, less cost and effort, better design, etc.” [8]. Games are not just projects that finish whenever they are released, but they are products they are evolved with a reduced time to market.

In this new age, not only has the game platform changed, but also the way people play. In the last century, people tended to play alone or with only one partner at a time. Nowadays, thanks to Internet, we are able to play with an almost unlimited number of players at the same time. As a matter of fact, according to the ESA report, 54% of the most

* Corresponding author.
E-mail addresses: miguel@dsi.uclm.es (M.A. Teruel), n.condori.fernandez@udc.es, n.condori-fernandez@vu.nl (N. Condori-Fernandez), Elena.Navarro@uclm.es (E. Navarro), Pascual.Gonzalez@uclm.es (P. González), p.lago@vu.nl (P. Lago).

https://doi.org/10.1016/j.infsof.2018.02.008
Received 20 July 2016; Received in revised form 12 February 2018; Accepted 21 February 2018
Available online 24 February 2018
0950-5849/

frequent gamers play in multiplayer mode at least once a week. Indeed, it is common to play games where players are organized in groups to achieve collaborative game goals such as World of Warcraft's team quests [9]. Nevertheless, as has happened with other serious scenarios [10], these collaborative games have made it clear that there is a need for awareness, i.e. players need to know what is happening in the game space, who is connected, who the collaborators/enemies are, what they can do, etc. In other words, they must be aware of what is going on in the multiplayer game. *Gamespace Awareness* (GA) was developed to identify this awareness by compiling the awareness requirements of collaborative computer games [11] as a set of elements, along with questions that will help developers to identify them. GA has been empirically evaluated by using a survey filed in by 89 fourth-year Computer Science students. In this survey, the participants were asked whether the presence of the awareness information (GA awareness elements) would improve their enjoyment when playing two different games, namely a First-Person Shooter and a Real Time Strategy one. Finally, by analyzing the participants' answers, it was found out that most of the GA elements would improve the players' enjoyment when featured those games. However, it does not establish how much awareness information players should be provided with in order to not only play properly but enjoy the game as well. Therefore, the core contribution of this work is to assess the impact of the awareness level on a co-operative game, that is how much awareness information players perceive as satisfying and useful as well as help them to be effective and efficient.

In order to carry out this assessment we created an *ad hoc* co-operative action game that supports three different configurations of awareness. It is worth noting that co-operative games are a specialization of collaborative games where teams that players belong to are typically temporal [12], being players reassigned to a new team at the beginning of every match. This behavior is typical on trendy best-selling games such as Splatoon [13] or Call of Duty [14]. Each configuration of the developed game has a different awareness level, i.e. a different number of GA elements. The game score, time, player happiness, enjoyment and perceived usefulness of the GA elements of these three configurations were evaluated by means of a controlled experiment carried out in Amsterdam on 14 undergraduate students and replicated in Albacete on 29 undergraduate students to validate the original results. The awareness level was assessed by different empirical techniques such as analysis of game results and post-game surveys, as well as by facial analysis. The results revealed that the higher the awareness level, the better the game results. We also found that a high happiness level depended on both the players' expertise and their awareness level. However, happiness was not correlated with higher awareness. These results are broadly detailed in Section 5.4.

This work is organized as follows: after this Introduction, Section 2 describes *Gamespace Awareness*, the awareness interpretation that have been used throughout this work to evaluate the impact of the awareness levels. Next, Section 3 presents the related works. Section 3.3.3 describes the design of the experiment carried out using a co-operative game. The results are presented in Sections 5. Section 6 outlines our conclusions and future work. Finally, the appendix illustrates the detailed results of the experiment as well as the documents used during the external validation of the game.

2. Background: gamespace awareness

When playing a collaborative game, being aware of your collaborators (i.e. their locations, abilities, status, etc.) is paramount to achieving both your own and the shared game goals. Nevertheless, not only do we have to be aware of our collaborators, but also of information about ourselves, the game scenario, its mechanics and even about our rivals. This constitutes the main motivation of Gamespace Awareness (GA) [11]. GA is a collection of 40 awareness elements aimed at helping game developers and designers to gather together the awareness requirements of collaborative computer games in order to enable players to play together effectively. These elements are as follows:

As can be observed in Table 1, GA elements are classified into 3 different temporal categories (present, past and future) and a non-temporal one related to social and group dynamics. In order to help game practitioners to identify the awareness requirements of a collaborative game, GA also features a set of questions related to each awareness element (Table 1, “Specific questions” column). Finally, it is worth noting that GA does not provide guidance on how to implement each one of its elements, but leaves this decision to the game designers.

3. Related work

As was stated in the Introduction, Games is one of the most challenging and complex domains of software development because they cannot be just considered as finished projects once they are delivered. Rather the opposite, they should be considered as products that must be evolved with a reduced time to market in order to provide players with a unique user experience so that they tie up with the game during its full lifetime. For this aim, the development and assessment of a game should be carefully planned and examined, in order to provide players with the best game experience. Awareness is one of the keys for such success.

Considering that the main aim of this work is to evaluate the influence of the awareness level regarding the game experience, the related work is analyzed in the following two sections from two different perspectives. First, in Section 3.1, the most relevant papers related to the assessment of awareness are presented. Second, in Section 3.2 it is analyzed which approaches and metrics are more widely used for the assessment of games. Finally, in Section 3.3 a summary of the presented proposals is presented, and the novelty of this work is discussed.

3.1. Awareness assessment

As was stated in the Introduction, the majority of the most successful games have awareness as one of their main features because players need to know what is happening in the game space, who is connected, who the collaborators/enemies are, what they can do, etc. In other words, they must be aware of what is going on in the multiplayer game. Multiple studies have been carried out about awareness, providing a constellation of awareness interpretations. Among them, the most widely accepted interpretations are Collaboration Awareness [15], Situational Awareness [16], Workspace Awareness [17], Location Awareness [18], Context Awareness [19], Social Awareness [20], Activity Awareness [21], Knowledge Awareness [22] and Shared-Knowledge Awareness [23]. Other interpretations have also been defined that focus on specific domains but, as far as we know, the only awareness interpretation specifically defined for computer games is Gamespace Awareness (GA) [11], which was already presented in Section 2.

Workspace Awareness is the interpretation most widely used among the computer science community, since it focuses on Computer Supported Cooperative Work systems [24,25]. This has led to some researchers [26] to evaluate the impact of Workspace Awareness elements on a serious collaborative game. Researchers concluded that the teams who used the awareness-enabled version of the game obtained a higher score than those who used non-awareness-enabled versions. However, these researchers evaluated only the effect of awareness on the game score without considering awareness levels. Only other work, Khanal et al. [27], has evaluated the impact of awareness on serious games. Specifically, their aim was to compare a classic non-computer-assisted procedure with a virtual reality (VR) serious game for learning an Advanced Cardiac Life Support procedure. This game had two different configurations. The first configuration featured some limited feedback support (awareness), meanwhile the second one provided full feedback. Khanal et al. concluded that the VR version with full-

Table 1
Gamespace Awareness elements.

Concern		Category	Element	Specific questions
Time	**Present**	Who	Presence	Is anyone in the gamespace?
			Identity	Who is playing? Who does this avatar belong to? Who is available to collaborate with?
			Authorship	Who is doing that?
		What	Task	What are they doing? What is the difficulty of this task?
			Intention	What goal is that task part of?
			Object	What object are they working on? What object can I work with?
			Status	What are the players / avatars status? What are their feelings? What is the objects' status?
			Abilities	What are the avatars' abilities? What are they able to do with such abilities?
			Perception	What are the players perceiving? What can they perceive? (looking, touching, hearing…)
		Where	Location	Where are the players / avatars playing? Is it a physical or virtual location?
			Gaze	Where are the players looking at?
			Reach	Where can the avatars reach?
			Position	Where is an object? How near is it?
		How	Device	How do I use a certain device to interact?
	Past	How	Task history	How did that task happen?
			Object history	How did this object come to be in this state?
		When	Event history	When did that event happen? How often? Is there any network delay?
		Who	Presence history	Who was here and when?
		Where	Location history	Where has a player / avatar been?
			Position history	Where has an object been?
		What	Task history	What has a player / avatar been doing?
	Future	When	Next event	When will the next event happen? How often?
		Who	Next participant	Who will be the next participant?
		Where	Next location	Where will a player / avatar be?
			Next position	Where will an object be?
		What	Next task	What will it happen next?
			Next status	What will the players / avatar next status be? What will the next status of the object be?
			Next abilities	What will the avatar' abilities have? What will the avatars be able to do with such abilities?
Social & Group Dynamics		Who	Members	Who are the members of my group? Has anybody left the group?
			Other members	Who are the members of the other groups?
		What	Belonging	What group do I belong to?
			Role	What are my privileges within my group?
			Others' roles	What are the privileges of my group's members?
			Alliance	What is my relationship with others? (ally, neutral, foe…)
			Exposed information	What do the others know about me?
			Structure	Is there any structure within my group? How is it?
			Group goal	What are the goals of my group?
		When	Next appointment	When do I have to meet with my group?
		How	Inner communication	How should I communicate with each group member?
			Outer communication	How can I communicate with others?

feedback configuration provided a learning experience similar to face-to-face training according to their score.

Directly related to the kind of game we are dealing with in this work, namely co-operative games, several works can be found throughout the literature. It is worth noting that many of the existing proposals are mainly based on design patterns as a manner to pack and express the knowledge gathered by domain experts [28]. For instance, Rocha et al. [29] identified 6 design patterns for co-operative games by analyzing the cooperative game mechanics featured in several commercial games. In order to validate such patterns, they developed a game that featured and evaluated them regarding 2 variables, fun and need for cooperation. They concluded that their patterns are useful to design a fun and cooperative game experience. Seif El-Nasr et al. [30] supplemented such work with 7 more patterns identified by analyzing 14 commercial games. Moreover, they also proposed Cooperative Performance Metrics (CPMs) aimed at analyzing the co-operative nature of the 4 games that they evaluated their patterns with. They finally concluded that the design of effective co-operative patterns will impact the development of both educational and informal games. In order to analyze the aforementioned co-operative game patterns in remote gameplay, Beznosyk et al. [31] performed an experiment where 36 participants had to play remotely without communication. The goal of the experiment was to determine which patterns provided a more enjoyable experience for closely-coupled and loosely-coupled interaction. Besides, not only do they identified such patterns but they also concluded that even though such patterns might have been implemented in several manners, the results of the evaluation may be useful for further research in co-operative game design. Finally, Reuter et al. [12] presented a collection of game design patterns for collaborative players interaction. These patterns are also identified by analyzing several popular collaborative games. As a novelty, they classified the patterns into two dimensions, namely spatial and temporal. In this case, the evaluation is performed by assessing the applicability of such patterns to an already-existing serious game, concluding that players perceive them as appropriate.

Based on the analysis of 40 games, Toups et al. [32] performed a qualitative study to identify which game mechanics players use to communicate different from synchronous verbal communication. Hence, they obtained six trees of co-operative communication mechanic (CCM) techniques, resulting into 9 different mechanics. Nevertheless, no evaluation was performed to validate this proposal. Later, Vaddi et al. [33,34] did perform an evaluation of such CCMs. In this case, the two CCMs implemented in a commercial game were evaluated by means of an experiment participated by 40 subjects (playing in pairs). Such participants played the game by using CCMs, voice, CCMs + voice or none of them to achieve two milestones, being the achievement time the independent value. Authors concluded that the combination of CCMs and voice is paramount to achieve a proper team performance on co-operative games.

3.2. Assessing computer games

This subsection presents different articles related to the assessment of computer games. Different articles about the evaluation of user experience in computer games are explained in Section 3.2.1 describes. Besides, Section 3.2.2 focuses on the assessment of serious games.

3.2.1. Assessing the user experience in games

As any software product, in order to evaluate a video game from the very beginning the quality factors and metrics to be used must be identified [35,36]. When evaluating games, it is common to measure the game User Experience, a.k.a. Game Experience (GX). However, according to Bernhaupt [37], this GX "still misses a clear definition, especially when it comes to trying to measure the concept or related constructs or dimensions". Nevertheless, Bernhaupt [37] states that GX "focuses rather on positive emotions and emotional outcomes such as joy, fun and pride". Therefore, when measuring GX, players' emotions must be taken into account, along with the well-known components of usability, namely effectiveness, efficiency and satisfaction in a particular context of use [38]. Playability defined by Fabricatore et al. [35] as "the instantiation of the general concept of usability determined by understanding and controlling gameplay" is also an important factor for games evaluation. However, as Zhu et al. claim [39] confusion between playability and usability still remain unaddressed, being necessary additional experimentation that resolves their correspondences and/or differences.

Focusing directly on the evaluation of game enjoyment, Mekler et al. [40] systematically reviewed 87 quantitative studies on the enjoyment of video games. This study extracted interesting findings related to the evaluation method. First, 82 out of the 87 studies analyzed used subjective questionnaires to evaluate enjoyment, being this one the only technique used in 75 of them. In those studies revolving around gathering physiological data, the most widely used technique was facial electromyography [41], which measures muscle activity by detecting and amplifying electrical impulses generated by the facial muscle fibers when they are contracted. Other physiological techniques, such as electrodermal activity (measurement of skin conductance), electrocardiography or electroencephalography were also used. According to authors of those papers, these techniques cannot be directly related to enjoyment. However, another study [42] does confirm that these techniques can be used to measure other emotional aspects, like arousal, fear or stress. Cusveller et al. [43] also measured the skin conductance and heart rate of participants while playing three different games. They concluded that not understanding a game does not cause a difference in the stress level. They also found a relationship between stress level and skin conductance, but no relationship between stress level and heart rate was detected. Therefore, they concluded that skin conductance is a technique more appropriate to measure stress.

Other authors have considered the evaluation and development of games from a different perspective, by defining collection of heuristics that help to improve or provide some quality factor. For instance, PLAY [44] was created as a collection of 116 heuristics classified into the following categories: (i) gameplay, (ii) coolness/entertainment/humor/emotional immersion and (iii) usability & game mechanics. However, these heuristics are more related to the game design than to its evaluation, as the authors do not define a method for performing this evaluation. Paavilainen [45] collected a set of heuristics for designing and evaluating social games, but author did not present results of their evaluation. Pinelle et al. have defined another collection of heuristics [46] specifically to evaluate games usability. Authors analyzed 108 existing games from 6 different genres that led to identify twelve usability problems. These heuristics were then defined to guide developers in the evaluation. Authors also present some preliminary results of this set of heuristics that seemed promising.

3.2.2. Assessing serious games

Special attention has received the evaluation of *serious games* [47], that is, games created for achieving a primary purpose other than mere entertainment [48], including gamified applications [49–51], games for the eHealth domain [48,52,53], etc.

Effectiveness, efficiency and satisfaction are also usually evaluated in the context of such games exploiting different metrics. For instance, related to effectiveness, Khanal et al. [27] used *players score* as a metric to evaluate the impact of awareness regarding the learning experience provided by a serious game (see Section 3.1). Also related to effectiveness, Gresse von Wangenheim et al. [54] evaluated an educational game to teach how to measure software. The goal of this study was twofold. First, the authors wanted to assess whether the use of this game improved the learning effectiveness of the participants to define and execute measurement programs of project management. Second, they wanted to evaluate the game itself, i.e. its content, teaching method, duration and engagement. They used pre- and post-tests as the evaluation technique. Although the experiment did not show a significant difference in the improvement of learning, it did confirm that featuring awareness can improve educational games. The experimental results also contributed to the evolution of the game.

Regarding efficiency, by Vallejo et al. [55] have exploited *competition time* to evaluate a serious game developed for assessing the cognitive functions of Alzheimer's disease patients. Authors proposed 6 experimental tasks and measured both the elapsed time to complete them and the success rate, finding significant differences among these tasks regarding such metrics. Finally, the evaluation of the game confirmed it to be useful for evaluation patients cognitive functions.

Concerning enjoyment (*subjective satisfaction*), Fuchslocher et al. [56] used a *questionnaire* to evaluate a game for the intercommunication of teenagers during cancer treatment. During the evaluation, they used a configuration of the game that had explicit references to cancer and another without such references. They concluded that this last configuration provided a higher enjoyment and acceptance results. As regards enjoyment measurement on a gamified application, Herzig et al. [57] applied several game mechanics to SAP [58], the widely-known enterprise resource planning platform. They created a manufacturing scenario covering material management, sales & distribution and production planning. They provided 112 participants with a questionnaire with 162 items regarding the telepresence, interactivity, content and enjoyment constructs. Participants filled in such questionnaire after being exposed to such scenario. Authors concluded that the use of gamification in their prototype improved the analyzed constructs. In other words, gamification increases the quality of the work and improves organizational outcomes.

Besides, happiness (*objective satisfaction*) has been often measured by using Electroencephalogram (EEG) [43] or facial emotion recognition [52]. Regarding the former, Charisis et al. [59] proposed a serious game whose users have to reach and sustain certain affective states. In this game, authors use three configurations with different images to evoke affective states such as happiness, tenderness, sadness, fear and anger. However, authors did not provide experimental results, yet they claim that the game has received positive feedback from its users.

Finally, other factor also considered while evaluating serious games is *usefulness*. For instance, Wrzesien and Alcañiz Raya [60] used usefulness to evaluate a serious virtual world virtual class regarding a traditional classroom. Usefulness was evaluated by using a questionnaire and the results showed that the participants considered the virtual world artifacts more useful than the real-world ones. In addition to the constructs state above, Herzig et al.'s [57] also evaluated the perceived usefulness using the questionnaire offered by the evaluation model that they based their work on.

