UNIVERSITY OF
BIRMINGHAM

AN INVESTIGATION OF SENSORY AND BEHAVIOURAL FIDELITY IN GAMING TECHNOLOGIES TO SUPPORT ENHANCED PERCEPTION OF INTENT OF INSURGENTS

By

Mark E W Blyth

A thesis submitted to The University of Birmingham for the degree of

MASTER OF PHILOSOPHY

Department of Electronic, Electrical & Computing Engineering
School of Engineering and Physical Sciences
The University of Birmingham
May 2011

Abstract

The thesis was originally sponsored by HS-C (QinetiQ) to conduct research towards understanding insurgents ‗perception of intent‘ using different levels of fidelity. The phrase ‗perception of intent‘, while well-known in military circles, becomes difficult to define in a manner that allows to be tested experimentally.

The thesis therefore places focus on the ways in which observers (naïve and expert) explain what is happening in a scenario through detection of context-relevant features which have been designed into the scenario. The scenarios which are hypothesised and reflect no know real-world events are created using different fidelities achieved in NetLogo and VBS2. This is supported and complemented by understanding from SMEs knowledge, organisational visits, literature and the measurements of experiments. The analysis of experiment results focused on the three areas fidelity, provision of information and experts and included both statistical and text analyses. The findings thorough different measurement either returned a significant effects or non-effect creating mixture of results. Therefore fidelity, provision of information and experts have an observers perception but to exactly point which one area was too difficult suggesting that further research would be beneficial.

Dedication

The thesis is dedicated to all that have supported me, especially my wonderful loving wife Cristina for the many cups of tea.

Acknowledgement

This is to thank my Professor Chris Baber for his support, advice and spending several bank holidays reviewing my thesis; Professor Bob Stone, who worked endlessly to get the support of the MoD; Dr. James Knight for his expert guidance towards researching and understand statistical analysis; all my colleagues and fellow researcher in the HITS department.

CONTENTS PAGE

GLOSSARY

General Acronym

AA	America's Army
AI	Artificial Intelligence
ANA	Afghan National Army
ANP	Afghan National Police
BIS	Bohemia Interactive Simulations
CCTV	Closed Circuit Television
CSV	Comma Separated Values
CTC	Counter Terrorism Centre
DERA	Defence Evaluation Research Agency
DSTL	Defence Science and Technology Laboratory
EECE	Electronic, Electrical, and Computing Engineering
ESIM	Elaborated Social Identity Model
FOB	Forward Observation Base
MET	Metropolitan Police
MoD	Ministry of Defence
MOUT	Military Operations on Urbanized Terrain
MPhil	Master of Philosophy
NGO	Non-Government Organisation
N.I	Northern Ireland
HITS	Human Interface Technologies
HS-C	Haldane-Spearman Consortium
HSE	Health and Safety Executive
IED	Improvised Explosive Device
IET	Institute of Engineering
IDE	Integrated Development Environment
IRA	Irish Republican Army
ISAF	International Security Assistance Force
LN	Local National
OPTAG	Operational Training Advisory Group
PhD	Doctors of Philosophy
REME	Royal Electrical Mechanical Engineers
SDK	Software Development Kit
SIDE	Social Identity Model of Deindividuation Effects
SIED	Suicide Improvised Explosive Device
SMEs	Subject Matter Experts
UAV	Unmanned Aerial Vehicle
UoB	University of Birmingham
UDK	Unreal Developer Kit
U.K	United Kingdom
U.S	United States
VBS2	Virtual BattleSpace 2
VIED	Vehicle Improvised Explosive Device
WIZ	Weapon Investigation Report
WTO	World trade Organisation

LIST OF FIGURES AND TABLES

List of Figures

List of Tables

CHAPTER 1 INTRODUCTION

1.1 Preface

The University of Birmingham was awarded funding by the Haldane-Spearman Consortium (HS-C) to conduct research proposed by the University of Birmingham Electronic, Electrical, and Computing Engineering (EECE) Department. The proposal was to conduct research to address Task 9c entitled –*An Investigation of Sensory and Behavioural Fidelity in Gaming Technologies to Support Enhanced Perception of Intent of Insurgents"* for QinetiQ HS-C on behalf of the Defence Science and Technology Laboratory (DSTL) who are part of the Ministry of Defence (MoD)

HS-C was developed by QinetiQ and Quintec to bring together scientists, engineers, academics and consultants into a single group for Human Science that would create a new level of research capability. QinetiQ was formed after the government privatised DERA (Defence Evaluation Research Agency). QinetiQ is known for its domain knowledge expertise in the areas of defence, security and aerospace, which allows QinetiQ to work closely with the UK and US governments.

During the research period it became apparent that, due to the sensitive nature and an already over stretched British Armed Forces, it would be difficult for the Author to identify, contact and speak to sufficient subject matter experts (SMEs) to address the original research proposal. The original research proposal was written for the academic award of Doctor of Philosophy (PhD). However this was changed to a Master of Philosophy (MPhil) by request of the Author.

1.2 Background and Motivation

At the time when the proposal was written by the University of Birmingham's EECE department, the British Armed Forces was in need for research to be conducted at the University level to look at new means of training future or existing Military personnel who were preparing to go serve in current areas of conflict i.e. Iraq and Afghanistan. With growing cases of insurgent attacks the need to add new or existing features into the current training was becoming increasing important.

It is also important to provide some background on the Author as this has influenced the direction as well as the emphasis of technical aspects of this research paper. Prior to being accepted by the University of Birmingham for the research, the Author came from a varied industry background. The Author holds a Bachelor of Science degree, which he conducted part-time whilst working for Northrop Grumman Information Systems as component support engineer on the Airwave Police project. The degree was in Software Design and Networking; and the final year thesis was written on serious games design. The Author also served nine years in the Royal Navy as an Operator Mechanic and Above Water Warfare engineer.

The Author's main drivers for taking on this research was to bring the two passions of the Author together: designing of virtual environments, using knowledge gain from the Author's first thesis, and a passion as an ex-servicemen to give back any knowledge gained from the research to help in any way towards better training those who serve in an effort to address some of the challenges currently faced by the British Armed Forces.

1.3 Research Objectives and Overall Approach

The original research title emphasised ‗perception of intent'. This phrase, while well-known in military circles, becomes difficult to define in a manner that allows to be tested experimentally. The reason that the phrase is problematic is because it is not clear that ‗intent' of other agents can be directly perceived so much as inferred from the combination of a host of factors. This process of drawing inference from a combination of features implies a significant role for expertise and experience. Given some of the problems of access that has bedevilled this research, it was felt that the focus could initially be placed on the ways in which observers could explain what is happening in a scenario through detection of context-relevant features which have been designed into the scenario.

The objectives of this research are to better understand on how fidelity of training media impacts an observer's detection and more specifically to attempt to answer the following questions:

1. Can observers detect what is happening in a virtual environment?
   - 1a. Does the provision of a briefing help?
   - 1b. Does the type of briefing matter?

2. Does the level of fidelity in the virtual environment have an effect on how observers perceive?
   - 2a. Do responses become more accurate?
   - 2b. Will differences be sufficient to demonstrate that fidelity does make a difference?

3. Is there a difference between a naïve observer and one who has had experiences in observation, or intelligence gathering skills (expert)?

The Author ensured that the research considers both the technical and human aspects by taking the research proposal and segmenting it into two definable segments as illustrated in the flow diagram below (see Fig. 1.1).

