HAL
open science

Usability Evaluation: Commonalities and Discrepancies in Games and Safety Critical Systems

Philippe Palanque, Regina Bernhaupt, Marco Winckler

To cite this version:

Philippe Palanque, Regina Bernhaupt, Marco Winckler. Usability Evaluation: Commonalities and Discrepancies in Games and Safety Critical Systems. European Workshop on HCI Design and Evaluation, Mar 2011, Limassol, Cyprus. pp.45–53. hal-03647161

HAL Id: hal-03647161

https://hal.science/hal-03647161v1

Submitted on 20 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Usability Evaluation: Commonalities and Discrepancies in Games and Safety Critical Systems.

Philippe Palanque
IRIT, Groupe ICS
118 Route de Narbonne
Toulouse, 31062, France
palanque@irit.fr

Regina Bernhaupt
IRIT, Groupe ICS
118 Route de Narbonne
Toulouse, 31062, France
regina.bernhaupt@irit.fr

Marco Winckler
IRIT, Groupe ICS
118 Route de Narbonne
Toulouse, 31062, France
winckler@irit.fr

.copyright notice to be completed.

Abstract

This paper presents the basic principles of widely agreed upon practice for usability evaluation in two different domains: Games and Safety Critical Interactive Systems. The user interfaces (and their associated interaction techniques) proposed in the applications deployed in these two domains are usually very different while the objects and functions to be manipulated might be very similar. For instance in a flight simulator game a rear view of the aircraft is usually proposed (see Figure 1), this rear view is also proposed in flight simulators (see Figure 2) while in a real cockpit such representation is not possible (see cockpit capture Figure 7). This is surprising as the key point in flight simulator games is to propose an as-realistic-as-possible cockpit interface.

Based on an overview on the development processes and the respective usability evaluation methods used in these two domains, a comparison of the approaches showing commonalities and discrepancies is presented. Concluding remarks focus on how these two domains could fruitfully cooperate, by adopting and enhancing methods used in the two domains. Additionally, such comparison could provide foundations for trans-

Figure 1. Screen shot from piloting game (outside view of the aircraft)

Figure 2. A rear view of the VRS Hornet on Flight Simulator

sectorial usability evaluation methods identifying both ground for common shared tools and techniques and also identifying idiosyncrasies.

Keywords
Games, safety-critical systems, usability evaluation, development processes.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human-Factors, Reliability.

Introduction
Since early work on software quality from McCall in 1977 [8], usability has been explicitly identified (even within the software engineering community) as a key factor for the quality of software. However, for the evaluation of usability of any interactive system it is important to take into consideration the application domain as the application domains can heavily influence how the usability is evaluated. For instance, when evaluating usability in games it is important to take into account that games typically are not task based and that playing a game is (by definition [4]) non-productive and oriented towards entertainment. On the other hand, in domains where safety plays a more important role than usability, this preponderant factor influences usability evaluation and the interpretation of results: detected usability problems in a given user interface might not be resolved by improving the user interface, but by training the user to simply learn how to use the problematic interface in a safe and error-free manner.

Goal of this research is to compare the domains of safety-critical systems and games in terms of usability evaluation methods. We argue that both domains can benefit from each other by incorporating the knowledge available in the other domain into their usability evaluation approaches.

First, the two domains are described, giving an overview on the development processes used, the applied usability evaluation methods (including their need of technical infrastructure). Next the approaches in the domains are compared showing commonalities and discrepancies. Concluding remarks highlight potentialities for the two domains to learn from each other. Lastly, we outline a draft research agenda for a more global approach on trans-sectorial usability evaluation methods.

Games
For the development of games there is a common agreement in the community that successful games rely on an iterative development approach. Usability evaluation is an important aspect in games development: if a game is not usable (e.g. the interaction technology does not allow easy learning how to play the game), a game is typically not successful. There is a broad range of game development processes used, the majority of them having special evaluation phases (including usability evaluation).

Games Development Processes
When looking at the development of games, standard software engineering approaches like the waterfall-model or Boehm's spiral model have been described as useful for game development. In the majority of cases a successful game development company does not

Figure 3. Test bench for UAVs command and control system (Copyright © 1999-2011 by Aero-News Network, Inc. All rights reserved)

Figure 4. Microsoft playtest lab

Figure 5. Games usability testing in the field

release any information about the software development process used. Novak [9]reports on a mixture of concepts and methods that support iterative development as well as flexibility in the development process.

