Australian Government
Department of Defence
Defence Science and
Technology Organisation

DSTO

Lessons for the Construction of Military Simulators: A Comparison of Artificial Intelligence with Human-Controlled Actors

Jennifer Sandercock

DSTO-TR-1614

Australian Government
Department of Defence
Defence Science and
Technology Organisation

Lessons for the Construction of Military Simulators: A Comparison of Artificial Intelligence with Human-Controlled Actors

Jennifer Sandercock

Air Operations Division
Systems Sciences Laboratory

DSTO-TR-1614

ABSTRACT

Computer Generated Forces (CGF) are a significant component of military modelling and simulation. The creation of artificial agents with human-like decision-making is difficult. By using a Turing Test in a virtual environment that relates to military scenarios, areas in which CGF need to improve in order to help direct future studies are probed. The main areas determined for improvement are: environment awareness, human variance, persistence, vengeance, anticipation, learning and teaming. The CGF show consistent weakness in these areas across all virtual environments and should be considered in future studies. They present a technical challenge in adoption of CGF as surrogate players in computer supported simulation activity for training, mission rehersal, operations research and miliatary experimentation.

RELEASE LIMITATION

Approved for public release

Published by

DSTO Systems Sciences Laboratory
PO Box 1500
Edinburgh South Australia 5111 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

AR-013-185
September 2004

APPROVED FOR PUBLIC RELEASE

Lessons for the Construction of Military Simulators: A Comparison of Artificial Intelligence with Human-Controlled Actors

Executive Summary

Modelling and simulation form a core part of work undertaken in Air Operations Research branch within AOD. Computer Generated Forces (CGF) are widely used to act as opponents or teammates or perform other roles. Work is being done locally and internationally to improve intelligent agents and human-like decision-making. This study compared behaviour of human-controlled actors to those controlled by a computer.

Computer games have many parallels to military simulations. Using a cheap, off-the-shelf computer game (Quake III by *id Software*) provided an immersive environment to explore CGF with the additional benefit of allowing expert users to be obtained easily. Turing Tests were used to determine whether an experienced operator can distinguish between another human operator and a computer-controlled operator. Several versions of the Turing Test were used ranging from teamed and non-teamed versions, using an observer, allowing the human to act as a deceiver (pretend to be a computer-controlled actor), and a test involving only human actors.

The test results are tabulated to allow easy look-up of the differences between human-controlled and computer-controlled actors. Criteria for distinguishing actors were grouped into the following categories for ease of reporting: skill level, personality, persistence and vengeance, anticipation, learning, teaming and designed limitations and advantages. In many cases prior knowledge of the individual computer-controlled character and the human was required. The criteria that do not require the interrogator to know the individuals before the experiment are more widely applicable actor distinguishing criteria. These criteria were: environmental awareness, human variance, persistence, vengeance, anticipation, learning and teaming.

Many of the significant distinguishing features of computer-controlled actors are consistent across computer games and military simulations. Although the purpose of games is significantly different to military simulations, the computer-controlled actors encounter many of the same difficulties and apparent inexperience. The major areas for improvement to CGF are the same areas that require no prior knowledge of the actors.

This study identified areas in which current CGF technology needs to be improved to generate more human-like behaviour and decision-making. The possibility of using computer games as an adjunct to military simulations in future studies was demonstrated as viable given the success of this work.

Authors

Jennifer Sandercock
Air Operations Division

Jennifer Sandercock graduated from the University of Melbourne with a B.E./B.Sc in 2001. She majored in mechanical engineering, computer science and applied mathematics. Since joining DSTO in 2002, she has worked in AOD on military modelling for platform acquisition. Currently, she is studying intelligent agents for use in military simulations for operations research.

Contents

1. Introduction

Differences between humans and current computer generated actors in computer games and military simulations are explored in this report, that is: what is it about an opponent that allows you to determine whether the character is computer generated or a human-controlled actor (HCA)? By examining the differences we are better able to implement Computer Generated Forces, CGF.

Although military simulations and computer games have differences they also have much in common: both involve a combination of human-controlled and artificial actors in virtual environments. Differences between the areas include purpose. The main purpose of games is to entertain, whereas the purposes of military simulations can range across training, platform/system/human factors analysis, platform/system acquisition advice, and tactics and concepts of operations (CONOPs) evaluation and creation. The intended purpose changes the implementation [13]. For instance, visual appearance of a game is usually one of the most important aspects; therefore if an object is outside the player's field of view, then it is not modelled (both physics and visuals) to the same level of detail as an object that is within the player's view. In a military simulation it is important to obtain an adequate representation of the factors that are being studied; other factors can be modelled to a lower level of detail. This Level of Detail programming is consistent through time for the scenario and is not viewpoint dependent. Excepting the domain of training simulations, graphics are not necessarily important in military simulation. Sandercock, Papasimeon & Heinze [14] presented a more thorough discussion of differences and similarities of computer-controlled actors in games and military simulations.

The immersive three-dimensional worlds of computer games provide complex, exciting, stable, and cheap environments for research and development into the advanced capabilities required by computer generated forces [7]. Relative to military simulations, computer games are easily understandable and cheap to use. In order to study virtual environments human operators need to have expertise or familiarity with the environment. Compared to military models and simulations, it is much easier to acquire expertise in computer games and specifically first-person-shooter games, due to factors including good user interface. First-person shooters are tactical level scenarios that allow for easy comparison to military simulations. The lessons learnt about differences that exist between HCAs and agents in a computer game provide direct insight into what factors are most important for Computer Generated Forces (CGF) in military simulations.

Despite all the differences mentioned, games have been used to successfully study the real world, including:

- Human decision making (Microworlds [2][4])
- Environments for agent technology demonstration and development [8]
- Training environments for the development and retention of war fighting skills [3] [9]

The work presented here complements the technology development work in teamed agents [1] and command agents in LOD [13]. The work also informs technology requirements for distributed mission training and experimentation.

As a long-term goal effective CGF should be delivered to other defence simulation applications. Current differences between computer-controlled actors and humans in First Person Shooter (FPS) computer games are examined. This leads to insight into differences between CGF and humans. Similarities and differences of CGF to artificial actors in games are explored with a view of considering use of games in further studies and gaining a better understanding of areas for improvement of CGF.

2. Method

A modification of the Turing test has been used in this study to compare artificial and human-controlled actors (HCAs).

2.1 The Turing Test

In 1950, one of the founders of computer science, Alan Turing [18], boldly asked the question: "Can machines think?" Turing proposed the "imitation game", a game played by three people: a man, a woman and an interrogator. The interrogator is in a separate room and is allowed to ask questions via a computer terminal to determine who is the woman and who is the man. The man's purpose is to cause the interrogator to make the wrong identification. The woman's purpose is to be as truthful as possible. If a machine replaces the man, will the interrogator be able to tell the difference between the two players? Turing showed that the imitation game encapsulated the same problem whether it was played by three humans or by a human interrogator with a second human and a machine as deceiver. There is much discussion as to whether the Turing test is a true test of *intelligence*[1]; this will not be debated here.

