Moody Agents: Affect and Discourse during Learning in a Serious Game

Carol M. Forsyth[1,2], Arthur Graesser[2], Andrew M. Olney[2,] Keith Millis[3],

Breya Walker[2], and Zhiqiang Cai[2]

[1]Educational Testing Services
Rosedale Rd. Princeton NJ 08648
[2]The University of Memphis
Institute for Intelligent Systems
Innovation Drive, Memphis TN 38152
[3]Northern Illinois University
Psychology Building, Dekalb, IL 60115

Abstract. The current study investigated teacher emotions, student emotions, and discourse features in relation to learning in a serious game. The experiment consisted of 48 subjects participating in a 4-condition within-subjects counter-balanced pretest-interaction-posttest design. Participants interacted with a serious game teaching research methodology with natural language conversations between the human student and two artificial pedagogical agents. The discourse of the artificial pedagogical agents was manipulated to evoke student affective states. Student emotion was measured via affect grids and discourse features were measured with computational linguistics techniques. Results indicated that learner's arousal levels impacted learning and that language use is correlated with learning.

Keywords: discourse, serious game, emotion, Intelligent Tutoring Systems, learning

1 Introduction

The goal of the current investigation is to discover both affective and cognitive- discourse components influencing learner's experiences in a serious game known as *Operation ARIES!* (ARA, in the commercial web version; [1]). The game teaches students critical thinking in the domain of research methodology using natural language conversations between a human student and two or more artificial agents. In previous research, the system has been used to investigate both discourse and affect during learning [2,3] as it provides an atmosphere of learning tasks that are easily manipulated and controlled by researchers. The current investigation was designed to measure the impact of teacher (an artificial pedagogical agent) mood, or more generalized emotion, on student learning, human student emotions on learning, and the relationship between student discourse features with learning during interaction with this serious game.

adfa, p. 1, 2011.

1.1 Emotions and Learning

Recently, emotions have been studied on a moment-to-moment basis during the learning process with Intelligent Tutoring Systems, including but not limited to *Operation-ARIES!* [3,4]. Many previous studies focus on transient and discrete learning-centered emotions (e.g. confusion or surprise) whereas the current investigation focused on a broader and longer-lasting category of emotions, categorized by two separate dimensions of valence and arousal [5]. Valence describes the polarity of emotion, whereas arousal defines the intensity. Accordingly, learning-centered emotions can be categorized as having either a negative valence (i.e., frustration and boredom), a positive valence (flow/engagement and delight), or somewhere in-between, depending on context (e.g., confusion and surprise).

Affective states of teachers may influence student emotions, and therefore learning gains [6-8]. Some research suggests enthusiasm and other positive emotions increase teachers' effectiveness by increasing the efficiency of the instruction [7-8]. However, Sullins and colleagues [6] found teachers, in this case artificial agents, displaying negative affect, to be most conducive to learning. Thus, contradictory empirical evidence exists as to whether it is positively- or negatively- valenced teacher emotions that best facilitate learning [6-7].

Regardless of the teacher, recent studies suggest student emotions play a role in learning [3-4,7]. Students in a positive affective state may have greater cognitive flexibility, global processing capacities, creativity, and attention. Conversely, students in a negatively- valenced affective state have more analytical and focused processing (for a review,see [7]). These findings are supported by *the broaden and build theory* [9] that draws on a Darwinian evolutionary perspective. The broaden and build theory is based on the idea that negative emotions tend to require immediate reactions, whereas positive emotions allow one to view the surroundings (broaden) and accrue more resources in order to increase the probability of survival (build). Another possibility is that valence itself does not affect the functioning of cognitive processes. The *arousal theory* suggests that only the intensity (not the valence) of an emotional trigger is important [10]. However these findings may only be true for simple tasks [11], but not more intricate tasks such as creative problem-solving [7]. Simple vs. complex tasks may evoke shallow vs. deep learning, a characteristic that can be captured by discourse features.