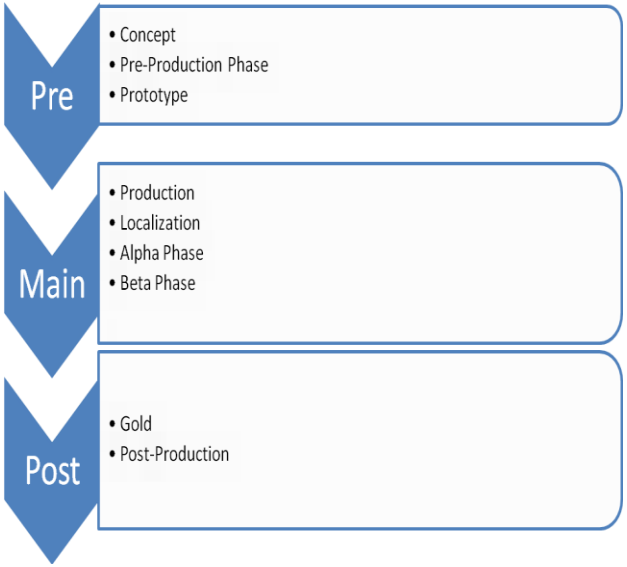

Figure 6 an example of game development process model (showing the 3 main phases)

Figure 6 presents an abstract view of a game development process. We have not represented the iterations that are of typical HCI practice (in the pre-phase), typical software engineering practice such as presented in the V model (in the main-phase) and closer to the agile processes (in the post-phase).

The following phases typically guide a game development process [9]:

(1) **Concept**: This phase is dedicated to the initial game idea and is devoted to producing a first concept document describing the game. The development team in this phase is typically small (e.g. consisting of designer, programmer, artist and producer).
(2) **Pre-Production Phase**: This phase includes the development of art style guides, production plans and first description of the game design and the technical design document.
(3) **Prototype**: Goal of this phase is a first working piece of software allowing to demonstrate key characteristics of the game and enabling to understand basic concepts related to the general user experience of the game ("Is the game fun to play?").
(4) **Production**: The production phase can range from few weeks development to years of programming. This phase can be structured additionally, following approaches like an increment to completion approach, a cascade approach or an "iterative- until you drop" approach [6].
(5) **Localization**: an important phase for games that will be delivered to different markets (countries) is the localization phase. In this phase game-play can be adjusted to suit the tastes of the market, to allow language translation and modifications due to local regulatory authorities.
(6) **Alpha-Phase**: This is the phase when a game is playable from start to finish, allowing different evaluation methods to be applied to better understand aspects like fun, playability and user experience.
(7) **Beta-Phase**: Main goal during this phase is normally to fix bugs. In terms of user experience in this phase lots of fine-tuning is necessary to improve the overall user experience. The beta-phase includes steps like certification or submission (the hardware-manufacturer of the proprietary platform will test the

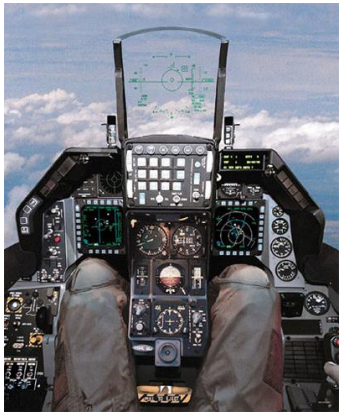

Figure 7. F-16 A/B fighter cockpit – inside view of the aircraft

game).
(8) **Gold**: In this phase the game is sent to be manufactured.
(9) **Post-Production**: In this phase subsequent versions of the game may be released (including patches and updated) and allows to improve the user experience of the game.

Games development has some additional special milestone (depending on the product developed). Especially the release candidate milestone is important for games that are produced for special consoles, as the release candidate is then tested and evaluated by the game console manufacturer.

Usability Evaluation in Games Development Processes

METHODS

For the usability evaluation there is a broad range of methods used, that are applied in the different phases of the game development. At the early stages of the development (concept phase) these methods are: focus groups, interviews, informal play-testing, questionnaires; at later stages, especially during the production phase these are play testing (including bio-metrical measurements), (semi-structured) interviews, observation, video coding, quantitative comparisons of gamers behaviours, questionnaires focusing on users attitudes or experiences or heuristic evaluation (including the usage of heuristics that focus additionally on the playability of the game).

Two usability evaluation methods from the entertainment and games domain have become a focus of attention in other areas (1) play testing and (2) beta-testing. Play tests are used during all phases of a game development, with the majority of usability and play testing performed once the alpha phase is accomplished until the beta-phase is finished. (Alpha is the milestone in games development were the game can be played the first time from the beginning to the end, still missing some branches in the game that are not finalized, or graphics that are not ready; Beta is milestone in the game development when no further developments and changes are made for the game play, and the only goal is to get rid of bugs). Beta-testing is a form of game-play evaluation that also includes usability evaluation. During Beta tests the game is distributed to (a selected community of) end users that are playing the game. These users (sometimes thousands of users) provide feedback in terms of usability (e.g. mapping of the input to the various game elements) as well as in terms of game-play (allowing for example the fine-tuning of game-rules based on player-death logging).

