Constructing the Virtual Section

Michael Barlow[¥]; Peter Morrison[¥]; Mathew Luck[¥]; Alistair Dickie [±]

(¥) Virtual Environments & Simulation Laboratory (VESL)
School of ITEE
UNSW @ ADFA
{spike,p.morrison,m.luck}@adfa.edu.au

(±) Land Warfare Development Centre
Australian Army
alistair.dickie@defence.gov.au

Abstract. This paper details the conduct and results arising from an experiment known as VISE (Virtual Infantry Section Experiment) performed at The Australian Army's Headline03. VISE employed a 1st Person Simulator (1PS) known as VBS (Virtual Battlefield System) to examine alternate section structures – 8-member, 9-member, and 12-member. Three classes of section combat scenarios were constructed within VBS – a (quick) attack; a deliberate attack; and the section in defence. A platoon of (28) soldiers from 5/7 RAR participated in the study – playing both the BLUEFOR (section) and OPFOR (enemy) in a number of repetitions of each scenario and section structure combination. Contrast of the 3 section structures on the basis of casualty MOE (e.g., Loss Exchange Ratio) found, on aggregate, that a larger section performed better than a smaller. When efficiency-based performance criteria were examined the 9-member section was found to be more efficient (quicker to complete and less rounds fired), with 8-member being next best and 12-member the least efficient of the three. Subjective evaluation resulting from survey of the soldiers indicated that all three section structures were adjudged equally capable of fulfilling section tasks. The larger section sizes were considered to possess more firepower but the 12-member section was seen as more unwieldy and less adaptable.

1. INTRODUCTION

This report briefly details the conduct and results arising from an experiment known as VISE (Virtual Infantry Section Experiment) performed at The Australian Army's Headline03. VISE employed a 1st Person Simulator (1PS) known as VBS (Virtual Battlefield System) to examine alternate section structures – 8-member, 9-member, and 12-member.

VISE was achieved through the efforts of Defence – the Land Warfare Development Centre (LWDC), industry – Bohemia Interactive Australia (BIA), and academia – the Virtual Environments & Simulation Laboratory (VESL) UNSW @ ADFA. LWDC defined the task and provided resources and participants. BIA provided a version of VBS configured for the experiment. VESL provided experimental design, implementation, conduct, and analysis.

As specified by the Force Development Group of LWDC, the goal of the Virtual Infantry Section Experiment was "*… to determine the most appropriate section structure for undertaking a number of standard section combat tasks.*" This is against a wider background of the Objective Force [1], & on-going reviews of the structure of the Interim Objective Force 2012, DSTO's current study of platoon and company Infantry compositions; and with the tacit understanding that VISE was an initial and limited study with as much emphasis on the viability of an experimental methodology that employed a 1PS - decisions about section sizes and composition have a flow-on effect that can impact Army at all levels.

COTS (Commercial Off The Shelf) games, in particular 1PS, have received considerable recent attention as to their potential as military tools [4],[2],[7]. 1PS appear to offer the prospect of greater fidelity in modelling close combat, in a number of environments ranging from urban through littoral and densely vegetated, than other forms of simulation. This is attributable to a number of factors that include rich 3D environments, sophisticated physics and world models; combined with human-in-the-loop technology such that each participant controls a single individual (soldier) in the simulation – introducing the complexity and richness of human behaviour and decision making. The use of COTS games for military purposes also poses risks, as their drive is for entertainment rather than simulation. [5] contains a more detailed discussion of the issues in use of a 1PS – both the risks and approaches necessary to optimise utility.

VISE was conducted as follows. The alternate section structures – 8-person, 9-person, and 12-person were designed and implemented in VBS. Then a set of military tasks – attack & defence – that a section might undertake were implemented as scenarios within VBS. Measures of Effectiveness were designed and a scalable data capture mechanism implemented in VBS to ensure

In SimTecT 2004 Conference Proceedings, pages 100-105, Canberra, May 2004

the MOE were secured. Members of an infantry platoon – 5/7 RAR – were then employed to 'play' the OPFOR (section) and OPFOR in the scenarios. Each combination of scenario and section structure was run a number of times to acquire a range of outcomes. This occurred across the duration of Headline03. At the end of this period the participants were surveyed and the scenario runs analysed.

