

The Design of Learning Material for Poor Comprehenders: Lessons Learnt from Experts

Vincenza Cofini^a, Fernando De la Prieta^b, Tania Di Mascio^c, Rosella Gennari^d, Ivana Marenzi^e, Pierpaolo Vittorini^{a*}

^aUniversity of L'Aquila, MeSVA, Italy

^bUniversity of Salamanca, Department of Computer Science and Automation Control, Spain

^cUniversity of L'Aquila, DISIM, Italy

^dFree University of Bozen-Bolzano, Faculty of Computer Science, Italy

^eUniversity of Hanover, L3S, Germany

*Corresponding author: pierpaolo.vittorini@univaq.it

Article history

Received :11 September 2012

Received in revised form :

21 February 2013

Accepted :15 April 2013



Abstract

TERENCE is an FP7 ICT European project that is developing an adaptive learning system for poor comprehenders and their educators. The learning material is made of stories and smart games for stimulating reading comprehension. The design of stories and smart games is also based on data collected from experts for the analysis of the context of use of the system, and is incrementally revised via evaluations of prototypes of stories and games, with domain experts of text comprehension or education as participants. In particular, since smart games are semi-automatically generated via artificial intelligence technologies, they contain mistakes that have to be fixed by experts of pedagogy before the games are given to learners. In this paper we focus on the design and evaluations of the TERENCE stories and smart games for poor comprehenders via lessons learnt with domain experts.

Keywords: Serious games; artificial intelligence; heuristics; user centred design; evidence based design; adaptive learning system

© 2013 Penerbit UTM Press. All rights reserved.

1.0 INTRODUCTION

1.1 Background

Nowadays, more and more children turn out to be poor (text) comprehenders: they demonstrate difficulties in deep text comprehension, despite well developed low-level cognitive skills like decoding of words, e.g., see [17]. There is clear evidence that reasoning skills like inference-making are very likely to be causally implicated in the development of deep text comprehension. In particular, experiments show that inference-making questions concerning challenging but not exceedingly difficult stories, together with adequate visual aids, are pedagogically effective in fostering the deep comprehension of stories. TERENCE is a Collaborative Project funded by the EC under the ICT Call 5 FP7- ICT-2009-5 that developed the first *adaptive learning system* (ALS) with stories and reading interventions in the form of smart games, all designed within a stimulation plan for poor comprehenders.

1.2 Rationale of This Paper

The TERENCE primary users are *learners*, hearing or deaf poor text comprehenders, from primary schools, older than 7 years and without decoding difficulties. The other types of main users of the TERENCE ALS are the learners' *educators*, e.g., their

teachers. The *learning material* is in English and in Italian, and is made of *stories*, collected into books, and of *smart games*, for drawing inferences about each story. The paper sketches the process of design of the learning material of TERENCE starting from the analysis of the learner requirements and stepping through evaluations with experts of pedagogy, psychology and linguistics. In [2], the reader can find information on the refined characterisation of the TERENCE learners for the ALS of TERENCE.

1.3 Outline of This Paper

The paper starts sketching the design methodologies chosen in TERENCE. Having laid the groundwork, the paper delves into the three main stages for the development of the learning material of TERENCE (and the TERENCE system in general). Firstly, it outlines the data gathering with learners, educators and domain experts for setting the requirements for the learning material and tasks, concentrating on the data gathering with experts. Secondly, it illustrates how the requirements lead to the design of the TERENCE learning material, e.g., of game instances via the TERENCE game framework. Finally, the paper presents the evaluations with experts of the designed learning material via specific usability methods. Evaluations, in turn, provide designers with inputs for revising the design process and the development of the learning material, in particular, for the

automated development of textual components of smart games via artificial intelligence technologies. A conclusive section recaps the lessons learnt and paves the way for future work.

