This is the pre-print version of: Verpoorten, D., Castaigne, J.-L., Westera, W., & Specht, M. (2012). A quest for meta-learning gains in a physics serious game. *Education and Information Technologies*. Advanced online publication: 10.1007/s10639-012-9219-7

A quest for meta-learning gains in a physics serious game

Dominique Verpoorten, Wim Westera, and Marcus Specht

Abstract
This paper describes how a short, repeated and structured opportunity to evaluate one's learning was integrated in the storyline of a serious game in order to stimulate the development of a meta-cognitive skill: the ability to self-assess the degree of confidence in own answers. An empirical validation of the approach took place with an early version of the game. The results from 28 college pupils deliver an uncommon pattern: while the cognitive benefits – the acquisition of academic knowledge in optics – are mixed up, the meta-cognitive gains present a raising tendency. The experiment also demonstrates that meta-cognition does not necessarily hamper the game flow, if certain conditions, discussed in the paper, are met.

1. Gaming and thinking

Today's educational literature is prone to grant virtues to games for supporting learning (Mitchell & Savill-Smith, 2004). However, questions remain about their potential to train transferable reflective skills (Bopp, 2006; Mac Farlane, Sparrowhawk, & Heald, 2002), which are considered as key leverage points in a lifelong learning society (Claxton, 2006; EuropeanCommission, 2006; Rychen & Salganik, 2003). At first sight the awareness and training of these second-order mental processes seem to entail stop-and-think episodes. If taking a step backwards is the hallmark of reflection, it can sound discordant with or even antagonistic to the immersive characteristics of games[1], at least adventure games. It is therefore not surprising that the few examples of deliberate training of reflection reported in the serious game literature are connected with logics/strategy games, to which introspective pauses are inherent. For instance, Anderson (2002) reports about accounts of sixth grade students playing a game named "Stock Market" designed to help children become familiar with how financial transactions function. One female player says, "This game makes me think how to think". What this statement reveals is that this young learner is beginning to understand the real key to learning; she was engaged in meta-cognition using a game. Saldana (2004) has enriched a "Master Mind" game to assess and exert thinking skills with 3 levels of assistance: support of the meta-cognitive processes internal to each step of task (planning, control, revision), scaffolding of the main steps composing the whole task, modelling of the task solution process.
In contrast to the aforementioned examples, this article depicts an attempt to harness opportunities to reflect to an educational adventure game. It also provides a first empirical evaluation of the effects of this instructional feature on both the understanding of the to-be-learnt concepts – here properties of the light – and the enhancement of a specific reflective

[1] Westera, Nadolski, Hummel, and Wopereis (2008, p. 2) rightly summarize this perceived tension: "Especially in higher education, the mental mode of learning which reflects profundity, reflection, concentration and perseverance seems to conflict with the mental mode of gaming which is commonly associated with amusement, fun and relaxation".

skill: ascertaining the confidence in the quality of one's answers, as explained in the next section.

2. Confidence degrees

In an assessment based only on identifying correct and incorrect answers there is little information available for both teacher and learner other than right or wrong (Leclercq, 1982). Adding confidence degrees to evaluation leads to refined considerations about learning and teaching. For instance what conclusion should teachers raise when 95 % of their learners succeed answering a question? What other conclusion if those students only produce a mean confidence of 10% for their correct answer? Teachers might reconsider their teaching as not completed despite the 95% of success at the test. Corrective behaviours can also benefit from the externalization of confidence. For instance, a wrong answer given along with a confidence degree of 10% is better than the same wrong answer with 90% confidence attached. The students in the latter case convey two erroneous information: one related to their knowledge and one related to themselves (their belief in their answer's rightness). This situation may be considered as dangerous as students will trust what they think they know. These examples suggest that learning does not move someone from total ignorance to perfect knowledge. Often people already have some knowledge or representation about what is taught, even if these representation or knowledge are misleading. So evaluation should not be limited to either knowledge (viz. correct answer) and ignorance (viz. incorrect answer). As De Finetti (1965, p. 109) states: "Partial information exists. To detect it is necessary and feasible (…) It is only subjective probability that can give an objective meaning to every response and scoring method". This is the meaning of Fig. 1 which associates a measure of knowledge (obtained through multiple choice questions, Y axis) with a confidence degree (chosen out of a scale of 6 degrees of certitude, X axis). The output is a "spectral distribution of knowledge" (Hunt, 1993; Jans & Leclercq, 1999). On the left side, the wrong answers are distributed by the confidence degree (from 100% down to 0%) given by the learner. In the middle (grey area) are the unanswered questions. On the right side are the correct answers, also distributed by confidence but ranking from left to right from 0% to 100%. Each rectangle defines a type of relation to knowledge: a) red rectangle: dangerous knowledge (wrong answer/high confidence), b) orange rectangle: unawareness (wrong answer, low confidence), c) blue rectangle: mid knowledge (right answer, low confidence), and d) green rectangle: usable knowledge (right answer, high confidence). Compared to the usual "correct/not correct" feedback, such a view on students' performance allows a refined diagnosis about the relevant kind of remediation (cognitive and/or meta-cognitive)

