Serious Games for Language Learning: How Much Game, How Much AI?

W. Lewis Johnson, Hannes Vilhjalmsson, & Stacy Marsella
Center for Advanced Research in Technology for Education
USC / Information Sciences Institute, 4676 Admiralty Way, Marina del Rey, CA 90292

Abstract. Modern computer games show potential not just for engaging and entertaining users, but also in promoting learning. Game designers employ a range of techniques to promote long-term user engagement and motivation. These techniques are increasingly being employed in so-called *serious games*, games that have non-entertainment purposes such as education or training. Although such games share the goal of AIED of promoting deep learner engagement with subject matter, the techniques employed are very different. Can AIED technologies complement and enhance serious game design techniques, or does good serious game design render AIED techniques superfluous? This paper explores these questions in the context of the Tactical Language Training System (TLTS), a program that supports rapid acquisition of foreign language and cultural skills. The TLTS combines game design principles and game development tools with learner modelling, pedagogical agents, and pedagogical dramas. Learners carry out missions in a simulated game world, interacting with non-player characters. A virtual aide assists the learners if they run into difficulties, and gives performance feedback in the context of preparatory exercises. Artificial intelligence plays a key role in controlling the behaviour of the non-player characters in the game; intelligent tutoring provides supplementary scaffolding.

Introduction

In the early days of intelligent tutoring system (ITS) research, intelligent tutors were conceived of not just as aids for academic problem solving, but as supports for interactive games. For example, Sleeman and Brown's seminal book, *Intelligent Tutoring Systems*, included two papers on tutors that interacted with learners in the context of games: the WEST tutor [1] and the Wumpus tutor [5]. It was recognized that games can be a powerful vehicle for learning, and that artificial intelligence could amplify the learning outcomes of games, e.g., by scaffolding novice game players or by reinforcing the concepts underlying game play.

Fast forward to 2005. Computer games have become a huge industry, a pastime that most college students engage in [12]. In their striving for commercial success, game developers have come up with a set of design principles that promote deep, persistent engagement, as well as learning [17]. Education researchers are now seeking to understand these principles, so that they can understand how to make education more effective [4]. There is increasing interest in *serious games*, programs that obey solid game design principles but whose purpose is other than to entertain, e.g., to educate or train [20]. Meanwhile, with a few exceptions (e.g., [2, 6]), very little current work in AI in education focuses on games.

This paper examines the question of what role artificial intelligence should play in serious games, in order to promote learning. The artificial intelligence techniques used must support the learning-promoting features of the game, otherwise they may be superfluous or

even counterproductive. These issues are discussed in the context of the Tactical Language Training System (TLTS), a serious game for learning foreign language and culture.

1. Overview of the Tactical Language Training System

The language courses delivered using the TLTS have a strong task-based focus; they give people enough knowledge of language and culture to enable them to carry out particular tasks in a foreign country, such as introducing yourself, obtaining directions, and meeting with local officials. The current curricula address the needs of military personnel, however the same method could be applied to any course that focuses on the skills needed to cope with specific situations, e.g., booking hotel rooms or meeting with business clients. Two training courses have been developed so far: Tactical Levantine Arabic, for the Arabic dialect spoken in the Levant, and Tactical Iraqi, for Iraqi Arabic Dialect.

Figure 1. Views of the Tactical Language Training System

Figure 2. Arcade Game in the TLTS

The TLTS includes the following main components [8]. The Mission Game (Figure 1, left side) is an interactive story-based 3D game where learners practice carrying out the mission. Here the player's character, center, is introducing himself to an Iraqi man in a café, so that he can ask him where the local leader might be found. The player is accompanied by an aide character (middle left), who can offer suggestions of what to do if the player gets stuck. The Skill Builder (Figure 1, right) is a set of interactive exercises focused on the target skills and tasks, in which learners practice saying words and phrases, and engaging in simple conversations. A virtual tutor evaluates the learner's speech and gives feedback on errors, while providing encouragement and attempting to overcome learner negative affectivity [10]. A speech-enabled Arcade Game gives learners further practice in speaking words and phrases

(Figure 2). Finally, there is an adaptive hypertext glossary that shows the vocabulary in each lesson, and explains the grammatical structure of the phrases being learned.

The TLTS has been evaluated multiple times with representative learners, through an iterative formative evaluation process [11]. The evaluations provide evidence that the game format motivates learners who otherwise would be reluctant to study a difficult language such as Arabic. A significant amount of content is being developed, which by July 2005 should be able to support around 80 hours of interaction for Iraqi Arabic and somewhat less for Levantine Arabic. Multiple military training centers have volunteered to serve as test sites.