■2.0 THE DESIGN METHODOLOGIES

The TERENCE system is developed following the evidence-based design (EBD) [21] and the user-centred design (UCD) [13] so as to attain pedagogical effectiveness and usability, respectively. For producing a *pedagogically effective system*, EBD stresses the need of basing the system design on empirical evidence, also gathered from *domain experts*, namely, diagnosis or stimulation plan experts for poor text comprehension and linguists. For producing a *usable system*, UCD places the users at the centre of the design process, and iteratively designs the system starting with the analysis of the context of use and revising prototypes of the system through evaluations with users or UCD experts.

In TERENCE, the analysis of the context of use leads/led to the specification of requirements concerning:

- (1) the characteristics of the users, such as knowledge, skills, experience;
- (2) the learning tasks and their organisation into a stimulation plan;
- (3) the environment.

The analysis was done via data gathering activities with the TERENCE learners, educators and domain experts for the learners. Prototypes (of components) of the TERENCE ALS were then incrementally developed and progressively evaluated with users and domain experts. In the following section we focus on the context of use analysis of the characteristics of the users and of the learning tasks carried out by domain experts. Afterwards, we show how that leads/led to the design of stories and smart games, and afterwards to their expert-based evaluations.

■3.0 CONTEXT OF USE ANALYSIS

3.1 Introduction

The context of use analysis, run at the beginning of the project, considered the characteristics of the TERENCE users and learning tasks. In this section, firstly we sketch the *design* and the *participants* of the data gathering activities, and then we focus on the *main results* for the design of the TERENCE learning material, focusing on stories and smart games.

3.2 Data Gathering Design and Participants

The data were collected in the UK and in Italy between September and December 2010, with c.a 60 educators, c.a 10 domain experts of diagnosis or stimulation plan experts for poor text comprehension, and c.a 500 learners. Additional details can be found in [17]. The data collection methods were UCD contextual inquiries with domain experts and educators, and game-based field studies with primary-school classes.

3.3 The Main Results

In this section, we focus on the data gathered for designing the stories and smart games of TERENCE, as well as on the resulting main requirements.

3.3.1 Requirements for Stories

The data gathered for the TERENCE stories include (1) text features relevant for the text comprehension of the TERENCE learners, (2) story features relevant for the satisfaction of the learners, and (3) story illustration constraints.

As for (1), domain experts of the diagnosis of poor comprehension specified the reading skills of the TERENCE learners according to the available evidence in their literature. The gathered data can be clustered into three areas:

- (1) the area of global consistency skills for integrating textual information necessary for the comprehension of the entire story,
- (2) the area of local cohesion skills for correlating information not distant in the text, e.g., the proper understanding of cohesive devices like temporal connectives that signal temporal relations between events in the same sentence or paragraph,
- (3) the area of lexicon and grammar skills, e.g., the ability of inferring the meaning of unknown words or polysemes from contextual clues.

That yields that the TERENCE stories should be organised into four main levels of difficulty:

- (1) at the top level should be stories that are challenging for global coherence, local cohesion, lexicon and grammar;
- (2) at the top-intermediate level should be stories that are challenging for local cohesion, lexicon and grammar;
- (3) at the entry-intermediate level should be stories that are challenging for lexicon and grammar;
- (4) at the entry level should be stories that are simplified for all the aforementioned features.

As for (2), the data mainly result from game-based field studies with the TERENCE learners and contextual inquiries with teachers. The main results are concerned with the plot and characters that are appealing for nowadays' children. The main requirements are that the plot should privilege the narrative part, include fantasy elements, and focus on actions for the younger learners and on emotions for the older learners. Moreover, the characters should be mainly children of the same age range as the TERENCE learners.

As for (3), the main data come from domain experts of stimulation plans, requiring that illustrations alone should not convey information that is the focus of the TERENCE smart games. That was translated into the requirements that the temporal flow and causal relations between relevant events should not be passed by story illustrations alone, which should only decorate and introduce the main characters of the stories in the main story episodes. Table 1 recaps such requirements for stories.

3.3.2 Requirements for Smart Games

The initial data for the design of smart games were reading interventions used by educators in class. Such interventions mainly refer to the stages outlined and scheduled in Figure 1: first reading, mainly aloud, in class; secondly, the analysis of the text, mainly via inference-making question-answering; thirdly, the emotional interpretation of the text; finally, the rewarding stage with children with reading difficulties, like deaf children.