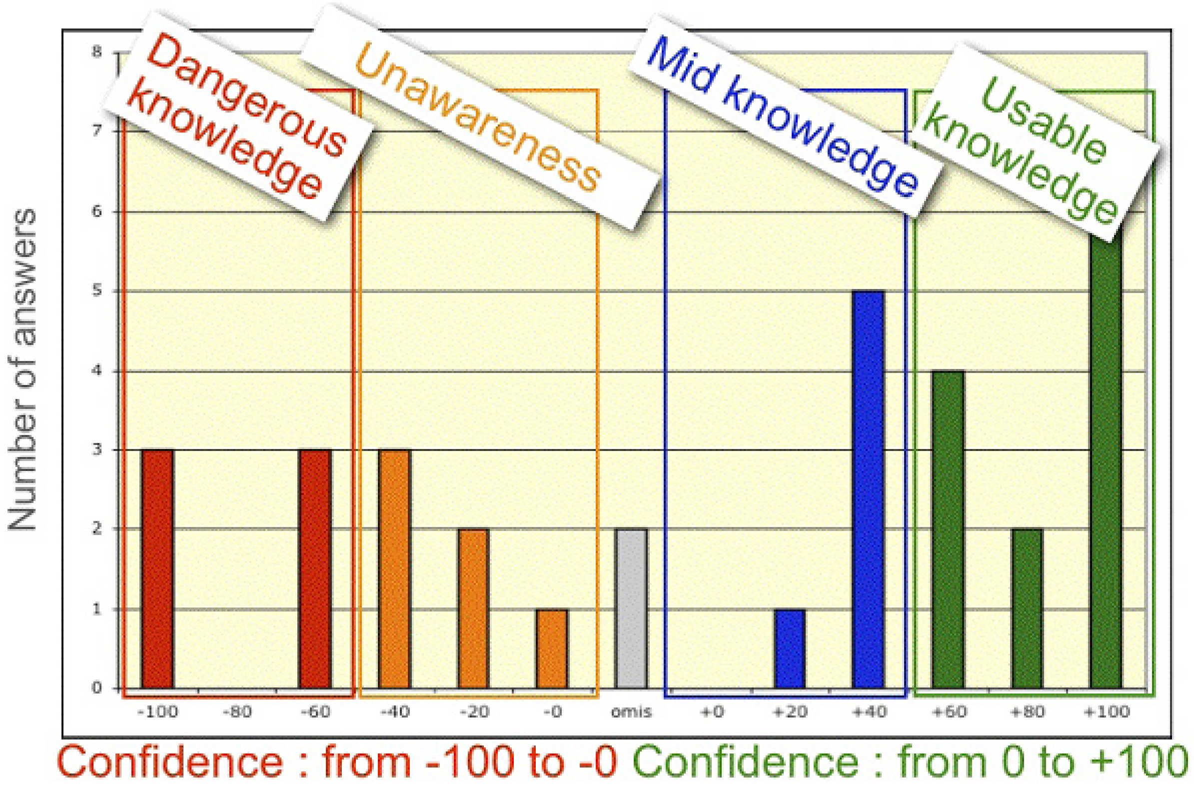

Figure 1: A spectral distribution of knowledge intersperses cognitive (right/wrong answers) and meta-cognitive (confidence degrees) information about the learner

In this study, the confidence ratings embedded in the Elektra game are conceptualized as "reflection amplifiers" (Verpoorten, Westera, & Specht, 2011). This idiom refers to compact, structured and repeated reflection affordances displayed during learning in order to make aspects of it deliberate objects of attention. Reflection amplifiers feature clear-cut reflective operations interlaced with the cognitive processes at work for the completion of a first-order learning task. The underpinning assumption tied to reflection amplifiers is that by continuously interpreting their actions in terms of personal relationship to knowledge (here, the confidence in own answers), learners develop an increased awareness of and an intensified presence to the learning process itself.