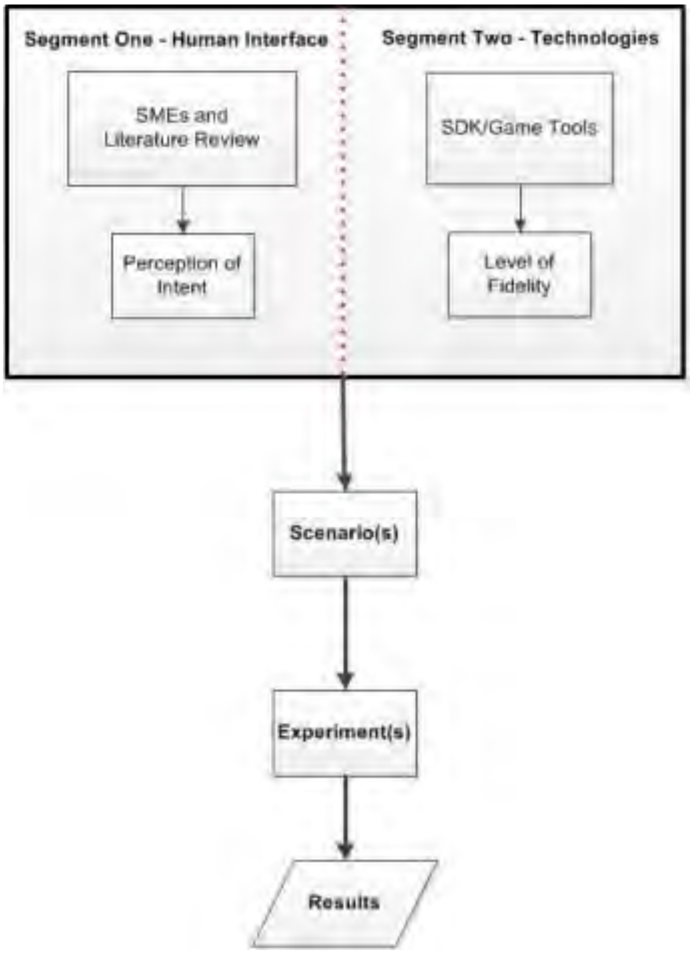

Fig. 1.1 Research Proposal Flow Diagram

Segment one focuses on the human interface aspects and segment two investigates the technical area of the research. The human interface segment looked at behaviour psychology and included both a literature review and discussions with subject matter experts (SMEs). Segment two, investigated current game tools and technologies, also known as software development kits (SDKs), as well as looking into whether these tools and technologies are currently used within the MoD.

Information gathered helped implement a virtual environment test bed to evaluate how well these behaviours can be implemented given the current technology and to go one stage further by introducing one, or a small number, of insurgents. Using inputs from SMEs (individuals with some crowd handling/ observation experience in the Middle East) the virtual environment could also be used to script a small number of scenarios to replace videos, either where they do not exist or where they are used to test aspects of the evolving intent taxonomy.

The virtual environment was defined as a computer-generated scenario, which is trying to replicate a real world scenario. The scenarios in this research were hypothesised and as such do not make reference to or depict any event that may have occurred in the real world.

In order to create appropriate levels of fidelity, the research considered a range of games development applications or other forms of simulation of human behaviour. The additional challenge was not merely the representation of human behaviour, but also the manipulation of the environment to provide sufficient variation for the observer to detect. The scenarios consist of both the developed simulations and supplementary textual material for briefing the observer.

These scenarios were used to test observers (experiments). The experiments results were analysed in the hope they provide additional insights and a better understanding of how information fidelity and impacts on detection.

1.3.1 Research questions defined

Each of the research questions draws upon information gathered in each segment of research: technologies, human interfaces and the results of the experiments based on scenarios created in the virtual environment. As such it is important to provide additional detail on the focus of each question to better understand how they are interrelated.

Starting with the first research question:

- Can observers perceive what is happening in a virtual environment?

The focus of this first question is on whether observers are able to detect what is happening when shown a scenario in a virtual environment. From the main question two sub-questions are asked: does the provision of a brief assist the observation and does it change depending on the type of brief?

It hypothesised that the observer would require some form of briefing or intelligence that would guide their search or that they would use previous experience to decide what features should be attended to.

This leads into the second question, which is:

- Does the level of fidelity in the virtual environment have an effect on how observers perceive?

Research conducted by (Stone, 2008) suggests that there are two types of fidelity: *"Physical or engineering fidelity"* and *"Psychological fidelity".* "Physical" fidelity refers to the simulated image content and quality as opposed to their real-world counterparts. "Psychological" fidelity is defined by how well the simulation transfers information and or skills. Purely looking at ―psychological” fidelity, both Fig. 1.2 and 1.3 are the same as they both show a helicopter. If the type of helicopter is required to be conveyed then there is a difference in the level of fidelity between both images.

This definition is similar to that found in research for the US military: (Alexander et al., 2005) suggest that *“physical fidelity is the degree to which the physical simulation looks, sounds, and feels like”* and that the that the *“psychological fidelity is degree in which the simulation replicates psychological factors experienced in the real-world environment”.*

The thesis refers to fidelity as the ―physical” fidelity when comparing between simulators used in the experiments, unless otherwise stated, and is determined by the format in which the virtual environment has been developed. In this thesis it was either a 2 dimensional (2D) or a 3 dimensional (3D) format. The formats are derived from the game tools selected, a process that is discussed in greater detail Chapter 3.

Fig. 1.2 2-Dimension Helicopter

[redacted])

[redacted]

Fig. 1.3 3-Dimension Helicopter

[redacted]

Therefore the sub-questions to question two focuses on whether observers give a different response when shown a 2D environment compared with a 3D environment and, if so can it be concluded that fidelity is the reason for the difference.

The third question builds on the above and focuses on the observers themselves:

- Is there a difference between a naïve observer and one who has had experiences in observation, or intelligence gathering skills (expert)?

This question looks at whether experts are able to notice things differently in a virtual environment and if so will this ability be show in their responses.

The second and third questions in particular could help define how virtual environments are designed and shed some light on how fidelity can make a different when asking an observer to make decisions in the virtual world. The conclusions could have implications on whether significant investment into high fidelity virtual worlds is required, and whether input from experts plays an important role in the design and development of such virtual environments.

1.4 Thesis Structure

The structure of the thesis follows the core elements of the research proposal flow diagram described in Section 1.3. The objectives of each chapter are described below.

Chapter 2 Literature Review: the literature review will try to find if any research has been conducted on insurgent behaviours in crowds and to understand how experts make decisions when observing or looking for patterns in crowds also speaking. It will also look at more generic research in crowd observation that has been conducted. This will help identify a scenario in which the virtual environment can be developed, and hopefully help identify any key behaviour that can be associated with crowds.

Chapter 3 Software Development Kits: this chapter covers a number of game technologies that can be used to create a virtual environment. It is not meant to be an exhaustive analysis of each tool but explains the rationale behind the selection of the tools used to conduct the experiments.

Chapter 4 Observed Activities: this chapter follows on from the information gathered in the literature review and software development kits and is an account of the SME meetings and observations that the Author has conducted over the research period. To gain an understanding of how SMEs conducted observational tasks it was necessary to attend various organisations. Specifically this chapter covers SMEs who have served in the British Army and have conducted patrols and visits to selected military and civilian organizations.

Chapter 5 Experiment I: this chapter describes how the Author designed, developed and conducted an initial experiment (experiment I). The initial experiment was conducted with all naïve observers. Observers were shown the same scene using two different levels of fidelity and four different crowd sizes. In all cases one individual‗s actions was not in line with the rest of the crowd. The observers were then provided with five possible explanations of what was happening and asked to record how strongly they felt each explanation was an accurate depiction of what was happening.

Chapter 6 Experiment II: this chapter describes how the Author refined the design and running of the second experiment. Similarly to experiment I, two levels of fidelity were used. However unlike experiment I, three different scenes were run simultaneously and each scene had four different camera angles. Each observer was first asked to describe what he/she thought was happening in each scene without any background information and then asked to repeat the experiment with specific background information or looking for someone acting ‗suspiciously'. The eye movements of observers were tracked to analyse which ‗camera' each observer focused on and how long for. Experiment II was conducted with both naïve and expert observers, allowing for a comparison to be made.

Chapter 7 Conclusion and Recommendations: this chapter concludes the research by revisiting the three research questions in light of the research and experiment findings. It also highlights how this work could be improved on in possible future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The starting point for this research was an investigation into whether observers can detect ―unusual behaviour" in crowds. Initial discussions with subject matter experts (SMEs) suggested that observers tend to be positioned at a distance from the crowd, and tend not to have a long time to make decisions. Observation could be performed using technology such as closed circuit television (CCTV), using binoculars, or from a vantage point such as a rooftop or the back of a vehicle. The observer would need to detect such activity that could require further attention. Thus an initial assumption is that there are patterns of crowd behaviour that are ‗normal' (i.e. to be expected given the circumstances). The role of the observer in this situation would therefore be to determine when behaviours deviate from the ‗norm'. However in order for the observation to correspond to ―intent", it is necessary to spot potential deviations from what is deemed to ‗normal' before they happen, i.e. to predict such behaviours.

This chapter provides a summary of the literature reviewed in an effort to better understand crowd behaviour and intent. It firstly focuses on the fact that there is no consistent definition of crowds. It then provides a summary of crowd behaviour theories and types. The parameters used to define crowds and crowd types are insightful because they highlight the significance of understanding the context. Without context, norms are difficult to define. And without accepted norms, it is hard to find deviations. As the original aim of this research was into military training, part of this chapter is dedicated to research conducted by others into crowds within a military or policing context. Finally this chapter looks at the importance of experience.