Table 1 Requirements for stories

ID	What for	Description
s1	Difficulty levels	Stories should be layered into difficulty levels for the TERENCE learners, namely, <ul style="list-style-type: none"> – top: stories challenging for global consistency, local cohesion, grammar and lexicon; – top-intermediate: stories challenging for local cohesion, grammar and lexicon; – entry-intermediate: stories challenging for grammar and lexicon; – entry: stories simplified for all the aforementioned aspects.
s2	Plot	The plot should have a predominant narrative part with fantasy elements, favouring actions for younger learners and emotions for older learners.
s3	Characters	The main characters should resemble the TERENCE learners.
s4	Illustrations	Illustrations should not represent the temporal flow or causal relations between relevant events of the story.

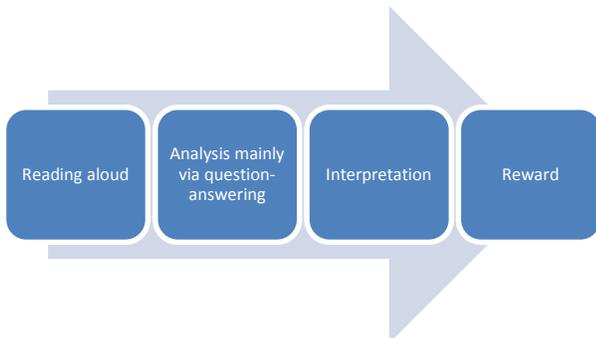


Figure 1 Reading interventions and their scheduling by teachers

In turn, the analysis stage can be broken down into the following interventions, done in class, mainly using question-answering and drawing:

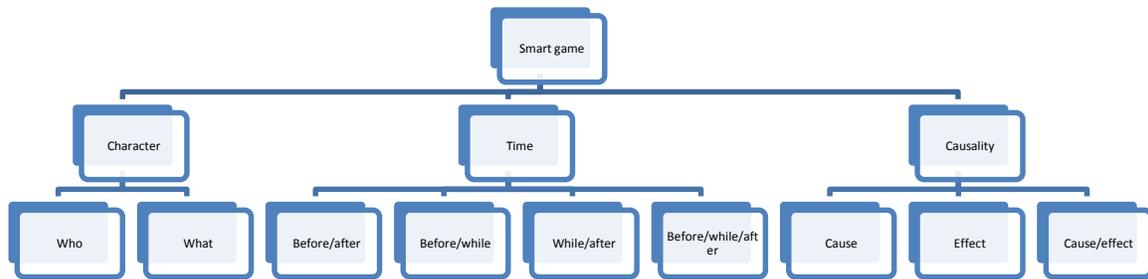


Figure 2 The TERENCE smart game taxonomy

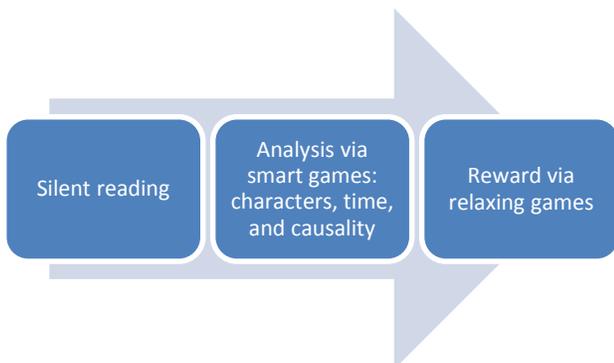


Figure 3 Reading interventions selected and scheduled by experts

The analysis interventions selected for the smart games were thus hierarchically organised and scheduled into difficulty macro-

- (1) the entire story is discussed with learners, analysing the unknown vocabulary and paraphrasing its text;
- (2) the story is broken down into its sequence of episodes, if possible, referring to the story grammar, that is, the story setting, the initiating episode, the culminating episode, the resolving episode, and the final episode;
- (3) finally, the time, space and characters of the story episodes are analysed together.