3. Research questions

In an exploratory study, 28 college pupils trained cognitive (academic knowledge in optics) and meta-cognitive (confidence degrees) skills by playing a version of the game Elektra. The whole experiment was guided by two research questions: a) how can a reflection amplifiers be reasonably implemented in the concrete of a learning game? b) what will be the effect of such an instructional feature respectively on the game play and on learning? With regard to the research question b, it must be noted that the influence of a confidence degree rating tool was difficult to ascertain beforehand due to possible ambivalent effects. On the one hand, a reflection amplifier represents a reflective pause in the learning process. As such, it can be perceived as a game play breaker. If explicit calls to reflection harm storytelling and immersion, there is a risk to decrease learners' motivation, one of the main lever of learning in games, according to their proponents (Egenfeldt-Nielsen, 2011). With less motivation, players may not exert sufficient effort to engage in learning. On the other hand, the reflection amplifier is designed in such a way that it minimizes the disruption (it represents a rather short episode of reflection) and is integrated in the game flow (gaining confidence in own answers is part of the hero's missions. See section "Storytelling aspect"). So, this instructional feature can also turn to be useful to the support of the first-order learning task (for a similar dilemma with another reflection amplifier, self-explanation, see (Mayer & Johnson, 2010)).

4. Method

4.1 Context

The experiment took place in the context of the European project Elektra. The goal of the project was to develop the demonstrator of a state-of-the-art 3D adventure game teaching physics according to national curricula. The demonstrator was conceived for afternoon market and targeted 13+ students (www.elektra-project.org).

4.2 Sample

Data were collected from 28 pupils from a college in Thiais, France (mean age = 14 years old, male/female = 58/42%). One can ask if meta-cognition, and especially its self-assessment component, is usable as such for teenagers of this age. In brief, major work in the field consider that the components of meta-cognitive monitoring and control do not significantly differ between adults and 10 years old children. Below the age of 10, meta-cognitive processes evolve with age. For instance, Flavell, Friedrichs, and Hoyt (1970) provided evidence of significant correlation between predicted and actual memory span in children from the 4th grade but no significant correlation was found below that age, including at nursery and kindergarden. Schneider (2008) observed unrealistic performances prediction in young children and outlined 3 reasons: 1) insufficient meta-cognitive knowledge: young children do not monitor their memory activities or lack in understanding about the interplay of relevant factors, 2) predominance of wishful thinking over analytical expectations: children's predictions reflect their desires, and 3) belief in the power of effort: the mere fact of spending time on a task induces the prediction of success. Duell's findings (1986) brought further evidence that as children get older they demonstrate more awareness of their thinking processes.

4.3 Type of game

Elektra was designed as a typical first person adventure thriller game wherein a character named George had to rescue Lisa and her uncle Leo, a researcher, who were kidnapped by a villain secret society. Whilst the plot was set the day of the next solar eclipse in Europe in the year 2026, the rescue operation undertaken by George partly immersed him in the world of the Renaissance and its scientific achievements[2]. To save his friends (and incidentally the earth), George had to confront with specific concepts from a eighth grade physics course and to get acquainted with them. Yet, using this knowledge was a condition to move forward in his quest. Learning occurred through various modes of engagement with notions, ranging from hearing or reading to freely experimenting. After finding a magic hour glass, George found himself in company of the ghost of Galileo Galilei (Fig. 2, a) who observed and tutored him while he was busy with the physics experimentations. Elektra developed only a

[2] The trailer of the game is available at: http://player.vimeo.com/video/24224447?title=0&byline=0&portrait=0&color=ff2e90

demonstrator of the game, viz. the opening sequence and the first secret room that George encountered on the track of the evil kidnappers.

4.4 Apparatus

Story-wise, the game element the current empirical study was concerned with was located in the basement of uncle Leo's villa. It presented as a device that allows balls of different materials rolling down a slope (Fig. 2). The goal was to make the marble fall into a hole (Fig. 2, b). To succeed, the learner had to alter the trajectories of the balls by adjusting a magnet (Fig. 2, c) and/or a fan (Fig. 2, d) with sliders (Fig. 2, e, shuffled force between 1 to 100). By contrast – and this was the main knowledge to acquire at this experimentation table -, a laser ray (Fig. 2, f) could not be influenced by such external forces.