2.2 Defining crowds and crowd types

During the initial phases of the research it became clear that little pre-existing research existed on generic ‗patterns of crowd behaviours'; indeed *"crowds and their behaviours are a relatively understudied area in psychology"* (Challenger et al., 2009a, d). This limitation in research becomes particularly critical when either considering the training of observers (as in this thesis) or in the development of new technology. There seems to be a shift towards automating CCTV to look for known ‗behaviours'. The February edition of The Engineering and Technology Magazine (Institute of Engineering (IET) monthly magazine) looked at how companies are developing ―intelligent cameras" (Courtney, 2011) to assist an observer who may monitor multiple cameras at one time. Whilst the trend toward automation seems to be the next step forward in crowd detection, there remains a need to improve our understanding of crowd behaviour.

Table 2.1 looks at crowd planning in the context of different types of spectator events. He suggest that there are 11 different types of crowds, and defines them as shown in the table below (see Table 1). It is important to consider these as part of the thesis, because it highlights that in-depth crowd understanding is inevitably contextual and that therefore a good understanding of the context is fundamental to the quality of the work.

Table 2.1 Types of Crowds

(Berlonghi, 1995)

Crowd Type	Explanation	Intent
Ambulatory	A crowd walking in or out of a venue, to and from parking area or walking to the toilets or concessions	Gain access to a specific location
Disability or Limited	A crowd of people that in some way is limited or restrictive in their movement. Their lack or ability to walk, see, hear or speak may require planning than is provided for all other spectators.	Gain access to a specific location
Cohesive or Spectator	A crowd watching the activities of an event or at the scene of an accident. Its primary character is the fact that people are interested in watching something specific that they came to see.	Obtain view of event
Expressive or Revellous	A crowd which is involved in some sort of emotional release which can include cheering, movement in unison, celebrating, dancing, chanting or singing.	Participating in collective celebration
Participatory	A crowd of people involve in actual activities of an event. Sometimes these people may be professional performers or athletes. At other times the people attending the event are participating in an actual sport, such as a marathon.	Participating in event
Aggressive or Hostile	A crowd which is one that is becoming verbally aggressive towards or disregarding the instructions of a ticket taker, usher or security personnel. This crowd can get threatening, rowdy and is very open to lawlessness.	Acting against instructions

Crowd Type	Explanation	Intent
Demonstrator	A crowd that is organised to some degree by some established leadership and whose actions may include picketing, marching, chanting or demonstrating at a particular location and for a specific purpose.	Making their collective view known
Escaping or Trampling	A crowd that is attempting to escape from danger either of an actual or imagined threat to life. This includes crowds involved in an organised evacuation procedure and a panic mob pushing and shoving with no order whatsoever.	Gain access to a specific location
Dense or Suffocating	A crowd in which one individual's physical movement is rapidly becoming less likely and possible due to the density of the crowd. People are attempting to move, but they are either swept along with the movement of the crowd or are falling on top of each other. The result of this compression of people is fatalities and serious injuries due to suffocation.	Gain access to a specific location
Rushing or Looting	A crowd whose principal purpose is to obtain, acquire or steal something. This includes rushing to get to the most preferred seats autographs or actually stealing property. This very often results in fatalities, serious injuries and considerable property damage.	Gain access to valuable assets
Violent	A crowd that is attacking, terrorising and rioting with complete disregard for laws and rights of others.	Causing harm to others

What Table 2.1 suggests is that all crowd types are not the same. Just like mushrooms are not all alike and some can kill.

(Berlonghi, 1995) was also concerned by the lack of understanding crowds. This concern is also shared by research conducted in 2009 for the Cabinet Office Emergency Planning College by the University of Leeds into *"Understanding Crowd Behaviours".* This research acknowledges that there were gaps in the College's understanding of crowd behaviours. As such, the research focuses on: addressing these gaps in understanding what a crowd is; the behaviours that can emerge from a crowd; and how to best manage a crowd in certain situations. The report conducted a thorough review of available literature, interviewed SMEs who have conducted crowd managements and looked at technology that could produce similar scenarios for training purposes.

The report tries to set out the criteria that define what a ‗crowd‘ is. It states, *"there is no single, agreed, detailed definition of „a crowd‟"* and highlights the need for a more definitive meaning of crowds for research purpose. The generic dictionary definition of a crowd is: *"a large number of persons gathered so closely together as to press upon or impede each other; a throng, a dense multitude"* (OED, 2010). However a crowd, as defined by the Emergency Planning College, also needs to have distinguishing factors to attribute it as a crowd (Challenger et al., 2009a, d). These factors are:

- size;
- density;
- time;
- ‗collectivity‘; and
- novelty.

Size is simply defined as a *"considerable size"* and density refers to where crowd elements co-dwell in an area with a sufficient *"density distribution"*. What is suggested is that both size and density work together to contribute to a crowd. For example, if 50 people gathered in a large field, then this would not look like a crowd. However if the same number of people gathered in a small garden, then this could be seen as a crowd. The Health and Safe Executive (HSE) publishes guidance on managing crowd safely, including looking at crowd density when setting up a venue (i.e. sports venue or concert). Time refers to individuals who congregate in precise place, for a precise purpose and for a *"measured amount of time"*. An example of this is individuals attending a football match or waiting at a train station. Collectivity is defined as the cohesion of the crowd through *"social identity, goals, interest, and behaviours"*. Collectivity therefore defines the ‗crowd‘ not as a ‗physical crowd‘ but as a *"psychological crowd".* For example this could be a group of football supporters, i.e. all individuals that are collectively by the supporting the same team. The final element to defining a crowd is novelty. It is suggested that novelty is the way that individuals can join together in *"unfamiliar and ambiguous situations"* and are able to link up and behave as a formed crowd ―*without any prior awareness or communication of group norms and values to guide their behaviours*" (Reicher, 1996a, J. C. Turner, 1982). A recent example of ‗novelty‘ is the gathering of UK students in London to protest against increased University tuition fees.

2.3 Crowd Behaviour Theories

Several psychological theories have been applied over the years to understanding types of individual and group behaviours in crowds. (Challenger et al., 2009d) compiled a list of these theories. The report suggests a list of ‗key psychological theories‘ in crowd behaviour taken from various researched literatures (see Table 2). It is worth noting these theories as part of understanding crowds because the wide range shows how complex the subject is. Furthermore understanding the theories contributed to the Author‘s understanding of ‗crowds‘ and helped define the content/approach to the scenarios used in the experiments. However, each theory will not be explained in depth nor criticised in this thesis as this has already been conducted by the Emergency Planning College. Please note that the brief explanation provided is only a summary and is written in the Author‘s own words.

Table 2.2 List of Crowd Theories

(Challenger et al., 2009d)

Theories	Brief Explanation	Relation to "Leeds" Categories
Classic theories (Group Mind, Freud) (Le Bon, 1896) (Freud, 1951)	This theory suggests that an individual loses his own identity which allows him/her to become anonymous within the crowd. With this anonymity, the crowds gains boldness and defiance.	Collectivity
De-individuation (Festinger et al., 1952) (Zimbardo, 1970) cited in (Challenger et al., 2009d) (Diener, 1989) (Prentice-Dunn and Rogers, 1989)	This theory suggests that the individual‘s normal behaviours become diminished in the crowd. The reduction of his/her own self restraint is lost towards the crowd‘s behaviour. It also suggests that the guise of anonymity increases as the group grows.	Collectivity
Social Loafing (Latane et al., 1979)	The theory behind social loafing looks at the extent of an individual‘s energy when in a group compared to when the individual is isolated. The theory suggests that individuals will exert less when in a group, as it is seen the group is working towards a joint goal.	Size
Minimal Group Paradigm (Tajfel et al., 1971) (Tajfel, 1978)	This theory suggests that in-groups (small groups) are created through a bias towards those not in their group (social categorisation). Individuals inside this in-group become united because of a continual measuring of differences with the outer-group. It suggests this helps to enhance ones confidence in the in-group to a level of predominance status.	Collectivity; Novelty
Self-Categorisation (J. C. Turner, 1985) cited in (Challenger et al., 2009d) (J. C. Turner et al., 1987)	This theory suggests that an individual will choose a group by indentifying with it by the way the individual categorises the norms and values. This is how the individual chooses an in-group or defines what an out-group is. This also incorporates the theory that as an individual becomes more influenced by the	Collectivity