2. METHODOLOGY

VISE followed a methodology of specification (LWDC of section structures and combat missions), design, implementation (in VBS and external software), experimentation, and analysis.

The LWDC stipulated three different section structures, differing in their size (8, 9, & 12) and equipment. The 9-person section corresponded to that currently used by the Australian Army – three groups (command, assault, and support) of 3, with the assault and support groups each possessing a machinegun (F89). The 8-person section was split into two fire teams of 4 each. One fire team possessed two machineguns (an F89 and a GPMG), while the other possessed one (F89). The 12-person section was divided into three fire teams of 4 members each. One fire team possessed two machineguns (F89 and a GPMG), while the other two fire teams each had a F89. Other soldiers in each of the sections were equipped with F88s and 2 fragmentation grenades. There was also a scattering of grenade launchers (2 for 8 and 9 person sections; 3 for 12 person section) and light anti-tank weapons (2 for 8 and 9 person sections; 3 for 12 person section).

LWDC specified three classes of scenarios on which the three section structures were to be evaluated. In each case the basic task for the section was specified together with the size and composition of the opposed force. They were a *quick attack* against an enemy listening post (OP) of 3; a deliberate attack (DA) with fire support (8 members) against an enemy of 5; and a section in defence (SD) against an enemy of 10.

A single Deliberate Attack (DA) scenario was constructed, consisting of a farm compound in a large clearing (approximately 100 metres on a side) surrounded on all sides by sparse jungle vegetation (many palm trees but relatively sparse ground cover). Two Outpost (OP) scenarios were constructed – one in medium vegetation (MV), and one in denser tropical vegetation (DV). Both were constructed with the OP on a raised feature within several hundred metres of a road or track that the OP was likely observing. The Section Defence (SD) scenario was situated on a raised hill with good command of all immediate (out to 200 metres in most cases) approaches, though lightly forested further out, and hence providing cover for the manoeuvre of the attacking opposed force. As for the OP scenarios, defensive emplacements were constructed for the section through a combination of natural cover such as rocks and fallen trees, as well as camouflaged sandbagged positions.

Five Measures of Effectiveness (MOE) were designed in order to quantify the performance of the sections for each run. Three were casualty oriented – Loss Exchange Ratio (number of enemy killed for each section member killed), Casualty Exchange Ratio (number of enemy killed or wounded for each section member killed or wounded), and Damage Exchange Ratio (amount of damage done to enemy for each point of damage received). Two were efficiency oriented – Mean Effective Rounds Per Enemy Death (number of rounds fired per enemy killed) and Effective Duration (length of engagement).

To support the acquisition of these MOE a robust and scalable client-server data capture facility was implemented in VBS [5]. This facility captured shot, hit, movement, and facing data for all world objects (all participants).

A number of ordnance and equipment changes were required in order to support the experiment – 4x scopes, Steyr with GLA (grenade launcher), the addition of larger and more vegetated terrain, fire rates and dispersion for weapons to match their real-world counter-parts and realistic simulations of (damage and suppression from) explosive effects. These were all implemented [5].

A platoon of 28 soldiers from 5/7 RAR participated in the study by playing (controlling) the soldiers in each scenario – both friendly (section and any fire support) and enemy. In all 108 successful experimental runs of scenarios (3 sections x 4 scenarios x 9 repetitions) were made across the 2-week period of Headline03, with number of participants in a run varying from 11 (8 person section against 3 enemy in OP) up to 25 (12 person section plus 8 member fire support against enemy of 5 in DA).

In order to ensure a core competency with the VBS tool before experiments started, all participants were given training with the tool. Training consisted of approximately 8 hours (a day and part of the following morning) familiarisation and testing ranging from individual skills to work as a section. Training included basic weapon and movement skills, navigation, advanced weapons, and working together as a team. Before experimentation commenced all participants were 'qualified' by completing firing and movement tests to a minimum standard [5].