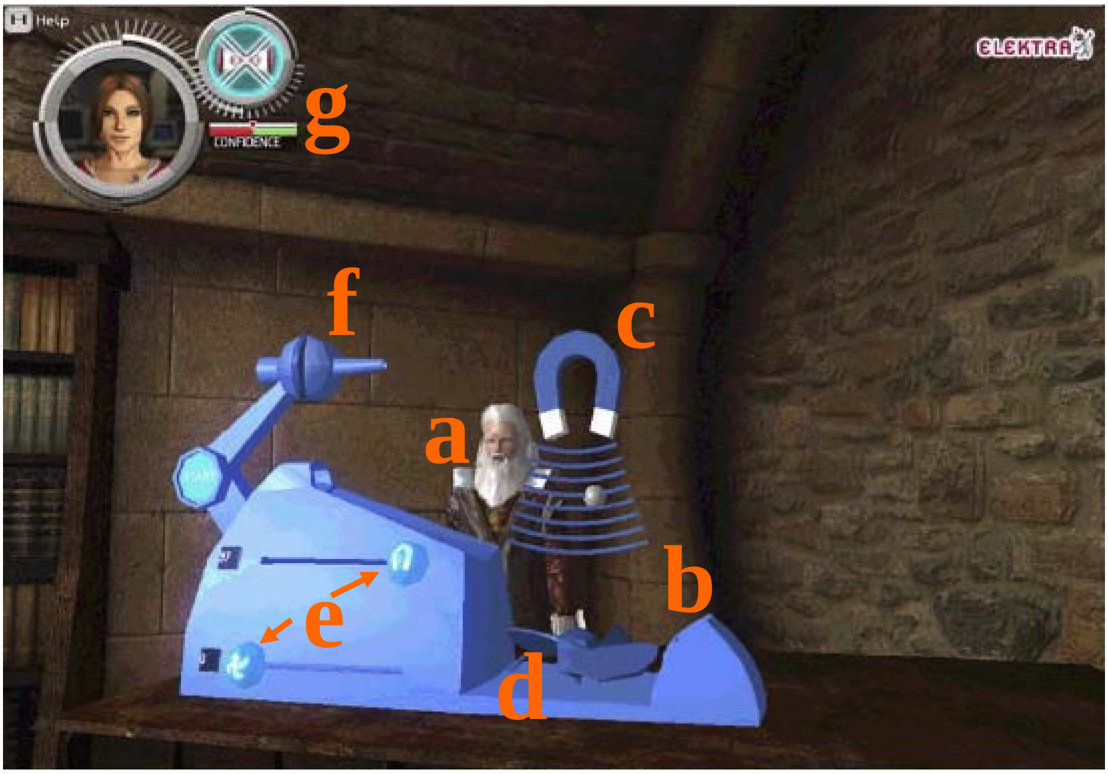

Figure 2: The "slope device" used in the experiment was dedicated to the exploration of light properties

After the magnet/fan arrangements, George had to give, through an additional slider, his degree of confidence that the marble would fall in the hole. Throughout the whole game, the status of his confidence in his actions was mirrored to him via a "smart indicator" (Glahn, Specht, & Koper, 2007), called "certimeter" (for "certitude meter") in the narrative (Fig. 2, g). The certimeter actually displayed the mean confidence degree of the successful trials, as computed real-time by the system. Colours went from red to light red between 0% and 50% mean confidence and from light green to green from 50% to 100% mean confidence (see section "Storytelling aspects" for the connection between the confidence indicator and the plot). The joint setting of the confidence slider and the confidence indicator supported a visual and systematic coordination between the game play and the evolution of the meta-cognitive skill.

4.5 Learning aspects

The learning goals of Elektra drew on the usual distinction between specific skills (confined to a domain, here: optics) and generic skills (domain-independent, transferable, here: confidence ranking) [3].

4.5.1 Cognitive goal

The main learning goal of the experimentation table shown in Fig. 2 was to support the understanding that light propagated in straight lines, as opposed to the curved trajectories of other objects when they were under the influence of forces (wind, magnet, gravity).

4.5.2 Meta-cognitive goal

The setting also pursued a meta-cognitive objective: to develop the awareness of players regarding the confidence that they had in their prevision about the trajectory of the balls and of the light.

4.6 Storytelling aspects

Cognitive and meta-cognitive learning goals harnessed to the slope device underwent a careful integration in the storyline (Moser, 2000). The acquired knowledge about light properties was needed to move further in the adventure: for instance, learners had to use what they had learnt at the experimentation table to unlock a door by exactly hitting a small light sensor with a laser beam. Meta-cognitive gains were also rewarded from a game play perspective: George had to gain the trust of Galileo and this trust evolved on the basis of his good use of confidence degrees. Indeed, George had to succeed at discovering the different influences of the fan and the magnet (and their lack of effect on the light ray) but he had also to reach green level score on the certimeter meaning that Galileo could trust him when he said that he was sure of his knowledge or when he said that he had doubts (a transversal skill and a condition for any scientific work).

4.7 Procedure

Participants filled in the pre-questionnaire, were briefed about the game, confidence degrees and the slope device. They played it 20 min. on average and took the post-questionnaire. During the game, a marble appeared at the top of the slope. The player had to manage the fan and the magnet with the sliders to get the ball into the hole. Learners could make as many attempts as they wished. However, in order to release the marble and observe the effect of the adjustments made, players had to state their degree of confidence that the ball would land in the hole with this configuration of forces. As soon as the confidence degree was provided on the dedicated slider[4], the marble started rolling the slope, then felt through the air under the influence of gravity, and if applicable, under the influence of the magnet and/or airstream, as fixed by the player. Players then saw whether or not they managed to hit the target. They received right after 2 pieces of feedback: one related to the success/failure of the task (e.g

[3] The selection of the pedagogical endeavours of the game came on top of several work packages: European curricula comparisons, breakdown of identified skills in various granularity levels, distribution of the retained skills in the entire game play, multiple checks of the consistency between learning objectives, methods and evaluation. For details on this intensive preparatory work, see Petit, Castaigne, and Verpoorten (2007).