As a standard, the Turing Test requires the machine to appear to be intelligent or *human-like,* not that it actually is intelligent [20]. The requirement for military simulations is that Computer Generated Forces (CGF) should be suitable for the study. For training operators, such as pilots, opponents in the real world are currently human, so the simulation requires the artificial actor to appear to be human. Considering the case of military simulations whose purpose is development and evaluation of new weapons and tactics: current real-world operators of friendly and opposing forces are human; therefore human-like behaviour of employment of weapons and tactics is appropriate. Since it is sufficient that CGF appear to be human, they should also be able to pass a variant of the Turing Test.

There are many ways that a Turing Test can be implemented. An example of a Turing Test to measure CGF is: take 5 active squadron fighter pilots, place each pilot in a simulator and do air combat against an enemy fighter or fighters. By observing their performance against CGF and human opponents it would be possible to form answers to the following questions:

- Could the pilots tell the difference between CGF and human opponents?
- How were pilots able to distinguish between human and CGF?

Another approach is to use computer games instead. Computer games are cheap, and due to ease of use experts are not as difficult to obtain. Computer-controlled actors in

---

[1] Dale Thomas claims that the Turing test is a naïve, biased and limited view of intelligence [15]. His criticisms do not apply to the experimental method that will be used here:

- Thomas asks why human intelligence is the upper bound - military simulations attempt to mimic human behaviour, so it is logical to test for human intelligence.
- Thomas also questions the need for mastery of the English language, the tests here are expertise in a virtual domain where action rather than language is important.

some game genres are called Bots and can exhibit complex behaviour. In computer games the overall goal of the game is entertainment, and a goal of Bots is to contribute to the suspension of disbelief. This is can only be defined within the context of a specific game [7] and is even more vague than the goal of CGF in military simulations, which is to provide "human-like behaviour". Only the illusion of human-like behaviour is needed in games and the effort is put into the illusion, not into the accuracy or competence of the underlying behaviours[7]. Nevertheless, Bots do exhibit similar behaviours to both humans and CGF, so by exploring behaviour in the context of a Turing Test it is possible to transfer overall deficiencies in Bots to CGF.

An example of a modern day Turing Test using computer games is: consider an virtual environment with some (from one to many) HCAs and Bots playing against each other in a first person shooter (FPS), such as Quake III (see Appendix A: for an introduction to FPS games). The following questions can be posed:

- Could the interrogators tell the difference between a HCA and a Bot?
- How did interrogators distinguish between HCA and Bot?

Other tests can easily be performed to answer additional questions:

- Does participation in the test bias results?
- If the human pretends to be a Bot and the Bot pretends to be a human, can the interrogator differentiate correctly?
- Can interrogators distinguish individuals in a group of HCAs?
- Do personality differences in the real world emerge in the virtual world?

A Turing Test using Quake III makes apparent deficiencies and strengths in Bots. Humans are considered to be more intelligent and therefore better at problem solving than Bots; Bots are considered to have better firing accuracy. The aim is to determine if it is these factors that allow an interrogator to distinguish between a Bot and a HCA. Some other factors investigated include personalities and fighting skill level. As a comparison humans vs. humans were examined to determine if human personalities emerge in the gameplay.

2.2 Experimental Method

Computer games have been used in prior work for the military, such as Operation Flashpoint [9] in Australia, and Full Spectrum Warrior [19] in the USA. Laird [5] and Norling [10][11] have used Quake II as a research test-bed. Quake has also been used as a vehicle for team competition in university leagues. We chose to use Quake III since it is an established game, we had experts in the gameplay and it is typical of first-person shooter games.

The original Quake game, from *id Software*, game was released in 1996 and was extremely popular, allowing for the follow up releases of Quake II (1997) and Quake III (2000). Using the Quake III game engine, there are four different types of multiplayer games that can be played, not all of which are used for the tests done here. Quake III was constructed as a game rather than a Turing Test vehicle. Therefore, the goal of the Bots is not to deceive human players into believing them to be a HCA; it is to create a *worthy* opponent. The Bot just needs to put up a good fight, and then lose convincingly, so that the player feels a sense of achievement [7].

In some cases the human played the role of deceiver, i.e. pretending to be a Bot, in order to draw out more complex criteria for distinguishing human controlled actors. Bots were chosen to match their "personality" (weapon choice and skill level) to the person being interrogated, see Figure 1 for the Bots used in the tests. Terrain used was a non-standard map in Quake III [21], eliminating any unfair advantages to players who already knew the map.

Figure 1: On-screen representation (avatar) of Bots used during testing (L-R): Crash, Doom, Gorre, Lucy, Hossman, Bones, Visor [23]. Note: either a human or a computer can control avatars.

See Table 1, below, for explanation of the type of tests carried out in order to explore the importance of human interaction, observers and teams. The gameplay heading groups those factors that were inherent in the setup of Quake III itself. The experimental artefacts are those factors that were imposed in order to extract the experimental data. An observer is a human not participating in the gameplay, but watching through the screen. In some test types the HCA took on the role of a deceiver, i.e. pretending to be a Bot. Four main versions of tests were considered: teamed, non-teamed, interrogator as an observer and human only. After every test each interrogator was questioned individually to determine their perceived dimensions of difference between Bot and HCA.

Table 1: Type of tests to explore facets of Bots compared to HCAs

Test Type	Aims & Questions	Gameplay				Experimental Artefacts	
		Actors			Teamed?	Interrogator	Human as deceiver?
		Computer controlled	Human controlled	Observer			
Teamed	• Aim: Examine effects of teams. • Q: If one human works out identity of other human, does their behaviour change?				✓		×
Non-teamed, double interrogator	• Aim: Examine if the Turing Test can be passed. • Q: Does behaviour change when both humans are interrogators and being interrogated?				×		×
Non-teamed, single interrogator	• Aim: Allow humans to have unique roles, either deceiver or interrogator. • Q: By concentrating on their role, is it easier to distinguish the HCA?				×		×
Non-teamed, human as deceiver	• Aim: Make the test closer to the original concept of the Turing Test – where the computer played the role of deceiver to trick the interrogator into thinking it was human. Quake III was not set up as a test-bed the Bots do not attempt to deceive. Allow human to take on the role of deceiver • Q: Will the interrogator need to look more closely for criteria to distinguish HCA?				×	OR	✓
Interrogator as observer	• Aim: Eliminate any bias connected with the interrogator participating in the game • Q: Does non-participation make distinguishing easier or harder? • Q: Does it make a difference that the human being interrogated does not know where the interrogator is, so their behaviour cannot change?				×		✓
Human only	• Aim to explore personality issues • Q: Do real-world personalities migrate to the virtual world?	NONE			×		N/A

3. Results and Analysis

See Table 2, below, for the individual tests carried out (for a further explanation of the type of tests examined cross-reference to Table 1).