Theories	Brief Explanation	Relation to "Leeds" Categories
	in-group, the individual gains a greater social identity with the in-group.	
Game Theory (Berk, 1972a, b, 1974b) (Berk, 1974a) cited in (Challenger et al., 2009d)	This theory suggests that there are five key steps toward crowd behaviour and that a *"rational calculus"* and be drawn on from using the *"principals of Decision Theory"* (Berk, 1974b). The five steps are: • The crowd member will look for facts. • It is likely that the event can be determined from the facts found. • A selection of behavioural choices will be listed. • An arrangement of favourable potential outcome of substitute choices is established. • A direction of something being done is decided upon, aimed at increasing rewards whilst reducing costs. This theory suggests that if an individual‘s behaviour is seen by the group/crowd as being accepted and supported, then the individual will carry out this behaviour. This is because the individual, as suggested in step 5 above, knows that the reward is the support of the group and will override any personal costs of performing the behaviour.	Novelty
Social Facilitation (Zanjoc, 1965) cited in (Challenger et al., 2009d) (Cottrell, 1972) cited in (Challenger et al., 2009d)	This theory suggests that individuals become stimulated when conducting tasks in the company of other people.	Collectivity
Emergent Norm (R. H. Turner, 1964) (R. H Turner and Killian, 1957)	This theory tries to expand on how a group‘s combined action is controlled by the norms that arise from inside the crowd. What is suggested is that when crowds form for an event there are no norm behaviours. However, as ‗social interaction‘ begins to take place, the crowd tries to define the situation (and this can be seen as establishing ‗behavioural norms‘). As the crowd begins to establish its norms, the norms begin to take power over the crowd members. What is also suggested is that these norms usually arise from a member who sticks out from inside the crowd. It is	Collectivity

Theories	Brief Explanation	Relation to "Leeds" Categories
	believed that as these norms grow it becomes harder for those in the crowd not to conform to them.	
Social Identity (Tajfel, 1978, Tajfel and Turner, 1979) cited in (Challenger et al., 2009d)	This theory suggests that as individuals *"we are continually involved in the process of categorisation, identifying and comparing"* (Tajfel, 1978, Tajfel and Turner, 1979) cited in (Challenger et al., 2009d). The process of Categorisation suggests that we as individuals look at others to categorise ourselves and them. The process of identification is how individuals affiliate themselves with particular groups, *"known as in-groups who share a sense of identity and belonging"* (Challenger et al., 2009d) The final process is that of comparison. It suggests that the individual compares their in-group with other groups *"known as out-groups"* (Challenger et al., 2009d). These out-groups are seen to be unconnected to the individual's in-group.	Collectivity; Novelty

Whilst there are varying points of views on crowd behaviours, there are some commonalities. In the simplest term then a crowd's behaviour is generally developed through individuals who, through social identity, are united towards a specific group, such as their love for a specific sports team. This then suggests that as the individual becomes de-individualised, or through multiple identities, the crowd/group's social behaviour becomes more unified towards a particular social identity.

However, (Challenger et al., 2009a, d) also suggests that crowds have groups within groups. Multiple behaviours may become apparent. Hence a single type of behaviour cannot be defined for one type of crowd. The table below summarises combined crowd theories.

Table 2.3 List of Combined Crowd Theories

(Challenger et al., 2009d)

Combined Theories	Brief Explanation
Social Identity Model of Crowd Behaviour (Reicher, 1984b, 1987) (Reicher, 1984a) cited in (Challenger et al., 2009d) (Reicher and Levine, 1994b) (Reicher and levine, 1994a) cited in (Challenger et al., 2009d)	Extends the Social Identity & Self-categorisation Theories. This model suggests that instead of an individual's identity and behaviours the individual's identity and behaviour shifts to a more to a social crowd identity and behaviour. This way the individual is not losing control of him but is shifting toward the social identity of the group values and beliefs.
Elaborated Social Identity Model of Crowd Behaviour (ESIM) (Reicher, 1996a) (Reicher, 1996b, c, 1997) cited in (Challenger et al., 2009d) (Drury and Reicher, 1999) (Stott and Drury, 1999) cited in (Challenger et al., 2009d)	This theory is based on the principles of Social Identity & Self-categorisation and extends the social identity model of crowd behaviour. Based on the same theories as above. This model suggests is that a group never forms on its own but is created when another group is present. It also suggests that the context of group action is taken based on the action of one group versus another other group. This would therefore suggest that the action of the police would have an effect on the action of the crowd and vice-versa.
Social Identity Model of De-individuation Effects (SIDE) (Reicher et al., 1995) cited in (Challenger et al., 2009d) (O. Klein et al., 2007)	This theory expands on the Social Identity and De-individuation theories. This model redefines de-individuation by suggesting that an individual can have numerous way of defining himself/herself. What the SIDE model suggests that there is difference between how a person acts (action as an individual) as an individual and the individual actions in the group. It would suggest then that individuals can define themselves thorough a varied number of levels. The example suggest that as individuals we will continually try to identify with other individuals or draw a comparison with other individuals and this is also done in a social identity, how as an individual we see are self in the group, differ from those in the group and how we differ from other groups.

Combined Theories	Brief Explanation
Place Script Theory (Donald and Canter, 1992)	This theory suggests that individuals generally adhere to rules. These rules guide an individual's behaviours and can be legal, prescribed or built upon previous experience, perceptions and ones expectations. What is suggested by this theory is that these rules are also scripts or schemas that help individuals define and describe behaviours of others in a defined environment. It then suggests that these scripts become associated with specific places. The individual begins to cultivate a precise ‗schema' for these environments. A familiarisation of the rules and behaviour associated with that environment takes place. An example is the way an individual might enter a building or exit a build everyday (probably the same).

The literature summarised above suggests that no one model or theory would give a clear framework for defining ‗a crowd' and understanding its behaviours. It is also important to note that there are limitations to the theories and models. For example, models like ESIM are purely based around the police and the crowd and no further research has been conducted to see if the ESIM model works with different crowds. The Place Script theory implies that people become so immersed in the schemas that they would carry on running through the script even if an emergency happens. However, having an understanding of these theories helps the development of a contextual scenario to experiment from, especially when looking for that ‗individual' in the crowd who's behaviour is not conforming to the social bonding behaviours of that crowd.

Little emphasis is given to the size, density and time ―needs" categories in these theories. This may simply be because the theories focus on causal factors of behaviours. For the observer of a crowd, such causal factors can, at best, only be inferred from the behaviour and might have little relevance to identifying the crowd's behavioural intent.

2.4 Crowds within a Military and Police Context

2.4.1 Military

Research conducted towards a more specific crowd context was undertaken for the U.S Military due to the numerous non-combatant missions involving their personnel being involved with crowds. This work was of great interest yet not so forthcoming when trying to access data from the researchers. The research was conducted to integrate crowds into existing U.S Military federates (simulations). The term ‗federate‘ derives from the U.S Military set of collaborating simulations (R. C Gaskins et al., 2008). The research of interest was known as the ‗Crowd Federate‘. The ‗Crowd Federate‘ was developed because of a growing need for training that represent hostile environments where both civilian and military people co-exist.

Gaskins and his co-researchers formed the psychological studies approach to the ‗Crowd Federate‘. There approach was to conduct three separate studies: firstly naturalistic observation, secondly conducting interviews with SMEs and thirdly a survey research. (Ryland C. Gaskins et al., 2004, R. C Gaskins et al., 2008) hoped that from this study crucial parameters for crowd behaviours could be taken into account, along with random values for the type of behaviours, that would arise given a ―particular context”. This ―particular context” is interesting as what (Gaskin et al. 2008) is not trying to do is define what is ‗normal‘ for each event but what each event holds in particular to the context of that event.

Gaskin and his co-researches began using naturalistic observation techniques to measure human behaviour of three events: the 1999 World Trade Organisation (WTO) protest (in Seattle), the 2004 anti-war protest in New York and Military Operations on Urban Terrain (MOUT) training exercise involving crowds.

The Gaskin *et al*. 2004, 2008 studies provide a list of ―fifty-five behaviours” that students were required to observe video of the WTO event, each student had to indicate which behaviours where happening during a 10 minute interval from the list of behaviours. The students were provided with training prior to carrying out the observation on the WTO video. The training gave the characteristic of the behaviour and at anytime students could ask questions to clarify any items. The first hour of observation was conducted as a group and students verbally indicated when behaviours emerged. The two other hours where conducted as individuals.

This research showed the means and percentage of behaviours that occurred the most in the crowd and the degree of aggression throughout the protest timeline. The most frequent behaviour found from the World Trade Organisation meeting in Seattle was described as ―crowds standing on elevated structures” at 6.5%, followed by ―yelling, shouting, raising flags” (in combination) at 4.4%, and ―filmed by the media and chanting” at 4.3%. Aggressive behaviour was rare but, when it did occur, it was the most destructive. Such behaviour included glass throwing, fighting, jumping on moving vehicles, looting and blocking the road by lying down (Gaskins *et al.* 2004).