3.3. Summary and discussion

This subsection presents different game mechanics proposals and their comparison with regarding GA. Moreover, several evaluation metrics used in different works are also illustrated. Finally, the metrics used in this study are summarized in Section 3.3.3.

Table 2
Comparison among works related to awareness assessment.

Authors	Year	Proposal	Elements	Items per element	Games to evaluate the proposal	Participants
Rocha et al. [29]	2008	Design patterns for co-operative games	6	Name, description and implementation in commercial games	1 ad-hoc game	16
Seif El-Nasr et al. [30]	2010	Design patterns for co-operative games	7	Name, description and implementation in commercial games	4 commercial games	60
Beznosyk et al. [31]	2012	Evaluation of previous patterns	–	–	6 ad-hoc games	36
Reuter et al. [12]	2014	Design patterns for collaborative player interactions	9	Category, name, description (and examples), consequences, how to use the pattern and relations with other patterns	1 ad-hoc game	–
Toups et al. [32]	2014	Co-operative communication mechanics	10	Category, name, description and implementation in commercial games (exemplar mechanics)	–	–
Khanal et al. [27]	2014	Advanced Cardiac Life Support training game	–	–	1 serious game	148
Vaddi et al. [33,34]	2016	Evaluation of Co-operative communication mechanics	2	Name, description and implementation in the evaluated game	1 commercial game	40
Teruel et al. [11]	2016	Awareness interpretation for collaborative games	40	Category, name, specific question to identify the awareness need and implementation in commercial games	2 commercial games	89

3.3.1. *Game mechanics proposals and gamespace awareness*

Table 2 shows some of the most relevant proposals about game mechanics including also GA, the one this work is based on. Considering such proposals, several conclusions can be drawn. First, GA is the only one that has been developed by synthetizing different awareness interpretations from different domains, including digital entertainment, not only analyzing existing games. Second, it comprises a larger number of awareness elements for co-operative games than the already existing ones, thus enabling game designers to wonder more awareness-related questions while creating a game. Third, GA is the second that has been validated with a higher number of participants, only surpassed by Khanal [27], although the main goal of such proposal was not to measure the impact of awareness levels. Fourth, it is worth noting that the only two proposals considering the temporal dimension are Reuter et al.'s [12] and GA. However, the authors of the former only take into account the duration of the interaction, instead of when it was (or will be) performed. Furthermore, none of the proposals distinguishes a non-temporal category regarding social aspects of the game. Last but not least, it must be highlighted that GA has been defined to be used in the requirements stage of the game development process rather than during the design, where the analyzed proposals of game mechanics are usually applied. Indeed, it is complementary to the existing proposals.

It must be also emphasized that this work revolves around the concept of awareness level, that is, a collection of awareness elements that are offered to players through a specific configuration of a game (see *The Game: Tanks! Mod featuring Gamespace Awareness in* Section 4.2.2). As far as we know, this work along with that presented by Khanal et al.'s [27] are the only ones that evaluate which level of awareness should be used while creating a co-operative game. Nevertheless, in [27] only two configurations are used (full and basic awareness), thus skipping the configuration for medium awareness. Moreover, their aim was not to evaluate the impact of the awareness level, but to compare the learning experience provided by a virtual reality environment regarding to a traditional one. Besides, in such work, only the score was used as evaluation metric. The rest of analyzed proposals focused either on the assessment of the awareness elements identified in the proposal (Section 3.1) or the assessment of games (Section 3.2) but not on the awareness level to be provided.

3.3.2. *Evaluation metrics*

As far as evaluation metrics are concerned, several approaches and metrics have been used in the literature to assess games (Section 3.2). Table 3 presents a summary of the analyzed proposals along with the metrics used for the evaluation. It is worth noting that this paper has also been included in this table for comparison purposes. Moreover, Table 3 also includes the number of different game configurations that were used to evaluate each game. In this sense, our proposal and Charisis et al.'s [59] are the ones dealing with more game configurations. Looking at Table 3, it can be stated that the exploitation of score, time, enjoyment, happiness, and usefulness have been already used for evaluating games (either serious or entertainment-related ones) in previous works. Therefore, their usage emerges as appropriate in the context of this work.

3.3.3. *Summary of elements used in this paper*

Summarizing this section, Table 4 depicts the different metrics that are used throughout this paper. Moreover, the dependent variable that these metrics are related to, as well as the artifact and procedure used to their measurement are also specified in this table. Additional explanation about these elements is presented in Section 4, where the experiment conducted is presented.

4. Assessing the impact of gamespace awareness on collaborative games

As awareness is considered a cornerstone in the development of

Table 3
Works using different evaluation metrics.

Authors	Year	Domain	Metrics	Configs.
Gresse von Wangenheim et al. [54]	2009	Software measurement	Learning effectiveness	1
Wrzesien and Alcañiz Raya [60]	2010	Serious Virtual World	Usefulness, enjoyment, intention to participate, educational value, engagement, motivation	1
Fuchslocher et al. [56]	2011	Intercommunication of teenagers during Cancer treatment	Enjoyment, affect, flow, immersion	2
Herzig et al. [57]	2012	Gamification	Flow, perceived usefulness, enjoyment, ease of use	1
Khanal et al. [27]	2014	Advanced Cardiac Life Support	Score	2
Charisis et al. [59]	2015	Reaching and sustaining affective states	Happiness, tenderness, sadness, fear, anger, valence, arousal	3
Vallejo et al. [55]	2017	Alzheimer's disease	Time, inactive time, percentage of success	1
Teruel et al. (this work)	–	Co-operative action game	Score, time, happiness, enjoyment, usefulness	3

collaborative games [46,61], it seems only reasonable to expect that the more awareness the players are provided with, the better the *Game Experience* (GX). However, is this statement just an opinion or a fact? In order to answer this question an experiment was conducted whose goal was defined by the Goal-Question-Metric template [62] as follows: *analyze* the awareness level *for the purpose of* assessing its influence *with respect to* the GX *from the viewpoint* of game researchers *in the context of* undergraduate students.

4.1. Research questions

Although GX has no clear definition, it involves players positive emotions [37], so that it is usually assessed by the typical usability components (effectiveness, efficiency and satisfaction) will be used, but considering players' emotions as well. According to Bevan et al. [38], satisfaction is defined in the new version of ISO/DIS 9241-11 [63] as "the extent to which attitudes related to the use of a system, product or service and the emotional and physiological effects arising from their use are positive or negative" [38]. Therefore, we considered the importance of the emotions [38], and especially the positive ones when a game is evaluated [37]. We assessed user satisfaction by two different methods: i) a survey to obtain user satisfaction subjectively by asking the participants about their enjoyment sensation after gameplay; ii) face analysis software was also used to this factor objectively. This technique is reliable because users can hardly hide their real impressions during gameplay, since they are totally focused on the game. We also measured the usage and perceived usefulness of the awareness elements of the game to complement this GX analysis. This led us to define the following research questions that drove the design of the experiment:

- RQ1: How does the awareness level affect the *effectiveness* of players when playing a co-operative game?
- RQ2: How does the awareness level affect the *efficiency* of players when playing a co-operative game?
- RQ3: How does the awareness level affect the *satisfaction* (objective and subjective) of players when playing a co-operative game?
- RQ4: How does the awareness level affect the players' *perceived usefulness of the awareness elements* when playing a co-operative game?

In order to answer these questions, a metric was identified for each one as follows:

- Players' *score* was considered to assess the effectiveness of the awareness level (RQ1).
- Competition *time* was measured to assess the efficiency of the awareness level (RQ2).
- Players' *enjoyment* was surveyed to assess the subjective satisfaction of the awareness level (RQ3).
- Players' *happiness* was analyzed to assess the objective satisfaction of the awareness level (RQ3).
- Awareness elements' *usefulness* was evaluated to assess the perceived usefulness of the awareness elements i.e. how useful the players consider these elements (RQ4).

Fig. 1 summarizes the different metrics that this experiment took into account. In view of the above definition, the main features of the experiment, which was carried out following the guidelines proposed in [64] and [65] are shown in Table 5.

4.2. Design

This subsection will present the experimental design. Hence, Section 4.2.1 will briefly introduce such design and next, Section 4.2.2 will present the experimental design main steps in detail.

4.2.1. Experiment overview

The experiment took place in Amsterdam (The Netherlands) in December 2015 and was replicated in Albacete (Spain) in January 2016. Roughly, our experiment consisted in measuring the impact of awareness level on Game Experience. With this aim, Tanks!-Mod was developed to offer three different configurations: low, medium and high-awareness. The experimental subjects were divided in groups of 4 people each one and played in 2-player teams. The flowchart of this experiment can be seen in Fig. 2. In the following, each of the

Table 4
Elements that are being used for this research.

Metric	Dependent variable	Artifact	Measurement procedure
Score	Effectiveness	Video recording of gameplay	Observation of the final results after the 1st, 2nd and 3rd match
Time	Efficiency	Video recording of gameplay	Observation of the elapsed time after the 1st, 2nd and 3rd match
Happiness	Satisfaction	Video recording of participant' faces	Automated facial expression analysis with 3rd party software
Enjoyment	Satisfaction	Post-test questionnaire	Rating of the 3 different game configurations played
Usefulness	Perceived usefulness of the awareness elements for achieving the game goal	Post-test questionnaire	Rating of the awareness elements included in the game

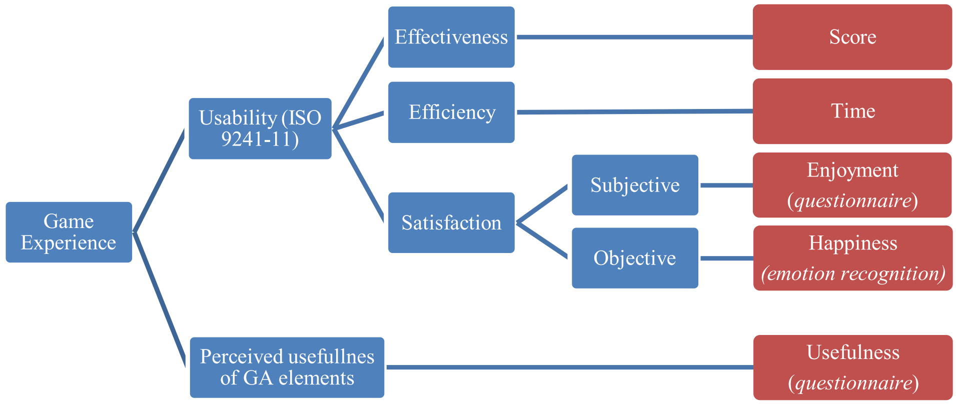


Fig. 1. Quality Factors (dependent variables) and Metrics used in the experiment.

Table 5
Main features of the experiment.

Null hypothesis	H0A: The different levels of awareness have the same result for the game score of the participants. H1A: ¬H0A H0B: The different levels of awareness have the same result for the time of the game competitions. H1B: ¬H0B H0C: The different levels of awareness have the same result for the enjoyment of the participants. H1C: ¬H0C H0D: The different levels of awareness have the same result for the happiness of the participants. H1D: ¬H0D H0E: The different awareness elements are perceived equally useful by the players to achieve the game goal. H1E: ¬H0E	
Dependent variables and related metrics	Effectiveness: Game score (Sco) Efficiency: Game time (Tim) Satisfaction: Players' enjoyment (Enj) Satisfaction: Players' happiness (Hap) Perceived usefulness of the awareness elements for achieving the game goal: Usefulness (Use)	
Independent variables	The action game Tanks!-Mod. Three treatments: low-awareness level (c1), medium-awareness level (c2), high-awareness level (c3) Players expertise level (beginners, normal, advanced)	
Locations	Vrije Universiteit Amsterdam (The Netherlands)	University of Castilla – La Mancha (Albacete, Spain)
Date	December 2015	January 2016
Subjects	14 Computer Science students	29 Computer Science students

experiment steps will be presented.

4.2.2. Experiment's main steps

The different steps in the experiment will be presented in this section. Regarding the pre-test stage, at the beginning of each experimental session, the instructor asked the participants to sign an informed consent form and introduced the experiment with beamer slides. The game used in this experiment was implemented to be as straightforward as possible. Therefore, in order to make players aware of the game mechanics in advance, an introductory session was performed. In this session, the game controls, goal and mechanics were carefully explained. Moreover, demonstrative videos of game sessions were shown to make the players familiar with the look-and-feel of the game. First, the instructor explained the distribution of the teams, then described the game mechanics, including the goal, controls and power-ups and, finally, answered any doubts. As two participants did not show up in Amsterdam and one in Albacete, we used artificial-intelligence-controlled tanks in those game sessions with absent participants.

As far as the test stage is concerned, the game sessions lasted around 19.5 min (6.5 min per game configuration) and no participants dropped out. Moreover, during these 19.5 min, the participants played a minimum of 15 game rounds (an average of 21 rounds per player), taking into account that each awareness configuration was played until one team achieved 5 victories.

As for the post-test stage, at the end of the game sessions, the participants filled in the survey in around 5 min each. At the end, the participants received a reward. Those from Amsterdam received a snack, a drink and a free digital copy of the game. In Albacete they were given a credit of 0.25 points towards the final marks of the Software Engineering II course. We explained to the participants that the rewards would be given regardless of the game score or performance in order not to affect the experiment's results.

Game score (*Sco*), game time (*Tim*), players' enjoyment (*Enj*), players' happiness (*Hap*), and usefulness (*Use*) were the dependent variables. The *score* and *time* of each game competition was recorded and analyzed with a between-subject design. During the first game competition, the players in Team 2 would have a technical advantage, since their awareness information would be far better than that provided to Team 1 (Table 6). In the second competition, both teams had the same awareness information and Team 1 had an advantage in the third. The relationship among the metrics and the research questions stated in Section 4.1 is shown in Table 7.

The Game: Tanks! Mod featuring gamespace awareness: In order to measure the impact of GA on player experience while playing a cooperative game, it was decided to develop an *ad hoc* game.[1] However, instead of developing a game from scratch, we used one that had been developed during the Unity Europe Conference in 2015 [66]. *Tanks!* was initially a local multiplayer non-collaborative comic-style action game with a tridimensional orthographic view (Fig. 3). This game was chosen for two reasons; first, its source code is public so we could create a game mod freely [67] and second, it was created with Unity [68], a powerful multiplatform game engine with which we were familiar.

Tanks! mechanics were really simple. Each player controlled a tank and fired shells at the enemy until its health level was zero. The game was over when a player won 5 rounds. However, it could only be played on one system among multiple players, i.e. collaboration was quite limited, and awareness of other players was limited to the health level and the players' name. We therefore extended the game so that it could be played by up to 8 players using the Unity UNET services [69]. This

[1] The game (binaries and source code) can be downloaded from http://www.materuel.com/Tanks/.

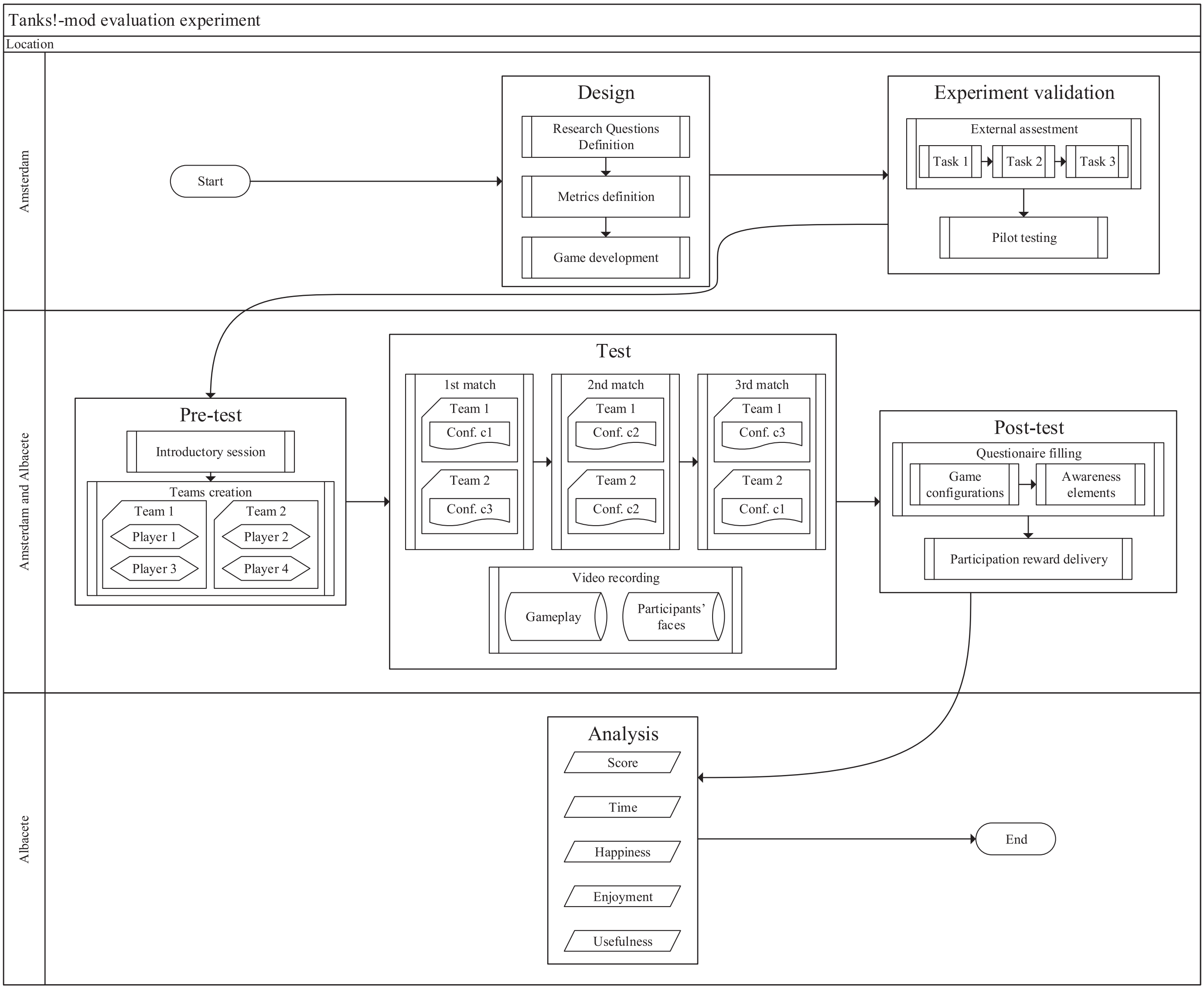


Fig. 2. Experiment flowchart.

new version, called *Tanks!-Mod,* was developed as a serverless network game with a matchmaking service to create game competition regardless of the players' location. It was also enriched by adding several power-ups to improve the players' tanks as follows:

- *Hearth*: Tank health is restored
- *Yellow gas tin*: Firing range is doubled for 15 s.
- *Red gas tin*: Speed is doubled for 15 s.
- *Blue gas tin*: Fire power is doubled for 15 s.
- *Green gas tin*: Invincibility for 15 s.