At the end of the experiment participants completed a survey that included questions to ascertain their opinion of the three different section structures by a number of criteria. The criteria were: capability to complete a mission;

flexibility and adaptability; firepower; situational awareness of own section members; and situational awareness of enemy. Participants rated each section structure for each criteria with a value from 1 (woefully inadequate, impossible, etc.) through to 10 (perfect, trivially easy, etc.).

3. RESULTS

The mean (across all 9 runs of each combination) Loss Exchange Ratio for each scenario and section structure combination is found in Table 1. Figure 1 is a bar-plot in which the three section structures are directly compared by averaging the LER across the four different scenarios modelled. Figure 2 is a plot in which the results for different scenario types are highlighted, by each one being represented by its own line.

Table 1: Mean LER (Loss Exchange Ratio) figures for the different scenario and section structure combinations. Each value in a cell is the average across 9 runs of that particular scenario and section structure combination.

	8 person	9 person	12 person
OP (MV)	1.80	1.17	2.06
OP (DV)	0.90	1.59	2.45
SD	0.85	0.97	1.54
DA	1.54	2.25	1.36
mean	1.27	1.49	1.85

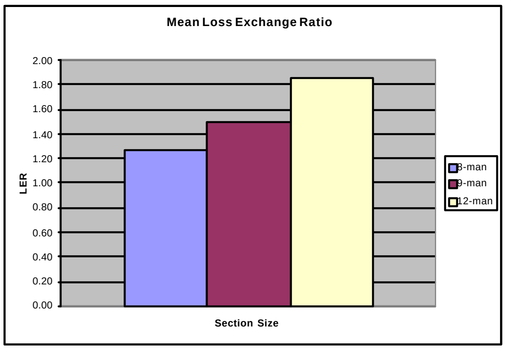

Figure 1: Plot of mean LER (Loss Exchange Ratio) for each of the three section structures - averaged across all (nine) runs of all (four) scenarios.

Figure 1 shows what appears to be a clear trend of increasing LER with increasing section size – indicating a more effective section as it increases in size. However the averaging across scenarios, and across runs for scenarios serves to mask a high degree of variability. Figure 2 gives some indication of that variability where it is seen that the relationship found in figure 1 (LER 12-member > LER 9-member > LER 8-member) does not hold true for all four scenarios.

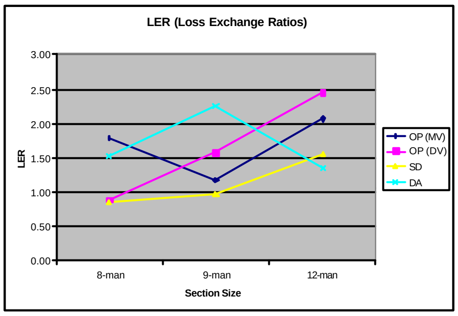

Figure 2: Plot of mean LER (Loss Exchange Ratio) for each section structure and scenario combination. The results for each scenario are drawn as a connected line (from 8-member to 9-member, to 12-member).

A 2-factor ANOVA (Analysis of Variance) test failed to show statistical significance for the differences between the different section structures. This was due to the extremely high variability between runs of single scenario and section structure combinations. As an (admittedly extreme – this was one of the 2 aberrant cases) example the deaths suffered by BLUEFOR in the 9 runs of the 12-member DA scenario were as follows: {12, 1, 0, 2, 0, 6, 9, 1, 2}. That is, losses ranged from 12 down to zero. Against such within category variability – statistical significance for differences in effect due to different section structures cannot be upheld.

Table 2 shows the mean and standard deviation of the losses suffered by BLUEFOR as a function of both scenario and section size. The table illustrates the high degree of variability (the std. dev.) found in certain scenario and section structure combinations. Several, but not all of the high variability cases appear to have explanations in the experimental configuration. For instance 9-person OP(MV) was the first set of experiments run and based on observation of participants it appeared that not all were fully 'at home' with the 1PS and operating as a virtual section (e.g., little use was made of suppressive fire or explosive ordnance in several of the runs). Similarly, a number of 'friendly fire' incidences marked the 12-member section runs – paralleling the subjective evaluation results that maintaining situational awareness of friendlies was difficult in the 12-member section (particularly when further complicated by an additional fire support team – the DA scenario).

