Secret Agents, Alien Spies, and a Quest to Save the World: Engaging Students in Scientific Reasoning and Critical Thinking through Operation ARIES!

Heather A. Butler — *Claremont Graduate University*
Carol Forsyth — *The University of Memphis*
Diane F. Halpern — *Claremont McKenna College*
Arthur Graesser — *The University of Memphis*
Keith Millis — *Northern Illinois University*

Operation ARIES! is an interactive educational game that teaching critical thinking and scientific reasoning. On one hand Operation ARIES! is an educational tool that was developed by experts in the learning sciences, but on the other hand it is a fantasy game with a science-fiction plot and students compete against other players using the knowledge they learn. Operation ARIES! was designed for advanced high school students, undergraduate university and college students, and soldiers in the military. Students learn the material by engaging in interactive conversations with agents that adapt to their skill level, by applying the acquired knowledge in critiques of ecologically valid case studies, and by generating their own questions about incomplete research while interrogating suspected alien scientists.

The game can be adopted and implemented by educators in various ways. First, it is appropriate for introductory courses in science (e.g., psychology, biology, and chemistry), critical thinking, and research methods. Second, students can complete Operation ARIES! in 15-25 hours as an in-class activity, as a homework assignment, or combination of the two. Third, because the game includes different modules or levels, it is possible for educators to pick and choose which will be most applicable to their needs. Operation ARIES! is published by Pearson's Education. The purpose of the present chapter is to introduce interested readers to the science behind Operation ARIES!

Operation ARIES!

Your mission: to expose the aliens who endeavor to take over Earth by stealing our natural resources, spreading bad science, and lulling mankind into mindless consumerism. These aliens must be stopped. As an agent with the Federal Bureau of Science, you will receive the latest training methods to spot aliens posing as human scientists, you will identify the flaws in research from a variety of fields, and you will interrogate suspected alien spies. This is the science-fiction plot behind Operation ARIES!, an Intelligent Tutoring System (ITS) that teaches scientific reasoning and critical thinking skills. Student players become Federal Bureau of Science agents-in-training charged with defending Earth from aliens who are intent on destroying it. To defend Earth, student agents must learn the principles of the scientific method and critical thinking. The student players then must use these principles to evaluate case studies and interrogate suspected alien scientists.

Figure 1. Operation ARIES! Logo.

The importance of critical thinking skills cannot be understated. In a recent survey, 81% of employers listed critical thinking skills as a top priority for new employees (AAC&U, 2010). The United States Bureau of Labor Statistics for 2010-2011 reports that "knowledge workers" or "symbol analysts" are in high demand. As such, workers who can complete multi-step operations, manipulate abstract symbols, understand complex ideas, acquire new accurate information efficiently, and exhibit flexible thinking will be highly employable, and necessary for the economic and political progress of any nation.

Critical thinking has been defined in various ways, but researchers generally agree that critical thinking achieves a desired outcome by thinking rationally in a goal-oriented fashion (e.g., Ennis, 1993; Halpern, 2003; Moseley et al. 2005; Sternberg, Roediger, & Halpern, 2007). Halpern (2003) defined critical thinking as: the use of those cognitive skills or strategies that increase the probability of a desirable outcome. It is used to describe thinking that is purposeful, reasoned, and goal directed--the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions, when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task. (p. 6).

Critical thinking is both a collection of learned skills and a disposition towards information (Halpern, 2003). Critical thinkers are more flexible, more deliberate information processors (or more intentional in how they process information), more informed decision makers, and are more persistent than non-critical thinkers who may accept as truth (without questioning the validity of) any and all information presented to them. A wealth of evidence suggest that critical thinking skills can be learned and critical thinking disposition can be developed (for reviews, see Chance, 1986; Halpern, 2003; Moseley et al. 2005; Nisbett, 1992). Educators can teach critical thinking explicitly in the form of critical thinking or problem solving courses, and can teach critical thinking in other courses either explicitly or imbedded in the coursework (Marin & Halpern, 2010). One aspect of critical thinking that Operation ARIES! targets is scientific reasoning.