Table 2: Tests completed in Quake III

		Gameplay				Experimental Artefacts		RESULT
Test No.	Test Type	Actors			Teams?	Interrogator	Human as deceiver?	Identified actors correctly?
		Computer controlled	Human controlled	Observer				
1	Teamed	💻💻💻 💻💻💻	🧍🧍		✓	🧍🧍	×	✓
2	Non-teamed	💻💻💻 💻💻💻	🧍🧍		×	🧍🧍	×	✓
3		💻💻💻 💻💻💻	🧍🧍		×	🧍🧍	✓	×
4		💻💻💻 💻💻💻	🧍🧍		×	🧍	✓	✓*
5		💻💻💻 💻💻💻	🧍🧍🧍		×	🧍🧍 #	✓	×
6		💻💻💻 💻💻💻	🧍🧍		×	🧍	✓	✓
7		💻💻💻 💻💻💻	🧍🧍		×	🧍	✓	✓
8	Interrogator as observer	💻💻💻 💻💻💻	🧍	👁	×	👁	✓	✓
9		💻💻💻 💻💻💻	🧍	👁	×	👁	✓	×
10	Human only	NONE	🧍🧍🧍 🧍🧍		×	🧍🧍🧍🧍 🧍	N/A	In some cases
								6/9 correct

Questioning of interrogators during preliminary studies found that the differences between Bots and HCAs were so significant that interrogators were easily able to detect the HCAs. These effects were due to individuals using drastically different perceived behaviours leading to an inference of considerable differences in skill level attributes between the Bots and HCAs.

- When HCA is *more* skilled than Bot, HCA will use:
  - Anticipation
  - Stealth
  - No warning shots
- When HCA is *less* skilled than Bot, HCA will not be able to use:
  - Strafing
  - Firing whilst running backwards
  - Jumping
  - Multi-tasking

* Interrogator stood still and observed reactions

Two interrogators deciding independently

These differences are important in themselves, however, in order to look at the finer criteria for distinction, a Bot avatar with a similar skill to the HCA should be chosen. When the HCA demonstrates known Bot characteristics (i.e. acts as the deceiver) and plays with or against Bots of similar fighting styles to the human, the interrogator will be forced to look into more depth to determine identity. These deceptive techniques were used in some of the tests; see Table 2 for details.

During the knowledge elicitation process from the experiments the reasons or criteria why interrogators thought that an actor was human or a Bot were noted. Each criterion formed a two rule statements with qualifiers, for example:

- A player is human-controlled if walks forward whilst firing - however some Bots also exhibit this behaviour
- A player is a Bot if runs backwards most of the time - however skilled humans will also demonstrate this behaviour

Table 3, on page 10, shows the criteria used by interrogators to distinguish between HCAs and Bots as determined in interviews with the interrogators. The presentation style of the table maps to the intuitive thought process reported as being used by many of the participants.

The author grouped these reasons or criteria for clarity into the following categories:

- Skill level
- Personality
- Persistence and vengeance
- Anticipation
- Learning
- Teaming
- Bot limitations and advantages

Most of the criteria used in Table 3 have a qualifying statement that causes the statement not to be true in certain circumstances. Those distinguishing criteria without qualifying statements are true across all scenarios and human or computer controlled actors in Quake III. In order to gain a better understanding of which criteria are the strongest indicators of HCA or Bot a statistical approach such as discriminant function analysis or logistic regression could be used. Table 3 allows us to consider questions like: if a HCA is highly skilled they will demonstrate a given set of behaviours. This could then be used as a comparison table to improve or decrease behaviours exhibited by the Bots in order to create a more enjoyable experience or a better training environment for the HCA. In Unreal Tournament, by Epic Mega Games, the Bots can be set so that they automatically adjust to player skill level.

3.1 Teamed Version

In the teamed version of the game the HCA was required to give orders to the Bots as well as participate in the competition. Bots did not always follow the orders and effort was required by the human trying to correct them, thereby distracting them from their interrogation or fighting roles. Often it was too difficult to coordinate with the Bots so the HCAs moved about on their own; whereas Bots moved about in groups, allowing the interrogator to easily determine identity.

3.2 Non-Teamed Versions

The non-teamed games with the interrogator playing in the game directly generated new criteria to distinguish humans. There were more Bots that could be the opposing HCA, and therefore more time and more criteria were required in order to be certain of identity. Interviews with the interrogators established the reasons for their classification of human controlled or Bot. Each interrogator used several methods for classification; these combine to form Table 3. When the HCA and the Bots were of similar skill level, identification was difficult. On the other hand when the HCA and the Bots were poorly matched, identification became easier.

3.3 Interrogator as an Observer Version

Results were similar when the interrogator was an observer and did not participate in the fighting. The interrogator was not required to have any prior expertise or knowledge of the gameplay, which meant that skill of the interrogator didn't affect the results. As an observer the interrogator can choose whom to follow on the screen, or wait in one spot, allowing repetitive mistakes of Bots to become evident. The Turing Test did not necessarily become easier for the interrogator however it eliminated any possible bias of the person being interrogated to alter their behaviour when known to be fighting a human. None of the data collected allows us to infer that it is or is not a bias to have the interrogator participate directly in the test. Therefore, if this question is considered important further studies need to be completed.

Table 3 Criteria used to distinguish human-controller actors from Bots in Quake III virtual environment

Type of Criteria	Player is *human-controlled* if	Qualifier	Player is *Bot* if	Qualifier
Skill Level	Easy/hard to kill	Depends on relative skill levels	Easy/hard to kill	Depends on relative skill levels
	Stands still whilst firing at enemy	Some Bots don't move much	Jumps or crouches whilst firing.	Skilled humans do this
	Walks forward whilst firing	Some Bots do this	Runs backwards most of the time	Skilled humans do this
	Doesn't strafe effectively	Only inexperienced humans do this	Uses strafing	Skilled humans do this
	Cannot multi-task	Only inexperienced humans do this	Can multi-task	Skilled humans do this
	Doesn't use weapons requiring high aim accuracy	Only inexperienced humans do this	Uses weapon requiring high aim accuracy	Depends on Bot weapon preferences
	Standing still in a corridor	Bots do not stand still for extended periods of time.	Continuously moving	Skilled humans do this
	High/low number of kills	Depends on relative skill levels	High/low number of kills	Depends on relative skill levels
	Accidentally kills themselves multiple times	Only inexperienced humans do this		
Personality	Combats from a distance & from behind	Depends on fighting style	Stays close during combat; rarely attacks from behind	Some humans do this
	Stands still in strategic positions, e.g. sniping	Depends on human's preferences	Rarely stays in one position for a long time	Bots do not snipe for extended periods of time
	Exhibits characteristics unlike avatar, eg skill level, weapon choice, jumping/crouching	Some players have similar characteristics to Bots	Exhibits characteristics consistent with Bot personality	Some humans have similar characteristics to Bots
	Follows paths that are "creative" and spontaneous	Get same effect with random or erratic Bot movement	Erratic in a predictable way (follows repetitive patterns/paths in terrain)	Some humans that know the terrain do this
	Reacts to same situation in different ways		Reacts to situations in the exact same way.	
	Avatar is of the same gender as the human	Some humans do not match avatar to their gender	Avatar is of opposite gender to human	Some humans do not match avatar to their gender
	Does an exciting activity	Some Bots do this	Doesn't do anything out of the ordinary	Some humans are repetitive
Persistence & Vengeance	Persists with chase until kill achieved		Gives up chase after a given distance	
	Will attack a stationary player until kill achieved		Runs past stationary player or gives up attack before achieved kill	
Anticipation	Anticipates opponent's behaviour: runs away from battle and waits until for other player to give chase		Can't anticipate opponent's behaviour: runs from battle to get more health/ ammunition, doesn't expect opponent to follow	
Learning	Watches others to learn sniping points to use and to check for		Unable to learn from watching others	
	Doesn't make same mistake in same location		Makes a mistake at the same location every time	
Teaming	Is leader (humans do not take orders from a Bot)	Bots can lead other Bots	Follows other Bots	Bots can lead other Bots