The same observation techniques were used during an anti-war demonstration in New York. This time, however, a difference was observed in the crowds‘ behaviours from those observed at the Seattle WTO ―sign and banner carrying” had a 60% incidence, with ―flag-waving” at 10%. The remainder of the behaviours, including chanting, observers not participating and general ―yelling and shouting" had a combined incidence of 11% (Gaskins *et al.* 2004). Although the figures are not identical from the two events, it could be assumed that Gaskins was trying to show that no two events produce the same behavioural outcome although contextually both could be seen as the same crowd of protestors. In other words, crowds will behave differently even in the same context.

Gaskin et al. (2008) indicated that reliability increased during the 5 hours of observation made by the students. Reliability was determined by using standard deviation measurements. This is significant because it could imply that, with training and a set of classifications, it is possible for observers to become better at detecting particular behaviours in a crowd.

The natural observation also allowed Gaskin to survey SMEs who had returned from operational theatre using both open ended and Likert scale questions. The survey objective was to understand what types of behaviours appeared in operational areas, and if any similarities could be linked with the WTO work. The survey would also look at how crowds were best dealt with and how peace was brought back to the situation, how much experience the soldiers had with crowds and finally examining if any of factors found from the WTO where present.

What the individuals saw as causes towards crowds forming was from *"payment to Iraqis, political protest, religious protest and protest due to an arrest".* Within crowds unrest cues where noticed by the individuals; when pay was being dealt with aggression arise at military police control points, political protest was largely non-aggressive but a vocal approached was used and the religious crowds where aggressive and used physical methods to gain attention. Other determinants where also found in the crowds, the existence of instigators, weapons, eagerness to take risk, peacekeeping becoming aggressive, alcohol, drugs, crowd size, approval of violence in the crowd, an undertaking towards the cause, and organised crowd leadership.

Although this was a small concentrate analysis, Gaskin et al. (2004, 2008) shows the growing need for research to be conducted into understanding types of crowd behaviours in a Military situation where the context may not be similar to those in a public realm. This is the only study that closely resembles the research of the Author's. Gaskin never talks in-depth to how (if any) the research from the WTO had any similarities to the findings of the SMEs and research survey. Further work pertaining Gaskin recommendation, seem to never evolve and it can only be assumed that this work may have been published in a non-public domain.

2.4.2 Policing

Some key factors that do fit in with (Ryland C. Gaskins et al., 2004, R. C Gaskins et al., 2008) was (Bessel and Emsley, 2000) who created a general flashpoint model when looking at policing during civil unrest. This framework is based on six interdependent levels. The levels are briefly explained here as:

- ―Structural”: the social inequalities, including subjective deprivation, how people feel, and gain resentment towards an order.
- ―Political/Ideological”: how governing parties, media, and government organisations look towards or portray an opposing group. An example of this could be the Irish Republican Army (IRA).
- ―Cultural”: the differences between "Us" and ―Them”, knowing the differences and understanding the ―behavioural norms”. This was also noticed by Gaskin.
- ―Contextual”: understanding the situation through known history where unrest may have arisen and know negatives towards those enforcing.
- ―Situational”: the ―spatial” or ―social” cause of a situation going from order to disorder. This is reflects how police, police through spaces either confining parties to controlled area or allowing a softer approach to allow the parties to self-police (police presence at a distance).
- ―Interactional”: the calibre of the relationship on the ground between the protesters and the police. This is can be measured by the levels of ―respect, cooperation, restraint or provocation”.

The Author has concluded from this work that not all events have the same degree of aggression and this will undoubtedly be a problem when trying to identify a generic framework. However, using (Bessel and Emsley, 2000)'s understanding of flashpoints may help indentify when a crowd, or a person, is intent on creating unrest. What can be seen from the crowd behaviours is that different factors may give rise to a crowd formation (such as poor pay, feeling of unjustness or fighting for a cause). However, not all crowd behaviours will turn to unrest. For example chanting football supporters, groups attending an indoor music concerts or a free event in a park outdoors will not normally produce crowd unrest. What can be gathered from the literature is that behaviour that may lead to ‗intent' from within a crowd seems to be based on the context/scenario of that crowd, this means that the context may create enough key factors, or known flashpoints, to alert the observer towards further investigation (to find the intent).

Research conduct by (Hoggett and Clifford, 2010) at the University of Liverpool looked at crowd psychology, public order and the current police training and training towards policing football crowds (UK Police). (Hoggett and Clifford, 2010) suggest that Police training lacks a clear focus towards *―crowd theories”* when training the police therefore potential creating a basis attitude towards the crowds. It also suggests that a lack of research and literature has been conduct into police training.

An interesting finding from (Hoggett and Clifford, 2010) was that during training towards public order, trainees found it difficult to near impossible to be able to identify or even separate hooligans from the crowd.

As these were trainee student police, the observations carried out by (Hoggett and Clifford, 2010) could also suggest that without an adequate level of experience or understanding of crowds it is hard to notice potential key changes that may alter the behaviour of the crowd and to respond appropriately. (Hoggett and Clifford, 2010) watched trainee officers work through crowd control scenarios. The trainees where told through intelligence that a group members of a football team where arriving at the same time as a rival football team members. The trainees suggested that it would be easier to remove the entire crowd of supporters and then to look through the crowd for potential trouble makers. However this could result in a flashpoint, like (Bessel and Emsley, 2000), whereby the group feels they are being treated unfairly by the police. The potentially counterproductive approach suggested by the trainees is with (Hoggett and Clifford, 2010) views that training in the English and Welsh Police forces regarding understanding crowd theories was considerably out-of-date and that there was a need to consider modern scientific approaches to crowds.

The challenge of effective training was further highlighted when the Author spoke to an SME at the Metropolitan Police CCTV control room. The SME suggested that training on ‘what to look for’ was not given but built on from experience over time. This experience would then be passed on either via briefings or word of mouth to more junior or novice operators.

2.5 The Importance of Experience

As already mentioned above, distinguishing different types of crowds is central to understanding what is happening in a crowd. (Berlonghi, 1995) states that *“distinguishing allows us to become more effective and competent at a particular activity”.* The Author likens this statement to the difference between novice and expert understanding. As such gaining an insight into differences between how an expert notices something compared to a novice, was the logical next step of this research.

2.5.1 Decision Making Models

There is a suggestion that observation requires the observer to spot deviation from the ‘normal’. This requires both an appreciation of what constitutes normal, which could be situation-dependent, and how deviations are spotted. The definition of ‘normal’ is likely to involve schema (Bartlett, 1932), which can be considered to be a collection of knowledge, gathered through experience, which applies to a specific situation.

(G. Klein, 1999) looked at trying to understand how an expert makes decision and devised a theory toward getting that information out. (G. Klein, 1999) suggests experts in a specific field have knowledge that needed to be extracted, and therefore developed a frame work known as the *"Recognition-Primed Decision Model"* (RPD) to obtain an understanding of how experts make decisions. RPD uses schema as patterns for detecting/selecting features. Fig. 2.1 illustrates RPD.