[4] A literature review and an empirical pre-study (23 participants) was carried out in order to identify the most relevant metrics to express confidence (for an in-depth discussion of this topic, see Castaigne, 2007). Based on the different arguments, it was decided that the confidence slider would be graduated with the values 0, 20, 40, 60, 80, and 100% confidence.

"Well done. You noticed that the magnet has no influence on the aluminium ball and you controlled well the power of the fan") and one related to the confidence evaluation (e.g. "You did well with this marble but you indicated a confidence degree of 20%. You should trust yourself more"). Both feedbacks were given as textual monitoring pop-up. The "certimeter" was updated after each feedback. Players exercised with 5 marbles made successively of iron, plastic, wood, aluminium, granite.

4.8 Measure instruments

Cognitive and meta-cognitive performance were measured with a test inserted in the game and with a pre/post-test comparison.

4.8.1 Intermediate test (within the game)

The test came after the players succeeded in throwing the third marble (wood) in the hole. It was designed as a formative test that brought a contrast to the reflection triggered by the manipulations around the previous marbles. The test presented as 3 visual multiple choice questions (Figure 3) probing successively the presumed effect of the fan, the magnet and the combination of both on the laser ray. Students gave their answer by clicking on the visual representing in their view the correct trajectory. For each answer, they were asked as usual to indicate their confidence degree. After the test, players went on with the final two marbles in an identical manner to the previous marbles.

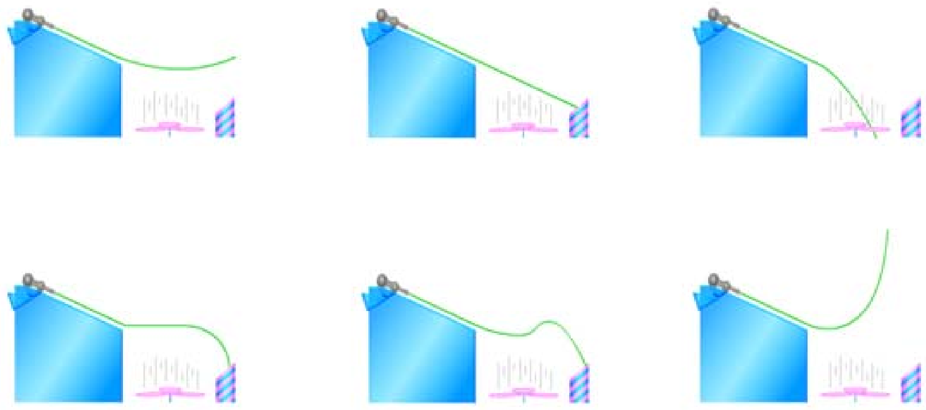

Figure 3: To perform the intermediate test, embedded in the game, players had to click on the visual giving the correct representation of how light would propagates in a given situation (here with the fan activated)

4.8.2 Pre/post-test

Before and after the gaming session, pupils answered with paper and pencil to questions about the influence of wind and magnet on a stone marble, an aluminium marble and light (Table 2).

5. Results

5.1 Whole game

Students performed an average of 4 trials with each ball before achieving success. The overall mean confidence degree (all attempts, all pupils, N=28) was 54%.

5.2 Intermediate test

The relationship between right answers and mean confidence degrees were summarized in Table 1 for the intermediate test (3 multiple choice questions with visuals as answers, cfr. Fig. 3). Results showed a steady progression in the confidence that students had in the rightness of their answers. While the average confidence associated to the trials with the marbles (54%, see above) and to the first multiple-choice question (55%) were still in the mid-knowledge rectangle (cfr. Fig. 1), it toppled over the usable knowledge rectangle for the two last questions.

Table 1 – Results of the intermediate test

	N	**MCQ#1**	**MCQ#2**	**MCQ#3**
Correct answers	28	19	14	22
Mean degree of confidence	28	55%	64,3%	83,6%

5.3 Pre/post-test

Results at the pre and post-test were wrapped up in a single table to sharpen comparison (Table 2). Arrows indicated whether knowledge improvement or deterioration occurred (in terms of percentages of correct and incorrect answers.