Type of Criteria	Player is *human-controlled* if	Qualifier	Player is *Bot* if	Qualifier
Bot Limitations & Advantages	Turns towards a dead-end in terrain	Initial effect with new terrain	Erratic in a predictable way (returns to points in room already traversed)	
	Has low/high aim accuracy	Depends on relative skill between avatar and human & pre-set Bot skill	Has low/high aim accuracy	Depends on skill level at which Bot has been set
	Heads towards PowerUps first, when entering a room		Does not go to PowerUp immediately	
	Runs forward to reach PowerUp	Some Bots will do this	Runs backwards to exact point where PowerUp is, then stops	Experienced humans with good knowledge of the terrain do this.
	Doesn't give warning before firing		Gives warning before firing (designed limitation)	
	Not able to use chat functions whilst fighting		Able to use chat functions easily[2]	Experienced humans with useful menus do this
	No rear-vision capability		Has rear-vision capability	Depends on skill level at which Bot has been set
	Keeps running when fired upon	Depends on current health/ammunition & experience at matching sound to location	Turns to attack in the right direction when fired at, even if from great distance	
	Rarely fires from *exactly* the same position	Unless sniping	Moves to the same waypoint to attack enemy	
	Difficult to find since explores entire terrain		Stays to the main areas of terrain	
	Doesn't use straight line down corridors		Always runs along the middle of corridors	

3.4 Human Only Version

This test was different from the others since it was not a Turing Test; all players were HCAs and knew that there were no Bots in the test. All participants knew each other in the real world, but they did not know the associated avatar names in the test. The five participants had varying knowledge of the other participants' game personality, varying from no knowledge to many hours of playing against each other. Those participants with greatest knowledge of game personalities were most easily able to distinguish individuals within the test.

This test with only HCAs allowed differences between individual humans and also from Bots to be explored. For instance, anticipation and learning were used more extensively when known to be combating another HCAs, this lead to these characteristics being included in Table 3. It was possible to determine differences between individual humans due to skill level (eg number of kills), weapon use, fighting style (eg jumping), and aggressiveness.

Aggressiveness is a behaviour that is also present in the real world, and so to some degree behaviour differences in the real world did extend to the virtual world, although aggressiveness in a game and in real life are not necessarily linked. In Quake III aggressive behaviour was perceived when an actor persisted with an attack or continuously attacked humans of known lower skill level. Non-aggressiveness could

[2] This feature in the Bots was turned off, in order to eliminate this effect.

be inferred when an actor ran away during conflict from humans of known lower skill level in order to allow them time to find more health. It may follow that aggressiveness is related to confidence and skill in the game: more skilled players are more certain to win a fight, and therefore more likely to enter fights, and therefore are perceived as aggressive. To examine the effect of aggressiveness across real and virtual worlds, a psychology test could be performed in the real world and then compared to virtual personality behaviours. In order to differentiate between humans at all, prior knowledge of the individual's personality and technique was required.

4. Discussion of Results

This study has revealed differences between Quake III Bots and HCAs. Interrogators were successfully able to determine the difference between Bots and HCAs in more than 50% of all cases (see result column in Table 2), including when the human was a deceiver pretending to be a Bot. Success at distinguishing identity means the Turing Test was failed. This is not an unexpected result. The real value of the work presented here lies in the identification of significant or interesting criteria for distinguishing HCAs from Bots. The criteria found can be grouped into: skill differences (which affected movement ability), personality, persistence to goals, vengeance, anticipation, learning, teaming, and design limitations or advantages of the Bots.

4.1 Skill & Personality Criteria

Skill levels between Bots and HCAs have significant impact on the ability to distinguish actors. When a Bot's skill level is increased, it is increased uniformly with a percentage, so that a Bot's other strengths and weaknesses are consistent, such as multi-tasking. Often the increased skill is firing more often and more accurately, faster movement, and 360° vision. In contrast, there is a much greater breadth for human skill levels, see

Figure 2. Inexperienced players are unable to exhibit complex movement (such as jumping and strafing), and use chat functions, which even the simplest Bot can implement. On the other hand experienced players can beat the most difficult Bots.

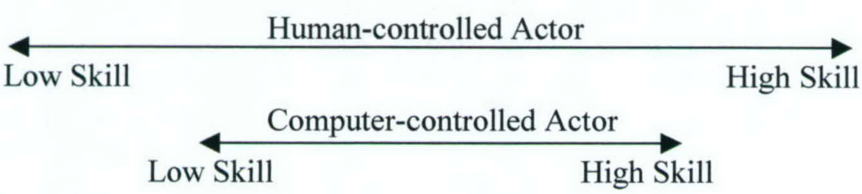

Figure 2: Breadth of skills – human- vs. computer-controlled

Many of the criteria in Table 3 require that characteristics (especially the skill level) of both actors are already known to make it easier to determine the HCA from the Bots. In a single-player game at a given difficulty level of Quake III the player encounters Bots that are increasingly difficult to defeat, thereby giving the computer-controlled actors a ranking compared to other Bots. Each computer-controlled Bot also has an associated weapon preference. When a human player of high skill uses an avatar with an associated low skill or vice versa, it is much easier to distinguish between human and computer. A human can be easily recognised when their weapon choices do not match the computer's choices, either by using only one weapon or a different set of weapons. If interrogators who did not know the avatar personalities were used in the tests, or if an entirely new set of avatars was created, then this situation would not exist. By using avatars with similar fighting styles to the HCA, the challenge is reduced, and more important factors can be examined.

Other personality criteria used by interrogators to determine the differences between actors (not always with the correct result) were:

- Creativity - Bots were seen to follow repetitive paths
- Matching gender to known HCA
- Ordinariness or Novelty- if the actor did nothing interesting
- Lack of individual variance - an individual reacted exactly the same in most situations

4.2 Persistence, Vengeance, Anticipation & Learning Criteria

When the interrogator does not know either human or Bot personalities, distinguishing criteria that are independent of qualifying statements must be used. Such criteria include persistence and vengeance. Compared to Bots, HCAs will chase after an opponent for a much longer amount of time and will generally persist at firing at an opponent until they are killed, even if the opponent is not reacting to the attack. In one test an interrogator stayed in one point on the terrain and watched the interactions of the opponents: Bots would run straight past the interrogator, or would start firing but would break off before a kill had been achieved. Using this criteria allowed the interrogator to easily identify the players.

It was unclear whether the interrogators and the HCAs changed their behaviour when their opponent was known to be human-controlled. Since humans generally like *gloating* over opponents, in some circumstances humans will be more likely to persist until a kill is reached or be vengeful if the opponent is known to be an HCA. More studies would need to be carried out in order to determine whether this was a significant factor in differences between humans and Bots. In the tests where the interrogator was an observer this behaviour was eliminated, since the HCA was always fighting Bots and was unable to tell when the interrogator was watching them and therefore could not change their behaviour when known to be interacting with a human.