These power-ups are activated when tanks run over them and their effects are cumulative (i.e. a player can have several active power-ups at the same time). These power-ups are randomly placed in the game world every 30 s. Finally, regarding the winning condition, the original one (be the first player to win 5 rounds) was adapted to the new co-operative functionality. Therefore, a team will win the game if they win 5 rounds. It is worth noting that is not necessary to win these 5 rounds sequentially. Moreover, if there is a tie, none of the teams will be considered as the round winner.

Finally, two additional features were developed for *Tanks!-Mod.*

Table 6
Distribution of players, teams, tanks and game configurations.

Player	Player 1	Player 2	Player 3	Player 4
Team	Team 1 (blue)	Team 2 (red)	Team 1 (blue)	Team 2 (red)
Tank				
Sequence in which configurations are played	c1, c2. c3	c3, c2, c1	c1, c2, c3	c3, c2, c1

Table 7
Relationship among metrics and research questions.

Metrics	Research question
Game score (Sco)	RQ1: How does the awareness level affect the effectiveness of players when playing a co-operative game?
Game time (Tim)	RQ2: How does the awareness level affect the efficiency of players when playing a co-operative game?
Players' enjoyment (Enj)	RQ3: How does the awareness level affect the satisfaction (objective and subjective) of players when playing a co-operative game?
Players' happiness (Hap)	RQ3: How does the awareness level affect the satisfaction (objective and subjective) of players when playing a co-operative game?
Usefulness (Use)	RQ4: How does the awareness level affect the players' perceived usefulness of the awareness elements when playing a co-operative game?

Fig. 3. Screenshot of the original *Tanks!* Game.

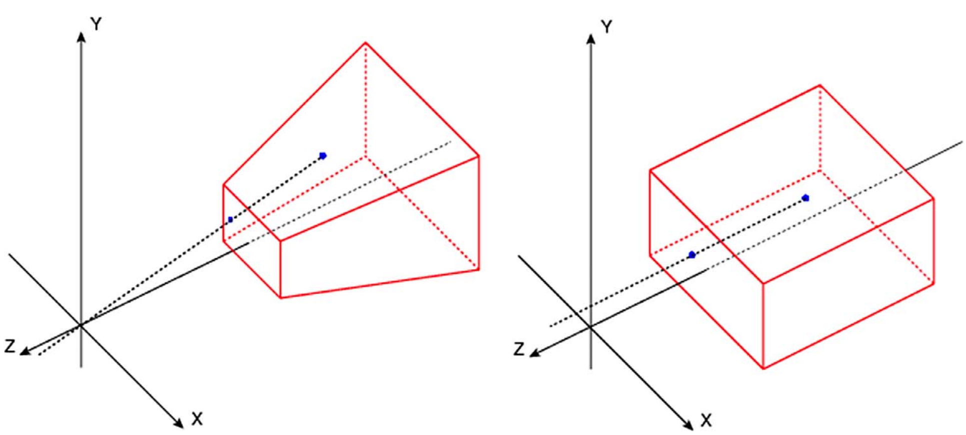


Fig. 4. Perspective (left) and orthographic (right) cameras.

First, the orthographic aerial camera (Fig. 3) was replaced by a third-person perspective (Fig. 5) to immerse the player deeper into the game. Fig. 4 shows a comparison between the game cameras. Secondly, an artificial intelligence system was implemented so that competitions can be created without requiring the same number of players as tanks. The difficulty level of the computer-controlled tanks can be set to *easy, normal* or *hard,* according to the players' expertise. A screenshot of *Tanks!-Mod* can be seen in Fig. 5.

To detect any possible flaw, it was pilot tested by four post-doctoral students of the Vrije Universiteit Amsterdam. Although these participants did not find any issue regarding the experiment itself, they considered that the physical effects of being impacted by a shell were irritating, so it was decided to reduce the effect of the explosions. Indeed, this was a collateral effect of changing the game view from 2D to 3D. In the original game, when the shell impacted the tank, it was rotated and moved a few meters away from the position where the impact was received. This fact barely changed the orthographic aerial camera of the 2D. However, using a third-person perspective camera, the rotation of the tank caused by the shell impact produced a rapid movement and rotation of the camera rapidly, making several players feel dizzy and uneasy. Several other minor bugs involving collisions with game objects were also solved after analyzing the videos of the recording sessions.

As stated before, the independent variable of the experiment will be

Fig. 5. Screenshot of the new *Tanks!-Mod*.

the game configuration being played, i.e. the awareness level. An *awareness level* can be defined as the combination of awareness elements that are enabled in a specific configuration of a game. In our case, three different awareness levels were considered, namely *low, medium* and *high*. These levels were defined as game configurations c1, c2 and c3 of Tanks!, respectively. The *low-awareness* configuration (c1) included only the essential GA elements required to play. For the *medium-awareness* configuration we had the invaluable help of a professional game studio, *Radical Graphics*, whose developers defined the optimal combination of GA elements (see *External assessment of the game in* Section 4.2.2 to obtain more details regarding the involvement of this company in the game). Using this combination, we obtained a medium-awareness configuration optimized by professional game developers for the likely players of our game. It is worth noting that these developers had never played the game before, but they had developed similar types of game. They also checked that each GA element was implemented in the right manner. For the *high-awareness* configuration (c3), all the possible GA elements were enabled (i.e. all those implemented as described in *External assessment of the game in* Section 4.2.2). All in all, c1 includes just the minimum awareness to play the game properly, c2 has the optimal awareness for beginner players according to Radical Graphics and c3 includes all the awareness elements that could be implemented for our game. The elements implemented in each configuration of the game are shown in Table 8. Furthermore, Fig. 6 shows different screenshots of the game where each awareness level is being applied.

Finally, in order to activate the different awareness level, the game features a setup system based of XML configuration files. Fig. 7 shows an example of one of these files. As can be seen, there are four different XML elements corresponding to the four Gamespace Awareness concerns. The different awareness elements implemented in the game are represented as attributes of such XML elements. Moreover, the XML file *TanksMod* also includes the following set of attributes:

- *Configuration*: it is used to display on the screen textual information regarding the game configuration that is being used, as well as other debugging information.
- *Tanks*: it indicates the number of tanks that are going to participate in each match. If the number of players is lower than such number, the game will fill the remaining player slots with AI-controlled tanks.
- *Difficulty*: in the event that AI-controlled tanks were deployed into the game field, this attribute will configure them. Hence, this attribute defines the number of bullets per second that AI-controlled tanks will fire (easy = 2-s-wait between firing bullets, normal = 1-s and hard = 0.5-s).

External assessment of the game: Although most of the GA elements were implemented in *Tanks!-Mod*, some were left out as unsuitable. For instance, *Next participant* was not implemented, since all the players started the game simultaneously. However, in order to assess the suitability of the GA elements, we requested the assistance of the Dutch game studio, Radical Graphics.[2] Moreover, they also defined game configuration c2 (see Table 8). Details about the meeting are presented in Table 9.

This studio was chosen because its members have years of experience on developing games. They have participated in the development of top-selling mainstream games such as Killzone, Gears of War, Mass Effect or League of Legends. The meeting started with a 20-min presentation about Gamespace Awareness and Tanks!-mod in order to make participants aware of what we were working on. Next, the rest of the meeting was performed in a structured manner. In that sense, we proposed three different tasks to the participants from Radical Graphics. In each task, they had to answer one of the following questions (the documents we used to capture the task's results are presented in Appendix C):

- Task 1 – Optimal number and combination of awareness element: What GA elements would they activate in the game for each type of player regarding his / her expertise (beginners, normal and advanced players)?
- Task 2 – Not implemented awareness elements: Should we add any new GA elements or features not included in the current release?
- Task 3 – Alternative implementation of current awareness elements: What is the suitability of the GA elements, i.e. would they have implemented any of them in a different manner?

In answer to the first question, regarding the activation or deactivation of GA elements according to players' expertise [70], they created

[2] http://www.radicalgraphics.nl/.

Table 8
Summary of the implemented GA elements in each game configuration.

Game configuration	Presence	Identity	Authorship	Task	Status	Abilities	Location	Gaze	Reach	Position	Device	Task History	Event History	Next Event	Next Abilities	Members	Other Members	Belonging	Group Goal	Inner Communication	Outer Communication
c1		X																			
c2		X		X	X	X	X	X		X				X		X	X	X	X	X	
c3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

three different configurations of GA elements (i.e. the combination of element to be activated simultaneously) for each type of player. In our experiment, configuration c2 (see Section 4.2) is based on the combination of GA elements for beginners, since the participants in our experiment had never played the game before. As far as the implementation of the awareness elements is concerned, i.e. how to visualize the awareness information [71], they stated that they would not have implemented any of the current elements in a different manner.

Regarding the addition of new awareness elements (second question), they did consider that the *Presence* GA element should be implemented. Since there was no communication among players, they reckoned that a chat would be paramount to playing collaboratively. This led us to implement *Inner communication* and *Outer communication*. The GA elements finally implemented in the game can be seen in Table 10. However, for this experiment, three different GA configurations were created to evaluate the impact of the awareness level. This distribution of awareness elements per configuration is shown in Fig. 9 (Section 4.2).

As far as the third question is concerned, the correctness of the implementation of each GA elements was widely discussed. Indeed, not only we assessed the correct implementation of each GA element (which was the main goal for the third question), but we also checked their layout over the user interface. As a result, two GA elements were affected. First and foremost, the Status element was supposed to be implemented as a bar in the bottom-left corner of the screen. However, they considered that it was not totally adequate for a fast action game, because it was forcing the players to constantly look at the corner of the screen and, therefore, leaving unattended the most relevant area (screen's center). This made us re-implement and relocate the health bar around the players' tanks. The same happened with the abilities element, which was supposed to be located in the same corner (over the health bar), so it was placed over the players' tanks (see Table 10).

Subjects description: The participants in this experiment were Computer Science students who were used to playing computer games. However, to ensure that they had at least the basic experience with games similar to the one being evaluated, we asked them about their experience in their experiment registration form and used this information to create more balanced game competitions to avoid opposing weak players with experienced ones. When playing the game, these participants would be distributed into 2-versus-2 teams of 4 players. As Table 6 shows, for the players in Team 1 the awareness level was raised with each competition, with the opposite for players of Team 2, to cancel out the learning effect.

We then proceeded to recruit experimental subjects by a number of different methods. In Amsterdam, we resorted to flyers and posters, advertisements on the noticeboards of the Faculty of Sciences of the Vrije Universiteit Amsterdam, and the students' mail distribution list. In Albacete we exploited the e-learning system of the University of Castilla – La Mancha. In order to sign up for the experiment, each participant had to fill in an online registration form specifying their expertise in computer games, as well as the time frame they had available. In this way we acquired 14 participants in Amsterdam and 29 in Albacete, respectively.

Environment: For the hardware environment, the game was run on high-end computers (Intel Core i7 4790K, 16GB RAM, NVidia GeForce 970) so that the computer environment did not harm the player enjoyment. The computers were arranged in a single row to facilitate interaction between the instructor and subjects, but separated by screens to avoid interaction among the participants (see Fig. 8).

Questionnaires: For game *enjoyment* and *usefulness*, the participants were asked to fill in an online questionnaire. In this survey, the participants rated the three different game configurations to evaluate enjoyment (*Enj*) on a scale (see Fig. 9) of one star as minimum and 5 stars as maximum. This first part of the survey is related to research question RQ3. This 5-star scale was chosen because it is typically used for

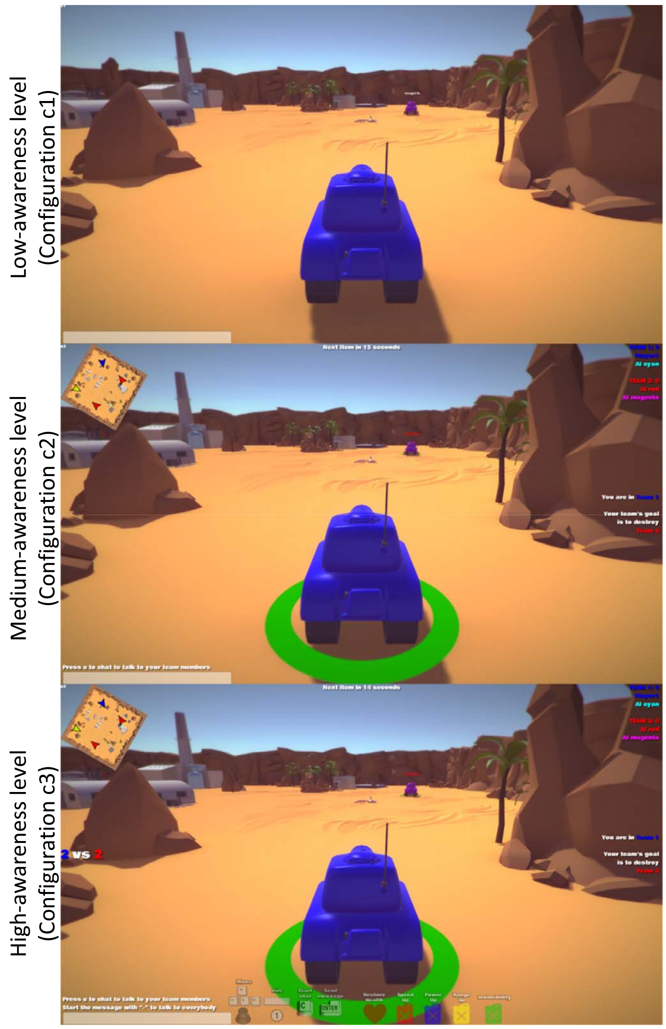


Fig. 6. Screenshots of an initial state of the game featuring different awareness levels.

evaluating and reviewing entertainment products (e.g., movies, books, games, etc.), and therefore, participants should be familiar with it.

We asked the participants to rate each configuration according to the order in which it was played, instead of using a specific name to avoid any potential source of bias. This method was selected instead of others, such as SUS questionnaire [72], because they evaluate non-fun-oriented software artifacts so that they do not have references to user enjoyment, which is paramount for the success of a game.

Moreover, the participants had to answer 21 additional questions about the awareness elements, so the questionnaire would have been too long to fill in after an intensive game session. These questions were related to the research question RQ4. All the implemented awareness elements were presented to the participants so that they could evaluate them according to the following scale [73] (see Fig. 10 for an example):

- I don't remember it (0)
- Not at all useful (1)
- Not very useful (2)
- Somewhat useful (3)
- Useful (4)
- Very useful (5)

Observational elements of the experiment: In order to measure player *satisfaction*, the players' face was filmed during the experiments and analyzed afterwards by means of the SDK face analysis application [74–76] (see Fig. 11). Facial expression has been used in previous experiments for emotion recognition in games [77]. Specifically, we used

```xml
<?xml version="1.0" encoding="utf-8"?>
<TanksMod Configuration="c2" Tanks="4" Difficulty="Normal">
  <!--
    WhoAuthorship requires WhatTask
    WhereGaze requires WhereLocation
  -->
  <Present
    WhoPresence="false"
    WhoIdentity="true"
    WhoAuthorship="false"
    WhatTask="true"
    WhatStatus="true"
    WhatAbilities="true"
    WhereLocation="true"
    WhereGaze="true"
    WhereReach="false"
    WherePosition="true"
    HowDevice="false"
  />
  <!--
    WhatTaskHistory requires WhatTask
    WhenEventHistory requires WhatTask
  -->
  <Past
    WhatTaskHistory="false"
    WhenEventHistory="false"
  />
  <Future
    WhatNextEvent="true"
    WhatNextAbilities="false"
  />
  <SocialGroupDynamics
    WhoMembers="true"
    WhoOtherMembers="true"
    WhatBelonging="true"
    WhatGroupGoal="true"
    HowInnerCommunication="true"
    HowOuterCommunication="false"
  />
</TanksMod>
```

Fig. 7. Configuration file for medium-awareness level (c2).

the *happy* output value to measure happiness when playing each configuration of the game. Thus, experiment had a within-subjects design with repeated measures.