Scientific reasoning is a process to ask and answer questions about the world. It involves understanding the scientific method and using that understanding to conduct research and evaluate the research of others. However, scientists are not the only people in need of good scientific reasoning skills. We live in a world that is more complex and technical with each passing day and we are constantly bombarded with pseudoscientific persuasive messages. Consumers today may encounter hate websites disguised as science or may encounter advertising that offers scientific-sounding explanations for a products' effectiveness. Without a background in research methods or scientific reasoning skills, consumers could be susceptible to the false claims made by the media. At best, consumers could be wasting money on ineffective products; at worst, they could be putting their health or wellbeing in jeopardy. Consumers need to arm themselves with an arsenal of good thinking skills in order to guard against such claims and persuasive messages, and Operational ARIES! provides a solution to that need in a highly engaging game.

Operation ARIES! is an agent-based Intelligent Tutoring System (ITS) that uses artificial pedagogical agents and natural language processing to teach the student aspects of the scientific method. Operation ARIES! utilizes an architecture similar to that of Autotutor (Graesser, Wiemer-Hastings, Wiemer-Hastings, Kreuz, & the Tutoring Research Group, 1999; Graesser, Person, Harter, & Tutoring Research Group, 2001), an ITS which instructs students on physics (VanLehn, Graesser, Jackson, Jordan, Olney, & Rose, 2007) and computer literacy skills (Graesser, Lu, Jackson, Mitchell, Ventura, Olney, & Louwerse, 2004) through mixed-initiative dialogue. This type of dialogue allows either the artificial agent or the human to direct the flow of the conversation. Autotutor produced learning gains comparable to one-on-one human tutoring (Graesser, Chipman, Haynes, & Olney, 2005; Graesser, Lu et al., 2004; Graesser, et al., 2001; VanLehn, et al., 2007).

Operation ARIES! employs the scientific principles of learning and serious games. Students are engaged in the material using the pedagogical principles of active learning, immediate feedback, dialog interactivity, multimedia effects, distributed practice, and transfer of learning. "Hard fun" (Papert, 1980) is another principle that may increase the success of educational or epistemic games (Schaeffer, 2006). This phenomenon describes the resulting sense of satisfaction students experience after struggling to understand a difficult topic. That is, the student's enjoyment should increase as the game moves from teaching basic declarative knowledge in the first module to the use of this knowledge in the analysis of ecologically valid cases in the later modules.

Operation ARIES! entails three modules: a training course, the evaluation of case studies, and the interrogation of suspected alien spies. The story-line continues to twist and turn through the first two modules ending in the fantastical climax and surprising resolution in the final module. In a similar fashion, the curriculum builds on itself and becomes increasingly more challenging across the three modules. The first module, the training course, teaches scientific concepts that are shared among the fields of psychology, biology, and chemistry. Student players learn twenty-one important scientific concepts including both the definition and function of specific topics such as theories, hypotheses, falsification, operational definitions, independent variables, and dependent variables (see Table 1 for a complete list).

The content of the training course is delivered through an e-book, multiple choice questions, and natural language conversations. The e-book contains

Table 1

A list of the topics presented in Operation ARIES!

Topics Presented in Operation ARIES!
Theories and the Experimental Method
Hypothesis
Science and Pseudoscience
Operational Definitions
Independent Variables and Participant Variables
The Dependent Variable
Dependent Variables: Reliability, Accuracy, and Precision
Dependent Variables: Validity
The Dependent Variable: Objective Scoring
Replication of Results
Experimental Control
Control Groups
Random Assignment to Groups
Subject Bias
Attrition and Mortality
Representative Samples
Sample Size
Experimenter Bias
Conflict of Interest
Causality vs. Correlation
Drawing Conclusions: Generalizability

illustrated chapters written from an alien's perspective and includes several vivid examples of each concept, as well as humorous descriptions of common misconceptions (see Figure 2 for a sample page taken from the training module). Perceived student control is an aspect present in many successful video games because it allegedly increases engagement. Therefore, Operation ARIES! allows students to have control over whether or not they read the chapter. Advanced students can skip chapters by demonstrating a proficient standard of performance, or they can choose to refresh their memory of the concept by reading the chapter. This autonomy and adaptation reduces the frustration and boredom associated with having to relearn concepts familiar to the learner. Affective states, such as boredom, are associated with poor learning and poor behavioral outcomes (Baker, D'Mello, Rodrigo, & Graesser, 2010) and are detrimental to student learning and engagement.