HCAs showed anticipatory behaviours when playing against other humans. For instance, an actor would run away from a battle, and then lie in wait for their opponent to run after them - when playing against Bots, this tactic would not work since Bots will not give chase. Another tactic was to wait with weapon aimed at a PowerUp (a desirable item) since it was certain that an opponent would cross this point. Laird confirms that expert human players attempt to anticipate the actions of their opponents [8].

Quake III Bots are unable to learn from mistakes, whereas the HCAs learnt where sniping points were and looked before moving into the line of fire. Also some Bots would always make mistakes such as falling into a pit at the exact same location, and could never learn that the terrain was dangerous at that point.

4.3 Teaming Criteria

When examining game versions with teaming amongst Bots and HCAs it is clear that the Bots are much worse than a human in terms of skill levels:

- Bots do not follow orders well
- HCAs find it difficult to follow a Bot's seemingly erratic path through rooms
- Bots do not allocate tasks to team members effectively
- Other teamed behaviour (such as covering fire) is flawed

In order to examine this area further, Bots need to have significant improvement to make their teamed behaviour more human-like.

Work being done internationally to address these teamed challenges includes:

- The teamcore research group at the University of Southern California is exploring the agent-human interactions in teams [15][16]
- Air Operations Division [1] in DSTO
- Land Operations Division [13] in DSTO
- Work by Agent Oriented Software in the area of Simple Teams [5]

4.4 Design Limitations and Advantages of Bots

There are many design limitations and advantages that are inherent in the Bots. Some of the limitations are purposefully created in order to give the human a chance or a more challenging opponent. For example, Bots often give some warning (either sound, or a warning shot) before firing a shot that will kill the opponent.

Some limitations are due to natural limitations in the tight time schedules during development of a computer game and the nature of computer games played in real-time. For instance, Bots follow given waypoints to navigate through terrain, which means that:

- They always run down the centre of a corridor;
- Attack from the same point every time;
- Miss PowerUps and other players whilst running to another waypoint;
- Run backwards to the exact location of a PowerUp;
- Stop and stay within well-defined areas of large terrains

These behaviours lead to the Bots seeming to be "erratic in a predictable way". The reason that Bots exhibit this behaviour is due to the fact that the terrain is very complex for non-visual based entities, such as Bots. Therefore waypoints for each terrain are pre-programmed for the Bots in order to save time in development and time whilst the game is running. In some computer games, the Bots "watch" which paths the HCAs follow, and learn to take these new paths. So after playing for a long time the Bots improve by being able to go anywhere on the terrain.

In order to overcome some of these forced disadvantages and create a challenging opponent, Bots in Quake III have been given 360° vision and almost 100% firing accuracy in the difficult settings, such as "Hardcore" and "Nightmare!"[3] For instance, Bots are able to turn 180° around and shoot in the correct direction instantly, whilst a HCA will turn, visually locate target and then aim.

Laird did some informal studies altering the Bots' decision speed, tactics, aggressiveness, and aiming skill and determining if this changed the player perceived humanness of the Bot. Laird found that bots with extremely accurate aiming or extremely fast (<25 msec) decision speed appeared less human than ones with less accurate aiming skills and slower (100 msec) decision speed [8]. These results confirm the findings from this study, whereby Bots with superior skill (in terms of aim accuracy and speed) were more easily distinguished from the HCAs by the interrogator.

[3] These traits are not consistent across all FPS games.

5. Discussion of Implications

5.1 Differences between Human-Controlled Actors and Bots

When the interrogator knows information about the human-controller, the Bot personalities, and the skill level that the Bots are set to prior to the test, then the easiest way to distinguish Bots is to examine skill levels and personality. However, in a "true" Turing Test the individuals do not know each other prior to the test and these methods cannot be used.

Designed Bot limitations and advantages, such as firing accuracy and waypoint following are easy to notice if you know to look for these factors. If they were reduced then more complex criteria would need to be used. The Human Agent Virtual Environment (HAVE) project [12] aims to eliminate pre-scripted waypoints by labelling the environment with tags that state the name of the object and action possibilities. This will allow for a more "level playing field" between artificial actors and human actors.

In order to be absolutely certain that the decision between HCA or Bot is correct, one needs to look at the criteria that have no qualifiers: persistence, vengeance, learning, anticipation and teaming. Although not all humans are able to exhibit these behaviours fully, Bots in Quake III were unable to demonstrate any of these behaviours at all. As the AI for Bots improves and Bots begin to exhibit these behaviours the interrogator will need to examine the complex interactions between these criteria and not a single criterion by itself.

5.2 Differences between CGF and Bots

Each computer game and military simulation has different AI (Bots and CGF), and therefore some criteria will not be consistent across the full range of virtual environments. Quake III is only one of many first-person-shooters, each of which has their own unique combination of designed Bot limitations and advantages and also predominant fighting style. For instance, Bots in Unreal Tournament have a much more cautious fighting style: they shoot, hide behind an object and then come out to shoot again. Therefore, just as all humans have their own fighting style, different games have their own predominant fighting style for all Bots. Each different type of military simulation would have a different perspective as well. In order to gain a better understanding of which distinguishing characteristics are consistent across all virtual environments, it would be important to do more extensive tests and analysis across the breadth of possible worlds.

Across the HCAs in all tests there was a wide variety in fighting styles. This leads to the belief that it is not valid to implement a uniform, average human as a CGF when a group of individual humans are being simulated. Humans are not the same and therefore CGF should not be the same. Within the computer game industry non-player characters (NPC) are often programmed with 5–10% randomness in order to simulate differences between humans. In military simulations for Operations Research (OR), repeatability is an essential part of every study, that is each run must be able to be re-run with the same starting conditions and achieve the same result in order to be able to perform validation. Consequently, it is not possible to include randomness in this way; Monte Carlo runs are used instead, where many runs are implemented with a known pseudo-randomness, generally across initial conditions.

Wray and Laird [22] have examined human variability in military simulations both across subjects (different people react to same situation in a different way), and within subjects (the same person will take different actions at different times in effectively the same situation). They have developed computational models of variability that attempt to show "realistic *individual* behaviour that must be sustained over the course of the scenario". Preliminary validation of their models has used automated comparisons of human and computer generated behaviour. This work could lead to introduction of more individual Bot personalities and therefore eliminate some of the sameness that allows Bots to be easily distinguishable from HCAs.

Validation is an important part of developing a good military model or simulation. The Turing Test is a form of validation to compare AI with human behaviours. If the artificial reasoning is required to demonstrate human behaviour then it must be compared against human reasoning. For instance, when examining output from a simulation involving CGF, as humans we look at the actions or decisions to determine whether a human would have made the same choices. "Dumb" decisions are easier to notice than "smart" decisions made by either humans or CGF. Humans often do not choose the most logical solution or the one they have been trained to take. When the CGF chooses an unexpected path, it is important to determine how the decision was made to verify the behaviour.