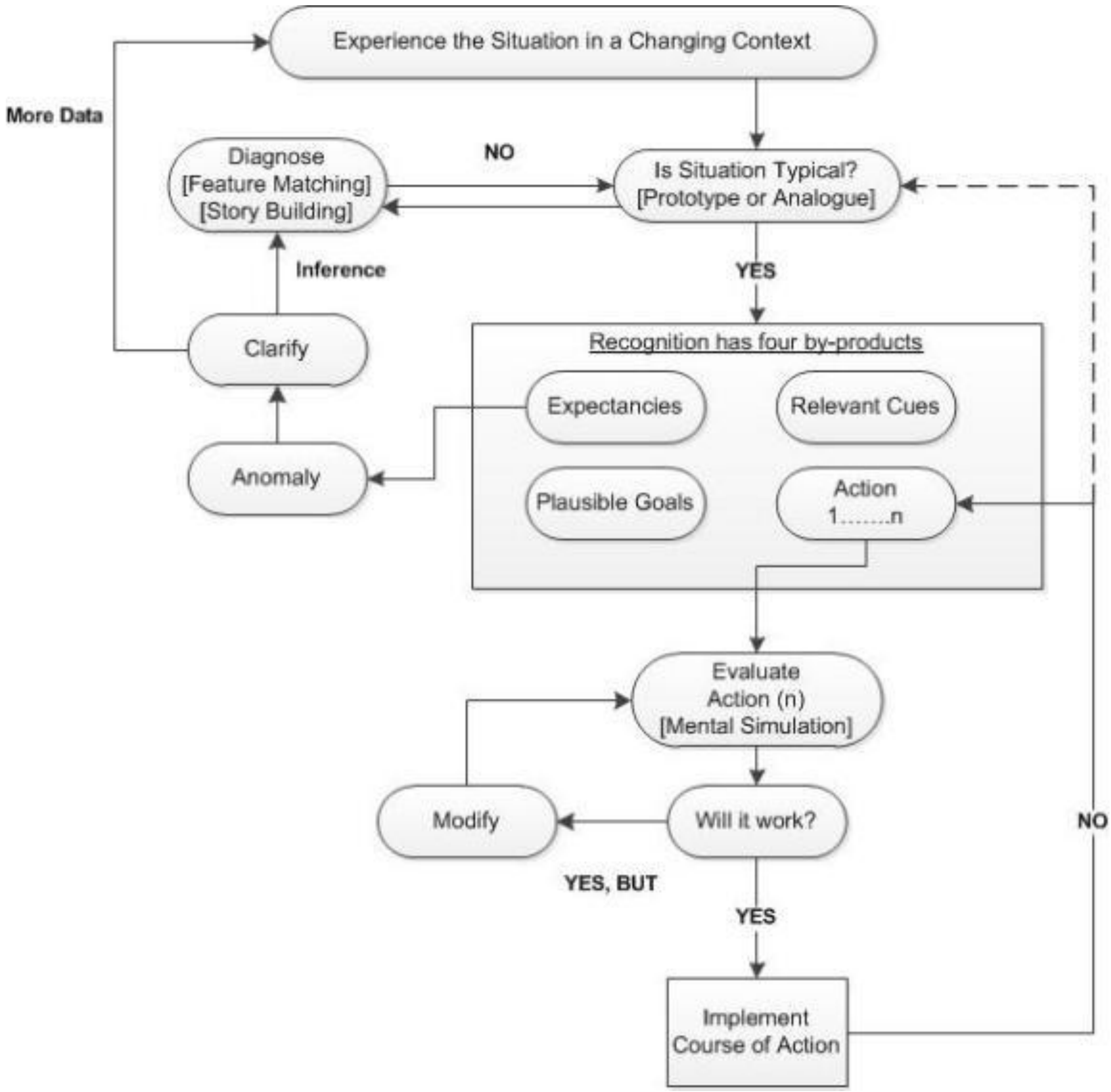

Fig. 2.1 Integrated Version of RPD Model

(Source: pg.91 (G. Klein, 1999))

(G. Klein, 1999) also suggest that the RPD model joins two processes together: the first is how *"decision makers"* measure a situation, and the second is how they evaluate the *"course of action"* mentally.

The RPD model fits into the Author‘s own experience as a Radar Plotter in the Royal Navy, has taught him that initially it is difficult to distinguish if the dot on the radar is an object or simply ‗weather effect‘ (e.g. sea spray). As the Radar Plotter becomes more experienced it becomes easier for him/her to distinguishing between weather effects on the radar and actual moving objects. Similarly, understanding what is happening in a crowd is central to being able to effectively stop problems before they arise. Therefore Author‘s own experience supports the theory that through experience and overtime a person builds mental maps or ‗schemas‘ on how events have happened in the past and solution to resolving or noticing key factors become easier to make.

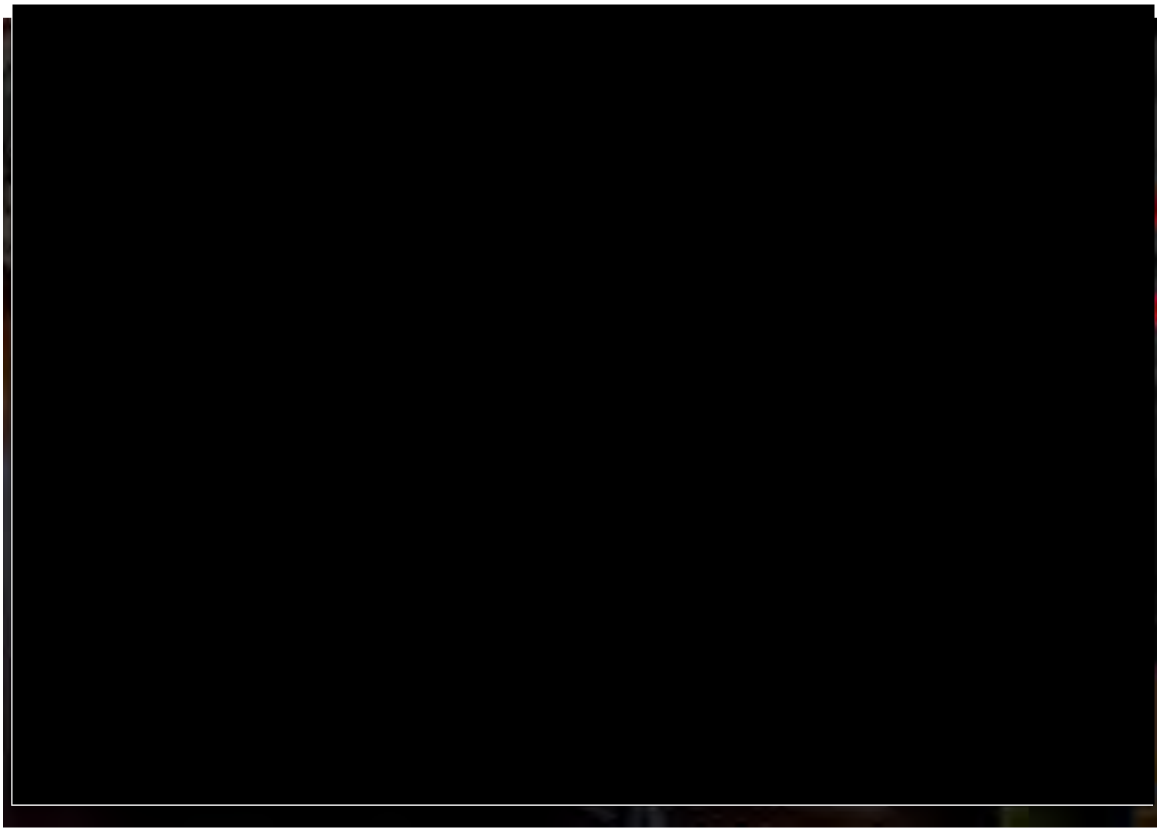

Fig. 2.2 Radar System used by the Author

(Source: www.DefenceImages.mod.co.uk)

While ‗schemata‘ might provide a reasonable framework for describing the observer‘s motion of ‗nomal‘, it is difficult to operationalise. This difficulty arises from the fact that each observer, rather than undergoing specific training, tends to construct their own sense of ‗nomal‘ from their experience. This means that, while it is possible to extract knowledge from observers, for example by verbal discussion during normal work, it is not obvious how this will create a generic model of observation. If each operator relies on cues based on previous experience, then one would need a large number of observations to reflect such variation.

Working in Radar‘s also required levels of searching for specific threats during exercise (known intelligence threats/unknown threats). (Koenig et al., 1998) suggest that the process of inspections forms a *“visual search”* with *“decision making”*. This inspection is known as *“Serial Search”.* Similar to searching a radar plot, serial search is a free-field search in which a target can appear in any location. The problem with this type of searching is that it can potentially take longer. (Boersema and Zwaga, 1990) also found that giving searchers multiple targets to search for in a noisy background made the searchers less effective. This is was because the increase of other signals (i.e. the noisy background) increased the search time and reduced locating performance.

Similar to what Gary Klein suggest (RPD), (Moray, 1976) suggests that *“observers will construct statistical models of the spatial and temporal properties of his environment. The observer uses the models both to govern the decision he makes about the data obtained when makes an observation, and also to decide when and where to make observations.”*

The thesis will not delve deeply into searching techniques however they can prove insightful when comparing naïve and expert observers. This could also lead to an understanding of how expert‘s mental mapping (Johnson-Laird, 1983, Smyth et al., 1995) suggest this is *“we”* comprehend the world by building working models of it within us.

An alternative approach might be to present naïve observers with a basic scenario and then increase the amount of information that is given to them. It is suggested this could provide a means of comparing performance on an observation task as more situational relevant information is provided. This suggestion forms one of the main drivers behind Experiment I (see Section 5).

2.5.2 Cultural Awareness

The importance of cultural awareness is highlighted by research conducted by Gaskins whilst visiting a U.S Military Operations on Urbanized Terrain (MOUT) to carry out natural observation. Gaskin observed an absence of understanding cultural differences which could result in unnecessary escalation of situations in the crowd. Gaskin observed a lack of understanding that Iraqis are curious individuals by nature and needed to have eyes contact when talking to people. He noticed that individuals did not remove sunglasses when entering a town and that this could then lead to a misunderstanding of how the crowd is feeling and creating a flashpoint in the crowd. Gaskin‘s findings from the MOUT observations suggest that an individual‘s known ‗norms‘ on perceiving events could have an impact on how theses ‗norms‘ are perceived in different cultures.