At the conclusion of each e-book chapter, students complete a multiple-choice posttest and engage in tutoring "trialogs." A trialog is a conversation between two artificial pedagogical

Figure 2. An example page taken from the training module.

agents and a human student. In the context of this game, the students are tutored on each scientific concept in trialogs with two avatars, a virtual teacher (Dr. Quinn, an FBS handler) and a virtual student (Glass, a fellow agent-in-training). The human students interact with the avatars using natural language dialogue. To assess and guide the student to a deeper understanding, natural language processing tools recognize the human student input and respond with appropriate feedback, hints, prompts for information, assertions, and misconception correction. An example conversation appears in Appendix A.

Figure 3. Human students engage in interactive trialog with these animated agents.

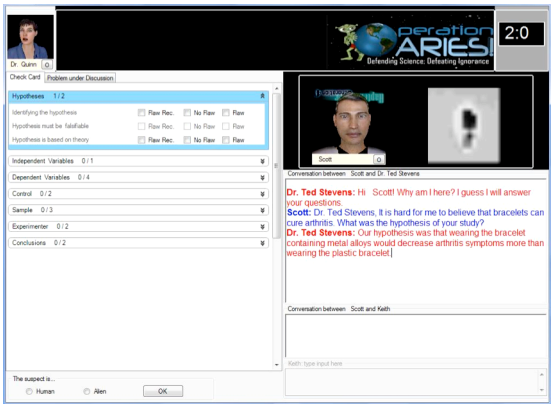
Effective tutors need to gauge and adapt to the student's current level of understanding. Proposed criteria for a successful adaptive tutor include choosing problems that specifically address the student's lacking knowledge and taking previous test scores into consideration (Graesser, D'Mello, & Cade, 2009). In compliance with these criteria, students are adaptively placed in one of three tutoring conditions by their scores on the previous multiple

choice tests. If the human students demonstrate a low-level understanding of the concept, they receive the *vicarious learning* trialog where they observe the virtual teacher tutoring the virtual student. Vicarious learning conditions have shown significant learning gains specifically for low prior-knowledge students (Driscoll, Craig, Gholson, Ventura, Hu, & Graesser, 2003). To maintain engagement, the students respond to the tutoring situation. For example, the virtual teacher might ask the human student whether the virtual student understands the concept or whether the virtual student's answer was correct. If the human students demonstrate a moderate understanding of the concept, they receive the *standard tutoring trialog*. For example, the virtual teacher might ask the human student to define the concept and scaffold the student by or with hints, prompts, feedback and misconception correction. If the students demonstrate good understanding of the concept, they interact with the *teachable agent trialog* and the human student tutors the virtual student. For example, the virtual student might tell the human student that they do not really understand the concept and offer an incorrect explanation. The human student would then have to explain to the virtual student what the concept is and why they were incorrect (see Figure 3 for an example of the animated agents). After completing the training course, the students graduate to a higher level of training where they apply their new skills to real cases.

Figure 4. An example of the case studies module.

The second module of Operation ARIES! is the case studies module (see Figure 4 for a an example of the case studies module). In this module, human students play against a sassy (and frequently obnoxious) avatar in a competitive game testing for transfer of the knowledge from the training module to ecologically valid cases. Each case describes published research in an ecologically valid way (e.g., a newspaper article written about a new study, a research abstract, a bloggers description or reaction to published research, etc.), and each case may contain flaws. The students and avatar compete to identify the flaws in the research, and earn points on the number of correctly identified flaws and lose points on the number of falsely identified flaws or missed flaws. A sophisticated algorithm ensures that the competition between the two players is close, thereby encouraging interest and engagement. Scaffolding is provided to struggling students who can purchase a list of potential flaws for points. As the module draws to a close, the plot of the alien invasion thickens and it becomes necessary for the human student (now a secret agent) to interrogate a series of suspected alien scientists.