5. Experimental results

This section analyzes the results of the experiment, which are detailed in Appendix A. Moreover, the different statistical tests we ran to validate the initial hypotheses are presented in Appendix B.

5.1. Game score

Regarding game score (*Sco*), it can be seen that the higher the awareness, the higher the score usually is. The results for configurations c1 and c3 for the average *Sco* were $\bar{x} = -1.38$, $\sigma = 3.33$ and $\bar{x} = 2.54$, $\sigma = 2.79$, respectively. This means the players usually won when playing with configuration c3 (high awareness) and tended to

Table 9
Meeting details.

Date	November 20th 2015, 16:00 – 19:00
Location	A-Lab[a], Overhoeksplein 2, Amsterdam (The Netherlands)
Participants	Miguel A. Teruel (University of Castilla – La Mancha, Researcher and Game developer) José M. Gómez de Lara (Radical Graphics, CTO) Manuel González (Radical Graphics, Creative director) Damir Kalbic (Radical Graphics, Game developer)

[a] http://www.a-lab.nl/.

Table 10
Implemented Gamespace Awareness elements.

lose with configuration c1. For c2, *Sco* was $\bar{x} = -0.62, \quad \sigma = 2.81$, which indicates that, at the same levels of awareness, the victory depends mostly on the players' skills. Fig. 12 gives the results for *Sco*, in which an outlier for game c3 in Albacete can be seen. In order to find any differences among the means, a repeated-measures ANOVA was performed, obtaining a *p*-value of 0.003. We can therefore conclude that there are differences among the means at the 0.05 level of significance (*p*-value of 0.473 for the sphericity test). Therefore, the null hypothesis H_{0A} is rejected, so the awareness level does not have the same effect on the game score of the participants.

5.2. Game time

Regarding game time (*Tim*), the players with configuration c2 needed slightly more time to finish the game ($\bar{x}$ = 7'04", σ = 3'58") than those using c1 and c3 ($\bar{x}$ = 6'26", σ = 2'54" for c1 and $\bar{x}$ = 6'23", σ = 2'32"). Fig. 13 summarizes the results (note that the time is measured in milliseconds). The similarity between the c1 and c3 results is coherent, because when a team was playing with c1, the rival team did so with c3, so the timing must tend to be the same. In order to assess whether there is a statistical difference among the configurations, a

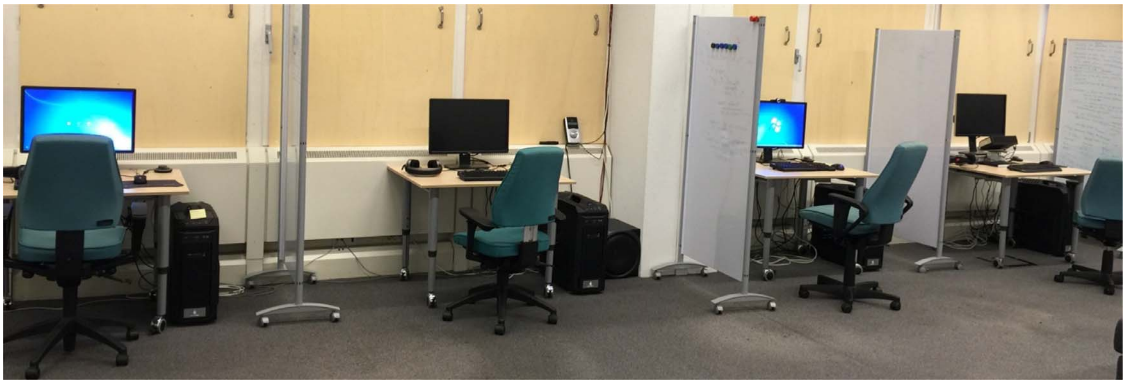


Fig. 8. Set-up of the experiment in Amsterdam.

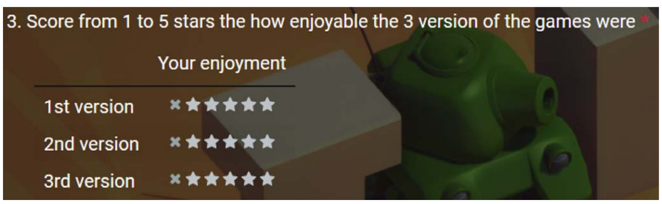


Fig. 9. Survey question to rate the game configurations.

Fig. 11. Example of the emotional analysis of a participant with CrowdSight [74].

Within-Subjects ANOVA was performed. On this occasion, the data did not pass the sphericity test (p-value ~ 0). Therefore, the repeated-measures ANOVA result was considered with a Greenhouse-Geisser correction [78,79], an adjustment method used when sphericity is not assured obtaining a p-value of 0.335. Consequently, we can conclude that there are no differences among the means at the 0.05 level of significance, so H_{0B} cannot be rejected. Hence, the awareness level has no effect on the game time.

5.3. Enjoyment

Next, regarding the subjective result for *Enj* (enjoyment), the configuration which obtained the highest score was c3 ($\bar{x}$=4.05, $\sigma = 0.77$), followed closely by c2 ($\bar{x}$=3.98, $\sigma = 1.04$) and finishing with c1, which obtained an unexpectedly poor result ($\bar{x}$=2.76, $\sigma = 1.20$). Further details can be seen in Table 11. These results for c2 and c3 are opposed to those obtained for the *Hap* objective variable, where c2 surpassed c3. To find any differences among these results, considering the game configuration as the within-subjects factor and the sequence among configurations (increasing or decreasing awareness) and players' expertise (according to their survey answer) as the between-subjects factors, a repeated-measures ANOVA was performed. With a p-value of ~ 0, we can conclude that there are differences among the means at the 0.05 level of significance (p-value of 0.804 for the sphericity test). Therefore, the null hypothesis H_{0C} is rejected, and the awareness level does not have the same effect on enjoyment. Besides, the combination of the configuration and the other between-subject factors does not significantly affect the results of *Enj*. However, an interesting conclusion can be drawn from these results: the players with a high degree of expertise prefer c2 rather than c3 (Fig. 14), i.e. they do not need too much awareness, since they rely on their own ability to win the game.

5.4. Happiness

As far as happiness (*Hap*) is concerned, c2 obtained the best score ($\bar{x}$=19.90%, $\sigma = 18.60\%$), followed by c3 ($\bar{x}$=16.42%, $\sigma = 18.60\%$) and c1 ($\bar{x}$=14.53%, $\sigma = 13.72\%$)). Again, a repeated-measures ANOVA was performed to assess whether there was a significant difference among these results, considering the game configuration as a within-subjects factor and the kind of awareness (ascending or descending) as a between-subject factor. With a p-value of 0.004, we can conclude that there are differences among the means at the 0.05 level of significance (p-value of 0.988 for the sphericity test). Consequently, the null hypothesis H_{0B} is rejected, so that the awareness level does not have the same effect on happiness. A p-value of 0.281 was obtained for the

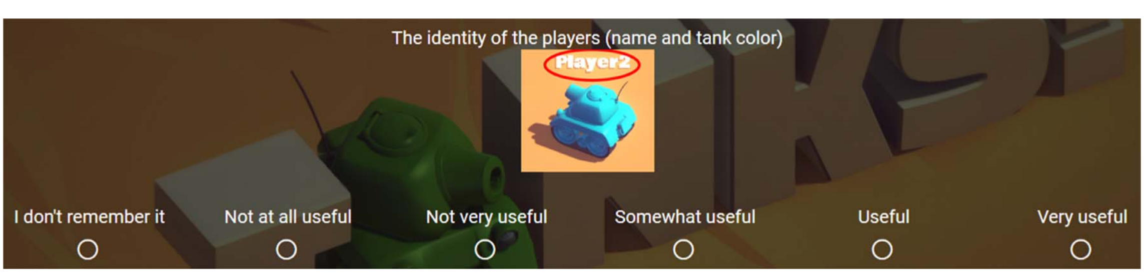


Fig. 10. Example of one of the questions regarding usefulness of GA elements.

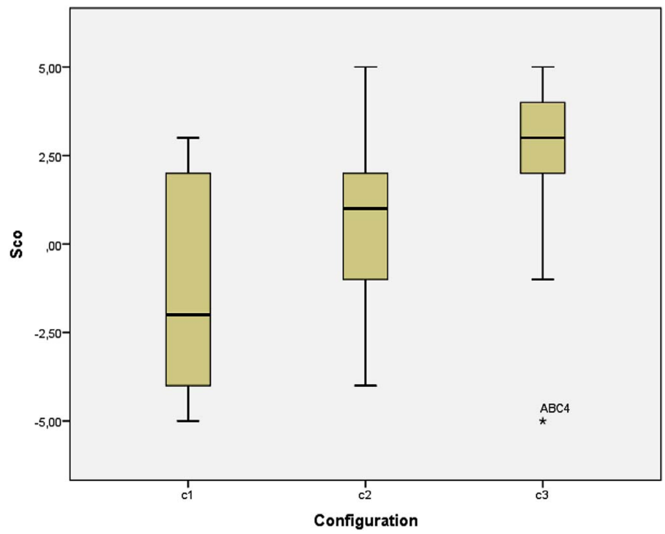


Fig. 12. Boxplot for *Sco*.

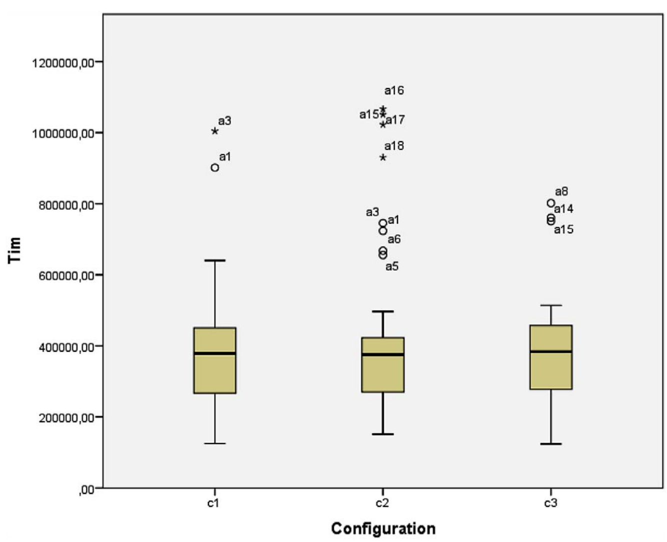


Fig. 13. Boxplot for *Tim*.

combination of both game configuration and increasing/decreasing awareness level (sequence), so there are no differences among the means when considering this combination of factors. These results led us to perform a one-way ANOVA of the kind of awareness to assess

Table 11
Descriptive statistics for *Enj*.

Expertise	Configuration	$\bar{x}$	σ	N
Advanced	c1	2.8462	1.40512	13
	c2	4.2308	1.09193	13
	c3	4.0000	0.81650	13
	Total	3.6923	1.25978	39
Normal	c1	2.7391	1.17618	23
	c2	3.8696	1.05763	23
	c3	4.0000	0.79772	23
	Total	3.5362	1.15783	69
Begginer	c1	2.6000	0.89443	5
	c2	3.8000	0.83666	5
	c3	4.4000	0.54772	5
	Total	3.6000	1.05560	15
Total	c1	2.7561	1.19959	41
	c2	3.9756	1.03653	41
	c3	4.0488	0.77302	41
	Total	3.5935	1.17234	123

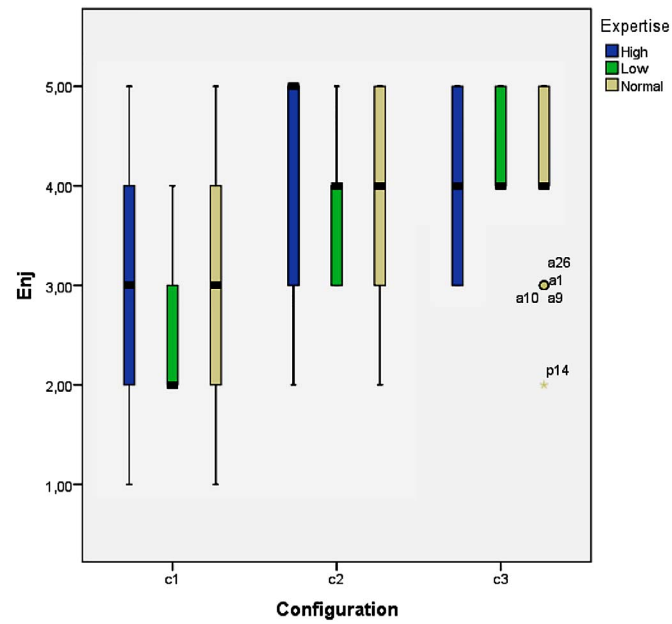


Fig. 14. Boxplot for *Enj*.

whether increasing or decreasing the awareness level would affect the participants' happiness. A *p*-value of 0.358 was obtained, indicating that the order of the game configurations does not affect the participants' happiness, then rejecting H_{0D}. However, although the standard deviations obtained were quite significant, this result can be considered normal, since the emotional state of the subjects could differ at the beginning of the experiment. For example, participant *p8* (Fig. 15) obtained considerably higher than average *Hap* results in the three game configurations. However, these results are in line with the tendency of the experiments (c2 > c3 > c1).

5.5. Usefulness

A wide range of results were obtained for *Use* (usefulness), *Status* (the circle representing the tanks' health level) being the most highly rated GA element with a *Use* result of $\bar{x}$=4.8, σ = 0.56 (close to 5, the

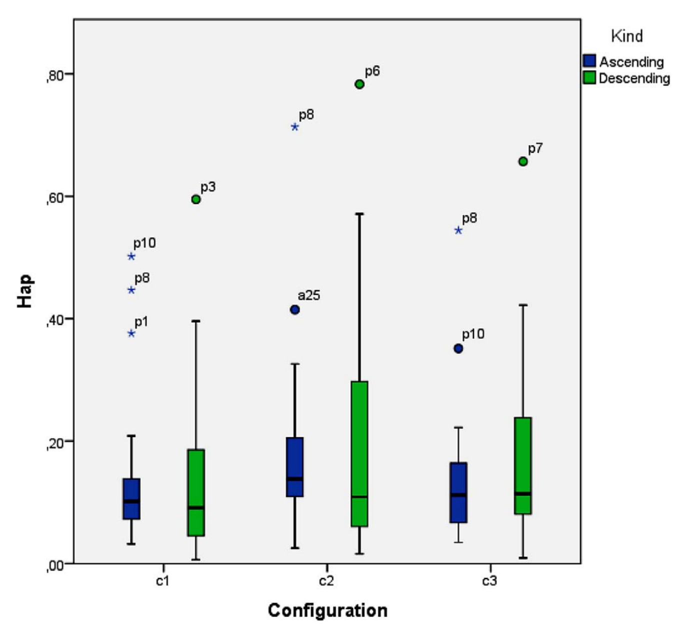


Fig. 15. Boxplot for *Hap*.

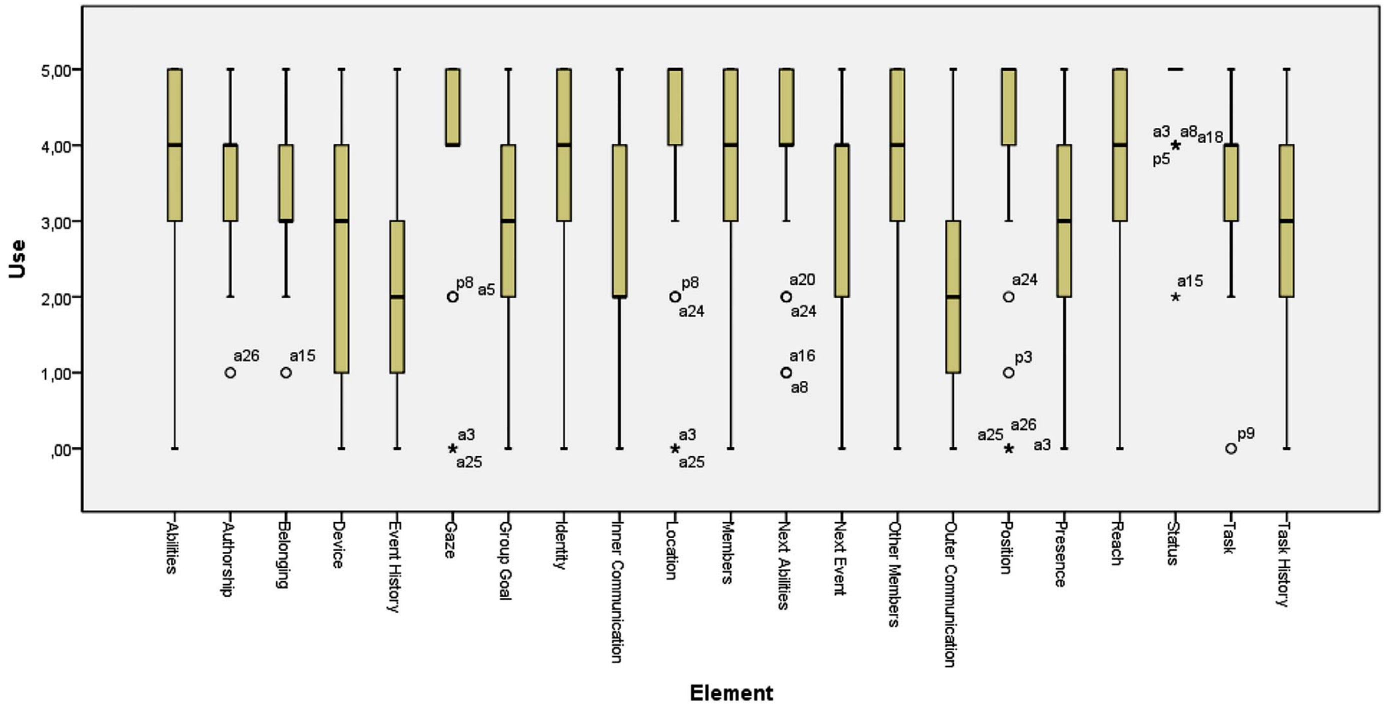


Fig. 16. Boxplot for *Use*.