Table 2: Mean and standard deviation (in brackets) of losses by BLUEFOR for the various scenario and section structure combinations. Certain combinations illustrate an extremely high variability in outcomes between individual runs.

	8-person	9-person	12-person
OP (MV)	1.7 (1.6)	2.6 (3.1)	1.5 (1.6)
OP (DV)	3.2 (2.9)	1.9 (1.8)	1.2 (1.1)
SD	7.2 (1.4)	6.9 (2.0)	5.8 (3.7)
DA	3.1 (2.3)	2.2 (1.6)	3.7 (4.3)

Tables 3 and 4 show the mean CER (Casualty Exchange Ratio) and DER (Damage Exchange Ratios) on section structure and scenario basis. It is worth noting the extremely high correlation between the LER, CER, and DER values. While perhaps not surprising it indicates that any one of the three is sufficient for capturing casualty based effectiveness.

The mean (across all 9 runs of each combination) Effective Rounds Per Enemy Death for each scenario and section structure combination is found in Table 5. Figure 3 is a bar-plot in which the three section structures are directly compared by averaging the ERPED across the four different scenarios modelled.

Table 3: Mean CER (Casualty Exchange Ratio) figures for the different scenario and section structure combinations. Each value in a cell is the average across 9 runs of that particular scenario and section structure combination.

	8 person	9 person	12 person
OP (MV)	1.29	1.17	1.65
OP (DV)	0.87	1.29	1.80
SD	0.96	0.98	1.42
DA	1.10	1.80	1.10
mean	1.05	1.31	1.49

Figure 3 shows a relationship where the 9-member section appears more *efficient* on the criteria of rounds fired per enemy killed (225) as opposed to 270 rounds for the 8-member section and 366 rounds for the 12-member. Table 5 shows this relationship holds for three of the four scenario cases – the one exception being SD, in which the 12-member section (on average) fired less rounds per enemy killed than the 8-member section.

Table 4: Mean DER (Damage Exchange Ratio) figures for the different scenario and section structure combinations. Each value in a cell is the average across 9 runs of that particular scenario and section structure combination.

	8 person	9 person	12 person
OP (MV)	1.47	1.17	1.80
OP (DV)	0.88	1.46	2.07
SD	0.89	0.97	1.49
DA	1.24	2.04	1.19
mean	1.12	1.41	1.64

The mean (across all 9 runs of each combination) Effective Duration for each scenario and section structure combination is found in Table 6. While the table reveals an average difference between the section structures such that the 9-person is the quickest to complete missions – around 10% faster than the 12-member – the difference is not statistically significant due to high degrees of variability. The fact that the participants were most familiar with the doctrine for a 9-person section may also explain any differences.

Table 5: Mean ERPED (Effective Rounds Per Enemy Death) figures for the different scenario and section structure combinations. Each value in a cell is the average across 9 runs of that particular scenario and section structure combination.

	8 person	9 person	12 person
OP (MV)	337	305	368
OP (DV)	275	211	538
SD	136	63	96
DA	334	320	463
mean	270	225	366

Table 7 summarises subjective evaluation of the three different section structures, on a number of criteria, by providing the mean and standard deviation for each. Several results are clear and noteworthy from the survey results. It is perhaps worth reiterating that these opinions were expressed after the soldiers had had two weeks of usage to familiarise themselves with the capabilities and composition of each of the three section structures.

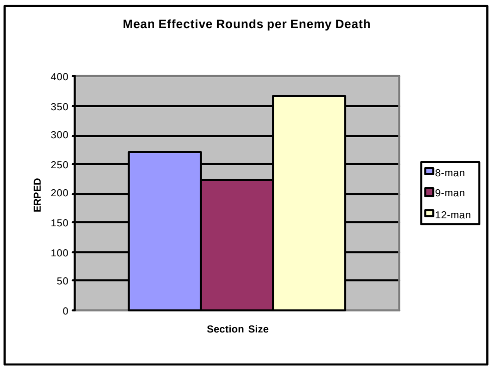

Figure 3: Plot of MERPED – Mean Effective Rounds Per Enemy Death – as averaged across all runs on the basis of section structure (8, 9 and 12-person).