Quake III is a game whose main purpose is to entertain and therefore it creates enjoyable opponents, not be an objectively correct simulation. By using Quake III as a test-bed, it has been possible to determine major differences between Bots and HCAs, which in turn gives insight into CGF in military simulations. Insight can be gained because experienced actors were examined in a combat situation, which is similar to many of the military simulations that are examined using CGF. So criteria found here are applicable to CGF as well, except for designed limitations and advantages.

5.3 Differences between CGF

There are many different forms of agents of computer-controlled actors. These forms are dependent on underpinning architecture. The Bots in Quake III are implemented using Finite State Machines. In Air Operations Research branch in Air Operations Division of DSTO, we examine heavyweight Beliefs-Desires-Intentions (BDI) agents. Although these implementations are dissimilar, the areas for improvement found in this study are the same for both agent architectures.

Solutions to teaming and human variance problems could lie in the realm of sophisticated BDI agents. However, Norling [10][11] found that the Quake engine was an unsuitable environment to plug-in these technologies due to inherent difficulties associated with the interface to external servers.

6. Lessons Learnt for Military Simulations

HCAs show persistence to a goal, and react to stationary objects more than their Bot counterparts. Anticipation was not obvious in all of the tests except the test with only HCAs. This may not have been the case if more experienced Quake III players were used or if each test had taken more time. The test involving only HCAs brought up the differences to Bots and to other HCAs in the areas of anticipation and learning. An interesting control would be to do a test where there were only HCAs for the interrogator to identify, or the converse where there were only Bots.

The results of this study contribute to experimental design for planned studies and the greater understanding of CGF required for agent studies in AOD. One such study is the Human-Agent Virtual Environment (HAVE) project [12]. The HAVE project is creating an environment in which both humans and computer-generated agents can interact on an even playing field: agents will have access to environmental information in a format that is easier for computers to interpret than the polygons that the human eye can see.

The study of Quake III found that the major areas in which Bots are still not "human-like" are: environment awareness (the use of pre-programmed waypoints), human behavioural variance (across different individuals and within each individual), persistence, vengeance, anticipation, learning and teaming. Although the purpose of Quake III is not the same as a military simulation, CGF are deficient in these areas as well. These areas provide a possible next step for future agent research. Other areas for future advancement include deception and recognition of intention.

Future studies should explore differences between military simulations and computer games with a view of using Bots instead of CGF where appropriate. Computer games need to be monitored for significant developments in Bots and other AI characters that will be useful to military simulations. This study has offered a way forward for future military simulation studies involving CGF.

7. Acknowledgements

Thank-you to all the people involved in the tests:

- Julia Chadwick
- Martin Cross
- David Graham
- Clint Heinze
- Michael Papasimeon
- Tim Patterson

Thanks also go to the reviewers of this report for their advice and patience.

8. References

[1] Appla, D., et al., "Modelling and Simulation of Teams for AI Operations Research", DSTO Technical Report, in publication

[2] Brehmer, B., Dörner, D., "Experiments with Computer-Simulated Microworlds: Escaping Both the Narrow Straits of the Laboratory and the Deep Blue Sea of the Field Study", *Computers in Human Behavior*, Vol. 9, pp 171-184, 1993

[3] Capps, M., McDowell, P., Zyda, M., "A Future for Entertainment-Defense Research Collaboration", IEEE Computer Graphics and Applications, January/February 2001

[4] Friman, H., Brehmer, B., "Using microworlds to study Intuitive Battle Dynamics: A concept for the future", in Proceedings of Command & Control & Research & Technology Symposium, US Naval War College, Rhode Island, USA, July 1999

[5] Hodgson, A., Rönnquist, R., Busetta, P., "Specification of Coordinated Agent Behavior (The SimpleTeam Approach)", IJCAI '99 Workshop on Team Modelling and Plan Recognition, Stockholm, Sweden, August 1999

[6] Laird, J.E., "It Knows What You're Going To Do: Adding Anticipation to a Quakebot", AAAI 2000 Spring Symposium on Artificial Intelligence and Interactive Entertainment, March 2000

[7] Laird, J.E., "An Exploration into Computer Games and Computer Generated Forces", The 8th Conference on Computer Generated Forces & Behavior Representation, Orlando, Florida, USA, May 2000.

[8] Laird, J.E., "Research in Human-level AI Using Computer Games", published in Communications of the ACM in January 2002. Also at: http://ai.eecs.umich.edu/people/laird/papers/CACM01.pdf

[9] Morrison, P., Barlow, M, "Child's Play? Coercing a COTS Game into a Military Experimentation Tool", Proceedings of SimTecT 2004 Conference, Canberra, Australia, May 2004

[10] Norling, E., Sonenberg, L., "Creating Interactive Characters with BDI Agents", in *Australian Workshop on Interactive Entertainment,* Sydney, Australia, February 2004

[11] Norling, E., "Capturing the Quake Player: Using a BDI Agent to Model Human Behaviour", in Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems, ed: Rosenschein, J., Sandholm, T., Woolridge, M., Yokoo, M., Melbourne, Australia, July 2003

[12] Papasimeon, M, HAVE Project, In preparation

[13] Rönnquist, R., Howden, A., Lucas, A., Vaughan, J., Connell, R., Goss, S., "Coordinated Team Behaviour: A Simple Teams Approach", in 2nd International Synthetic Environments Symposium, RMC Shrivenham, UK, October 1999

[14] Sandercock, J, Papasimeon, M., Heinze, C., "An Agent, a Bot and a CGF Walk Into a Bar...", Proceedings of SimTecT 2004 Conference, Canberra, Australia, May 2004

[15] Schurr, N., Scerri, P., and Tambe, M., "Coordination Advice: A Preliminary Investigation of Human Advice to Multiagent Teams", AAAI Spring Symposium, 2004

[16] Schurr, N., Okamoto, S., Maheswaran, R.T., Scerri, P., Tambe, M., "Evolution of a Teamwork Model", Book Chapter in *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Cambridge University Press, 2004

[17] Thomas, D., "New Paradigms in Artificial Intelligence", in *AI Game Programming Wisdom II*, Ed., Rabin, S., Charles River Media, USA, 2004

[18] Turing, A.M., "Computing and Machine Intelligence", *Mind*, 59:433-460, 1950

[19] Van Lent, M., Fisher, W., Mancuso, M., " An Explainable Artificial Intelligence System for Small-unit Tactical Behaviour", in Proceedings of 16th Conference on Innovative Applications of Artificial Intelligence, San Jose, USA, July 2004, In Publication

[20] Van Waveren, J.P., "The Quake III Arena Bot Thesis", University of Technology Delft, Faculty ITS, 2001

[21] Williams, P., "Kiss the Sky" Quake III map, http://www.planetquake.com/pjw

[22] Wray, R.E., Laird, J.E., "Variability in Human Behaviour Modelling for Military Simulations", Behaviour Representation in Modelling & Simulation Conference (BRIMS), May 2003

[23] http://www.planetquake.com/quake3/q3aguide/bots_a.shtml

[24] http://www.planetquake.com/pjw

Appendix A: Quake III Terminology

8.1 Quake III Overview

First Person Shooter (FPS) games are those where the player (you) is in the first-person viewpoint, as through you are actually there yourself, see Figure 3. The main action takes place in a kill or be killed battlefield. Quake III (*id Software*) was released in 2000 and is a typical FPS game.