Table 12
Descriptive statistics for *Use* grouped by GA element concern and configuration.

Concern	Configuration	$\bar{x}$	σ	N
Future	c2_c3	3.3902	1.49797	41
	c3	3.9512	1.07124	41
	Total	3.6707	1.32455	82
Past	c3	2.3537	1.31817	82
	Total	2.3537	1.31817	82
Present	c1_c2_c3	4.0000	1.32288	41
	c2_c3	4.0772	1.25768	246
	c3	3.3049	1.42018	164
	Total	3.7894	1.37193	451
Social	c2_c3	3.2049	1.37448	205
	c3	2.3415	1.45962	41
	Total	3.0610	1.42297	246
Total	c1_c2_c3	4.0000	1.32288	41
	c2_c3	3.6565	1.39168	492
	c3	3.0274	1.46806	328
	Total	3.4332	1.45371	861

theoretically perfect result) and *Event history* (the exact time when an action happened) the worst ($\bar{x}$=1.95, σ = 1.40).The rest of the results are given in Fig. 16. At a glance, it can be seen that there is a difference among the average values for *Use*, although a one-way ANOVA was performed to demonstrate this fact statistically. A *p*-value of ~0 was obtained, so we can conclude that there are differences among the means at the 0.05 level of significance. An additional analysis was performed by grouping the GA elements by their concern (present, past, future and social & group dynamics) and the configuration in which they are active, either all (c1_c2_c3), c2 and c3 (c2_c3) or only c3. In this case there are noteworthy differences between the best rated concerns (present and future) and the worst (past and social & group dynamics), as can be seen in Table 12 and Fig. 17. The elements active in all the configurations and those in c2 and c3 clearly surpassed the score for those in c3 only (Fig. 18). Once again, a one-way ANOVA was performed to assess these differences, obtaining a *p*-value of ~0 for concern and 0.001 for configuration. The null hypothesis H_{0E} is therefore rejected, so the different levels of awareness are not equally perceived as useful in achieving goals by the players.

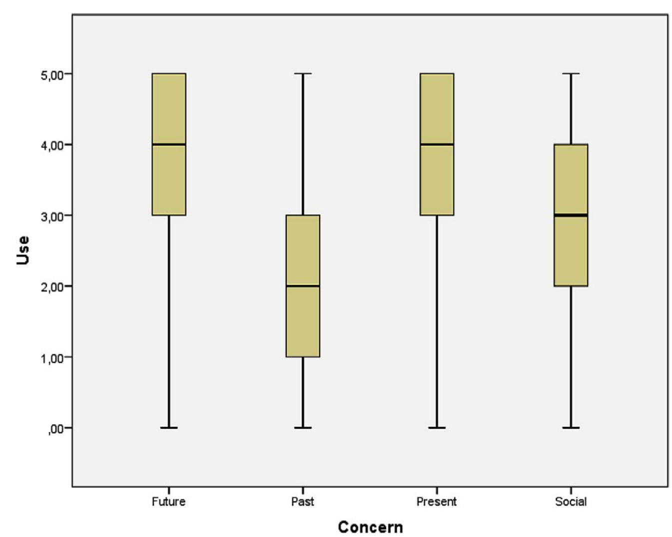


Fig. 17. Boxplot for *Use* grouped by GA element concern.

The results for each configuration are coherent, since the sole GA element in all the configurations (*Identity*) obtained the best result ($\bar{x}$ = 4.00, σ = 1.32), as this element is necessary to play the game effectively. The elements in c2 and c3 ($\bar{x}$ = 3.66, σ = 1.39) were in second place and those elements present in c3 only ($\bar{x}$ = 3.03, σ = 1.47) were third. The difference between these two last results confirms the appropriateness of the selection of GA elements for c2 by Radical Graphics, as their usefulness was 13.2% better than those in the other configurations. To conclude the analysis of *Use*, we analyzed the GA elements that were unknown to the participants (i.e. they answered “I don't remember it” in the survey). Four of these, *Authorship, Status, Next activities* and *Belonging* were recognized by all the participants (Fig. 19). Since they belong to different concerns and have different implementations (in graphical form for *Status* and colored textual form for the others, as shown in Fig. 19, we cannot find a common pattern among them. However, the elements which obtained the worst result,

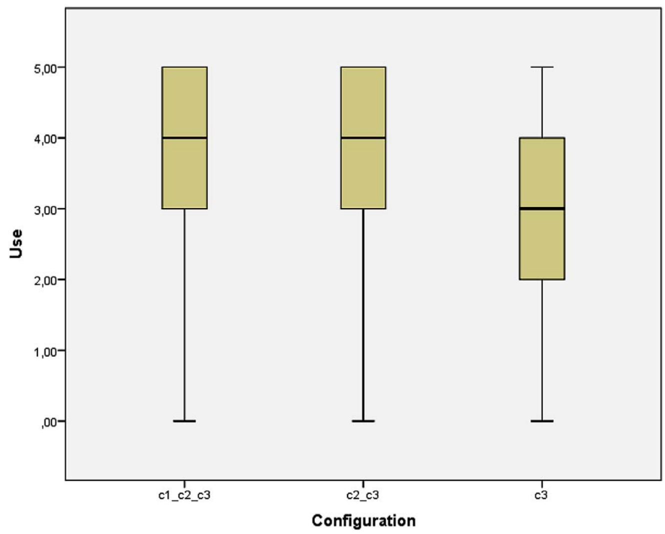


Fig. 18. Boxplot for *Use* grouped by the configuration in which each GA element is active.

Event history and *Outer communication* (unknown for 14.63% and 12.20% of the participants, respectively) were implemented as non-colored text. This explains the low score for these two GA elements, along with the fact that these elements were not crucial for an action game (i.e. it is important to talk to your allies and know what is happening, rather than talking to the enemy, and to know exactly when something happens).

5.6. Summary

To summarize the analysis of the different metrics and determine the configuration with the highest GX.

Table 13 gives the results for each pair configuration and metric. According to Bernhaupt [37], GX focuses on positive emotions and emotional outcomes such as joy, fun and pride. Therefore, the metrics that should be considered the most important for GX are *enjoyment* and *happiness* (i.e. subjective and objective satisfaction). These metrics measure how enjoyable the different configurations were for the players subjectively (enjoyment is measured through a subjective survey) and objectively (happiness is measured thought an objective software). For these metrics, the result for enjoyment was rather similar for c2 and c3, but the obtained happiness was considerably higher for c2. c1 obtained the worst results for both metrics.

Regarding *score* and *time*, these metrics can be considered important since they are related to the game goal (i.e. obtaining more points than your opponents). However, these metrics are not as important as the emotional ones, because winning easily and obtaining a high score quickly could be boring for players [80]. c3 obtained the highest result for *score*, which can be considered reasonable, because the more help the players receive, the better their chances of winning. According to Csikszentmihalyi [80] winning easily can be boring for the player and losing constantly can increase the anxiety level. Moreover, “a game that is too hard is frustrating, while too little challenge can be boring. In multi-player games, it is important that the game be fair, offering no player an intrinsic advantage” [81]. Therefore, c2 can be considered as the most balanced configuration because its win/lose ratio was closer to zero. The results for game time were quite similar, c1 and c3 obtaining the lowest result. Nevertheless, unlike non-gaming software, taking a long time to finish a game is not a negative aspect, provided the player is having fun while doing so. Lastly, *usefulness* focused on measuring the different GA elements.

Table 13
Summary of results for the experiment metrics.

Metric	c1	c2	c3
Score	−1.38	0.40	1.80
Time	6′26″	7′04″	6′23″
Enjoyment	2.76	3.98	4.05
Happiness	14.53%	19.90%	16.42%
Usefulness	4.00 (1 GA elment)	3.66 (13 GA elements)	3.03 (21 GA elements)

Table 13 shows the average usefulness of the GA elements in each configuration. Although c1 was the best rated configuration, this could be considered unfair since each configuration had a different number of GA elements. This is especially relevant in the case of c1, whose one and only element (*Identity*) obtained a result of 4 out of 5. The *usefulness* results are given in Section 5.5.

In view of the results obtained, and considering those related to positive emotions as the most important components of GX, we can conclude that GX was highest in configuration c2, whose results for positive emotions were better. Furthermore, it provided a more balanced result regarding *score* (neither too hard not to easy). Finally, its GA elements were rated better than those in c3, showing that its GA element was useful to the players.

5.7. Threats to validity

In the following, the different threats to validity of our experiment will be analyzed. More specifically, it will be analyzed the four types of threats proposed by Wohlin et al., namely conclusion, internal, external, and construct validity [65].

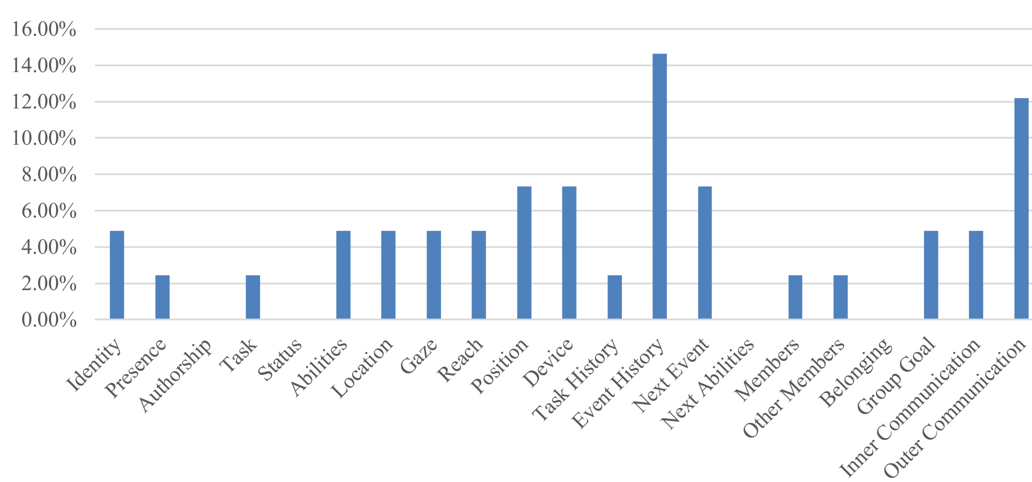


Fig. 19. GA elements that participants did not remember.

5.7.1. Conclusion validity

The threats to conclusion validity are factors which can lead to reach an incorrect conclusion about the observations. The five null and alternative hypotheses were contrasted by using ANOVA tests that indicated statistically significant *p*-values (at alpha = 0.05). Therefore, they supported the contrast of our initial hypotheses with enough degree of certainty.

5.7.2. Internal validity

Threats to internal validity are focused on the design of the study. More specifically, they are related to whether or not the results really do follow from the data. In order to assess such internal validity, the following threats to our independent variables are analyzed.

The first analyzed threat to the internal validity was instrumentation since erroneously designed instruments could affect the results of the experiment. Aimed at mitigating this threat, instructions, questionnaires and game settings were carefully reviewed and tested before and after running the pilot experiment. Moreover, in order to avoid a possible source of bias, two experienced game developers collaborated in this work by checking the suitability of the representation of each GA element as well as its position in the game screen. They also helped us to elaborate the questionnaire we used by taking into account the main features of the game used in the experiment.

Second, in order to mitigate the selection-maturation interaction threat, since subjects mature at different speeds, the teams were organized according to their experience level in action games, achieving balanced game competitions. For this aim, subjects were asked to provide their experience level regarding action games in the sing-up form. Taking into account their answers, we grouped them according their selected level (high, normal or low).

Third, mortality happens when participants drop out the experiment. In our study, we tried to avoid this threat by scheduling the participation of the subjects according their availability. Despite sending several reminders via email, 2 participants from Amsterdam and 1 from Albacete did not show up, even though they had confirmed their participation. In order to mitigate this threat, an artificial-intelligence-controlled player was used in those game sessions with absent participants. The AI tanks were configured according to the absent participant experience level (high, normal or low).

Fourth, it is worth noting that all subjects were rewarded simply for participating in the experiment. We clarify this issue at the beginning of the experiment by reminding the participants that they will receive the reward regardless of their game performance. Although this *general compensation* may equally affect the control and treatment groups, it helped us to reduce the dropout rates because participants simply could have given up at any time of the experiment and still receive the reward.

Finally, the number of subjects that participated in the experiment was large enough (≥ 30), according to the central limit theory [82], necessary to obtain a statistically significant set of data. Therefore, the conditions were valid to carry out the ANOVA tests.

5.7.3. Construct validity

The threats to construct validity represent the extent to which the experiment settings or the measures chosen reflect the construct under study. Therefore, for the research questions that guide the design of this study we considered the usability constructs defined by the ISO 9241-11 standard (effectiveness, efficiency and satisfaction), as well as other recommendations [37] coming from the games community. The metrics used to answer such research questions were defined after a review of the appropriate literature. For instance, player satisfaction was measured through two complementary measurement methods: a questionnaire based on a 5-star rating scale, and a facial expression analysis of the participants while interact with the three different game configurations.

Regarding the usage of 5-star rating scales, they are widely used for assessing customer satisfaction in terms of enjoyment (e.g. Yahoo, Amazon, Skype, etc.). Moreover, our study also took the advantage of an existing commercial and available tool for deriving emotional measures from facial expressions in a more objective and reliable way [74–76].

In order to measure the usefulness of the 21 awareness elements distributed along the three different game configurations, an online questionnaire based on a six-ordinal scale was also conducted. Although a discrete scale has been used to evaluate each one of the awareness elements, the results have been presented as if it were a continuous variable because the data have been summarized according to the dimension and the configuration the elements belongs to.

5.7.4. External validity

Threats to external validity relate to whether or not the claims for the generality of the results are justified. The experiment was performed by using just one game belonging to the action genre. Therefore, the results of this experiment cannot be generalized for any type of game beyond action games similar to the evaluated one. However, we consider that the design of the experiment may be used to other similar experiments with other games belonging to different genres in order to step forward the generalization of results.

Besides, the results of our experiment cannot be generalized to real settings, since most of our participants were computer science students (with similar backgrounds), who participated voluntarily in the experiments. However, we consider that the three configurations of our game are sufficiently representative, as they were created to suit the expertise and experience of real gamers.

6. Conclusions and future work

This work contains an assessment of the impact of Gamespace Awareness elements on several aspects of the game and on the players involved. We analyzed the effect of high and low awareness levels on the game score, the players' happiness and enjoyment and the usefulness of GA elements to achieve the game's goal. For this, a co-operative action game with three different awareness configurations was developed. Configuration c1 contained the basic awareness level necessary to play the game, c2 had the "ideal" GA elements selected by an external company and configuration c3 had the maximum possible GA elements. The experiment was performed initially in Amsterdam, with the participation of 14 Computer Science undergraduates and replicated in Albacete with 29 new participants.

The analysis of the results provided several interesting conclusions. Regarding the game score (RQ1), the two teams involved in each competition played with all the game configurations, but in a different order: Team 1 experienced an increasing awareness level c1, c2, c3, and Team 2 a decreasing level c3, c2, c1. The analysis of the results concluded that a higher awareness level gives the players a technical advantage and their scores were higher for this reason. As far as the competition time is concerned (RQ2), those players with configuration c2 needed slightly more time to finish the game than those with configurations c1 and c3. However, there was no significant difference among the means, so the awareness level had no effect on game time.

Regarding participants' satisfaction (RQ3), the participants were recorded in videos that were subsequently processed with facial analysis software. This automated analysis concluded that the players were happier when playing with c2, followed by c3 and, then, c1. In addition, the order in which the awareness was provided (i.e. ascending or descending) did not affect the happiness level. Furthermore, it can be concluded that this happiness did not depend directly on winning, but on playing the proper game configuration. In other words, happiness was highest when playing with configuration c2, regardless of the competition result. Game enjoyment was measured by means of a post-game survey. Unlike the previous results, the participants considered that c3 was the game they enjoyed most, followed by c2 and c1.

However, a complementary analysis revealed that this result is not applicable to expert players, who preferred c2. In other words, a high awareness level would help beginners, but could annoy or distract experienced players, thus leading to cognitive overload and interruptions due to the excess of awareness information.

This survey helped us to measure the usefulness of GA elements for achieving the game's goal (RQ4). According to the survey results, the most useful GA elements were related to the present and the future, as could be expected for an action game, since the most important information is what is happening or will happen, rather than what has happened in the past.