Such interventions were filtered by domain experts of stimulation plans, in line with the EBD. As a result of that sieve, for instance, drama exercises or other interventions meant for stimulating the learners’ empathy with the story characters were left out and the analysis was given a predominant role. See [11] for further details.

levels, according to their main pedagogical goal in the stimulation plan, as follows:

- at the entry macro-level, interventions focusing on characters, divided into the following levels: concerning who the agent of a story event is (who), and what a character in the story does (what);
- at the intermediate macro-level, interventions focusing on time, divided into five levels for reasoning about sequential and non-sequential temporal relations between events of the story;
- at the top macro-level, interventions focusing on causality, divided into three levels for reasoning about diverse causal-temporal relations between events.

That leads to the layering of smart games of Figure 2. Besides setting the difficulty levels, the domain experts for the stimulation plan set requirements for the actions the learner can perform while playing with smart games, that is, the learner should be allowed to:

- (a) skip a game,
- (b) choose a correct solution,
- (c) or choose wrong solutions.

Moreover, the stimulation plan sets constraints on the learner's actions: the learner should be allowed to choose a wrong solution until the correct solution "becomes obvious"; however, the learner must improve on the time s/he takes in resolving smart games, also because of organisational constraints on the stimulation that is done in class as a school activity, which in turns constrains how long the learner can spend on resolving smart games for a story.

The experts of the plan also recommend to show the learner his/her progress and diverse types of feedback: a yes or no clear feedback, explanation bits in case of a wrong solution, showing the solution before leaving the game, and soliciting the learner to make a choice if the learner is inactive for a too long time.

The domain experts' stimulation plan also organises the reading and playing within TERENCE in sessions as follows, see also Figure 3. In a session, each learner has to:

- (1) read a TERENCE story silently, in class, as part of his/her school activities,
- (2) resolve the related smart games,
- (3) and, finally, be rewarded with relaxing games according to his/her performances with smart games.

Table 2 recaps all such requirements.

The following section delves into how the design of the stories and smart games was carried out using the listed requirements.

Table 2 Requirements for smart games

ID	What for	Description
r1	Difficulty levels	Smart games should be layered into difficulty levels for the TERENCE learners, namely, <ul style="list-style-type: none"> – entry: character games; – intermediate: time games; – top: causality games.
r2	Scheduling of reading and playing	First silent reading of a story, then playing with smart games for the story, finally playing with relaxing games
r3	Learner actions	The learner should be allowed to skip a game, choose a correct solution or choose wrong solutions
r4	Constraints on actions	The learner should be allowed to choose a wrong solution until the correct solution "becomes obvious" but within a fixed resolution time for the smart game
r5	Progress and feedback	The learner should be informed on his/her progresses, on the correctness of his/her resolution and what is wrong or correct in it, and be solicited to give a resolution (within the resolution time)

■4.0 THE DESIGN OF STORIES

Stories in TERENCE were written by authors for children, considering the s2 and s3 requirements in Table 1. Domain experts of psycholinguists for poor comprehenders and linguists simplified the TERENCE stories into three difficulty levels, in line with the above s1 requirement for stories.

The stories were then stored in TERENCE into the following four main levels of difficulty for the TERENCE learners.

- Level 4 is the original story provided by the professional writer, not simplified.
- Level 3 simplifies the global coherence of the original story, making explicit the information necessary to understand the general meaning of the story.
- Level 2, simplified at the global level, is further simplified at the local cohesion level, to increase the logical connection between close sentences.
- Level 1 is the easiest level; the text previously simplified at global and local level, is further simplified in its lexicon and grammar, e.g., by using more common words.

Afterwards, the stories in four levels were illustrated by professional illustrators taking care of the s4 requirement.

■5.0 THE DESIGN OF SMART GAMES

5.1 Introduction

According to [1], a game should specify the following data: the instructions and overall goal of the game, the initial state of the

game, the termination state, the actions allowed for the players, and the maximum allowed time per action (if foreseen). For specifying the data for the TERENCE smart games, we analysed the requirements illustrated above. Then we abstracted and structured the data into the TERENCE game framework. The framework was used to automatise the development of *textual components of smart games*, in brief, textual smart games. These were automatically generated from flat stories, via artificial intelligence technologies, by populating the XML-based framework. See [7].