1.2 Discourse Comprehension

Psychological theories of comprehension have identified multiple representations and strategies corresponding to multiple levels of discourse comprehension ([14] for a review). In the current investigation, we use a taxonomy developed by Graesser and colleagues [15, 16] including five levels of discourse processes ranging from shallow to deep level comprehension. These levels can be captured in a computational linguistic tool known as Coh-Metrix [16,17]. The 5 levels measure comprehension on a continuum of shallow to deep comprehension include *word concreteness* (words), *syntactic simplicity* (syntax), *referential cohesion* (textbase), *deep cohesion* (situation model), and *narrativity* (genre). These 5 measures are derived from a principal component analysis of over 100 discourse features measured in Coh-Metrix [18]. In addition to these 5 measures corresponding to the framework, a composite measure of formality may potentially be indicative of cognitive strategies and comprehension. All scores are represented on a scale from 0 to 1 indicating presence of the corresponding features within the text. For example, a score of 0 for word concreteness would imply that many abstract words such as “imagine” were used whereas a score of 1 means that many concrete words (e.g. “participant”) are present in the text. In the current study, the discourse features corresponding to these levels of comprehension, along with both the teacher and student emotions, are investigated in relation to learning.

2 Experiment

The purpose of the current experiment is to investigate the relationships between teacher emotions, student emotions, and discourse features with learning in the environment of a serious game. Specifically, we were testing the hypothesis that teacher emotion (positive or negative) affect student learning gains. The second hypothesis was that student emotions impact learning gains. Finally, we predicted that student discourse features associated with shallow (e.g. repeating concrete words) vs. deep comprehension (e.g. conceptual understanding) may correlate with learning and provide insight to findings from the analyses on emotions.

2.1 Design

The emotions-oriented hypotheses were tested using a 4 condition, within-subjects, counter-balanced pretest-interaction-posttest design. Participants interacted with *OperationARIES!* by reading an e-book, answering multiple-choice questions, and conversing in natural language dialog about four topics on scientific methodology with two artificial agents. Three of the conditions were created by manipulating the affective language (i.e. positive, negative, or neutral) used by the pedagogical agents during the natural language conversations, and a fourth served as a control where learners only read the e-book. The assignment of condition (i.e. positive, negative, neutral or control) was counter-balanced across participants. However, the order of the topics remained constant across all 4 conditions. Thus, each participant received each mood condition for one of the topics (counterbalanced across participants and topics) and one chapter served as a control (i.e., text only). Students were asked to report affect 5 times throughout the experiment to gauge the students’ affective states. The initial report was obtained before interaction to gain a baseline, followed by sequential reports after each of the 4 topics within the system. Learning gains were assessed using a pretest and posttest consisting of multiple-choice and open-ended questions. Finally, discourse features of the student input during interaction with *OperationARIES!* were analyzed using the computational linguistic tool, Coh-Metrix.

2.2 Participants

The study included 48 (N=48) undergraduate and graduate students at a state university in Tennessee. Out of the total number of participants, 4 were graduate students and the other 44 were undergraduate students. Subjects were recruited using the university subject pool as well as through flyers and word of mouth. Two types of compensation were offered: A monetary reward ($15 for completion of the entire experiment) or course credit (2 hours course credit towards an Introduction to Psychology course). Only 13% of the students requested the monetary compensation ($15), whereas the others preferred 2 hours of credit towards an Introduction to Psychology course.

2.3 Materials

Agent-Human Interaction. *Operation ARIES!* [1] is an adaptive Intelligent Tutoring System with game-like features that teaches students the scientific method through natural language conversations. Throughout the game, a narrative about aliens invading the world and propagating bad science is delivered via e-mails, conversation, and videos. There are three interactive modules of ARIES, each focusing on a different type of knowledge acquisition (i.e. didactic, applied, applied question generation). However, the current experiment only focused on the first module, the Training module, which teaches the basic declarative knowledge about the definition, importance, and an example of each of 21 topics of research methodology. In the current experiment, students only interacted with the system across 4 topics of research methodology.

Students completed interaction with this module in 4 distinct phases. For each chapter, students first read an e-book (Phase 1), then answered 6 multiple-choice (MC) questions (Phase 2), and then conversed with the two pedagogical agents in a trialog, or three-way conversation (Phase 3) immediately after each of the final three MC items. Although the original game adaptively places students into one of three pedagogical modes (vicarious learning, standard tutoring, teachable agent) depending on their prior-knowledge, the game was altered in the current experiment to force all students regardless of prior-knowledge level into the intermediate mode of standard tutoring for the sake of consistency. Therefore we will only address the standard tutoring mode in relation to the current study. This mode of standard tutoring consisted of the pedagogical agents asking the student a specific question about the current topic and scaffolding the students to help them articulate a pre-determined ideal answer referred to as an expectation.