Some of Gaskin‘s work was conducted during a time when a limited training on cultural awareness was given before going to operational areas. The Author understands this may be of lower relevance now as an SME informed him he was given training in cultural awareness and that the British Army provide classes to those who are going to serve in operational areas the relevant cultural awareness training is conduct by Afghan nationals (meeting held a Defence School of Languages). However, the Author still believes that Gaskin's concern is still important as misunderstandings could alter a situation.

2.5.3 Findings from individuals experienced in dealing with crowds

(Challenger et al., 2009a, d) conducted a review of a number of event held in the U.K and spoke to those who managed or had experience in dealing with crowd events. The events ranged from Glastonbury to Hogmanay. What Challenger et al. tried to find was if there was any similarities towards the type and behaviour from the crowd. However what (Challenger et al., 2009d) found was that there is no one type of crowd and that behaviour of the crowd depended on several factors. These factors can range from time of day to location. What (Challenger et al., 2009d) suggest is that there is a need to draw on the knowledge and experience from previous events to gain an understanding of how the crowd may behave. It is also suggested that there is a range of potential behaviour that can rise from a crowd, ranging from passive to full out aggressive protest. What this means is that it is difficult to understand what a crowd will do and creates a problem to understanding that the crowd may have groups of behaviours within one crowd. Therefore if a framework towards crowd understanding is to be understood then their needs to be flexibility in to deal with the different types of crowds and the behaviours.

The SMEs that (Challenger et al., 2009d) spoke to suggested that individuals in the crowd feel empowered and are willing to do things in that crowd cover that they would not do if on their own. The SMEs speak about the different crowds that are normally dealt with by the Police force of the U.K. The most common is the *"Crowd Marches and Demonstrations"*. The SMEs define typical behaviours that are common to the crowd. The *"Crowd Marches and Demonstrations"* seem to hold multiple behaviours, from law abiding individuals, difficult individuals; those who want to cause slight disorder to what the SMEs call the *"professionals"* who main aim is to cause disruption and *"rent-a-mob"* have no interest in the protest but just want to commit disruptions.

"Sporting events" such as football, rugby and cricket also typically have crowds. With football come almost *"territorial rights".* These rights can be found both inside the grounds and out, and the crowds may feel they have right to protect this territory. Rugby is seen as a family sport and therefore is less likely to have a territorial dispute. Cricket has seen a change. Generally classed as a gentlemen sport with little trouble; the SMEs have seen a rise in trouble at international fixtures fuelled by the *"younger binge drinking groups in fancy dress".* The Olympics are also mentioned. It is suggested that that Olympic crowds are generally happy and contain a wide mix

of individuals. However, even here there is potential for issues to arise if individuals need to queue too long or cannot gain access into the desired event.

Comparatively, the last few crowd types seemed to raise little concern from the SMEs. Crowds in concerts, celebrations seem to be typically unproblematic, with the most problematic scenarios relating to intoxication at concerts or celebration.

Train station crowds comprise of two types, those who are season individuals who know exactly where he/she is going and the more relaxed *"leisure traveller",* who through observation is unsure of where to go and move in a more stop start motions.

2.6 Summary of Findings from Literature Review

Despite acknowledging that there has been limited research into understanding crowds and crowd behaviour, the Author identified many useful insights through the literature review. The findings included a wide variety of theories on factors that need to be considered when looking at crowds and the factors that influence the observer's ability to effectively recognise and act on information provided (Fig. 2.2 Crowd and Observer).

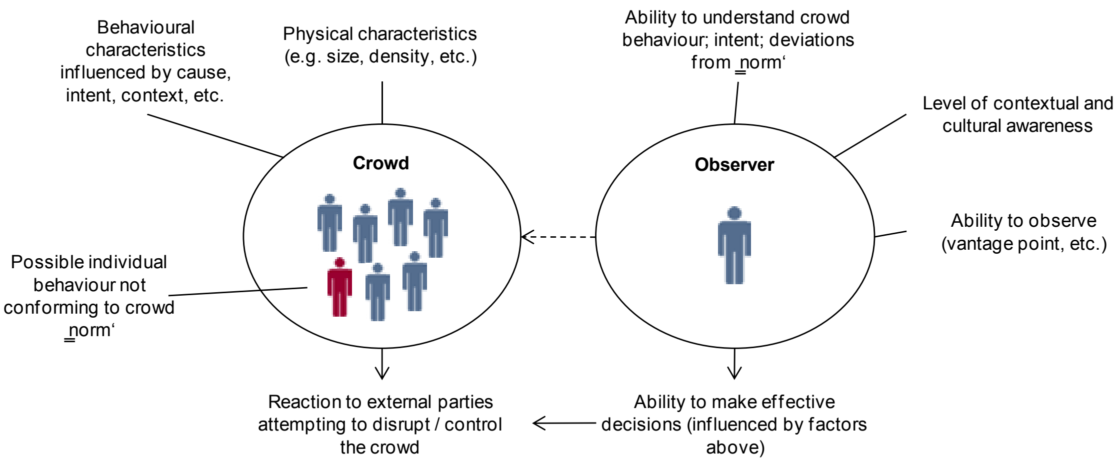

Fig. 2.3 Crowd and Observer

- There are five distinguishing factors of crowds: size, density, time, collectively and novelty.
- Most theories on crowd behaviour seek to explain behaviour as a form of collective decision making and make little reference to context and distinguishing factors
- However, SMEs and Military/Policing research finds that crowd behaviour varies according to context.

- Experience plays a key role in being able to understand crowd behaviours, intent and any deviations from the ‗norm'.
- Inexperienced or inappropriately trained observers may make ineffective decisions and negatively influence the crowd/cause adverse reactions

The key findings from the literature review described above relate to question one of thesis (Crowd Behaviour) and question three (Experience) of the thesis. Little emphasis has been give to literature regarding question two (Fidelity) as this will be explored in Chapter 3.

CHAPTER 3

SOFTWARE DEVELOPER KITS

3.1 Introduction

In order to answer the thesis questions, it is necessary to research and define what levels of fidelity in game technology were required, available and practical to use for the experiments. The chapter starts with a short recount of literature reviewed regarding perception of causality as this influenced the Author‘s understanding of how much fidelity was required. The research also investigated agent-based simulation tools, relevant literature on the importance of crowd psychology in simulations and the challenges of creating high fidelity simulations. Finally it looks at a selection of software development kits (SDK) and explains which two SDKs were selected for the research experiments.

Please note that this chapter is not intended to be an exhaustive summary of all the potential game development tools that are on the market.

3.2 Perception of Causality

An understanding of ―perception of causality" is needed to understand what is needed for a person to perceive if an agent/character was behaving or moving in a particular manner in a virtual environment. (Michotte, 1963) is seen as a pioneer in this area of research. Michotte believes that ―certain physical events give an immediate causal impression". This is shown through objects and motion. Michotte‘s theory is based on what is seen as a phenomenon when observers see two objects moving. The belief that an observer, thorough no prior knowledge, perceives that the objects in the experiment creates causality.

(Scholl and Tremoulet, 2000) conducted a study into perception of causality and animacy, looking at previous work that had been done during the 1900s. In its most basic form, perceptual of causality consists of two objects (A) and (B). Depending on the object‘s position, trajectory, distance and speed from each other, a cause can be seen - i.e. the phenomenon of perception of causality. Animacy is created thorough (A) interaction with (B).

For example, (A) begins to move towards (B), (A) reaches (B) and stops, then (B) begins to move away from (A) always along the same trajectory. In another example by (Scholl and Tremoulet, 2000) object (A) moves towards object (B) and stops, (B) moves, but this time before (A) gets to

(B), (B) starts to move quickly away from (A) this cycle continues again. What (Scholl and Tremoulet, 2000) believe is that in the first example —A causes the motion of B”, whereas the second example is seen as —A and B as alive”. Thus generating what is seen as perception of causality and animacy (cause through animacy).

What is also interesting is the lack of detail needed to create perception of causality. A simple object (such as a circle) and a plain background is all that is needed. (Scholl and Tremoulet, 2000) states

> —One especially intriguing aspect of these phenomena is how simple and spare the stimuli can be, with displays only containing only a few small-moving 2-D geometric shapes”

(Scholl and Nakayama, 2002) see perception of causality as a ―higher level conceptual of the world” that has not interference on ones intentions, beliefs and cultural differences. However this is not the same if grouping is introduced (Choi and Scholl, 2004).