The TERENCE game framework is based on similar frameworks found in the literature [1]. However, the TERENCE framework is more specific for puzzle-casual games, and is highly structured for enabling the design of the graphical layout of games and their automated development, sketched above. In the remainder, we present the main fields of the framework for the smart games, and their game mechanics.

5.2 Common Fields

The goal of the games is in relation to the TERENCE stimulation plan. That means reasoning about characters and their role in events, temporal relations or causality relations between events, according to the level of the game. See the *r1* requirement in Table 1.

The *instructions* for the game are specialised as follows: (i) questions specific to the game instance; (ii) motivational questions, usually related to the learner's avatar; (iii) questions concerning the rules.

The *central event* of the game is the main event of the story around which the game is designed and upon which the attention of the learner is drawn. The choices available to the learner may change from one state of the game to the other. The game

solutions consist of the choices or their combination that form a correct choice for the game (correct), and those that do not (wrong), in relation to the central event.

The *feedback* for the game is specialised into several types of feedback:

- (1) a consistency feedback (yes, the solution is correct; no, the solution is wrong),
- (2) an interaction feedback for training the user to the game's interaction gestures,
- (3) an explanatory feedback for finding a correct solution (for correct) or for spotting what is wrong in the current solution (for wrong),
- (4) a solution feedback that is a message consisting of the correct solution.

Smart points are the points a learner can gain in a smart game and show his/her progress through the game. These points can be calculated using the IRT [3, 12].

The framework specifies the *states of the avatar* in the gameplay. The states are two: one showing happiness, when the learner chooses the correct solution, and the other showing disappointment, when the learner chooses the wrong solution.

Finally, the *resolution time* of the framework is how long the learner can spend on a specific game. See the related *r4* requirement in Table 2.

Figure 5 shows a before-while smart game instance as it is displayed in the GUI that was developed within the TERENCE project, so to give a concrete feeling of the smart games design.



Figure 5 A before-while smart game instance

5.3 Game Mechanics

At a high level, the plan sets requirements for the actions that the learner can take, the states the system can be in, and constraints on them for all smart games, thereby setting the chore mechanics of the smart games.

Actions. Firstly, the stimulation plan sets that the learner should be allowed to choose no solution, choose a correct solution, or choose a wrong solution. See the *r3* requirement in Table 1. This means that the main actions the learner can take are as follows:

- no solution, that is, the learner chooses no solutions or no exit options;
- wrong, the learner chooses the wrong solution;
- correct, the learner chooses the correct solution;
- skip, the learner chooses an exit option.

Constraints. The stimulation plan sets time constraints on actions, so that the learner is allowed to choose a wrong solution until the correct solution “becomes obvious”. See the *r4* requirement in Table 1. That means, in particular, that the probability of guessing a correct solution for the game sets the maximum number of attempts that learners have at their disposal for choosing wrong solutions in the game. Moreover, there is a resolution time for each smart game that depends on the smart points a learner can gain.

States. The plan also recommends diverse types of feedback if the learner makes a wrong choice and still the learner can play the game: first, a no-consistency feedback for signalling that the solution is wrong, and then an explanatory feedback. Finally, the plan suggests a solution feedback, that is, to display the solution in case the learner chooses no solution within the resolution time, or the number of wrong solutions overcomes the wrong attempts' limit. See the *r5* requirement in Table 1. Given all that, the main states the system can be in are as follows:

- the initial state, in which the learner's score *s* and the resolution time *t* are set to 0, the smart points for the learner are computed as a function of the learner ability in the game, all the choices are set as available, and the number of wrong answers is set to 0;
- a terminal state reachable via a correct action, in which a yes-consistency feedback is given, the score is displayed and the avatar is in the happy status;
- a terminal state reachable via a skip action, in which the solution feedback is given, the null score is displayed and the avatar is in the displeased status;
- a state, reachable via a wrong action, in which a no-consistency feedback is given, an explanatory feedback is given, the set of available choices is updated, and the number of wrong answers is updated;
- a terminal state reachable via a wrong action, in which the no-consistency feedback is given, the solution feedback is given, the null score is displayed and the avatar is in the displeased status.