Table 2 – Results of the pre/post questionnaires (achievement)

	Stone marble		Light		Aluminium marble	
	PRE	POST	PRE	POST	PRE	POST
Wind	C*: 45% I**:55% ↗	C: 77% I: 23%	C: 100.0% I: 0.0% ↘	C: 93% I: 7%	C: 47% I: 53% ↗	C: 68% I: 32%
Magnet	C: 97% I: 3% ↘	C: 70.0% I: 30.0%	C: 51.6% I: 48.4% ↗	C: 71.4% I: 28.6%	C: 100% I: 0% ↘	C: 75% I: 25%

** C = Correct answers ** I = Incorrect answers*

Pupils enhanced their performance when answering if wind influenced the trajectory of a stone marble (45.2% became76.7%), when answering if magnetism influenced trajectory of light (51.6% became 71.4%), when answering if wind influenced trajectory of aluminium’s object (46.7% became 67.9%). On the opposite, there was a decrease of performance when asking if magnetism influenced trajectory of a stone (96.8% became 70.0%) or aluminium ball (100% became 75%), if wind influenced the trajectory of light (100% became 93%).

6. Discussion and further work

Three main findings emerge from the current research. Each contributes to a specific research field.

6.1 Contribution to research on confidence degrees

Previous work in the field of confidence ranking have generally noted that learners tend to overestimate the quality of their answers, especially in areas where their skills and knowledge bases are weak (Kruger & Dunning, 1999). In other words, it has been regularly observed that students don't know enough to recognize that they lack sufficient knowledge for accurate self-assessment. The pattern observed in this experiment does not show extremely high confidence degrees but a progression towards higher level when good answers are given. It advocates for a growing consciousness of the connection between rightness and certainty. After all, it would have been possible that pupils only focus on reaching the target with the marble, neglecting the reflection on their actions and disregarding or using superficially the confidence slider and the uncovering of their tacit confidence it invites to. The fair level of engagement with their meta-cognition can be imputed to convergent factors. During the general introduction to the experiment, students were shortly explained why gauging their confidence matter. This might have been an important contributor to the quality of confidence rankings. Yet, it is generally acknowledged that the rationale given for the usefulness of meta-cognitive interventions is a success factor thereof (Bannert, 2006). From an instructional game design viewpoint, it is plausible that the encapsulation of confidence degree, both in local challenges and at the global storytelling level, contributed to them being taken seriously. Lastly, the brevity of the reflective enactments (following a salient feature of reflection amplifiers) is another aspect that probably played a positive role, challenging the idea that a reflection is necessarily a long-lasting operation. Further investigation is needed to disentangle the respective influence of these factors.

6.2 Contribution to the integration of reflection in games

In her effort to understand the interplay and relationships between different kinds of learning environments and methods, Laurillard (1993) emphasizes that standard classroom, lecture and exercise techniques can lack in context, interactivity, and the ability to experiment freely. Conversely, games offer these features but have their own shortcoming in that they might be weak at providing students with opportunities to initiate reflection and to describe their conceptual knowledge[5]. Harteveld, Guimaraes, Mayer, and Bidarra (2007, p. 132) note in the same vein: "Games offer almost no opportunity for reflection as players are completely immersed into the game. Reflection is important to go from specific spontaneous concepts toward abstract scientific concepts. Reflection can be stimulated by an instructor, but it could be a valuable addition if it was somehow included into the game". Elektra somehow

[5] Each method has its shortcomings, but the aggregate combines their strengths to cover the individual weaknesses. Variety of learning methods and approaches is not only a matter of motivation enhancement but also of epistemology (Verpoorten, Poumay, & Leclercq, 2007). It is probably for this reason that learning games should more often be contextualized within a larger learning sequence and not conceived as stand-alone vectors of learning, as recommended by Quinn (2005, p. 14), "I do not believe that these engaging learning experiences of games will (or should be expected to), by themselves, lead to learning. I advocate discussion around the experience, and connecting learner actions to the underlying concept. As yet, computers are not quite capable of supporting such dialogue. Self-directed learners may be capable of facilitating their own reflection, but it's not the way to bet (though I believe strongly that meta-learning, or learning to learn, is a key leverage point for the future). So although such gaming environments are not sufficient, they are necessary; we need engaging experiences to motivate learners to attend to the content, give them rich practice opportunities, and provide fodder for discussion and refinement of their understanding". The same emphasis is put by De Freitas (2006, p. 11). By curiosity, players of Elektra were asked whether they would prefer gaming before or after a lecture on light properties. Results gave an even proportion of "before" and "after" (Verpoorten, Glahn, Chatti, Westera, & Specht, 2011)

challenges views on game that consider this medium as inappropriate for reflective pauses. When thinking episodes are carefully crafted, when they are kept short and active and when they make sense for the next steps of the game play, it seems that they can bring an added value without destroying the "flow of optimal experience" (Csikszentmihalyi, 1990). How to strike the right balance between action and thought remains however a complex question. On the action side, Kiili (2004, p. 16) states that "Ambitions to design engaging educational games have probably often failed because educational aspects have displaced game play". But conversely, on the teacher's side, it is legitimate to raise the question of the extent to which the storytelling should take the precedence over the examination of the task at hand and the conscious internalization of conditions of success, possibly at the expense of learning and met-learning. Effective trade-offs is a research topic that deserves additional inquiry (Kim, Park, & Baek, 2009).