Firstly, there is no significant difference in the soldier's rating of the 3 sections as to their capability to complete a mission. Scores fell within 0.2 of each other, with relatively similar standard deviation also; indicating effectively the exact same rating for all 3 sections.

In the area of flexibility and adaptability of the section during an encounter there is a significant difference (5% confidence interval). That is that the 12-member section was seen as less flexible and adaptable than the other two (for which no appreciable difference in rating exists).

Table 6: Mean ED (Effective Duration – in seconds) figures for the different scenario and section structure combinations. Each value in a cell is the average across 9 runs of that particular scenario and section structure combination.

	8 person	9 person	12 person
OP (MV)	273	316	180
OP (DV)	392	300	271
SD	409	355	585
DA	415	444	531
mean	372	353	392

In the area of firepower the 12-member section was, not surprisingly, rated as far higher than the other two section structures. Indeed, one participant wrote, in the area for comments, that *"…[the 12-man section]…perhaps has too much firepower."* The 9-member section was rated as possessing significantly more firepower (1% confidence interval) than the 8-member; despite the 8-member possessing an additional machinegun. It appears that in the participants' minds, the additional member counted for more.

Situational awareness of the disposition of own section members was one area in which ratings were tied (inversely) to section size – the larger the section the lower the rating. The differences were statistically significant with perhaps the most interesting result being just how low the 12-member section was rated (a mean of 5.5); perhaps indicating a loss of perception of the section as a single entity. This tends to support and parallel some of the results observed in the objective measures where some friendly fire incidents occurred for the 12-member section. Ratings of situational awareness of enemy disposition were not significantly different between the three section structures.

Table 7: Mean (and standard deviation – bracketed) participant ratings of the three section structures on a number of criteria. Low scores indicate inadequacy, with high scores indicating ease (e.g., situational awareness) or competence.

	8 person	9 person	12 person
Ability to complete mission	7.9 (1.6)	7.9 (1.2)	7.7 (1.6)
Flexibility & Adaptability	8.2 (1.7)	7.9 (1.3)	7.0 (2.0)
Firepower	7.1 (1.7)	7.8 (1.2)	9.7 (0.5)
Situation awareness – friendlies	8.4 (1.2)	7.6 (1.1)	5.5 (1.6)
Situation awareness - enemy	7.5 (1.4)	7.3 (1.4)	6.7 (2.4)

4. DISCUSSION

Beyond the investigation of alternate section structures, the experiment described here has entailed the establishment of a methodology for the use of 1st Person Simulator for military experimentation. That methodology has incorporated a design, implementation, experimental conduct, data collection & analysis tailored about the unique features and properties of a 1PS. The experiment has, to our knowledge, been a first for the ADF and the LWDC.

The choice of a 1st Person Simulator for military purposes appears to be an avenue worthy of considerable and serious further investigation. This claim is founded both on the positive reaction use of VBS received not only from the experimental participants and numerous visitors to the experiment during Headline03, but because of the solid data collected in this first experiment and the potential new forms of investigation and application that such a tool opens up.

Hence there is cause for cautious optimism in the pursuit and further application of 1PS. Caution, because a key ingredient necessary for confidence in the results arising from the utilisation of tools such as VBS is currently weak or lacking. That is *verification* and *validation* of the core algorithms and models that underlie the simulation. While the current set of results obtained often match intuition, one cause for optimism, a more rigorous analysis of the simulator(s) itself can only strengthen results obtained with such systems. Beyond a 'simple' analysis of the core simulation engine features (for instance the way damage is assigned, movement speeds, models of round penetration) a deeper analysis of the way that disparity between real-world and simulated environment (e.g., the fact that no foliage moves in the wind makes detection of enemy motion easier, or the fact that all users shoot as well as they are able to place a mouse cursor on a location on the screen, or that there is no 'fear of death' amongst users caused by a round passing close by, etc.) influences outcome and the success of doctrine appears a very worthy pursuit. Verification and validation, or at least a deeper understanding of the impact of simulator model versus reality; should be given high priority (see [3] and [5] for current work in this area).