6.0 EXPERT-BASED EVALUATIONS OF THE LEARNING MATERIAL

6.1 Introduction

Evaluating the TERENCE learning material and, more in general, the TERENCE system with domain experts via UCD analytic evaluations serves to catch and early remove possible usability and pedagogical effectiveness issues that could badly impact evaluations with users.

This section reports two main evaluation studies with domain experts: one, very briefly, for the story texts written into four levels; the other, more extensively, for the textual components of smart games automatically developed via artificial intelligent technologies. We prefer to focus on the latter one as it gives valuable feedback to the automated development process. For details, see [8].

6.2 Evaluation and Revision of Story Texts

The goal of the evaluation with experts of poor comprehends, hearing or deaf, is to assess the original story texts and their rewritings. Let us recall the following: the original story, written by an author, is at level 4; a story at level N-1 is simpler for a specific reading comprehension skill than the same story at level

N, where N is either 4, 3 or 2. The specific goals of this evaluation with experts were then to evaluate whether:

- G1.** the story at level N–1 is easier than that the same story at level N,
- G2.** the story at level N–1 is simpler for the considered reading comprehension skill than the same story at level N,
- G3.** the story at level 1 is comprehensible for deaf readers of the intended age range,
- G4.** the story at level 2 is comprehensible for hearing poor comprehenders of the intended age range.

The choice of using one evaluation method over another is established in relation to the material to evaluate, to the experts involved in the evaluation, as well as to the time constraints and the available resources of the project. In the case of the TERENCE story levels, being the involved evaluators all domain experts, the evaluation methods were expert reviews and heuristic evaluations.

The evaluation results, in relation to the above goals, were as follows.

Results for G1. The stories at level N–1 are easier than those at level N at least in 87% of the cases (worst case).

Results for G2. The stories at level N–1 are simpler for the considered reading skill than those at level N in at least 67% of the cases (worst case).

Results for G3. The stories at level 1 are suitable for deaf learners of the intended age range in 20% of the cases.

Results for G4. The stories at levels 3 and 2 are 100% suitable for hearing poor comprehenders of the intended age range .

The results are thus positive in general. The stories at level 1, however, do not seem to be always suitable for deaf learners, given the heterogeneity of deaf learners. This means that the stories at level 1 require additional simplifications.

6.3 Evaluation and Revision of Textual Smart Games

The goal of the evaluation and revision of textual components of smart games was to control the automated generation and ensure the formal, technical, and contents correctness of all the components of the games for the stimulation plan of TERENCE.

The evaluators were education experts or practitioners, with knowledge of the TERENCE system and of its smart games, but without any ICT skills and not involved in the design of the TERENCE smart games. Given that, evaluators were provided with guidelines and an ad-hoc authoring tool for conducting the manual revision of textual components of smart games. Details can be found in [19].

The revision work was divided into 3 main steps.

- Formal revision, i.e., correction of grammatical and syntactic errors in the text, correction of punctuation, correction of verbs (present tense, active form), correction of referential expressions, check of sentence length and structure.
- Substantial technical revision, i.e., check of the game identifier, correction of the main question texts, correction of the text of solutions, selection of new solutions.
- Construction of causality game instances from scratch because these were not automatically generated, i.e., proposal of textual instances, check-out of proposals, uploading.

The results of the revision were as follows. In the revision of who-games, it sometimes happened that the solutions proposed were not consistent with the automatically generated who-question. For instance, in a WHO-game instance, the generated question was “Who is curious?”. To make the necessary corrections, we had to (i) choose a new character for each solution, and (ii) verify that it was the correct/wrong solution.

We also had to take into account that each choice should be properly understood by all learners (hearing, deaf). Therefore, the changes were done by avoiding difficult referential expressions, paying attention to the spatial distribution in the text and the kind of characters, and to not facilitate the reader in selecting the correct solution.