2. Using AI Design to Support Game Design

The premise of the serious game approach to learning is that well designed games promote learner states that are conducive to learning. Serious game developers adhere to a number of common design principles that tend to yield desirable interaction modes and learner states [4, 17]. Some of these principles are commonplace in AIED systems, particularly those that employ a goal-based-scenario approach [19]; others are less common, and may appear new to AIED developers. Game AI can play a critical role in implementing these principles. Game AI is a major research area in its own right, which goes beyond the scope of this paper (see [13] for an overview). In educational serious games, the challenges are to make sure the game AI supports educational objectives, and to introduce other educational AI functions as needed without compromising game design principles, in order maximize learning.

2.1. Gameplay

According to Prensky, one of the foremost characteristics of good games is good gameplay. "*Gameplay is all the activities and strategies game designers employ to get and keep the player engaged and motivated to complete each level and an entire game.*" [18] Good gameplay does not come from the game graphics, but from the continual decision making and action that engages the learner and keeps him or her motivated to continue.

There are two aspects of gameplay: engaging users moment by moment, and relating current game actions to future objectives. In good moment-by-moment gameplay, each action or decision tends to naturally lead to the next action or decision, putting the player in a psychological state of flow [3]. Moment-by-moment gameplay is realized in the Mission Game as follows. The actions in the Mission Game that relate to the target tasks (namely, face-to-face communication) are embedded in a larger sequence of navigation, exploration, and information gathering activities that learners engage in as they carry out their mission. When the learner is engaged in a conversation with a nonplayer character, there is a give and take between the characters; the nonplayer characters respond both verbally and nonverbally to the learner's utterances, and may take initiative in the dialog. In the Arcade Game, there is a constant flow of action and reaction between the user's actions (issuing spoken Arabic commands to navigate through the game level and pick up objects) and the game's response (moving the game character as indicated by the spoken commands, scoring points for correctly uttered phrases and collected items, and immediately placing new objects in the game level). In the Mission Game, orientation toward future objectives occurs as the learner develops rapport with the local people, and obtains information from them relevant to the mission. In the Arcade Game this orientation occurs as learners seek to increase their overall game score and progress to the next game level.

One way that AI facilitates gameplay in the TLTS is by promoting rapid interaction with nonplayer characters. Speech recognition in the game contexts is designed to rapidly and robustly classify the intended meaning of each learner utterance, in a manner that reasonably tolerant of learner errors, at least as much as human native speakers would be [9]. Natural language processing is employed to generate possible dialog variants that a

learner might attempt to say during the game, but only at authoring time, to reduce the amount of game-time processing required on user input. The PsychSim package is then used to generate each character's responses to the learner's actions [21]. Pedagogical objectives are realized in PsychSim using an interactive pedagogical drama approach, by making sure that the nonplayer characters respond to aspects of learner communication that are pedagogically important (e.g., appropriate use of polite gesture and language).

On the other hand, the common use of AI in intelligent tutoring systems, to provide tutorial scaffolding, is carefully restricted in the TLTS. We avoid interrupting the gameplay with critiques of learner language. Such critiques are reserved for Skill Builder lessons and after-action review of learner performance.

2.2. Feedback

Good games provide users with feedback on their actions, so that they know how well they are doing and can seek to improve their performance. This has obvious relevance to serious games that motivate learners to improve their skills.

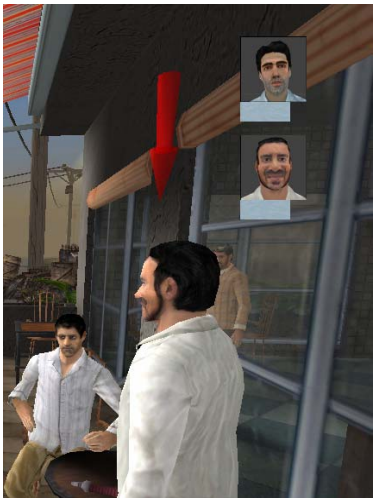

Figure 3. Close-up of the trust meters in the Mission Game

As we conducted formative evaluations of the TLTS, we frequently saw a need to improve feedback, and developed new feedback methods in response. For example, when learners carry out actions in the Mission Game that develop rapport with the local people (e.g., greet them and carry out proper introductions), they want to know if they are making progress. Some cues that people rely on in real life, such as the facial expressions of the people they are talking to, are not readily available in the game engine underlying TLTS (namely, Unreal Tournament 2003). We therefore developed an augmented view of the non-player characters' mental state, called a trust meter, shown in the upper right of Figure 3. The size of the grey bar under each character image grows and shrinks dynamically depending upon the current degree of trust that character has for the player. Note that this lessens the need for intelligent coaching on the subject of establishing trust, since learners can recognize when their actions are failing to establish trust.