In the standard tutoring mode, the pedagogical conversation with the artificial agents was launched after the MC questions by asking an initial question to students. If the student gave an answer to the initial question that was very close to the expectation, then the tutor gave positive feedback and moved forward in the conversation. However, if the student's answer was vague or incomplete, the tutor launched into a scaffolding dialogue consisting of pumps ("Tell me more"), feedback, hints, prompts, and correcting misconceptions if necessary. When all else failed, *OperationARIES!* simply told the student the right answer with an assertion.

Every question asked by the pedagogical agents had a corresponding expectation or ideal answer (the two terms are used interchangeably within the game). For example, a question requiring such an answer could be, "Why are operational definitions important?" The expectation is "Operational definitions are important because they allow a particular variable to be reliably recognized, measured and understood by all researchers." The human students' input is compared to the expectation using a combination of Latent Semantic Analysis [19], regular expressions [20], and word overlap metrics that adjusts for word frequency. This method of computing semantic matches in *OperationARIES!* is not significantly different from semantic similarity judgments between two humans [21].

For the purposes of this experiment, the affective display of the generalized mood of the agents was altered by changing the curriculum scripts, which is the predetermined speech of the two pedagogical agents. Both the teacher agent and student agent displayed the same valence of emotional content (i.e. positive, negative, or neutral) at the same time. So, during a "positive" chapter within the learning session, both the teacher and student agents exhibited a positively-valenced affective state. During a "negative" chapter, both displayed a negatively-valenced affective state. For each chapter, the emotional display or mood of the agents remained either positive, negative, or neutral for both agents throughout an entire chapter.

The affectively-valenced discourse of the artificial agents were changed using the Linguistic Inquiry and Word Count lexicon (LIWC; [12]) which has been used to measure persistent affective states[13]. This lexicon has numerous words associated with positive and negative affective states. For example, words with positive- valence include "happy", "curious", and "awesome" whereas negatively-valenced words include "bored" and "sad." Therefore, with *OperationARIES!* an example statement in the negative condition would be, "No. You are incorrect. Let's just go over the importance of these dull things one more time. Why do we need to have operational definitions?" As the reader may notice, the manipulation did not include a change in feedback as the goal of the study was not to provide false feedback but rather to manipulate the affective expression of the teacher agents.

The altered agent speech covered three chapters of material with three separate chapters each designated to one mood condition (i.e., positive, negative, or neutral). The participants were exposed to a fourth chapter of content through the E-book only. This chapter served as a control for the counter-balanced conversations with the artificial agents. In order to return the student to a baseline of emotion between the within-subjects conditions, each chapter began with the student answering 6 multiple-choice questions about the topic and reading a summarized chapter of the E-book within *OperationARIES!*.

Assessment of Learning Gains. Two similar, but not identical, versions of a learning gains assessment were developed using multiple-choice and short answer questions. The learning gains assessments were counter-balanced, so both versions of the assessment were used as a pretest or a posttest that assessed learning gains. Each test consisted of a total of 32 questions, 8 questions per topic. Both the pretest and the posttest were manually graded in accordance with the associated rubric. Inter-rater reliability was established on a by-item basis for the short-answer questions (r >.70). Then proportional learning gains (PLG) were calculated using the formula [(posttest-prettest)/(1-pretest)], [22]. This value was used to measure initial learning gains from interaction with the game. However, in conducting linear mixed effects models, the posttest was the dependent variable and the pretest was treated as a covariate. This was a necessary compromise as proportional learning gains often produce highly negative values when calculated on the item level.