The evaluation and review of time games was an even more challenging task, because it was necessary to locate the events and their temporal relations across the entire text. Table 3 shows the corrections performed on a before-after time game.

Table 3 Example pre/post revisions

Solution	Pre-revision	Post-revision
AFTER	to thank	The inhabitants of the land of “pi” thank Jasmine
BEFORE	Louis leads the electrician the wires	No change
WRONG	All manage to split the fairly rubbish without difficult calculations needs	All manage to divide garbage in the right way

Each evaluator had the task of filling in a diary in spreadsheet format, made up of 33 fields, specifying the changes made in every revised game. This diary allowed the monitoring of all activities and their analyses.

A total of 250 game instances were reviewed, with respect to 25 stories, with the highest proportion of time before-after game instances (30%). The average review times were estimated based on data reported by the evaluators. They were lower for reviewing who-game instances, and higher for time game instances. On average, for a set of games, i.e., one game for each level, it was necessary to work for circa 76 minutes. By considering the need for reading the related story, filling the excel, the average time of each evaluator to finalise a set of games was equal to circa 90 minutes, i.e., approximately 15 minutes per game instance, as capped in Table 4.

Table 4 Details about the revised games

GAME	n	%	Average time
WHO	25	10.00	10,6
WHAT	34	13.60	12
BEFORE/AFTER	74	29.60	12,8
BEFORE/WHILE	41	16.40	12,8
WHILE/AFTER	42	16.80	12,8
BEFORE/WHILE/AFTER	34	13.60	14,8
Total	250	100.00	12,7

Only in 6% of the cases it was necessary to change the automatically generated central event. In 72% of games the text of the event was corrected. The total number of changes (of both entities or choice events) was 120. The changes were necessary especially for the wrong choices, not sufficiently plausible as distractors, with 54 total changes.

The work of developing from scratch causality game instances was instead longer. Overall, 75 causality game

instances were created. The average time spent for their development was equal to 23 minutes per game instance. The total work, also including game loading and final review, was about 30 minutes for each causality game instance.

7.0 CONCLUSIONS

This paper sketches the design process of the learning material in the TERENCE TEL project, based on the knowledge of experts. The paper shows how the design process of the learning material were conducted moving from the learners' requirements, analysed for the context of use study by domain experts, and how the evaluation studies with domain experts allow for incrementally improving the learning material design. For instance, concerning smart games, the expert-based evaluation highlights the needs of improving the natural language generation of sentences for smart games and of heuristics for distractors more plausible for learners.

The on-going work is mainly devoted to the large-scale evaluation with learners of the TERENCE system. Future work foresees the refinement of the system design and of learning material in light of the results of the large-scale evaluation.

Acknowledgments

The authors' work was supported by the TERENCE project, funded by the EC through the FP7 for RTD, Strategic Objective ICT-2009.4.2, ICT, TEL. The contents of the paper reflects only the authors' view and the EC is not liable for it. Gennari work was also funded through the CRESCO and DARE projects, financed by the Free University of Bozen-Bolzano and the Province of Bozen-Bolzano.