Casualty based MOE (LER, CER, DER) could not show a statistically significant difference between the three section structures. However a trend of increasing effectiveness with increasing section size (e.g., higher LER for 12-member section as opposed to 8-member section) was evident for the data (75% likelihood that the differences were not random). Lack of statistical significance is, in part, a result of the large variability in outcomes obtained for section structure and scenario combinations. Such marked variability perhaps points to the inherent non-linearities of the simulation, but regardless means that a prohibitively large number of runs of each scenario and section structure combination would be required to "prove" statistical significance. On the other hand it may be that other factors, such as doctrine or simulation fidelity in certain areas, are more important in influencing the outcome than the size and compositions of sections that were examined. There is some support for this hypothesis in the fact that the participant section leaders were unfamiliar with any doctrine for how to use a 12-member section. For instance, the GPMG was often not used appropriately to lay down a beaten-zone, but was often seen being used by the assault team.

One of the efficiency based MOE (rounds fired per enemy death) did show a significant difference between the section structures – with the 12-member section being considerably less efficient (more rounds fired) than either of the other two structures.

Subjective measures – participant opinions – of the three section structures indicated that all three structures were considered equally capable of completing missions. While the superior firepower of the 12-member section was recognised it was considered significantly less flexible and adaptable than the smaller sections; and maintaining situational awareness of section members was found to be particularly difficult.

Taken together the results of the experiments indicate weak but ultimately inconclusive evidence supporting the hypothesis that the larger (12-member) section is more effective in a casualty sense. However, the larger section is less efficient in terms of rounds fired to achieve outcome and participants rate the larger section as more unwieldy and less adaptable, as well as far more difficult for section members to maintain awareness of the disposition of other section members.

Space limitations preclude a more detailed analysis here of the data arising from VISE. Those wishing further data should consult [5] and [6]. Certainly the results obtained to-date suggest great potential for further and more refined experimentation employing the current methodology – exploration of other factors such as doctrine, or simulation fidelity that may have influenced the outcome; the decoupling of section size (8,9, and 12-person) from equipment differences (numbers of machineguns, LAWs etc.).

More generally, work is on-going in VESL (e.g., [3]) to further explore the military potential (commensurate risks & method of best employment) of COTS games and in particular 1PS.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the valuable contributions of Bohemia Interactive Studio (BIA), The Australian Army's Land Warfare Development Centre (LWDC), and the soldiers of 5/7 RAR who participated in the study.

REFERENCES

1. - (2003) "The Australian Army 2020 – the journey so far", *Australian Defence* Magazine, Vol. 11, Issue 10, October 2003, http://www.yaffa.com.au/defence/current/10-news1.htm. Viewed 5/2/2004.
2. Barlow M.; Morrison P. & Easton A. (2002) "1st Person Tactical Shooters: COTS Games with Military Training Potential", *Proceedings of SimTect2002*, Melbourne, May 2002.
3. Barlow M. , Luck M., Lewis E., Ford M., Cox R. (2004) "Factors in Team Performance in a Virtual Squad Environment", *in proceedings.*
4. Macedonea M. (2002), "Games Soldiers Play", *IEEE Spectrum OnLine*, 28 Feb 02
5. Morrison P. Barlow M. (2004) "Child's Play? Coercing a COTS Game into a Military Experimentation Tool", *in Proceedings*.
6. VESL (2003) *Report on the Headline03 Virtual Infantry Section Experiment*, submitted to Force Development Group, LWDC, November 2003.
7. Zyda M., Hiles J., Mayberry A., Wardynski C., Capps M., Osborn B., Shilling R., Robaszewski M., Davis M. (2003) "The MOVES Institute's Army Game Project: Entertainment R&D for Defence", *IEEE Computer Graphics and Applications*, Jan/Feb.