6.3 Contribution to an extended definition of learning performance

What is the learning performance in Elektra's gaming sequence? If the traditional definition in terms of mark at the test is favoured, the answer is blurred results, at best (cfr Table 2). The game has sometimes clarified and sometimes confused students' ideas about the influence of different forces. The conclusion is different with regard to the reflection-related instructional purpose: training one specific kind of intellectual habit, the ability to assess one's certitude/doubt about knowledge. On this meta-cognitive level, the study provides indications that the game produced effect (cfr. Table 1). Indeed, the progression of the mean confidence degree cannot be attributed with 100% confidence (yes, authors also…) to a growing intellectual mastery of the relationship between providing an answer and assessing own confidence. Testing effect or local influence of questions are other candidates. More research work will help to sharpen causalities.
Another natural extension of the current study could bear on the notion of "prudence", as opposed to confidence. Confidence concerns only good answers and the evolution in the rectangles "mid knowledge" and "usable knowledge". When such a progression occurs, students win in both landscapes: cognitive and meta-cognitive. A symmetric empirical study would be worth conducting on the failed answers and the confidence degrees attached to them (usually referred in the literature as "prudence"). Evidence of gains in prudence (for instance students leaving the "dangerous knowledge" area to enter in "unawareness" realm) **would somehow paradoxically reveal progress being made despite choosing a wrong answer!** It would revamp, at the age of learning games, the invitation of Piaget (1978) to distinguish between success and understanding, between progress visible at the test (in the case of prudence, students at the test are nihil) and intangible benefits (getting the grip on an essential intellectual skill: being conscious of own ignorance) which cannot be traced by the traditional modes of assessment and are not reflected into regular learning achievement measures.

7. Conclusion

This experiment, which for the first time makes use of confidence ranking as reflection amplifiers in an adventure game, point at a potential for this type of game to develop not only the ability to perform (cognitive goal), but also the understanding of conditions of success (meta-cognitive goal). In this way, the game and the confidence elicitation technique it contains has been put in the service of "intentional learning", that is mental "processes that have learning as a goal rather than an incidental outcome." (Bereiter & Scardamalia, 1989, p. 363).

8. Limitations

The sample of this study remains limited, as well as the extent of the evaluation of confidence after a rather short training. It must also be noted that, for scientific purpose, this experiment did use a trimmed version of the Elektra demonstrator. It is difficult to say whether the observed cognitive and meta-cognitive trends would be confirmed with students playing with the full-fledged game, including all game features planned but not implemented by the time of the study), that is: 3 different experimentation tables, cognitive and meta-cognitive feedback given by Galileo and not as in the present study in a textual form, an additional feedback about the confidence trend (over several answers), active support by Lisa (the young girl in the left corner of Fig. 2). Lastly, due to limited tracking features and available analysis resources, the study had to limit itself to between-subjects measures and to the measures obtained in the tests and not in the training phase.

9. References

Anderson, N. J. (2002). *The Role of Metacognition in Second Language Teaching and Learning*. ERIC Digest: ERIC Clearinghouse on Languages and Linguistics, 4646 40th Street N.W., Washington, DC 20016-1859. Tel: 202-362-0700. http://www.cal.org/ericcll/DIGEST.

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 4*, 359-375.

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 361-392). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bopp, M. (2006). Didactic Analysis of Digital Games and Game-Based Learning. In M. Pivec (Ed.), *Affective and Emotional Aspects of Human-Computer Interaction. Game-Based and Innovative Learning Approaches* (pp. 8-37). Amsterdam: IOS Press.

Castaigne, J.-L. (2007). *Game based learning and metacognition: working with quantitative metacognitive feed-back based on subjective probability without loosing the game play*. Paper presented at the 12th biennial Conference EARLI: Developing Potentials for Learning. http://earli2007.hu/nq/home/scientific_program/programme/proposal_view/&abstractid=1757.

Claxton, G. (2006). Expanding the Capacity to Learn: A new end for education?. Keynote speech given at British Educational Research Association Annual Conference, University of Warwick, 6-9 September 2005, .

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

De Finetti, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item. *British Journal of Mathematical and Statistical Psychology, 18*, 87-123.

de Freitas, S. (2006). *Learning in Immersive Worlds: A Review of Game-based Learning*. Bristol, England: JISC. http://www.jisc.ac.uk/media/documents/programmes elearninginnovation/gamingreport_v3.pdf

Duell, O. K. (1986). Metacognitive skills. In G. Phye & T. Andre (Eds.), *Cognitive Classroom Learning: Understanding, thinking and problem solving* (pp. 205-239). Orlando, FL: Academic Press.