2.4 Procedure

After completing an informed consent, all participants were randomly assigned to different materials that varied: 2 tests (pre/post) x 4 pedagogical agent moods (e.g., positive, negative, neutral, neutral text only) in a counterbalancing scheme. The assignment of the two tests (version A and version B) was counter-balanced as pretest and posttest. Thus, within the four possible mood conditions (e.g., positive, negative, neutral, neutral text only) and the counter-balanced learning gains assessments, there were 48 cells in this within-subjects design. After being randomly assigned to a specific group, each subject was first given a pretest followed by instructions on how to complete the affect grid which is a self-report measure of valence and arousal on a 2 dimensional grid. The researcher then gained a baseline affective state for each participant by asking each individual participant, "How do you feel participating in this experiment?" Students verbally replied and were asked to denote this emotion on the first affect grid. After completing the baseline affect grid, each participant interacted with *Operation ARIES!* in the respective assigned condition. During interaction with *Operation ARIES!*, students read an e-book on each chapter as well as answered multiple-choice questions before engaging in a natural language dialogue with the two pedagogical agents. In accordance with the given condition, the participants conversed with the agents displaying the information in a specific mood, i.e. positive, negative, or neutral. In the control condition, participants only read the e-book and answered multiple-choice questions but did not participate in a tutorial conversation with the agents. After completion of each topic, participants were asked to fill out the subsequent affect grid. This process occurred iteratively across the chapters. Upon completion of the interaction with *OperationARIES!*, the participants completed the posttest and were debriefed.

3 Analyses

Before testing our hypotheses, we first conducted a paired samples t-test to confirm that learning actually occurred. Students' proportional learning gains were significantly higher for the topics covered in conversational conditions when compared to those covered in the control (t (1,47)=1.85, p <.05, *Cohen's d* = .55). Next, a series of linear mixed models were conducted to assess the three main hypotheses. The mixed model approach allows for calculation of differences based on observations not participants, thus leveraging the within-subjects power of the design. Recall, the first hypothesis was that the affective state of the pedagogical agents would affect learning. The second hypothesis was that the human students' affective state would affect learning. The third hypothesis was that the student's discourse features would correlate with learning gains. For all three hypotheses, we wanted to generalize across individual differences between participants, topics taught in *OperationARIES!,* and the counter-balanced order of the tests. Thus, these three variables were treated as random factors in all of the following analyses.

A mixed effect model was conducted to determine whether differences in learning existed between the three conversational moods and the control. The model included four conditions and pretest as fixed effects, with participant, topic, and test version held as random factors was used to evaluate posttest scores. The model was not significantly different from the null model including pretest as a fixed factor and participant, topic, and test as random factors(X^2(3, N=192) = .50, p = .92). This result indicates that student learning was not altered by the artificial teacher's mood.

Next, a manipulation check was performed to ensure that the mood displayed by the pedagogical agents transferred to the participant so that analyses about student emotions and learning could be conducted. Before analyses were conducted, the values from the affect grid were computed from the baseline and standardized so that interactions between valence and arousal could be analyzed. Using a model including the four conditions as a fixed effect and participant, topic and test as random factors, the manipulation appeared to have indeed induced student affect. Specifically, the condition of the tutor (i.e., positive, negative, neutral, or control) had a significant main effect on self-reported valence of the student (F (3,189) = 5.63, p < .01). The model was significantly different from the null model including the random factors of participant, topic and test (X^2(3, N = 192) = 16.30, p < .001) with a difference in variance accounted for of 8% (R^2 = .076). Therefore, the pedagogical agent's mood accounts for 8% of the variance in student reported affective valence.

Post hoc analyses with a Tukey correction showed significant mood contagion for the negative and positive tutorial conditions, with a marginally significant difference from the neutral condition. Specifically, the negative condition showed an increase in negatively-valenced affect compared to the text (or control) condition (z (1,192) = -3.987, p < .001). The positive condition showed a significant increase in positively-valenced affect compared to the negative tutorial condition (z (1,191) = 2.65, p < .05), and the neutral tutorial condition showed a marginally significant increase in negative affect (z (1,191) = 2.310, p < .1).

The second hypothesis investigated was the relationships between the students' valence, arousal, the interaction between the two, and learning gains. A series of additive models were conducted to assess whether any relationships were additive or interactive. Models testing the fixed factors of valence and the interaction of valence and arousal along with the three random factors were non-significant compared to the null model (X^2 (1, N = 293) = .02, p = .88); (X^2(1, N = 192) = .54, p = .46), respectively. However, a full model of arousal, valence and pretest as fixed factors with participant, topic, and test as random factors was significantly different from the previous model (X^2 (1, N = 192) = 3.72, p = .05). This model suggests that arousal, not valence, significantly contributed to learning (F (1,191) = 4.20, p < .05); (F (1.191) = .14, p = .71), respectively. Specifically, valence to contributed near 0% of the variance (R^2 = .001) whereas arousal accounted for 1.2% of the variance (R^2 = .012).This relationship between arousal and learning was negative (t (1,191) = -2.02, p = .04). Therefore, the lower the intensity of affect reported by the student, the higher the learning gains.