2.3. Affordances

Another feature of good games is their simple, well-defined interfaces, designed to support the interaction between the user and the game. Even in games that attempt to create very realistic 3-D virtual worlds, designers will augment that reality in various ways to provide the user with "perceived affordances" [16], in essence cues that suggest or guide user actions. For example, in Figure 3 there is a red arrow above the head of one of the

characters that informs the user about which character in the virtual world is engaging them in the conversation. More generally, the Mission Game uses icons and highlighting of the screen to help regulate the dialog turn-taking between the learner and the characters. Although this augmented reality diverges from strict realism both in terms of the rendering of the scene and the mechanisms used to regulate dialog turn-taking in real-life, they better serve the goal of maintaining a fluid interaction between the learner and the non-player characters. Again, effective use of affordances lessens the need for intelligent coaching to advise learners on what actions to take.

2.4. Challenge

An important aspect of game design is ensuring that users experience a proper level of challenge. Gee argues that the user experience should be "pleasantly frustrating:" a challenge for the player, but not an insurmountable one [4]. The role of challenge in promoting intrinsic motivation is not limited to games, but has been noted by motivation researchers as relevant to all learning activities [14].

The TLTS is configurable to adjust the level of challenge of play. When beginners play in the Mission Game, they receive assists in the form of subtitles showing what the Arab characters as saying, both in transliteration and in English translation. Also, each Mission Game scene can be played at two levels of difficulty, either Novice or Experienced. At the Novice level the Arab characters are relatively tolerant of cultural gaffes, such as failing to show proper respect or failing to make proper introductions. At the Experienced level the Arab characters become suspicious more easily, and expect to be treated with respect. This is accomplished by having content authors construct examples of dialog at different levels of difficulty, and using THESPIAN [21] to train PsychSim models of nonplayer character behavior separately for each level of difficulty. Also, the degree of complexity of the language increases steadily as the learner progresses through Mission Game scenes and Arcade Game levels.

2.5. Fish tanks and sandboxes

Gee [4] points out that good games often provide "fish tanks" (stripped down versions of the real game, where gameplay complexity is limited) and "sandboxes" (versions of the game that have similar gameplay to the real game, but where there is less likelihood for things to go wrong). These help users to develop their skills to the point where they can meet the challenges of the full game.

Fish tank and sandbox modes are both provided by the TLTS. An interactive tutorial lets learners practice operating the game controls, and utter their first words of Arabic (/marHaba/ or /as-salaamu 9aleykum/, depending upon the dialect being studied). The Novice mode described above provides sandbox capability. In addition, simplified interactive dialogs with friendly game characters are inserted into the Skill Builder lessons. This enables the learner to practice their conversational skills in a controlled setting.

Finally, sandbox scaffolding is provided in the Mission Game in the form of the virtual aide who can assist if the learner gets stuck. For reasons described above, we avoided having the aide interrupt with tutorial feedback that disrupts gameplay. The aide does not intervene unless the learner repeatedly fails to say something appropriate (this is often a microphone problem that has nothing to do with the learner's actual speech). In this case, or when the learner explicitly requests help, the Pedagogical Agent that drives the animated aide's behavior queries PsychSim for an appropriate user action, and then explains how to perform or say that action in Arabic. PsychSim maintains an agent model of normative user behavior for this purpose, alongside its models of nonplayer behavior.

2.6. Story and character identification

An important aspect of modern serious games is their use of story and character to maintain user interest, and to encourage the user to identify with the game character. Gee [4] has noted that it is not necessary to use virtual reality displays in order to immerse game players in a game. Gamers tend to identify with the protagonist character that they are playing, in a game such as Lara Croft. This is evidenced by the fact that nonplayer characters address either the player's character or the player directly, without seeming contradiction. Identification between player and character is reinforced in the TLTS by the fact that the player speaks on behalf of his character as he plays the game. Feedback from TLTS users suggest that this effect could be enhanced by allowing users to choose their character's uniform, and by adjusting mission and instructional content to match the learner's job, and we plan to provide such customizability in future work.

The TLTS makes extensive use of story structure; the game scenes fit within an overall narrative. This helps maintain learner interest. Also, it is our intention in the TLTS to make it so that actions earlier in the game can have effects on game play later in the game. If for example the player does a good job of developing rapport with characters in the game, those characters are more likely to assist the player later on in the mission. This will help reinforce gameplay to orient the learner toward future game objectives.