Figure 5. An example of the interrogation module.



The third module is the interrogation module (see Figure 5 for an example of the interrogation module). In this module, students interrogate captured scientists about their research to uncover their true identity as either an alien or a human scientist. Whereas poorly conducted research indicates that the scientist is an alien posing as a human scientist, properly conducted research indicates that the scientist is really a human. Unlike the more complete descriptions provided in the case studies module, the research described in this module is incomplete. Accordingly, students must question the sometimes hostile scientists and determine when enough information has been gathered. Ultimately, the students decide whether the scientist is a human or an alien based on the quality of the research and earn points for correctly identifying alien scientists and loose points for false accusations.

The intricate plot of Operation ARIES! is threaded throughout the modules to promote engagement. Plot, agency, and emotions play an integral part in maintaining reader interest (Brewer &

Lichtenstein, 1981; Brewer & Ohtsuka, 1988). Surprise, suspense, and curiosity are three emotions that can be used to encourage reader interest and in turn to impact comprehension. For example, information can be released at a critical time in order to surprise the reader. This occurs several times in Operation ARIES! when, for example, the human student discovers that a fellow agent is having an inter-species love affair with an alien defector. Suspense can be built by withholding the outcome of a story from the reader. Graesser and Klettke (2001) pointed out that this literary tension can be extended over a period of time, and clever writers create false alarm episodes that draw-out the suspense of a plot. In some sense, the entire Operation ARIES! game is a battle of wits with aliens with the fate of the Earth hanging in the balance. In addition, smaller mysterious events are also dispersed throughout the experience to increase the suspense. Curiosity can be evoked by telling the reader the outcome before the reader knows the steps that led to the outcome. Operation ARIES! evokes curiosity by means of emergency "breaking news" reports and email messages. Students receive messages that lakes are disappearing and that strange metallic formations have begun to form all over the world, but they are not told how these events occurred, what these formations are, who made these events happen, or why they are happening. Thus, Operation ARIES! evokes a variety of emotions to maintain student interest and improves cognitive processing of the material.

Of interest to educators is that Operation ARIES! is also a valuable assessment tool. Educators can track student progress in a variety of ways. In the training module, educators can track the number of questions correctly answered after reading the chapter, the tutorial style during the tutoring trialogs, or the qualitative responses during the tutoring trialogs. In the case studies module, students compete against a sassy avatar for points that could indicate correctly applied knowledge. In the interrogation module, students accrue points, generate questions to ask the suspected alien scientists, and make dichotomous decisions about the validity of research. Thus, Operation ARIES! can provide educators with both quantitative and qualitative data to assess student learning.

Operation ARIES! is an innovative way to teach research methods and critical thinking. It utilizes the principles of learning and serious games to engage students in thinking critically about the information they consume. Operation ARIES! meets the needs of advanced high school students, undergraduate university and college students, and soldiers in the military. It has a variety of applications, but was designed to accompany introductory science, research methods, and critical thinking courses. Students learn the material by engaging in interactive trialogs that adapt to their skill level. Then they apply the learned knowledge by evaluating ecologically valid case studies and competing against other players. Finally, the students must generate their own questions about incomplete research while interrogating suspected alien scientists. Operation ARIES! can be used as an in-class activity, assigned as homework, or a combination of both. It takes between 15-25 hours to complete. Student progress can be assessed throughout the game yielding both qualitative and quantitative data. Thus, while Operation ARIES is an engaging game for students, it is also a valuable resource for educators who are devoted to teaching not only shallow but also applied knowledge which helps students reach a deeper-level of understanding.

References

Association of American Colleges & Universities. (2010). *Raising the bar: Employers' views on college learning in the wake of the economic downturn*. Retrieved from the AAC&U website: http://www.aacu.org/leap.

Baker, R. S. J. d., D'Mello, S. K., Rodrigo, Ma. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*, 223-241.