As this experiment was limited to action games, in order to properly assess the GA elements, more experiments should be performed with games belonging to other genres (e.g. RPG, FPS, sports, strategy, puzzle, etc.). Our main aim is to compare the results obtained for an action game with those for other games, in order to generalize our results. Another limitation was that most of our participants were males, mainly because they were Computer Science students. A complementary study will be carried out involving equal numbers of males and females with profiles not limited to Computer Science.

It is worth noting that the final goal of the experiment presented in this article is neither the awareness elements themselves nor their implementation. It really focuses on evaluating the need of specifying different awareness levels and, thus, determining how much information users can perceive as useful and use with effectiveness, efficiency and satisfaction, that is, we aim at generalizing the results to the population. For this reason, it was cared that the selection of the sample was representative of the population following Wohlin et al. [65]. This experiment is a new iteration for the design of Game Awareness over our previous work [11]. In such work it was presented the evaluation of GA by using a First-Person Shooter and a Real Time Strategy game. Therefore, this work could be used as a blueprint for designing how to measure the right level of awareness in other games by carrying out some adaptation such as the tasks to be carried out by gamers, metrics to be applied, etc.

As part of our on-going work, we are developing a framework [83] that will help game designers to identify the GA elements that new games of different genres should feature. This framework will also establish the GA elements to be implemented according to the players' expertise level. Once this framework is created, it would be integrated with a requirement engineering methodology [84,85] to deal with the requirements specification of a collaborative game. As the results obtained in the present study revealed that the players' satisfaction as measured by a questionnaire differed when analyzed by the users' emotions, a deeper study could be performed involving more users and other games belonging to different genres.

Finally, in this experiment, an eye-tracking device was used in order to assess whether the participants used certain awareness elements. However, we only had one eye-tracking device for the experiment in Amsterdam, and none for the one in Albacete. Therefore, we did not obtain enough data to perform a significant study. Because of this reason, as a future work, we will perform a wider study on awareness elements use by employing an eye tracker device. Such study could complement the analysis presented in this work regarding RQ4, i.e., the perceived usefulness of awareness elements.

Acknowledgments

This research was funded by the Spanish Ministry of Economy and Competitiveness and by EU FEDER funds under project grants TIN2016-79100-R and TIN2015-72931-EXP. It was also funded by the Spanish Ministry of Education, Culture and Sports, with an FPU scholarship (AP2010-0259). We would like to thank Mojca Lovrencak and Marco Otte from VU, Roberto Valenti from Sightcorp, José Gómez de Lara and Manuel González from Radical Graphics and Miguel Oliver from UCLM for their priceless support, as well as the volunteer participants for their contribution to the pilot and experiment.

Appendix A. – Experiment results

This appendix presents the results for game score (*Sco*), game time (*Tim*), players' enjoyment (*Enj*), happiness (*Hap*) and usefulness (*Use*) of GA elements for achieving the game goal. The *Sco* results are given as the average of the differences of the scores of each game, e.g. if a game ended with a score of 3–5 for Teams 1 and Teams 2, respectively, the result of that game would be −2 (3–5 = −2). The results are shown in Table 14 and Fig. 20. As can be observed, Team 1 was usually defeated when they used configuration c1 and won when using configuration c3.

Game time (*Tim*) is given in Table 15 and Fig. 21 as the average of the game duration in minutes and seconds. As can be seen, games played with configuration c2 took longer than those played with c1 and c3. The competitions played with c1 and c3 should have taken almost the same time and, as Fig. 21 shows, the results are consistent. When the blue team used c1, they were playing against the red team with configuration c3 (and vice versa).

Table 14
Results for *Sco*.

	c1	c2	c3
Amsterdam	−1.60	0.40	1.80
Albacete	−1.25	0.75	3.00
Total	**−1.38**	**0.62**	**2.54**

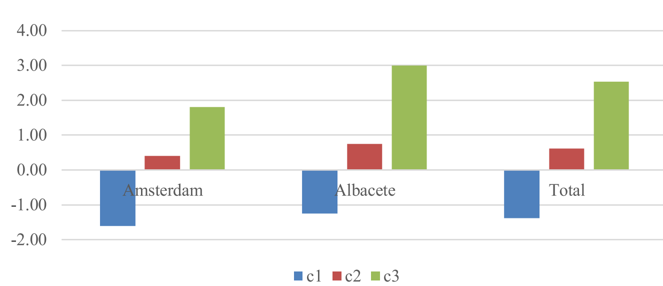


Fig. 20. Results for *Sco*.

As was stated in Section 4.2, the enjoyment of each one of the game configurations (*Enj*) was evaluated using a five-star scale (see Fig. 9). The results for this variable are given in Table 16 and Fig. 22. The results for *Enj* are shown according to the players' levels of expertise (low, normal and high), the configuration and number of participants in each category.

In order to measure the players' happiness (*Hap*) with certain configuration, we filmed their faces during the experiments. These videos were later

Table 15
Results for *Tim*.

	c1	c2	c3
Amsterdam	5′37″	6′31″	5′27″
Albacete	6′49″	7′20″	6′50″
Total	**6′26″**	**7′04″**	**6′23″**

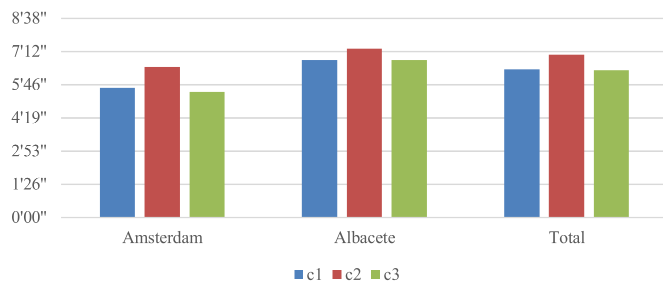


Fig. 21. Results for *Tim*.

Table 16
Results for *Enj* according to location, expertise and configuration.

	c1	c2	c3	Participants
Amsterdam - High	2.57	4.57	4.14	7
Amsterdam - Normal	3.00	3.50	4.00	4
Amsterdam - Low	3.00	4.00	4.00	3
Amsterdam - All	2.79	4.14	4.07	14
Albacete - High	3.17	3.83	3.83	6
Albacete - Normal	2.68	3.95	4.00	19
Albacete - Low	2.00	3.50	5.00	2
Albacete - All	2.74	3.89	4.04	27
Total - High	2.85	4.23	4.00	13
Total - Normal	2.74	3.87	4.00	23
Total - Low	2.60	3.80	4.40	5
Total - All	**2.76**	**3.98**	**4.05**	**41**

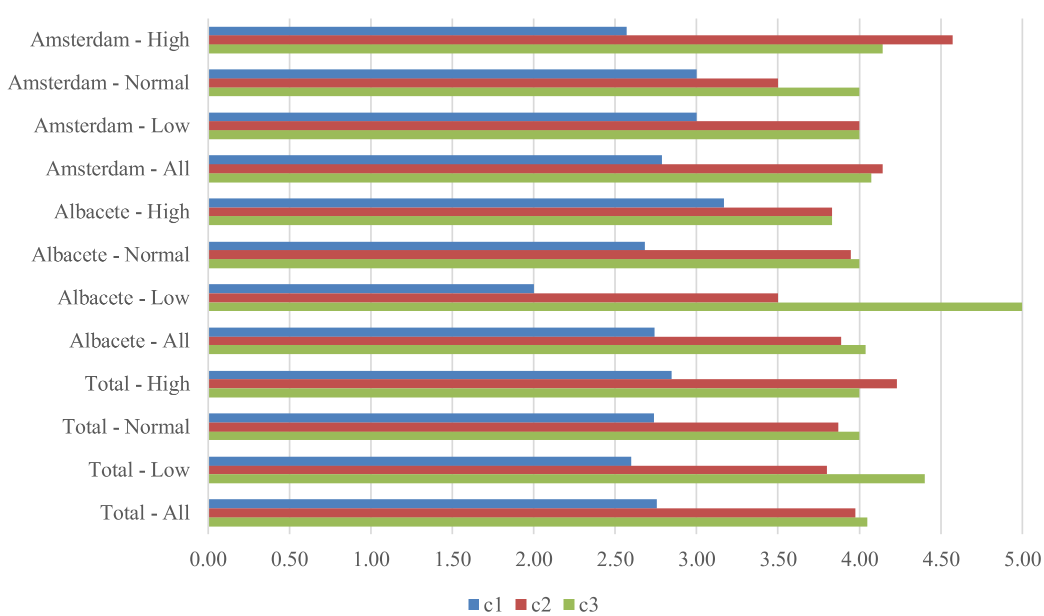


Fig. 22. Results for *Enj* according to location, expertise and configuration.

Table 17
Results for *Hap*.

	c1 (%)	c2 (%)	c3 (%)	Participants
Amsterdam - Increasing	25.96	25.86	22.02	7
Amsterdam - Decreasing	29.14	48.07	37.81	7
Amsterdam - All	27.55	36.96	29.91	14
Albacete - Increasing	9.47	14.80	10.37	15
Albacete - Decreasing	6.93	8.29	9.42	14
Albacete - All	8.25	11.66	9.91	29
Total - Increasing	14.72	18.32	14.08	22
Total - Decreasing	14.34	21.55	18.88	21
Total - All	**14.53**	**19.90**	**16.42**	**43**

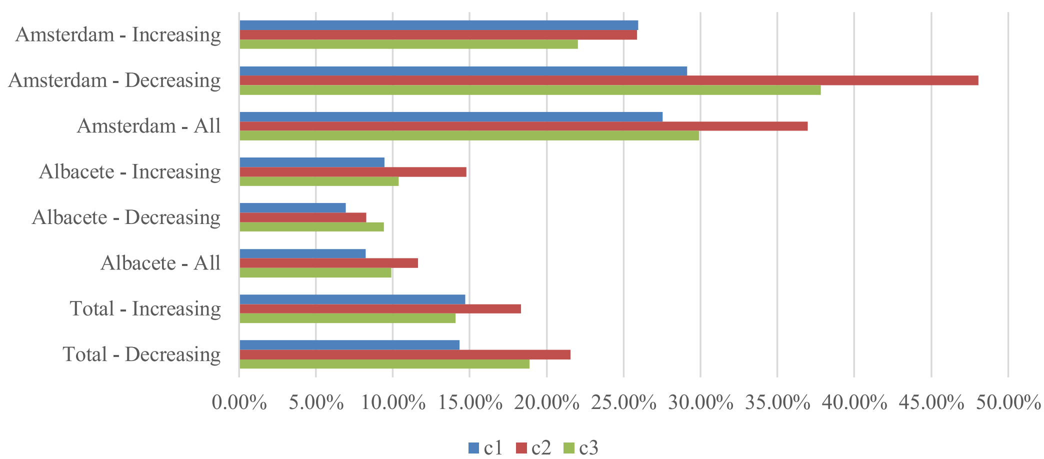


Fig. 23. Results for *Hap*.

analyzed with CrowdSight SDK [74]. These analyses provided us with 6 different emotional measurements (happiness, sadness, disgust, surprise, fear and anger) per video frame. For each of these emotions, the application returns a value of from 0 to 1. However, since our goal was to measure their happiness (i.e. their positive emotions [37], since this emotion is closely related to the player's GX), we only used this value. Finally, the average happiness was used as the measure of the enjoyment of each player. The collected results are shown in Table 17 and Fig. 23. Note that these results distinguish between players whose awareness level was raised and lowered between competitions. These figures depict one of the conclusions of this article, that is, too much or too few awareness affects users' enjoyment.

Finally, the results for the usefulness of GA elements for achieving the game goals (*Use*) are given in Table 18 and Fig. 24. For this last metric, only 41 results were obtained, since two participants did not fill in the survey. The scale used to measure this value is described in Section 4.2.

Table 18
Results for *use*.

	Amsterdam	Albacete	Total
Identity	4.21	3.89	**4.00**
Presence	2.86	3.30	**3.15**
Authorship	3.43	3.89	**3.73**
Task	4.07	3.11	**3.44**
Status	4.64	4.89	**4.80**
Abilities	3.29	4.00	**3.76**
Location	4.14	4.30	**4.24**
Gaze	4.29	4.11	**4.17**
Reach	3.36	3.74	**3.61**
Position	3.93	4.11	**4.05**
Device	2.36	2.93	**2.73**
Task history	2.86	2.70	**2.76**
Event history	1.86	2.00	**1.95**
Next event	3.21	3.48	**3.39**
Next abilities	3.64	4.11	**3.95**
Members	3.36	3.89	**3.71**
Other members	3.43	3.81	**3.68**
Belonging	3.07	3.59	**3.41**
Group goal	2.21	2.89	**2.66**
Inner communication	2.57	2.56	**2.56**
Outer communication	2.29	2.37	**2.34**

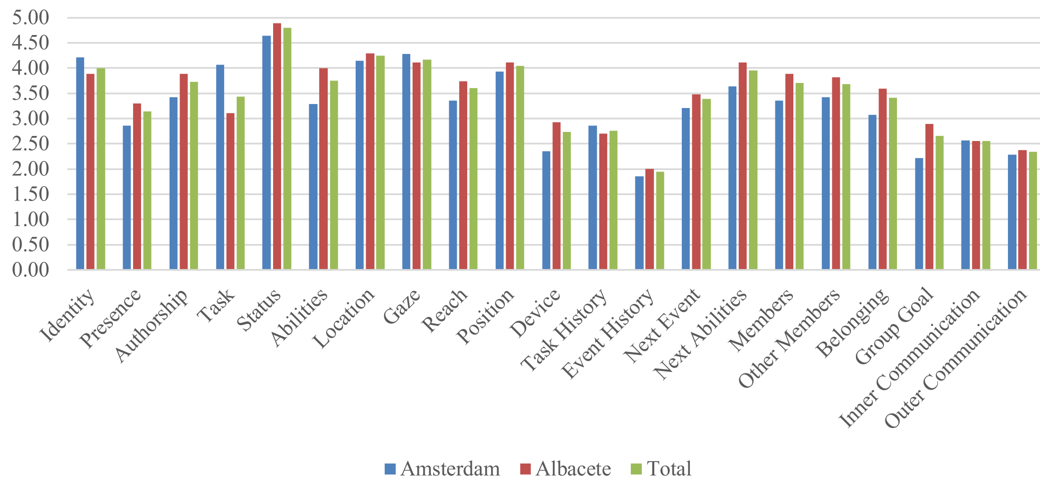


Fig. 24. Results for *usefulness by context-awareness element.*

Appendix B. – Results of statistical tests

This appendix shows the statistical tests ran in order to accept / reject our initial hypotheses (Tables 19–25).

Table 19
Tests of within-subjects effects for *Sco.*

Source		Type III sum of squares	df	Mean square	*F*	Sig.
Configuration	Sphericity Assumed	100.051	2	50.026	7.293	0.003
	Greenhouse-Geisser	100.051	1.774	56.391	7.293	0.005
Error(Configuration)	Sphericity Assumed	164.615	24	6.859		
	Greenhouse-Geisser	164.615	21.291	7.732		

Table 20
Tests of within-subjects effects for *Tim.*

Source		Type III Sum of Squares	df	Mean Square	*F*	Sig.
Configuration	Sphericity Assumed	3.7714E + 10	2	1.887E + 10	0.958	0.388
	Greenhouse-Geisser	3.7714E + 10	1.022	3.692E + 10	0.958	0.335
Error(Configuration)	Sphericity Assumed	1.655E + 12	84	1.970E + 10		
	Greenhouse-Geisser	1.655E + 12	42.936	3.854E + 10		

Table 21
Tests of within-subjects effects for *Enj.*

Source		Type III Sum of Squares	df	Mean Square	*F*	Sig.
Configuration	Sphericity Assumed	24.136	2	12.068	10.888	0.000
	Greenhouse-Geisser	24.136	1.975	12.221	10.888	0.000
Configuration * Sequence	Sphericity Assumed	1.379	2	0.689	0.622	0.540
	Greenhouse-Geisser	1.379	1.975	0.698	0.622	0.538
Configuration * Expertise	Sphericity Assumed	1.426	4	0.356	0.322	0.863
	Greenhouse-Geisser	1.426	3.950	0.361	0.322	0.860
Configuration * Sequence * Expertise	Sphericity Assumed	1.737	4	0.434	0.392	0.814
	Greenhouse-Geisser	1.737	3.950	0.440	0.392	0.812
Error(Configuration)	Sphericity Assumed	77.586	70	1.108		
	Greenhouse-Geisser	77.586	69.121	1.122		

Table 22
Tests of within-subjects effects for *Hap*.

Source		Type III Sum of Squares	df	Mean Square	*F*	Sig.
Configuration	Sphericity Assumed	0.064	2	0.032	5.776	0.004
	Greenhouse-Geisser	0.064	1.999	0.032	5.776	0.005
Configuration * Sequence	Sphericity Assumed	0.015	2	0.008	1.362	0.262
	Greenhouse-Geisser	0.015	1.999	0.008	1.362	0.262
Error(Configuration)	Sphericity Assumed	0.458	82	0.006		
	Greenhouse-Geisser	0.458	81.951	0.006		

Table 23
Tests of between-subjects effects for the kind of awareness.

Source	Type III sum of squares	df	Mean square	*F*	Sig.
Corrected model	0.021	1	0.021	0.850	0.358
Intercept	3.718	1	3.718	150.798	0.000
Kind	0.021	1	0.021	0.850	0.358
Error	3.131	127	0.025		
Total	6.858	129			
Corrected total	3.152	128			

Table 24
Tests of between-subjects effects for *use*.