2.7. Fun and learning orientation

One of the most important characteristics of a good game, of course, is that it be fun [17]. The fun element helps maintain learner interest and positive attitude, and promotes intrinsic motivation. Evaluations of the TLTS suggest that fun plays an important role in promoting a learning orientation. Consider for example the following quote from a test subject:

> Had I spent more time with the Skill Builder… I probably would have been able to shoot through [the MPE] with relatively little problem. This is kind of fun, the Skill Builder is not that fun. (laughs)

We cannot hope to make the drill and practice exercises of the Skill Builder fun, but if the game component of the learning environment is fun, then learners will engage in the other learning activities in the environment, as resources that help them develop knowledge and skills that are relevant to the game. Squire and Jenkins [22] make a similar observation regarding the serious games that they have developed at MIT.

Thus, the fun element of the games in TLTS sets the stage for serious study and practice in the Skill Builder, provided that learners understand how that study and practice can help them improve their game skills. We can and do apply a wide range of intelligent tutoring and learner modeling techniques in this context. Each Skill Builder lesson includes a variety of different lesson and exercise types. Passive Dialogs show typical dialogs between Arabic-speaking game characters, in a context that is similar to the task context that they are training for. Vocabulary pages introduce words and phrases, and give the learner practice is saying them. A disfluency analyzer analyzes the learner's speech for common pronunciation errors, and then provides coaching on those errors. The feedback is intended to motivate and encourage the learner [10]. The vocabulary pages first show both English translations and Arabic transliterations for the target utterances; these are immediately followed by pages in which the transliterations are removed, in order to make sure that the learner is committing the new vocabulary to memory. Utterance formation exercises require learners to think of an Arabic phrase to say in a particular context, and give them feedback as to whether or not the phrase was appropriate. Active dialogs are similar to passive dialogs, but where the learner plays the role of one of the characters in the conversation. Finally, learners complete a quiz consisting of similar exercise pages, to show that they have mastered the material. They are encouraged to retry these quizzes on subsequent days until they demonstrate full mastery.

One type of AIED processing that we have not found to be of great importance yet is curriculum sequencing. We expect TLTS learners to be motivated to improve their game

skills. To assist them in this activity, we provide them with a Skill Map, which shows them what skills they need to master in order to complete each Mission Game scene, and where to find relevant lesson materials in the Skill Builder. We plan to test this new capability in upcoming evaluations, to assess whether this alone is sufficient to provide guidance. If not, we will augment the Skill Map with automated assessments of whether or not they have demonstrated master of each skill, and recommendations of lessons to study or review.

Another role that fun element plays in TLTS, particularly in the Arcade Game, is to provide learners with a pleasant diversion from their study. Learners comment that they like being able to take a break from study in the Skill Builder and play a few levels in the Arcade Game. Yet even when they are taking a break in this fashion, they are still practicing their language use. The opportunity for learners to change pace in this way enables learners to spend many hours per day using TLTS without much boredom or fatigue, something that very few intelligent tutoring systems can claim.

3. Game Development

The TLTS makes use of an existing game engine from Epic Games called the Unreal Engine. A game engine refers to the set of simulation codes that does not directly specify the game's behaviour (game logic) or the contents of the game's environment (level data), but is responsible for visual and acoustic rendering as well as basic interaction such as navigation and object collision. Such engines are increasingly being employed by researchers as affordable and powerful simulation platforms [15]. What makes this technology especially appealing is that games, when purchased off-the-shelf, often include, free of charge, all the authoring tools necessary to create new game logic and level data. Serious games can therefore be crafted from games originally intended for entertainment, avoiding initial game engine development costs.

In TLTS we take a step further by interfacing the Unreal Engine with our own Mission Engine (ME) [8] through the Game Bots interface (http://www.planetunreal.com/gamebots/). The ME and its attached modules handle all of our game logic, including interaction with AI and advanced interfaces such as the speech recognizer. The ME is written in Python, which is a powerful scripting language gaining ground in game development, and reads in data such as descriptions of Skill Builder lessons and game scenes in XML format. This combination of scripting and XML processing enables flexible and rapid development, such as when we added Tactical Iraqi to the existing Levantine Arabic content. Being so heavily data driven, it is essential that we have a good set of data authoring tools. To this end, we have concentrated a good deal of our effort on streamlining the content authoring pipeline and designing tools that are intuitive and effective in the hands of non-programmers. This is important because the game design should not rest on the shoulders of programmers alone, but be a group effort where story writers and artists help enforce proper game design principles.