Brewer, W. F., & Lichtenstein, E. H. (1981). Event schemas, story schemas, and story grammars. In J. Long and A. Baddeley (Eds.), *Attention and performance IX* (pp. 363-379). Hillsdale, NJ: Erlbaum.

Brewer, W. F., & Ohtsuka, K. (1988). Story structure, characterization, just world organization, and reader affect in American and Hungarian short stories. *Poetics, 17*, 395-415.

Chance, P. (1986). *Thinking in the classroom: A survey of programs*. New York, NY: Teachers College, Columbia University.

Driscoll, D., Craig, S. D., Gholson, B., Ventura, M., Hu, X., & Graesser, A. (2003). Vicarious learning: Effects of overhearing dialog and monolog-like discourse in a virtual tutoring session. *Journal of Educational Computing Research*, *29*, 431-450.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into practice, 32*, 179-186.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions in Education, 48*, 612-618.

Graesser, A. C., D'Mello, S., & Cade, W. (2009). Instruction based on tutoring. In R. E. Mayer and P.A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 408-426). New York, NY: Routledge Press.

Graesser, A. C., & Klettke, B. (2001). Plot, agency, and a structural affect theory of literary short story comprehension. In S. Steen and D. Schram (Eds.), *Psychology and sociology of literature* (pp. 57-69). Amsterdam, Netherlands: John Benjamins.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers, 36*, 180-193.

Graesser, A. C., Person, N., Harter, D., & Tutoring Research Group. (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education, 12*, 257-279.

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & the Tutoring Research Group. (1999). Auto Tutor: A simulation of a human tutor. *Journal of Cognitive Systems Research, 1*, 35-51.

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Mahwah, NJ: Erlbaum.

Marin, L. M., & Halpern, D. F. (2010). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gain. *Thinking Skills and Creativity, 6*, 1-13.

Moseley, D., Baumfield, V., Elliott, J., Gregson, M., Higgins, S., Miller, J., & Newton D. P. (2005). *Frameworks for thinking: A handbook for teaching and learning*. Cambridge, UK: Cambridge University Press.

Nisbett, R. E. (1992*). Rules for reasoning*. Hillsdale, NJ: Erlbaum.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York, NY: Basic Books.

Schaeffer, D. W. (2006). *How computer games help children learn*. New York, NY: Macmillan.

Sternberg, R. J., Roediger, H. L., III., & Halpern, D. F. (Eds.). (2007). *Critical thinking in psychology*. New York, NY: Cambridge University Press.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, *31*, 3–62.

APPENDIX A: Example Script

This is an example script highlighting the scaffolding methods used by the artificial agents during interaction with a real human.

QUESTION: Glass: Okay, I know operational definitions are important in research, but I'm not sure why.

HINT: Glass: Let's just go over the importance of these things one more time. Why do we need to have operational definitions?
Human: so that they can be reliably measured and used by other researchers

NEUTRAL FEEDBACK: Glass: OK.

PUMP: Dr. Quinn: Um, can you add to that?
Human: it is important to create operational definitions because they provide a basis for
measuring variables and so that they can be reliably measured and used by other researchers in other studies.
Glass: That is not clear to me.

HINT: Glass: I thought operational definitions had something to do with telling others about your study. What could that be?
Human: how to measure a variable
Glass: I still don't understand.

PROMPT: Glass: Wow, it seems like operational definitions have something to do with being able to reliably do something. What is that something?
Human: measure a variable

NEUTRAL FEEDBACK: Dr. Quinn: Okay.

CORRECT PROMPT COMPLETION: Dr. Quinn: reliably recognize variables

PROMPT: Glass: I'm remembering something about how operational definitions make it so that a variable means the same thing to all researchers. That means that when other researchers use the same variable, it is what?
Human: reliable

POSITIVE FEEDBACK: Dr. Quinn: Great!
Dr. Quinn: it is reliable
Glass: Yeah, I understand.

SUMMARY: Glass: Operational definitions are important to consider because they allow a particular variable to be reliably recognized, measured, and understood by all researchers.