Source	Type III sum of squares	df	Mean square	*F*	Sig.
Corrected model	418.093	20	20.905	12.549	0.000
Intercept	10,148.590	1	10,148.590	6092.126	0.000
Element	418.093	20	20.905	12.549	0.000
Error	1399.317	840	1.666		
Total	11,966.000	861			
Corrected total	1817.410	860			

Table 25
Tests of between-subjects effects for *use* grouped by GA element concern and configuration.

Source	Type III sum of squares	df	Mean square	*F*	Sig.
Corrected model	284.104	7	40.586	22.579	0.000
Intercept	3998.819	1	3998.819	2224.601	0.000
Concern	159.333	3	53.111	29.546	0.000
Configuration	16.363	2	8.181	4.551	0.011
Concern * configuration	32.895	2	16.447	9.150	0.000
Error	1533.306	853	1.798		
Total	11,966.000	861			
Corrected total	1817.410	860			

Appendix C. – Documents from the meeting with Radical Graphics

This appendix presents the documents we used to lead the meeting with Radical Graphics as anticipated in *External assessment of the game in* Section 4.2.2 (Figs. 25–27).

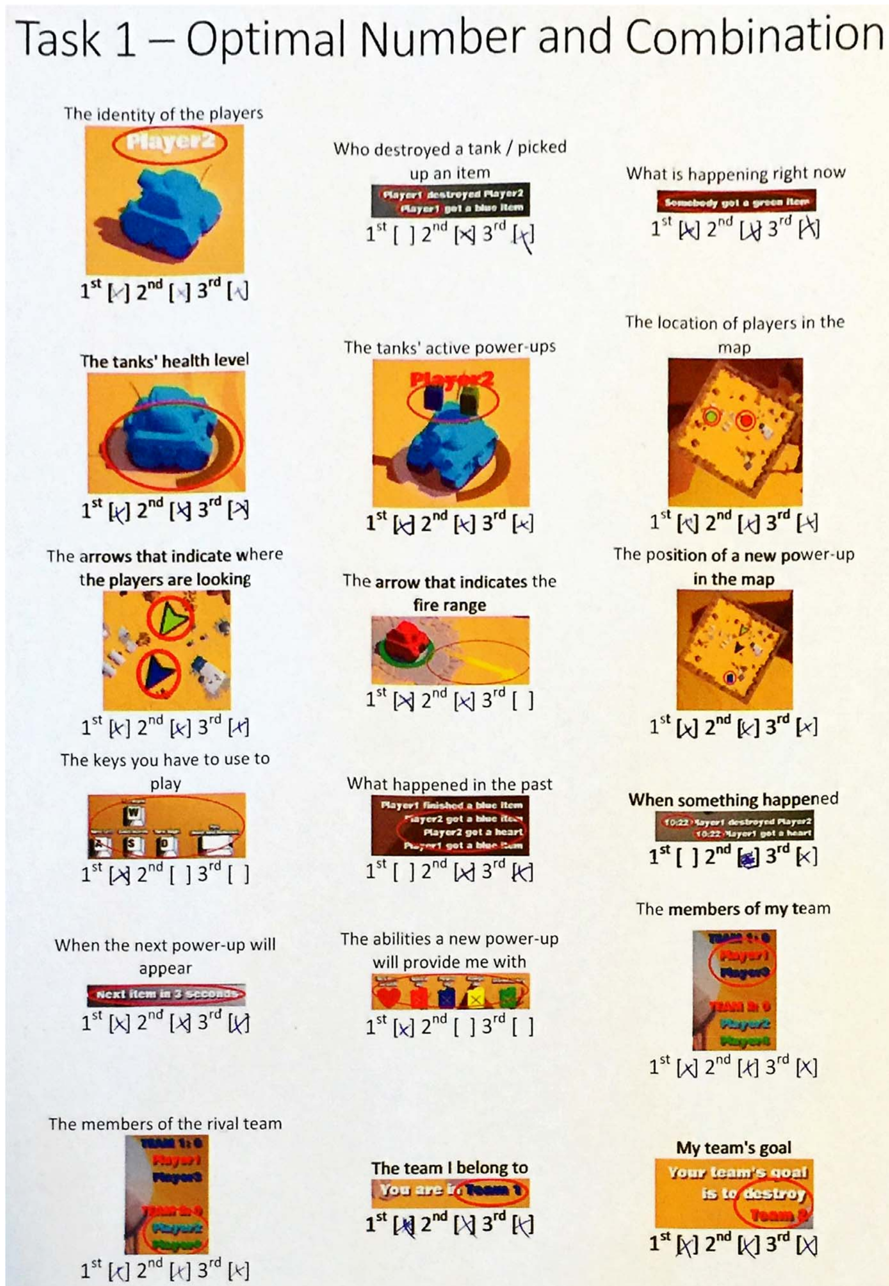
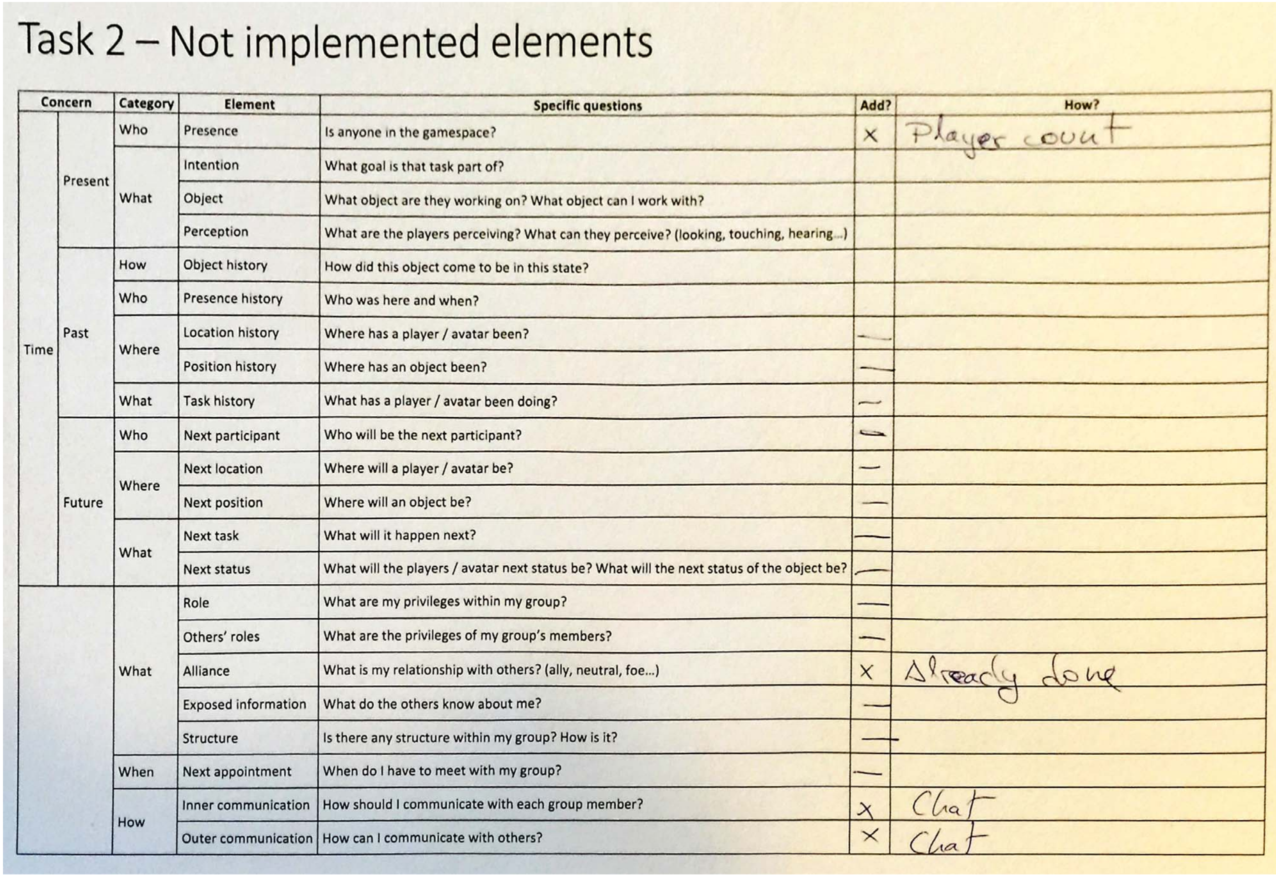


Fig. 25. Task 1 (Optimal number and combination of awareness element) results.



Task 2 – Not implemented elements

Concern		Category	Element	Specific questions	Add?	How?
Time	Present	Who	Presence	Is anyone in the gamespace?	X	Player count
		What	Intention	What goal is that task part of?		
			Object	What object are they working on? What object can I work with?		
			Perception	What are the players perceiving? What can they perceive? (looking, touching, hearing...)		
	Past	How	Object history	How did this object come to be in this state?		
		Who	Presence history	Who was here and when?		
		Where	Location history	Where has a player / avatar been?	—	
			Position history	Where has an object been?	—	
		What	Task history	What has a player / avatar been doing?	—	
	Future	Who	Next participant	Who will be the next participant?	—	
		Where	Next location	Where will a player / avatar be?	—	
			Next position	Where will an object be?	—	
		What	Next task	What will it happen next?	—	
			Next status	What will the players / avatar next status be? What will the next status of the object be?	—	
		What	Role	What are my privileges within my group?	—	
			Others' roles	What are the privileges of my group's members?	—	
			Alliance	What is my relationship with others? (ally, neutral, foe...)	X	Already done
			Exposed information	What do the others know about me?	—	
			Structure	Is there any structure within my group? How is it?	—	
		When	Next appointment	When do I have to meet with my group?	—	
		How	Inner communication	How should I communicate with each group member?	X	Chat
			Outer communication	How can I communicate with others?	X	Chat

Fig. 26. Task 2 (Not implemented awareness elements) results.

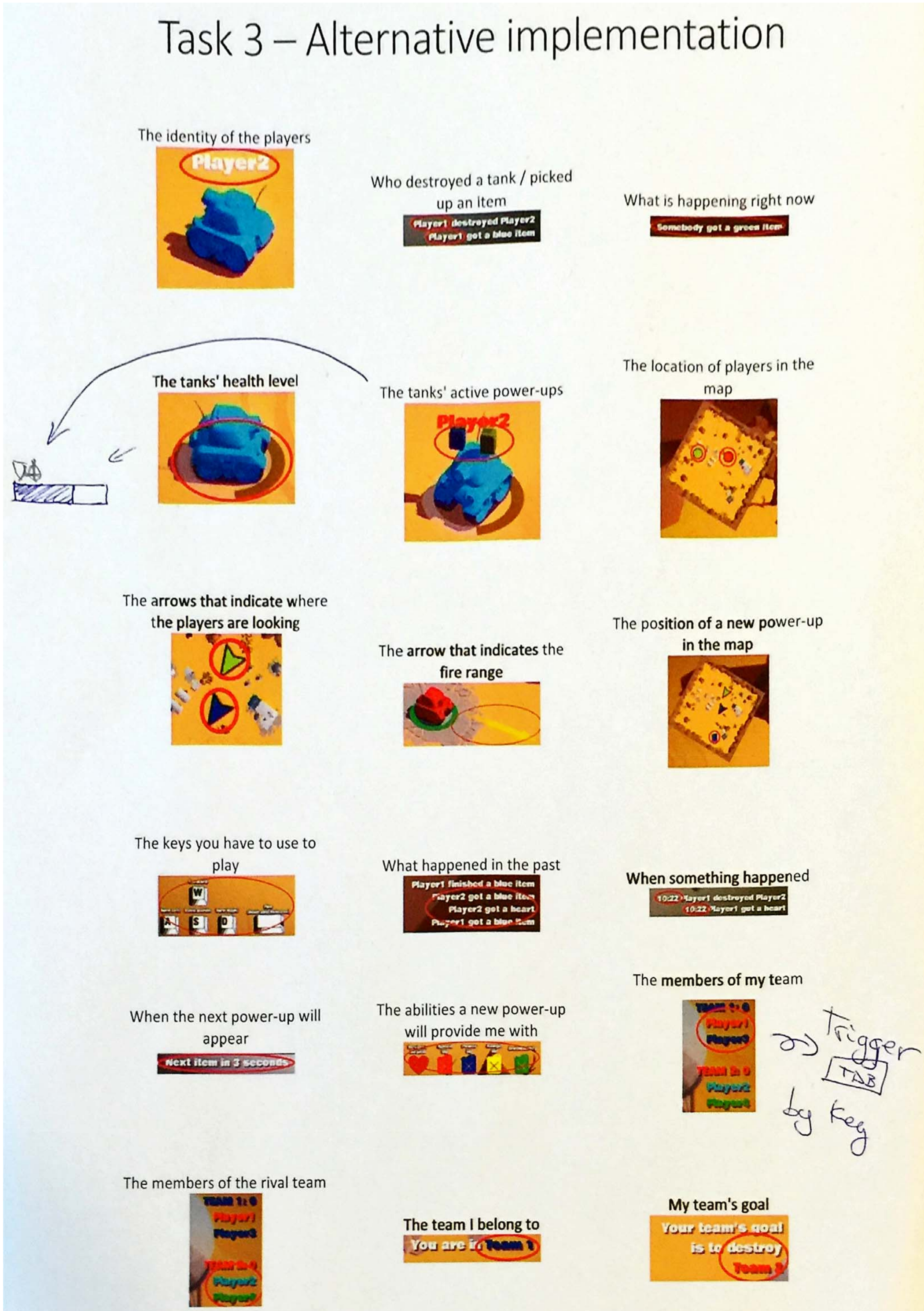


Fig. 27. Task 3 (Alternative implementation of current awareness elements) results.

References

[1] S. Kent, The Ultimate History of Video Games, 1st ed., Three Rivers Press, 2001.
[2] W. Day, Official Video Game & Pinball Book Of World Records, Sunstar, 1998.
[3] J. Watkins, L. Hjorth, I. Koskinen, Wising up: Revising mobile media in an age of smartphones, Continuum 26 (2012) 665–668, http://dx.doi.org/10.1080/10304312.2012.706456.
[4] Entertainment Software Association, Essential facts about the computer and video game industry, 2015. http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf, .
[5] P. Luckey, B.I. Trexler, G. England, J. McCauley, Virtual reality headset, USD701206 S1, 2014. https://www.google.com/patents/USD701206, .
[6] Oculus VR, Oculus Rift, http://www.oculusvr.com/, (2014) (accessed April 29, 2014).
[7] D.M. Ewalt, Virtual Reality Revenues Could Reach $21.8 Billion By 2020, Forbes (2015), http://goo.gl/cMAkN5.
[8] A. Ampatzoglou, I. Stamelos, Software engineering research for computer games: a systematic review, Inf. Softw. Technol. 52 (2010) 888–901, http://dx.doi.org/10.1016/j.infsof.2010.05.004.

[9] B. Nardi, J. Harris, Strangers and friends: collaborative play in world of warcraft, 20th Anniversary of the ACM Conference on Computer Supported Cooperative Work, ACM Press, BanffCanada, 2006, p. 149, , http://dx.doi.org/10.1145/1180875.1180898.

[10] M.H. Tran, G.K. Raikundalia, Y. Yang, Using an experimental study to develop group awareness support for real-time distributed collaborative writing, Inf. Softw. Technol. 48 (2006) 1006–1024, http://dx.doi.org/10.1016/j.infsof.2005.12.009.

[11] M.A. Teruel, E. Navarro, P. González, V. López-Jaquero, F. Montero, Applying thematic analysis to define an awareness interpretation for collaborative computer games, Inf. Softw. Technol. 74 (2016) 17–44, http://dx.doi.org/10.1016/j.infsof.2016.01.009.

[12] C. Reuter, V. Wendel, S. Göbel, R. Steinmetz, Game design patterns for collaborative player interactions, 2014 DiGRA Inter Conf. (DiGRA '14), 2014.

[13] Nintendo, Splatoon, http://splatoon.nintendo.com/, (2017) (accessed July 13, 2017).

[14] Activision Publishing Inc., Call of Duty, http://www.callofduty.com, (2017) (accessed July 13, 2017).

[15] R. Johansen, D. Sibbet, S. Benson, A. Martin, R. Mittman, P. Saffo, Leading Business Teams: How Teams Can Use Technology and Group Process Tools to Enhance Performance, 1st ed., Addison Wesley, 1991.

[16] M.R. Endsley, Toward a theory of situation awareness in dynamic systems, Hum. Factors J. Hum. Factors Ergon. Soc. 37 (1995) 32–64, http://dx.doi.org/10.1518/001872095779049543.

[17] C. Gutwin, S. Greenberg, A descriptive framework of workspace awareness for real-time groupware, Comput. Support. Coop. Work. 11 (2002) 411–446, http://dx.doi.org/10.1023/A:1021271517844.

[18] A. Dix, T. Rodden, N. Davies, J. Trevor, A. Friday, K. Palfreyman, Exploiting space and location as a design framework for interactive mobile systems, ACM Trans. Comput. Interact 7 (2000) 285–321, http://dx.doi.org/10.1145/355324.355325.

[19] B. Schilit, N. Adams, R. Want, Context-aware computing applications, First Workshop on Mobile Computing Systems Applications, IEEE, Santa CruzUSA, 1994, pp. 85–90, , http://dx.doi.org/10.1109/WMCSA.1994.16.