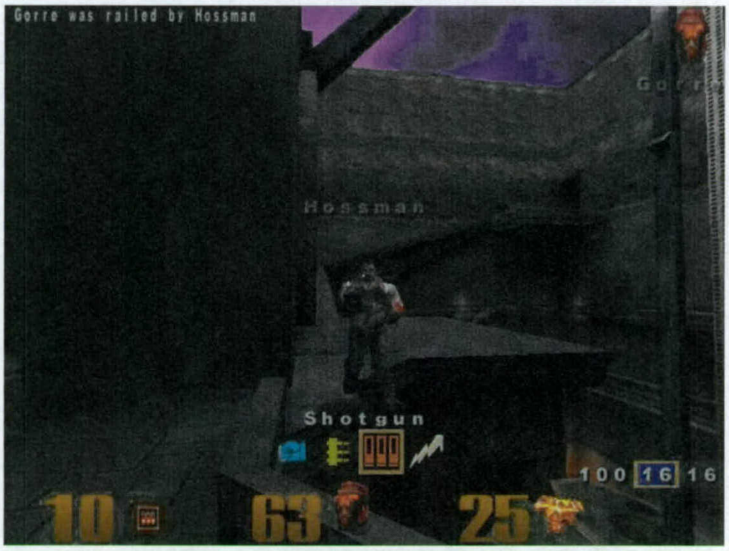

Figure 3: Weapon Choice & Viewpoint

You are in control of an avatar (a visual representation of your character) in the terrain, these avatars look the same whether a computer is controlling them or not. The computer-controlled avatars are called Bots. At the bottom of the screen (Figure 3) is the amount of ammunition for the selected weapon, your health (out of 100), and armour. Health decreases when opponents successfully hit you with their weapon fire. When your health reaches zero, you die, then you are regenerated or re-spawned as new at another point in the terrain. Killing someone is termed "fragging" in Quake III. The aim in Quake III is to frag the most people, and a score table is kept in order to keep track of rank, see Figure 4. The symbols to the far left-hand side of the score table represent HCA or not (this feature was hidden for the experiments).

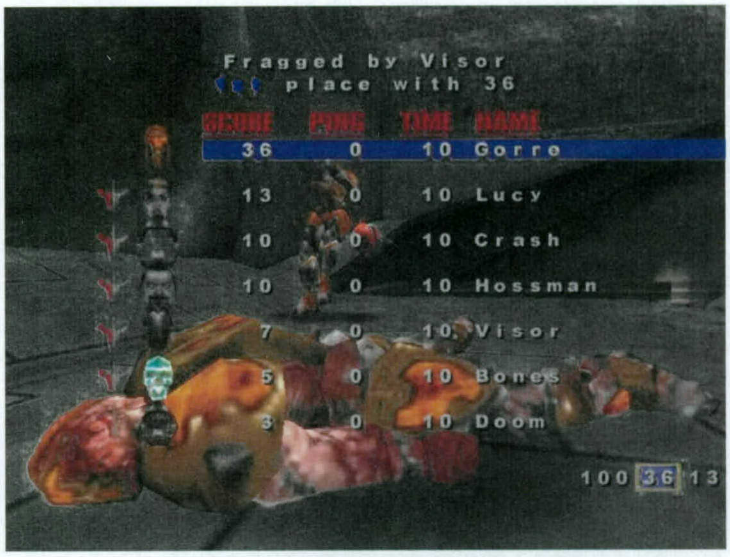

Figure 4: Score Table

8.2 Terrain

Gameplay takes place on terrain called "maps" in Quake III. Figure 5 shows part of the terrain used in the experiments, "Kiss the Sky", created by Pat Williams [24]. Features that can be in maps include: water, voids, teleporters, slime, platforms, pendulums, lifts, lava, gates, fog, doors and bounce pads. Note in Figure 5 that you can take on the role of a spectator or observer, who does not interact in the game, but merely watches.

Figure 5: The terrain used in the experiments, Kiss the Sky, showing multiple levels

8.3 PowerUps

Throughout the terrain are "PowerUps", which increase your health, ammunition or armour, see Figure 6. There are also special PowerUps such as Haste, Quad Damage (Figure 7), which give you an extra advantage over other opponents for a fixed period of time. When you begin you have a gauntlet and a machine gun, other weapons can be picked up by running over their symbol in the terrain. There is a diverse range of weapons available (Figure 8) each with set damage amount and area, reload time and range limitations, so that weapons with short range afflict large damage over a large area when close and long-range weapons afflict high damage over a small area.

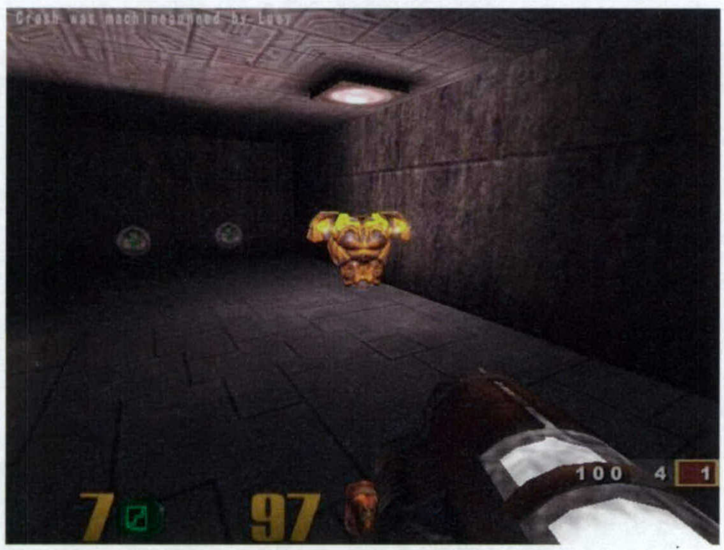

Figure 6: Health PowerUps and Armour

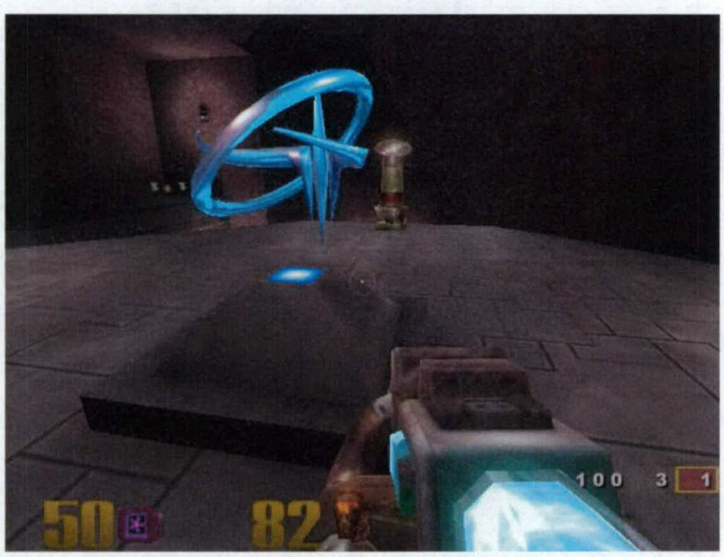

Figure 7: Quad Damage – when picked up gives you 4 times the amount of damage when using any weapon for a fixed amount of time

Figure 8: Assorted weapons available in Quake III with their associated ammunition symbol

8.4 Game Versions

There are four different game types in standard Quake III:

1. Free for All: This is the standard version as described above, where every player fights for the most points by themselves.
2. Team Death Match: Similar to a Free for All, however players are grouped into two teams (red and blue) and points gathered go to the group total. In teamed games humans can act as leaders and give orders to Bots. Orders are issued using a chat (typing) interface.
3. Tournament: A series of one-on-one matches where the winner plays the next person.
4. Capture the Flag: Two teams, each have their own base with a flag at the centre (Figure 9). The aim is to work as a team to capture the enemy's flag and bring it back to your base, whilst protecting your own flag.