Discourse Features. The analyses on the discourse of the students only applied to the conditions where students had tutorial conversations with the agents, thus excluding the control condition and reducing the number of observations (N = 144). In evaluating learning gains, first a model with pretest and the composite metric of formality as fixed factors and participant, topic and test version A or B (referred to as "test") as random factors was compared to a null model including just pretest and the three random factors, and found to be non-significant ($X^2(7, N = 144) = .46$, $p = .50$). Next, a model with all five of the components and pretest as fixed factors with participant, topic, and test as random factors was not significantly different from a null model including pretest and the random factors of participant, topic, and test ($X^2(11, N = 144) = 8.29$, $p = .14$). Although the full model with all 5 components was not significant, one of the components known as the word concreteness significantly correlated with learning. This component is indicative of words evoking a mental representation and more meaningful than abstract words. Therefore a trimmed model was tested with the fixed factors of word concreteness and pretest and topic, test, and participant as random factors was tested for effects on learning. This model was significantly different from a null model of pretest and the three random factors ($X^2(8, N = 144) = 4.69$, $p = .03$). The component of word concreteness had a significant main effect ($F(1,143) = 4.88$, $R^2 = .02$, $p < .05$) with a positive relationship with learning ($t(1,143) = 2.209$, $p < .05$). The principal component of concreteness might best represent the most shallow level of comprehension compared to the other components. This makes sense in the current context as students were learning the basic didactic knowledge [21].

4 Conclusions

Results indicated that there was no direct impact of the mood of the pedagogical agents on learning. On the other hand, student emotions did impact learning gains. Indeed students with lower arousal levels showed higher learning gains. Finally, the discourse analysis revealed word concreteness, which is indicative of shallow comprehension, is related to learning.

Teacher mood did not affect student learning. However, teacher mood did affect student affective state. Previous research has found contradictory yet significant findings as to the nature of the relationship between teacher moods and learning [6,7]. In fact, it may be the case that teacher mood can only have an indirect effect on learning via student affect. A future study may address this issue by examining moderators and mediators as well as exploring the impact of having a stronger manipulation such as altered feedback, body posture of the agent, or speech.

The result that arousal, not valence, impacts student learning may best be explained by the arousal theory. The theory suggests that intensity of emotions rather than valence contributes to learning on certain tasks [10], especially those shallow in nature [11]. The Training module of *OperationARIES!,* the module used in the current study, teaches shallow knowledge whereas subsequent modules within the game are required for deeper conceptual learning [23].The discourse analyses suggests that students were learning shallow knowledge during the interaction with the altered version of the game as well. Specifically, the component representing the use of meaningful, concrete words (a shallow level of comprehension) is the only component that significantly contributed to learning in this study. The corresponding comprehension framework suggests that this is indicative of didactic knowledge only. Therefore, it seems likely that a sweet spot of intensity, in accordance with the arousal theory, may be important for learning within this module of the game.

The nature and contributing factors while learning through interaction with an Intelligent Tutoring System such as *OperationARIES!* are varied and complex. Many factors not tested in this study may contribute to learning gains in such an environment. However, this study definitively shows a relationship between arousal and discourse features related to learning.

ACKNOWLEDGEMENTS.

This research was supported by the Institute for Education Sciences, U.S. Department of Education, through Grant R305B070349 to Northern Illinois University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