Egenfeldt-Nielsen, S. (2011). What Makes a Good Learning Game? - Going beyond edutainment. http://elearnmag.acm.org/archive.cfm?aid=1943210

European Commission. (2006). *Proposal for a recommendation of the European Parliament and of the Council on key competences for lifelong learning. COM(2005)548 final*. Brussels.
Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1*(4), 324-340.
Glahn, C., Specht, M., & Koper, R. (2007). Smart indicators on learning interactions. In E. Duval, R. Klamma & M. Wolpers (Eds.), *EC-TEL 2007 LNCS 4753* (pp. 56-70). Berlin, Heidelberg: Springer.
Harteveld, C., Guimarães, R., Mayer, I., & Bidarra, R. (2007). *Balancing pedagogy, game and realitycomponents within a unique serious game fortraining levee inspection.* Paper presented at the Second International Conference, Edutainment 2007.
Hunt, D. (1993). Human self-assessment : Theory and application to learning and testing. In D. Leclercq & J. Bruno (Eds.), *Item Banking : Interactive Testing and Self-Assessment* (Vol. F 112, pp. 177-189). Berlin: Springer Verlag.
Jans, V., & Leclercq, D. (1999). Mesurer l'effet de l’apprentissage a l’aide de l’analyse spectrale des performances. In C. Depover & B. Noel (Eds.), *Evaluation des compétences et des processus cognitifs* (pp. 303-317). Bruxelles: De Boeck.
Kiili, K. (2004). Digital game-based learning: Towards an experiential gaming model. *Internet and Higher Education, 8(1)*, 13-24.
Kim, B., Park, H., & Baek, Y. (2009). Not just fun, but serious strategies: Using meta-cognitive strategies in game-based learning. *Computers & Education, 52*(4), 800-810.
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.
Laurillard, D. (1993). Rethinking university teaching: a framework for the effective useof educational technology. London: Routledge.
Leclercq, D. (1982). Confidence Marking: its use in testing. In Postlethwaite & Choppin (Eds.), *Evaluation in Education* (Vol. 6, pp. 161-287). Oxford: Pergamon.
Mac Farlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games. An exploration by TEEM of the contribution which games can make to the education process*. Cambridge: Department for Education and Skills - TEEM project.
Mayer, R. E., & Johnson, C. I. (2010). Adding Instructional Features that Promote Learning in a Game-Like Environment. *Journal of Educational Computing Research, 42*(3), 241-265.
Mitchell, A., & Savill-Smith, C. (2004). *The use of computer and video games for learning. A review of literature*. UK: Learning and skills development agency.
Moser, R. (2000). *A methodology for the design of educational computer adventure games.* http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20021003.141152.
Petit, L., Castaigne, J.-L., & Verpoorten, D. (2007). *Checking the internal pedagogical consistency of a game learning situation: the Leclercq’s triple consistency triangle.* Paper presented at the Evete Conference 2007, Kaunas, Lithuania.
Piaget, J. (1978). *Success and Understanding*. Cambridge, Massachusettes. : Harvard University Press.
Quinn, C. (2005). *Engaging learning - Designing e-Learning simulation games*. San Francisco, CA: Pfeiffer.
Rychen, D. S., & Salganik, L. H. (2003). *Key Competencies for a Successful Life and a Well-Functioning Society*: Hogrefe Publishing.
Saldaña, D. (2004). Dynamic Master Mind. *School Psychology International, 25*(4), 422-438.

Schneider, W. (2008). The Development of Metacognitive Knowledge in Children and Adolescents: Major Trends and Implications for Education. *Mind, Brain, and Education, 2*(3), 114-121.

Verpoorten, D., Glahn, C., Chatti, A., Westera, W., & Specht, M. (2011). Self-Reported Learning Effects of a Tagging Activity Carried out in a Personal Learning Environment (PLE) by Secondary-School Pupils. *International Journal for Cross-Disciplinary Subjects in Education, 2(1)*, 276-284.

Verpoorten, D., Poumay, M., & Leclercq, D. (2007). The eight learning events model: A pedagogic conceptual tool supporting diversification of learning methods. *Interactive Learning Environments, 15*(2), 151-160.

Verpoorten, D., Westera, W., & Specht, M. (2011). Reflection Amplifiers in Online Courses: A Classification Framework. *Journal of Interactive Learning Research, 22*(2), 167-190.

Westera, W., Nadolski, R. J., Hummel, H. G. K., & Wopereis, I. G. J. H. (2008). Serious games for higher education: a framework for reducing design complexity. *Journal of Computer Assisted Learning, 24*(5), 420-432.