EQUIPMENT

Usability evaluation of games is conducted both in the lab (Figure 4) and in the field (Figure 5). To enable the evaluation of usability at early development stages prototypes are developed and used. To understand how usability and overall user experience and fun are related usability evaluation for games development today rely on sophisticated usability labs, including bio-physiological measurements, eye-tracking or motion sensors.

Safety-Critical Interactive Systems Development Processes

In safety critical systems several quality factors deeply influence the development process such as reliability, fault-tolerance or security. However, usability could/should be seen as a critical factor especially when

work such as [10] report that 80%-90% of accidents in industry are attributed to human error.

One key element of the development process of safety critical systems is the certification phase which involves an external body who is in charge of assessing the process that has been followed during the development of the system. Figure 8 presents such a process in which the last phase precisely concerns certification activities. It is important to note that, however, such phase is more prominent when the system involves "normal" citizens.

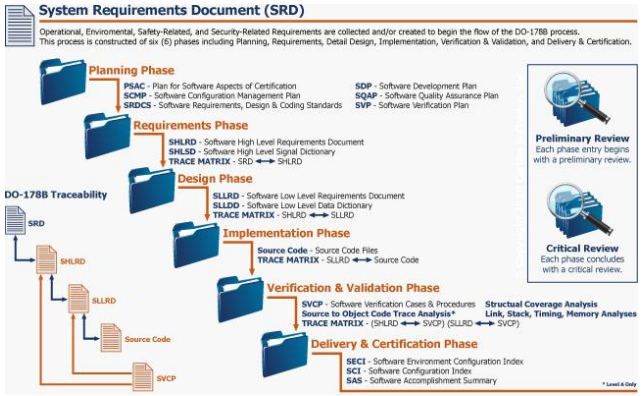

Figure 8. DO 178B [1] standard (no usability – no users)

Another typical aspect of the safety critical domain is the fact that standards processes are available and should be carefully followed. Figure 8 is defined in the DO 178-B standard from [1] while Figure 9 is from another standard IEC 61511 dealing specifically with safety issues [5].

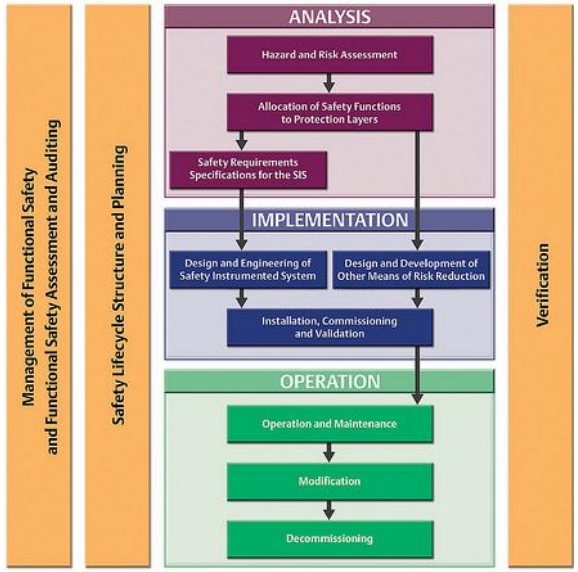

Figure 9. IEC 61511 [5] Process Safety (operations are part of the process)

Usability Evaluation in Safety-Critical Interactive Systems

Usability evaluation in safety critical systems is typically conducted as part of the development processes presented above. However, it is usually not an explicit phase as usability concerns are taken into account in more or less every phase.

From the scientific perspective, usability evaluation is scarcely described or even reported. From our experience, for instance in the case of civil aircrafts cockpit designs, usability and operability concerns are dealt with by means of expert users (called test pilots) who are involved from the beginning (architectural

design of the cockpit) to the very late final phases (design of the training of airlines pilots). This could be compared to participatory design approaches.

Using a simulator as a test-bed for design and testing of concepts and interfaces can be seen as a specific equipment or supporting usability studies. For instance, while designing the elements of a head-up display, scenarios can be defined and tested to assess the impact of a display freeze of the head-up display on the operations (i.e. the pilot's tasks).

To summarize a broad range of usability evaluation methods are used in this area but they typically do not refer and relate to the standard method descriptions advocated in the field of HCI.