5 References

1. Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A. C., Halpern, D.: Operation ARIES! A serious game for teaching scientific inquiry: Serious games and edutainment applications. In: M. Ma, A. Oikonomou, J. Lakhmi (eds.), pp.169-196. Springer-Verlag, London (2011)
2. Forsyth, C.M., Graesser,A.C. Pavlik, P., Cai, Z., Butler,H., Halpern, D.F., Millis,K.: OperationARIES! methods, mystery and mixed models: Discourse features predict affect in a serious game. J. Ed. D. Min. (2014)
3. D'Mello, S., Graessar, A.C.: Emotions during learning with AutoTutor: Adaptive technologies for training and education. In: P.J. Durlach, A. Lesgold. (eds.). Cambridge University Press (2012)
4. Conati, C., Maclaren, H.: Empirically building and evaluating a probabilistic model of user affect. User Mod. User-Ad. Inter. 19, 267-303 (2009).
5. Russell, J., Weiss, A. & Mendelsohn, J. A. Affect grid: A single-item scale of pleasure and arousal. J. of Pers. and Soc. Psych , 57, 493-5029(1989)
6. Sullins, J., Craig, S., & Graesser, A.C. (2009). Tough love: The influence of an agent's negative affect on students' learning. In V. Dimitrova, R. Mizoguchi, B. Du Boulay, B., & A.C. Graesser (Ed.), Artificial intelligence in education: Building learning systems that care: From knowledge representation to affective modeling (pp.677-679). Am-sterdam: IOS Press.
7. Isen, A.M. (2008). Positive affect and decision processes: Some recent theoretical developments with practical implications: Handbook of consumer psychology. In: C.Haugdevedt, P.M. Herr, F.R. Kardes (eds.) pp. 273-296. New York, Psychology Press (2008)

8. Franzel, A.C., Goetz, T., Ludtke, O., Pekrun, R., Sutton, R.E.: Emotion transmission in the classroom: Exploring the relationship between teacher and student enjoyment. J. Ed. Psy. 101, 705-716 (2009)
9. Frederickson, B. L.: The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. Amer. Psy. 56, 218–226 (2001)
10. Lang, P. J.: The emotion probe: Studies of motivation and attention. Amer. Psy. 50, 372–385 (1995)
11. Martindale, C.: Cognition and Consciousness. Dorsey Press, Illinois (1981)
12. Pennebaker, J., Francis, M.E., Booth R. J.: Linguistic inquiry and word count (LIWC):LIWC2001. Mahwah, New Jersey (2001)
13. Rude, S. S., Gortner, E., & Pennebaker, J. Language use of depressed and depression- vulnerable college students. Cog. & Emot., 18, 1121-1133 (2004).
14. Graesser, A. C., Forsyth, C.: Discourse comprehension: Oxford handbook of cognitive psychology In: D. Reisberg (eds.) Oxford University Press (2014)
15. Graesser, A. C., Millis, K. K., & Zwaan, R. A. Discourse comprehension. Annual Review of Psychology, 48, 163-189.(1997)
16. Graesser, A. C., McNamara , D. S.: Computational analyses of multilevel discourse comprehension. Top. Cog. Sci. (2011)
17. Graesser, A. C., McNamara, D. S., Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text?: Rethinking reading comprehension. In: A. P. Sweet, C. E. Snow (eds.) pp. 82-98. New York, Guilford (2003)
18. Graesser, A.C., McNamara, D.S., Kulikowich, J.: Coh-Metrix: Providing multilevel analyses of text characteristics. Ed. Res. 40, 223-234 (2011)
19. Landauer, T., McNamara, D. S., Dennis, S., Kintsch, W.: Handbook of Latent Semantic Analysis. Mahwah, New Jersey (2007)
20. Jurafsky,D., Martin, J.(2008). Speech and language processing. Prentice Hall, New Jersey (2008)
21. Cai, Z., Graesser, A.C., Forsyth, C., Burkett, C., Millis, K., Wallace, P., Halpern, D. Butler.: Trialog in ARIES: User input assessment in an intelligent tutoring system. In: 3rd IEEE International Conference on Intelligent Computing and Intelligent Systems, pp.429-433. IEEE Press, China (2011)
22. Jackson, G.T., Graesser, A.C., McNamara, D.S.: What students expect may have more impact than what they know or feel: Artificial intelligence in education; Building learning systems that care; From knowledge representation to affective modeling. In: V. Dimitrova, R. Mizoguchi, B. Du Boulay, A.C. Graesser (eds.) pp. 73-80. IOS Press (2009)
23. Forsyth, C.M., Graesser, A.C., Walker, B., Millis, K., Pavlik, P, Halpern, D. (2013). Didactic galactic: Types of knowledge learned in a serious game. In: 16th International Conference Artificial Intelligence in Education Lecture Notes in Computer Science, pp 832-835. Lec Berlin, Heidelberg: Springer Verlag (2013)