Figure 9: The Blue Flag in a Capture the Flag game

8.5 Advanced Player Skills

As a human player practices Quake III they become more proficient at gameplay. Advanced players use strafing: moving from side to side whilst firing at a stationary point. This is one of the first techniques that players will learn and is an indication of skill.

8.6 Further Information

For more information on Quake III look at:

- http://www.idsoftware.com
- http://www.planetquake3.net

DISTRIBUTION LIST

Lessons for the Construction of Military Simulators: A Comparison of Artificial Intelligence with Human-Controlled Actors

Jennifer Sandercock

AUSTRALIA

DEFENCE ORGANISATION

	No. of copies
S&T Program	
Chief Defence Scientist FAS Science Policy AS Science Corporate Management Director General Science Policy Development	shared copy
Counsellor Defence Science, London	Doc Data Sheet
Counsellor Defence Science, Washington	Doc Data Sheet
Scientific Adviser to MRDC, Thailand	Doc Data Sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	Doc Data Sht & Dist List
Scientific Adviser - Army	Doc Data Sht & Dist List
Air Force Scientific Adviser	1
Scientific Adviser to the DMO M&A	Doc Data Sht & Dist List
Scientific Adviser to the DMO ELL	Doc Data Sht & Dist List
Systems Sciences Laboratory	
Chief of Air Operations Division	Doc Data Sht & Dist List
Research Leader Air Operations Research	Doc Data Sht & Dist List
Head Operations Research Capability	1
Head Air Operational Performance Assessment	1
Task Manager	1
Author: Jennifer Sandercock	5
Clint Heinze	1
Michael Papasimeon	1
Julia Chadwick	1
Hal Harvey	1
David Cox	1
Chris Best	1
Peter Ryan	1
Lucien Zalcman	1
Information Sciences Laboratory	
Research Leader MIE	1
Greg Calbert	1
DSTO Library and Archives	
Library Fishermans Bend	Doc Data Sheet
Library Edinburgh	1
Defence Archives	1
Capability Systems Division	
Director General Maritime Development	Doc Data Sheet

Office of the Chief Information Officer

Deputy CIO	Doc Data Sheet
Director General Information Policy and Plans	Doc Data Sheet
AS Information Structures and Futures	Doc Data Sheet
AS Information Architecture and Management	Doc Data Sheet
Director General Australian Defence Simulation Office	1
Robert Carpenter	1

Strategy Group

Director General Military Strategy	Doc Data Sheet
Director General Preparedness	Doc Data Sheet

HQAST

SO (Science) (ASJIC)	Doc Data Sheet

Navy

SO (SCIENCE), COMAUSNAVSURFGRP, NSW Doc Data Sht & Dist List

Director General Navy Capability, Performance and Plans, Navy Headquarters Doc Data Sheet

Director General Navy Strategic Policy and Futures, Navy Headquarters Doc Data Sheet

Air Force

SO (Science) - Headquarters Air Combat Group, RAAF Base, Williamtown NSW 2314 Doc Data Sht & Exec Summ

Army

ABCA National Standardisation Officer, Land Warfare Development Sector, Puckapunyal e-mailed Doc Data Sheet

SO (Science), Deployable Joint Force Headquarters (DJFHQ) (L), Enoggera QLD Doc Data Sheet

SO (Science) - Land Headquarters (LHQ), Victoria Barracks NSW Doc Data & Exec Summ

Intelligence Program

DGSTA Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1 (PDF)
Assistant Secretary Corporate, Defence Imagery and Geospatial Organisation	Doc Data Sheet

Defence Materiel Organisation

Deputy CEO	Doc Data Sheet
Head Aerospace Systems Division	Doc Data Sheet
Head Maritime Systems Division	Doc Data Sheet
Chief Joint Logistics Command	Doc Data Sheet
Head Materiel Finance	Doc Data Sheet

Defence Libraries

Library Manager, DLS-Canberra	Doc Data Sheet

Library Manager, DLS - Sydney West	Doc Data Sheet

OTHER ORGANISATIONS

National Library of Australia	1
NASA (Canberra)	1

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy	
Library	1
Head of Aerospace and Mechanical Engineering	1
Serials Section (M list), Deakin University Library, Geelong, VIC	1
Hargrave Library, Monash University	Doc Data Sheet
Librarian, Flinders University	1

OUTSIDE AUSTRALIA

INTERNATIONAL DEFENCE INFORMATION CENTRES

US Defense Technical Information Center	2
UK Dstl Knowledge Services	2
Canada Defence Research Directorate R&D Knowledge & Information Management (DRDKIM)	1
NZ Defence Information Centre	1

ABSTRACTING AND INFORMATION ORGANISATIONS

Engineering Societies Library, US	1
Materials Information, Cambridge Scientific Abstracts, US	1
Documents Librarian, The Centre for Research Libraries, US	1

SPARES	5
Total number of copies: Printed 46 + PDF 1 =	**47**

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
Lessons for the Construction of Military Simulators: A Comparison of Artificial Intelligence with Human-Controlled Actors	Document (U) Title (U) Abstract (U)
4. AUTHOR(S)	5. CORPORATE AUTHOR
Jennifer Sandercock	Systems Sciences Laboratory PO Box 1500 Edinburgh South Australia 5111 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-1614	AR-013-185	Technical Report	September 2004

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
2004/1015570	LRR 03/226	DSTO	26	24

13. URL on the World Wide Web	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1614.pdf	Chief, Air Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DEFTEST DESCRIPTORS

artificial intelligence; computer simulation; intelligent agents (computer software); modelling; military training; virtual reality; human behavior; human-machine systems;

19. ABSTRACT

Computer Generated Forces (CGF) are a significant component of military modelling and simulation. The creation of artificial agents with human-like decision-making is difficult. By using a Turing Test in a virtual environment that relates to military scenarios, areas in which CGF need to improve in order to help direct future studies are probed. The main areas determined for improvement are: environment awareness, human variance, persistence, vengeance, anticipation, learning and teaming. The CGF show consistent weakness in these areas across all virtual environments and should be considered in future studies. They present a technical challenge in adoption of CGF as surrogate players in computer supported simulation activity for training, mission rehersal, operations research and miliatary experimentation.

Page classification: UNCLASSIFIED