[20] P. Dourish, *Re*-Space-ing place: "Place" and "Space" ten years on, ACM Conference Computer Supportive Cooperative Work, Alberta, Canada, ACM Press, 2006, pp. 299–308, , http://dx.doi.org/10.1145/1180875.1180921.

[21] J. Brewer, P. Dourish, Storied spaces: Cultural accounts of mobility, technology, and environmental knowing, Int. J. Hum. Comput. Stud. 66 (2008) 963–976, http://dx.doi.org/10.1016/j.ijhcs.2008.03.003.

[22] H. Ogata, K. Matsuura, Y. Yano, Sharlok: bridging learners through active knowledge awareness, IEEE International Conference on Systems Man and Cybernetics, Beijing, China, IEEE, 1996, pp. 601–606, , http://dx.doi.org/10.1109/ICSMC.1996.569860.

[23] C. Collazos, L. Guerrero, J. Pino, S. Ochoa, Introducing knowledge-shared awareness, in: M. Boumedine (Ed.), Information and Knowledge Sharing, ACTA Press, St. Thomas, USA, 2002, pp. 13–18.

[24] K. Schmidt, L. Bannon, Taking CSCW seriously, Comput. Support. Coop. Work. 1 (1992) 7–40, http://dx.doi.org/10.1007/BF00752449.

[25] G. Polančič, G. Jošt, M. Heričko, An experimental investigation comparing individual and collaborative work productivity when using desktop and cloud modeling tools, Empir. Softw. Eng. 20 (2015) 142–175, http://dx.doi.org/10.1007/s10664-013-9280-x.

[26] N. Nova, T. Wehrle, J. Goslin, Y. Bourquin, P. Dillenbourg, Collaboration in a multi-user game: impacts of an awareness tool on mutual modeling, Multimed. Tools Appl. 32 (2007) 161–183, http://dx.doi.org/10.1007/s11042-006-0065-8.

[27] P. Khanal, A. Vankipuram, A. Ashby, M. Vankipuram, A. Gupta, D. Drumm-Gurnee, K. Josey, L. Tinker, M. Smith, Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles, J. Biomed. Inform. 51 (2014) 49–59, http://dx.doi.org/10.1016/j.jbi.2014.04.005.

[28] S. Björk, Patterns in Game Design, 1st ed., Charles River Media, 2004.

[29] J.B. Rocha, S. Mascarenhas, R. Prada, Game Mechanics for Cooperative Games, Zo. Digit. Games., 2008, https://fenix.tecnico.ulisboa.pt/downloadFile/395138343981/artigo.pdf.

[30] M. Seif El-Nasr, B. Aghabeigi, D. Milam, M. Erfani, B. Lameman, H. Maygoli, S. Mah, Understanding and evaluating cooperative games, 28th International Conference on Human Factors in Computing Systems (CHI '10), Atlanta, USA, ACM Press, 2010, p. 253, , http://dx.doi.org/10.1145/1753326.1753363.

[31] A. Beznosyk, P. Quax, K. Coninx, W. Lamotte, The influence of cooperative game design patterns for remote play on player experience, 10th Asia Pacific Conference on Computer Human Interaction (APCHI '12), Matsue-city, Japan, ACM Press, 2012, p. 11, , http://dx.doi.org/10.1145/2350046.2350051.

[32] Z.O. Toups, J. Hammer, W.A. Hamilton, A. Jarrah, W. Graves, O. Garretson, A framework for cooperative communication game mechanics from grounded theory, 1st ACM SIGCHI Annual Symposium on Computer-Human Interaction Play (CHI Play '14), Toronto, Canada, ACM Press, 2014, pp. 257–266, , http://dx.doi.org/10.1145/2658537.2658681.

[33] D. Vaddi, R.R. Wehbe, Z.O. Toups, S.N. Stahlke, R. Koroluk, L.E. Nacke, Validating test chambers to study cooperative communication mechanics in portal 2, Annual Symposium on Computer-Human Interaction Play (CHI Play '15), London, UK, ACM Press, 2015, pp. 721–726, , http://dx.doi.org/10.1145/2793107.2810302.

[34] D. Vaddi, Z. Toups, I. Dolgov, R. Wehbe, L. Nacke, Investigating the impact of cooperative communication mechanics on player performance in portal 2, 42nd Graphics Interface Conference, Canadian Human-Computer Communications Society, Victoria, Canada, 2016, pp. 41–48, , http://dx.doi.org/10.20380/gi2016.06.

[35] C. Fabricatore, M. Nussbaum, R. Rosas, Playability in action videogames: a qualitative design model, Hum.-Comput. Interact 17 (2002) 311–368, http://dx.doi.org/10.1207/S15327051HCI1704_1.

[36] L. Caroux, K. Isbister, L. Le Bigot, N. Vibert, Player–video game interaction: a systematic review of current concepts, Comput. Hum. Behav. 48 (2015) 366–381, http://dx.doi.org/10.1016/j.chb.2015.01.066.

[37] R. Bernhaupt (Ed.), Game User Experience Evaluation, Springer International Publishing, Cham, 2015, , http://dx.doi.org/10.1007/978-3-319-15985-0.

[38] N. Bevan, J. Carter, S. Harker, ISO 9241-11 revised: what have we learnt about usability since 1998? 17th International Conference on Human-Computer Interaction (HCII 2015), 2015, pp. 143–151, , http://dx.doi.org/10.1007/978-3-319-20901-2_13.

[39] M. Zhu, X. Fang, Developing Playability Heuristics for Computer Games from Online Reviews, in: A. Marcus (Ed.), Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience, (DUXU 2014), Lecture Notes in Computer Science, vol 8517, Springer, Cham, 2014.

[40] E.D. Mekler, J.A. Bopp, A.N. Tuch, K. Opwis, A systematic review of quantitative studies on the enjoyment of digital entertainment games, Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, ACM, New York NY USA, 2014, pp. 927–936, , http://dx.doi.org/10.1145/2556288.2557078.

[41] G. Chanel, J.M. Kivikangas, N. Ravaja, Physiological compliance for social gaming analysis: cooperative versus competitive play, Interact. Comput. 24 (2012) 306–316, http://dx.doi.org/10.1016/j.intcom.2012.04.012.

[42] J. de Man, Analysing emotional video using consumer EEG hardware, in: M. Kurosu (Ed.), Human-Computer Interaction Advanced Interaction Modalities Techniques, Springer International Publishing, 2014, pp. 729–738, , http://dx.doi.org/10.1007/978-3-319-07230-2_69.

[43] J. Cusveller, C. Gerritsen, J. de Man, Evoking and measuring arousal in game settings, in: S. Göbel, J. Wiemeyer (Eds.), Games Training, Education, Health and Sports, Springer International Publishing, 2014, pp. 165–174, , http://dx.doi.org/10.1007/978-3-319-05972-3_17.

[44] H. Desurvire, C. Wiberg, Game usability heuristics (PLAY) for evaluating and designing better games: the next iteration, in: A.A. Ozok, P. Zaphiris (Eds.), Online Communities and Social Computing, Springer, Berlin Heidelberg, 2009, pp. 557–566, , http://dx.doi.org/10.1007/978-3-642-02774-1_60.

[45] J. Paavilainen, Critical review on video game evaluation heuristics: social games perspective, International Academic Conference on the Future of Game Design and Technology (Futureplay '10), ACM Press, VancouverCanada, 2010, p. 56, , http://dx.doi.org/10.1145/1920778.1920787.

[46] D. Pinelle, N. Wong, T. Stach, Heuristic evaluation for games: usability principles for video game design, SIGCHI Conference on Human Factors in Computing Systems, FlorenceItaly, 2008, pp. 1453–1462, , http://dx.doi.org/10.1145/1357054.1357282.

[47] A. Calderón, M. Ruiz, A systematic literature review on serious games evaluation: an application to software project management, Comput. Educ. 87 (2015) 396–422, http://dx.doi.org/10.1016/j.compedu.2015.07.011.

[48] D. Djaouti, J. Alvarez, J.-P. Jessel, Classifying Serious Games, in: P. Felicia (Ed.), Handbook of Research Improving Learning and Motivation Through Educational. Games : Multidisciplinary Approaches, IGI Global, 2011, pp. 118–136, , http://dx.doi.org/10.4018/978-1-60960-495-0.ch006.

[49] S. Deterding, M. Sicart, L. Nacke, K. O'Hara, D. Dixon, Gamification. using game-design elements in non-gaming contexts, 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems (CHIEA '11), ACM Press, VancouverCanada, 2011, pp. 24–25, , http://dx.doi.org/10.1145/1979742.1979575.

[50] J. Hamari, J. Koivisto, H. Sarsa, Does Gamification Work? – A Literature Review of Empirical Studies on Gamification, 2014 47th Hawaii International Conference on System Sciences, Waikoloa, USA, IEEE, 2014, pp. 3025–3034, , http://dx.doi.org/10.1109/HICSS.2014.377.

[51] T. Bosse, J. De Man, C. Gerritsen, Agent-based simulation as a tool for the design of a virtual training environment, 2014 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Warsaw, Poland, IEEE, 2014, pp. 40–47, , http://dx.doi.org/10.1109/WI-IAT.2014.148.

[52] M. García, M.A. Teruel, E. Navarro, P. González, A. Fernández-Caballero, A distributed tool to perform dynamic therapies for social cognitive deficit through avatars, in: S. Ochoa, P. Singh, J. Bravo (Eds.), Ubiquitous Computing and Ambient Intelligence, Springer International Publishing, 2017, pp. 731–741, , http://dx.doi.org/10.1007/978-3-319-67585-5_71.

[53] M.A. Teruel, M. Oliver, F. Montero, E. Navarro, P. González, Multisensory treatment of the hemispatial neglect by means of virtual reality and haptic techniques, in: J.M. Ferrández Vicente, J.R. Álvarez-Sánchez, F. de la, P. López, F.J. Toledo-Moreo, H. Adeli (Eds.), Artificial Computation in Biology and Medicine, Springer International Publishing, 2015, pp. 469–478, , http://dx.doi.org/10.1007/978-3-319-18914-7_49.

[54] C. Gresse von Wangenheim, M. Thiry, D. Kochanski, Empirical evaluation of an educational game on software measurement, Empir. Softw. Eng. 14 (2009) 418–452, http://dx.doi.org/10.1007/s10664-008-9092-6.

[55] V. Vallejo, P. Wyss, L. Rampa, A.V. Mitache, R.M. Müri, U.P. Mosimann, T. Nef, Evaluation of a novel serious game based assessment tool for patients with Alzheimer's disease, PLoS One 12 (2017) e0175999, , http://dx.doi.org/10.1371/journal.pone.0175999.

[56] A. Fuchslocher, K. Gerling, M. Masuch, N. Kramer, Evaluating social games for kids and teenagers diagnosed with cancer, IEEE 1st International Conference on Serious Games and Applications for Health, Washington DC, USA, IEEE, 2011, pp. 1–4, , http://dx.doi.org/10.1109/SeGAH.2011.6165463.

[57] P. Herzig, S. Strahringer, M. Ameling, Gamification of ERP systems-exploring gamification effects on user acceptance constructs, Multikonferenz Wirtschaftsinformatik, 2012, pp. 793–804 http://rzbl04.biblio.etc.tu-bs.de:8080/

docportal/servlets/MCRFileNodeServlet/DocPortal_derivate_00027481/Beitrag113.pdf.
[58] R. Zaidi, The Ultimate SAP User Guide: The Essential SAP Training Handbook For Consultants and Project Teams, Ecruiting Alternatives, Incorporated, 2015.
[59] V. Charisis, S. Hadjidimitriou, L. Hadjileontiadis, D. Uğurca, E. Yilmaz, EmoActivity - An EEG-based gamified emotion HCI for augmented artistic expression: the i-treasures paradigm, Universal Access in Human–Computer Interaction. Access to the Human Environment and Culture, Springer International Publishing Switzerland, 2015, pp. 29–40, http://dx.doi.org/10.1007/978-3-319-20687-5_4.
[60] M. Wrzesien, M. Alcañiz Raya, Learning in serious virtual worlds: evaluation of learning effectiveness and appeal to students in the E-Junior project, Comput. Educ. 55 (2010) 178–187, http://dx.doi.org/10.1016/j.compedu.2010.01.003.
[61] A. Rollings, D. Morris, Game Architecture and Design: A New Edition, 1st ed., New Riders, 2003.
[62] V.R. Basili, G. Caldiera, H.D. Rombach, The goal question metric approach, Encyclopedia of Software Engineering, Wiley, 1994, pp. 528–532 http://wwwagse-old.informatik.uni-kl.de/pubs/repository/basili94b/encyclo.gqm.pdf (accessed June 1, 2011).
[63] ISO/DIS, ISO/DIS 9241-11: Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts, 2015.
[64] B.A. Kitchenham, S.L. Pfleeger, L.M. Pickard, P.W. Jones, D.C. Hoaglin, K. El Emam, J. Rosenberg, Preliminary guidelines for empirical research in software engineering, IEEE Trans. Softw. Eng. 28 (2002) 721–734, http://dx.doi.org/10.1109/TSE.2002.1027796.
[65] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, A. Wesslén, Experimentation in Software Engineering, 1st ed., Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
[66] Unity Technologies, Tanks Tutorial, 2015. https://unity3d.com/learn/tutorials/projects/tanks-tutorial.
[67] W. Scacchi, Computer game mods, modders, modding, and the mod scene, First Monday 15 (2010), http://firstmonday.org/article/view/2965/2526.
[68] M. Menard, B. Wagstaff, Game Development with Unity, Updated, Cengage Learning, 2014.
[69] Unity Technologies, Multiplayer and Networking, http://docs.unity3d.com/Manual/UNet.html, (2015) (accessed February 2, 2016).
[70] A. Cockburn, C. Gutwin, J. Scarr, S. Malacria, Supporting novice to expert transitions in user interfaces, ACM Comput. Surv. 47 (2014) 1–36, http://dx.doi.org/10.1145/2659796.
[71] C.A. Collazos, L.A. Guerrero, M.A. Redondo, C. Bravo, Visualizing shared-knowledge awareness in collaborative learning processes, in: J.M. Haake, S.F. Ochoa, A. Cechich (Eds.), 13th International WorkShop, Springer-Verlag Berlin Heidelberg, Bariloche, Argentina, 2007, pp. 56–71.
[72] J. Brooke, SUS: a “quick and dirty” usability scale, in: P.W. Jordan, B. Thomas, B.A. Weerdmeester, A.L. McClelland (Eds.), Usability Evaluation in Industry, Taylor and Francis, London, 1996, pp. 189–194.
[73] P.W. Farris, N.T. Bendle, P.E. Pfeifer, D.J. Reibstein, Marketing Metrics: The Definitive Guide to Measuring Marketing Performance, 2nd ed., Pearson FT Press, 2010, http://www.amazon.com/Marketing-Metrics-Definitive-Measuring-Performance/dp/0137058292.
[74] H. Joho, J.M. Jose, R. Valenti, N. Sebe, Exploiting facial expressions for affective video summarisation, ACM International Conference on Image and Video Retrieval (CIVR '09), ,, ACM Press, AmsterdamThe Netherlands, 2009, , http://dx.doi.org/10.1145/1646396.1646435 31:1-31:8.
[75] A.A. Salah, T. Gevers (Eds.), Computer Analysis of Human Behavior, Springer, 2011.
[76] R. Valenti, N. Sebe, T. Gevers, I. Cohen, Machine Learning Techniques For Face Analysis, Springer, 2008.
[77] N. Heni, H. Hamam, Facial emotion detection of smartphone games users, 2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering, IEEE, 2015, pp. 1243–1247, , http://dx.doi.org/10.1109/CCECE.2015.7129456.
[78] A. Field, Discovering Statistics Using SPSS, 3rd ed., SAGE Publications Ltd, 2009.
[79] H. Abdi, The greenhouse-geisser correction, Encycl. Res. Des. SAGE Publications, Thousand Oaks, 2010.
[80] M. Csikszentmihalyi, Beyond Boredom and Anxiety, Jossey-Bass Publishers, 1975.
[81] F. Southey, G. Xiao, R.C. Holte, M. Trommelen, J. Buchanan, Semi-automated gameplay analysis by machine learning, Conference on Artificial Intelligence and Interactive Digital Entertainment, Marina del ReyUSA, 2005, pp. 123–128.
[82] C.M. Grinstead, J.L. Snell, Introduction to Probability, 2nd ed., American Mathematical Society, 2006.
[83] M.A. Teruel, E. Navarro, V. López-Jaquero, F. Montero, P. González, A comprehensive framework for modeling requirements of CSCW systems, J. Softw. Evol. Process. (2017), http://dx.doi.org/10.1002/smr.1858.
[84] M.A. Teruel, E. Navarro, V. López-Jaquero, F. Montero, J. Jaen, P. González, Analyzing the understandability of requirements engineering languages for CSCW systems: a family of experiments, Inf. Softw. Technol. 54 (2012) 1215–1228, http://dx.doi.org/10.1016/j.infsof.2012.06.001.
[85] M.A. Teruel, E. Navarro, V. López-Jaquero, F. Montero, P. González, A CSCW requirements engineering CASE tool: development and usability evaluation, Inf. Softw. Technol. 56 (2014) 922–949, http://dx.doi.org/10.1016/j.infsof.2014.02.009.