Equipment for Usability Evaluation
As for games, usability evaluation for safety critical systems is conducted both in the lab and in the field. But given the necessary focus on safety and reliability, evaluations in the lab are based on the extensive usage of very complex and realistic simulators (e.g. force feedback systems to simulate the behavior of physical equipment in the cockpit, vibrations, 3D movements of the cabin …). Figure 3 presents a realistic simulator for a UAV (unmanned aerial vehicle) control and command system embedding a physical model of the aircraft, weather, gravity ….

When usability evaluation is carried out in the field, very specific equipment has to be designed and embedded to gather information about the operations and the pilot. As for games, they can collect biometric measures about the crew but in addition, they have to be very robust to variations as the evaluation has to be as close as possible of the real operations. For instance, eye tracker devices must be able to track eyes movements both in the dark (while landing at night) and while facing sun (while departing East in the morning or landing West in the evening) [1].

Lastly, usability testing in safety critical systems exhibits the following constraints:

- Expensive users (retired pilot, current pilots from airlines, test pilots working for the manufacturer),
- Expensive real tests for a new designs/systems (50k€ one hour flight test to be included in the budget),
- Complex tests scenarios that are usually based both on normal and abnormal conditions (previous incidents/accidents, emergency situations, partial failure of systems …).

Summary of Similarities and Discrepancies

Usability is usually evaluated according to the time spent in training (costly training is perceived as poor design of the UI).

Games - Process	**Safety-Critical**
Development Process	
Process Model: iterative	Process Model: waterfall or V model
Specialized development phases include localization,	Specialized development phase include certification

internationalization, gold	
Special verification phase by game console manufacturer	Special customization phase of procedures by the airline
System should be of type walk-up and use (or with a very simple practice scenario included)	System is designed together with training material. Similarity with respect to other interfaces is relevant (to support cross-qualification of pilots).
Usability Evaluation Methods	
Usability can be less important than user experience	Safety and reliability are more important than usability (User experience is not an issue).
Usability evaluation is combined with user experience evaluation (e.g. playability tests)	Usability (performance and efficiency) is prominent for the definition and validation of operations
Beta-testing as a unique way to gather user feedback before the software is released. This generates a huge amount of data for improvements (e.g. balancing).	Incident and accident reporting activities constitute longitudinal studies and inform future designs and re-designs.

Concluding Remarks

Looking at the specificities of the two domains it is surprising to find a set of similarities: both domains share the approach that usability as a factor of the interaction with the system might be less important than other software quality factors. Both domains have a set of additional development phases and usability evaluation is typically quite heavy in terms of infrastructure. Being interested in the domain dependent adaptation of usability evaluation, the domain of safety-critical system and the domain of games are promising candidates to understand how usability evaluation methods must be adapted to better fit the domain.

We believe that safety critical systems and games represent the two extremes of a continuum of domains where usability evaluation methods have to be customized and adapted to fit their specificities.

In a nutshell and to provide a second adaptation of Landsburg's words as stated in [2]: "Usability wants us to die rich; UX wants us to die happy", we propose a new vision of it by saying “Usability wants us to die fit; UX wants us to die happy; Safety want us to die still alive”.

Acknowledgements

This work is partly funded by Airbus under the contract CIFRE PBO D08028747-788/2008, R&T CNES (National Space Studies Center) Tortuga R-S08/BS-0003-029 and ruwido.

References

[1] Dehais, F., Causse M. & Pastor J. Toward the definition of a pilot’s physiological state vector through

oculometry: a preliminary study in real flight conditions. In proceedings of HCI Aero 2010, ACM Digital Library.

[2] DO-178B (1992) Software Considerations in Airbone Systems and Equipment Certification. RTCA Inc, EUROCAE, December

[3] Hassenzahl M., Law E & Hvannberg E.T. (2006) User Experience – Towards a unified view. In proceedings of "User Experience – Towards a Unified View: Second International COST294-MAUSE Open Workshop", NordiCHI 2006. P. 1-3.

[4] Huizinga J (1950) Homo Ludens: A Study of the Play-Element in Culture. Beacon Press, Boston, MA

[5] IEC 61511 Functional safety - Safety instrumented systems for the process industry sector. Beuth Verlag GmbH, Berlin, 2003.

[6] Irish D (2005) The Game Producer's Handbook. Thomson Course Technology PRT.

[7] Landsburg S. (1993) "The Armchair Economist: Economics and Everyday Life", The Free Press, 241 p.

[8] Mc Call J (1977) Factors in software quality. (Ed.) General Electric.

[9] Novak J (2008) Game Development Essentials. Delmar Cengage Learning.

[10] Salminen S. And T. Tallberg, Human errors in fatal and serious occupational accidents in Finland, Ergonomics 39 (1996) (7), pp. 980–988