References

- [1] Adams. 2010. *Fundamentals of Game Design*. New Riders.
- [2] Alrifai, M., Di Mascio, T., Gennari, R., Vittorini, P., Melonio, A., and De la Prieta, F. 2012. The Learners' User Classes in the TERENCE Adaptive Learning System. IEEE 12th International conference on Advanced Learning Technologies (ICALT'12). 572–576. DOI: 10.1109/ICALT.2012.68.
- [3] Becker, F., Hamilton, K., Zimring, C., and Shepley, M. (Eds.). 2008. *A Practitioner's Guide to Evidence-Based Design*. The Center for Health Design, Inc. ISBN: 978-0974376387.
- [4] Brusilovsky, P. 2000. Adaptive Hypermedia: From Intelligent Tutoring Systems to Web-Based Education. 5th international Conference on Intelligent Tutoring Systems. Lecture Notes In Computer Science. 1839(1): 1–7.
- [5] Brusilovsky, P., Karagiannidis, C., Sampson, D. 2004. Layered Evaluation of Adaptive Learning Systems. *International Journal of Continuing Engineering Education and Lifelong Learning*. 14(4/5): 402–421.
- [6] Cain, K., Oakhill, J. V., Barnes, M. A., Bryant, P. E. 2001. Comprehension Skill Inference Making Ability and their Relation to Knowledge. *Memory and Cognition*. 29(6): 850–859.
- [7] Cofini, V., De la Prieta, F., Di Mascio, T., Gennari R., Vittorini, P. Design 2012. Smart Games with Context, Generate them with a Click, and Revise them with a GUI. *Advances in Distributed Computing and Artificial Intelligence Journal*. 3(6): 59–68.
- [8] Di Mascio, T., Gennari, R. Vittorini, P. 2012. Expert-based Evaluation: State of the Art and Results. Available at http://www.terenceproject.eu/c/document_library/get_file?p_l_id=16136&folderId=12950&name=DLFE-1946.pdf (accessed on 15 April, 2013).
- [9] Di Mascio, T., Gennari, R. Vittorini, P. 2012. Small-scale Evaluation: State of the Art and Results. Available at http://www.terenceproject.eu/c/document_library/get_file?p_l_id=16136&folderId=12950&name=DLFE-1944.pdf (accessed on 15 April, 2013).
- [10] Duffy, T.M., Cunningham, D.J. 1996. Constructivism: Implications for the Design and Delivery of Instruction. Jonassen, D. H. (Eds.). *Handbook of Research for Educational Communications and Technology*. New York: Simon and Schuster. 170–198.
- [11] Gennari, R. 2011. State of the Art and Design of Novel Intelligent Feedback. Available at http://www.terenceproject.eu/c/document_library/get_file?p_l_id=16136&folderId=12950&name=DLFE-1915.pdf (accessed on 15 April, 2013).
- [12] Gennari, R. and Alrifai, M. 2011. *Game Design Document*. Available at http://www.terenceproject.eu/c/document_library/get_file?p_l_id=16136&folderId=12950&name=DLFE-1908.pdf (accessed on 15 April, 2013).
- [13] Gulliksen, J., Goransson, B., Boivie, I., Blomkvist, S., Persson, J., and Cajanger, A. 2003. *Key Principles for User-centred Systems Design*. 22(6): 397–409.
- [14] Lyon, G., Fletcher, J., Barnes, M. 2003. *Child Psychopathology, chap. Learning Disabilities*. NY: The Guilford Press.
- [15] Marschark, M., Sapere, P., Convertino, C., Mayer, C., W.L., Sarchet, T. 2009. Are Deaf Students' Reading Challenges Really About Reading? *American Annals of the Deaf*. 154(4): 357–370.
- [16] Rizzo, R. 2009. Multimodal and Multimedia aspects of English Language Teaching and Studies in Italian Universities: An E-learning approach. IBIS, Como.
- [17] Slegers, K. and Gennari, R. 2011. State of the Art of Methods for User Analysis and Description of Context of use. Available at http://www.inf.unibz.it/~gennari/projects/terence/pub/D1_1_-_20110302.pdf (accessed on 15 April, 2013).
- [18] Yelland, N., Masters, J. 2007. Rethinking Scaffolding in the Information Age. *Computers and Education*. 48(3): 362–82.
- [19] Cofini, V., Di Mascio, T., Gennari, R., Vittorini, P.. 2013. The TERENCE Smart Games Revision Tool. In: 2nd International Workshop on Evidence-based Technology Enhanced Learning. *Advances in Intelligent Systems and Computing*. Springer. 218(3): 17–24.
- [20] Nanjappa, A., Grant, M.M. 2003. Constructing on Constructivism: the Role of Technology. *Electronic Journal for the Integration of Technology in Education*.
- [21] Ulrich, Roger S., et al. 2003. A Review of the Research Literature on Evidence-based Healthcare Design. *Health Environments Research and Design Journal*. 1(3): 1–13.