6. Conclusions

This paper has examined the methods that modern serious games employ to promote engagement and learning, and discusses the role of AIED technology within the context of such games. Serious games can support learning in a wide range of learners, including those who have little initial motivation to study the subject matter. They embody a range of design principles that appear to promote learning, although further evaluative research needs to be done to understand their effects on learning. The serious game context makes the job of the AIED development in many ways easier, since the game design assumes some of the responsibility for promoting learning. AIED development effort can then be focused towards using AI to promote instructive gameplay, managing the level of challenge of the user

experience, providing scaffolding selectively where needed, and supporting learners in their efforts to reflect on their play and improve their skills.

Acknowledgments

This project is part of the DARWARS Training Superiority Program of the Defense Advanced Research Projects Agency. The authors wish to acknowledge the contributions of the members of the Tactical Language Team.

References

[1] Burton, R. R., & Brown, J. S. (1982). An investigation of computer coaching for informal learning activities (pp. 79-98). In D. Sleeman and J. S. Brown (Ed.), *Intelligent Tutoring Systems*. New York: Academic Press.
[2] Conati, C. and Maclaren, H. (2004). Evaluating a probabilistic model of student affect. *Proceedings of ITS'04*, 55-66. Berlin: Springer-Verlag.
[3] Csikszentmihalhi, M. (1990). Flow: The psychology of optimal experience. New York: Harper Perennial.
[4] Gee, James Paul. *What Video Games Have to Teach Us about Learning and Literacy*. New York: Palgrave Macmillan, 2003.
[5] Goldstein, I.P. (1982). The Genetic Graph: a Representation for the Evolution of Procedural Knowledge. In D. Sleeman and J.S. Brown, editors, Intelligent Tutoring Systems, pages 51-78. Academic Press, London, 1982.
[6] Hall, L., Woods, S., Sobral D., Paiva, A., Dautenhahn K., Wolke, D. (2004). Designing Empathic Agents: Adults vs. Kids. ITS '04. Berlin: Springer-Verlag.
[7] Hill, R., Douglas, J., Gordon, A., Pighin, F., & van Velsin, M. (2003). Guided conversations about leadership: Mentoring with movies and interactive characters. Proc. of IAAI.
[8] Johnson W.L., Beal, C., Fowles-Winkler, A., Narayanan, S., Papachristou, D., Marsella, S., Vilhjálmsson, H. (2004). Tactical Language Training System: An interim report. ITS '04. Berlin: Springer-Verlag.
[9] Johnson, W.L., Marsella, S., Mote, N., Viljhalmsson, H, Narayanan, S., Choi, S. (2004). Tactical Language Training System: Supporting the rapid acquisition of foreign language and cultural skills. InSTIL/ICALL Symposium, Venice, Italy.
[10] Johnson, W.L., Wu, S., & Nouhi, Y. (2004). Socially intelligent pronunciation feedback for second language learning. ITS '04 Workshop on Social and Emotional Intelligence in Learning Environments.
[11] Johnson, W.L., Beal, C. (2005). Iterative evaluation of an intelligent game for language learning. Proc. of AIED 2005. Amsterdam: IOS Press.
[12] Jones, S. (2003). "Let the Games Begin: Gaming Technology and Entertainment among College Students," Pew Internet & American Life Project, http://www.pewinternet.org/report_display.asp?r=93.
[13] Laird, J. And vanLent, M. (2000). The role of AI in computer game genres. http://ai.eecs.umich.edu/people/laird/papers/book-chapter.htm
[14] Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J. Harackiewicz (Eds.), Intrinsic and extrinsic motivation: The search for optimal motivation and performance (pp. 257-307). San Diego: Academic Press.
[15] Lewis, C., Jacobson, J. (2002) Game Engines for Scientific Research, Communications of the ACM, January 2002.
[16] Norman, D. A. (1990). The design of everyday things. New York: Doubleday.
[17] Prensky, M. (2001). Digital game-based learning. New York: McGraw Hill
[18] Prensky, M. (2002). The Motivation of Gameplay: or, the REAL 21st century learning revolution. On The Horizon, Volume 10 No 1.
[19] Schank, R. & Cleary, C. (1995). *Engines for education*. http://engines4ed.org/hypermedia
[20] Serious Games Initiative. http://www.seriousgames.org.
[21] Si, M. & Marsella, S. (2005). THESPIAN: An architecture for interactive pedagogical drama. Proc. Of AIED 2005. Amsterdam: IOS Press.
[22] Squire, K. & Jenkins, H. (in press). Harnessing the power of games in education. *Insight*.