The use of perception of causality is seen as a link to the real world and how observers perceive. The research in this thesis is not going to compare real with simulated environments. But will conduct a level of fidelity in the simulated environment to see when observers notice/perceive changes. What strengthens the case for using perception of causality as a good baseline, is that the experiments are conducted using a simulated environment (slides or movies depicting simple moving objects)

Studying the perception of causality further (Choi and Scholl, 2004) wanted to investigate the effect that grouping and attention have on perceptual of causality. Through their experiments (Choi and Scholl, 2004) found that the perception of causality is not only sensitive —to local information within individual events, but that it can also take a wide variety of contextual information in account”. This could explain the findings that (Houghton and Baber, 2009) found when studying deviancy within flocks. The research was based on the simulated flocking behaviour developed by (Reynolds, 1987). The study took the three steering rules used to create flocking.

Houghton and Baber (2009) were able to create deviant behaviour in the flock by adjusting one or two of the three steering rules in the flock. What (Houghton and Baber, 2009) found by conducting the experiments was that there was a clear distinction between global and local observation. The experiment was broken into two variants cued and un-cued. Un-cued was the flock shown all in the same colour with numbers in the centre. Cued was the flock shown as the same colour except two, which were a different colours, and not numbers. What Houghton and Baber, (2009) found from the results that naïve observers where better at find deviant behaviour in cued conditions, except when collision and separation was present this was identified better in un-cued condition.

3.3 Agent-based Simulation Tools

Software used to develop crowd/pedestrian simulations relies on understanding ‗what' the crowd/pedestrian may do in a specified area. These simulations are known as agent-based simulation tools as an agent can be given different behaviours/rule to test for different results. These tools are used when designing areas where crowds tend to gather.

A report written by the Emergency Planning College (Challenger et al., 2009b, c, d) highlighted that there are different patterns to crowd movement. It listed that people in crowds will choose the straightest route or the best line of sight that reduces the number of directional changes required, but will keep a distance to avoid encountering with someone unless in a hurry or if a crowds size increases. Crowds do not regularly fills spaces equally but form into clumps. When a crowd is faced with an obstacle in which it reduces the pace, certain people in the crowd will grow agitated and counter flow system can develop in which a crowd will generate into self-organised lanes, to allow a smooth transition and avoid contact.

Table 4.1 briefly describes some existing agent-based simulation tools looked at by the Author.

Table 3.1 Agent-Based Simulation Tools

Agent-Based Simulation Tool	Brief Description
Simulex	Simulates escape movements of people through large complex buildings, spaces are defined in 2D floor plan.
Legion	Simulates and analyses crowd dynamics in evacuations from a constrained environment. i.e. football stadium
Myriad II	General purpose crowd analysis tool, which is three different environments but is incorporated into one.(Agent Based Analysis, Spatial Analysis and Network Analysis)
Mass Motion	Predictive tool which tries to determine ―what if" events. It concentrates on two types of crowds, commuter or evacuation and is typically used in engineering design for public transport, retail etc.
NetLogo	Multi-based agents and environments, runs normally in 2D; however also has a 3D viewing mode.
Exodus	Developed by the University of Greenwich, fire and safety engineering group. This simulation tool is used to understand evacuation and pedestrian dynamics.

Most agent-based simulators are being used to design and develop buildings and concentrate on how user friendly they are. For example how easy it is to identify the location of a ticket booth or exit, thereby providing useful information for designers to alter building designs before the building is actually built.

The simulation tools listed above do not look specifically at the ‗intent' of the crowd, nor an individual behaving different. This is not to suggest that these agent-based tools could not do such a task. Therefore the challenge for the author was that in order to use such tools, a framework of behaviour (set of rules) is needed and this was not easy to establish. Also not all tools were free to use or develop from and would incur a license fee.

NetLogo is explained in more detail as part of the Developer Kits in section 3.6 as it was seen as a contender for the experiments.

3.4 Crowd Psychology in Computer Simulations

Some interesting research in developing realistic crowd simulations has been conducted for the US Military, as the US Military is less involved conventional style warfare and used increasingly in peacekeeping operations. Therefore soldiers need to interact with the public in hostile places and be able to distinguish the difference between a friend and a foe.

Work carried out by Gaskins et al. (2004) and Petty et al. (2004) led to the development of a ―Crowd Federate", based on events in Somalia, where hostiles were ―mixed" with civilians, as well as other areas where the US military had encountered crowd unrest. Although the papers written on ―Crowd Federate" seemed to suggest the evolution of a potentially useful tool, an effort to obtain a demo or copy of the software to test has proved unsuccessful.

Gaskins et al. (2004) suggested that the problem with existing crowd behaviour models for simulators was a ―lack of psychological underpinning", yet their paper failed to address which behaviours the observers subjects' were given when they were attempting to detect which type of crowd behaviours were occurring during the three demonstration events. As such it could not be used to develop a framework that typifies what behaviours are generated in crowds.

(Pelachaud et al., 2007, Pelechano et al., 2005) followed a similar direction to the ―Crowd Federate", but focused on civilian evacuation. (Pelechano et al., 2005) believed that there is a need for each autonomous agent to have a psychological model that drives its "mind" to enable it to make its own decisions independently. What their work achieved was the development of a Multi-Agent Communication for Evacuation Simulation (MACE).

MACE uses an existing framework called PMFserv (Performance Moderator Function), which is a psychological model that is integrated into the MACE crowd simulation system. A system was developed that allowed agents to develop a ―way-finding process that allowed them to explore and learn the internal structure of a building". This differed from other simulations as it allowed agents to find their way around an environment with which they were previously unfamiliar. Agents can communicate and share information with each other, and a built-in behaviour base that defines the roles of each agent was added during an evacuation simulation.

3.5 Challenges of Creating High Fidelity Simulations

Research like (Braun et al., 2003, Soraia Raupp et al., 2005, Thalmann and Musse, 2007) look at simulation and the complexities in making them believable by understanding how to give agents their own behaviour to create a more realistic virtual environment. Most of the agent-based simulations are developed using a rule-base system whereby the agent has a list of rules. This is also the case for most video games. If you give an agent open-ended rules, the computation level would decrease as the more agents you have the more computer processing is required. Hence some video games and simulators try to decrease the level of detail in the virtual world. Using trees as an example, the level of details decreases in the distance to the point they become ―billboards". Similar to real billboards, billboards in a virtual world are non-dimensional walls (in this example with trees).

The Author gained valuable insights into creating scenarios whilst working at Rare™ Microsoft Game Studio. A senior software engineer suggested that to create a realistic crowd, the Author would needed the latest game technology and a team of three to four programmers for at least 6 months. It therefore became clear that the Author alone would not be able to creating a realistic crowd event without significant support and effort. However, this did not mean that a simpler solution could not be found.

3.6 Developer Kits

Software Developer Kits (SDKs) are developer tools used to develop application or video games for specific hardware and software. For example, Microsoft's Visual Studio SDK (also known as an Integrated Development Environment) allows developers to build project templates to create application, embedded programmes or XNA games. Another example is Google's android SDK. SDKs are normally licensed by the proprietor who only allows development for their own operating systems.

The SDK selected for the experiments plays a key part in addressing the level of fidelity (defined in Chapter 1) needed to make an observer aware (of intent) in a virtual environment.

A virtual environment is the world recreated into a digital media. Videogames are good examples of virtual environments. An example of this is Microsoft's flight Simulator X where a user can fly various types of aircraft around the world without leaving his/her chair.

In its simplest form, fidelity is dependent on how well the tools selected to create the virtual environments are at recreating a “real-world” scenario. A selection of tools considered by the Author for the experiment is described in the following sections.

3.6.1 Unreal Developer Kit

Epic Games developed the Unreal game engine. The game Unreal Tournament is a good example of what the Unreal engine can do. The Author considered the Unreal Engine as Epic Games also provide a developer environment with the Unreal Tournament game. The developer's environment is known as the Unreal Developer Kit (UDK) and allows anyone to start developing game levels. The user became a developer and creator of his/her videogame levels. Another interesting aspect of this development environment is that it has community of developers and forums which can be useful when developing. The UDK can now be accessed freely for developers who want to use the Unreal engine framework (UDK, 2010) Prior to this a videogame from Epic Games was need to access the UDK.

3.6.2 CRYTEK

Similarly to Epic Games, Crytek provide a free developer kit packaged within the FarCry videogame (the latest is Crysis) (CRYTEK, 2011). Developers to use the CryENGINE framework to develop game levels in a similar way to Unreal. This also became popular and has a devoted site for those who want to learn how to modify and developed their own game levels (CRYMOD, 2011). An illustration of the virtual environment and developer kit are shown in Fig. 3.1 and Fig. 3.2 below.

Fig. 3.1 Crytek Virtual Environment Crysis Videogame, (PC)

(Image from Crytek, Crysis Videogame, 2008)

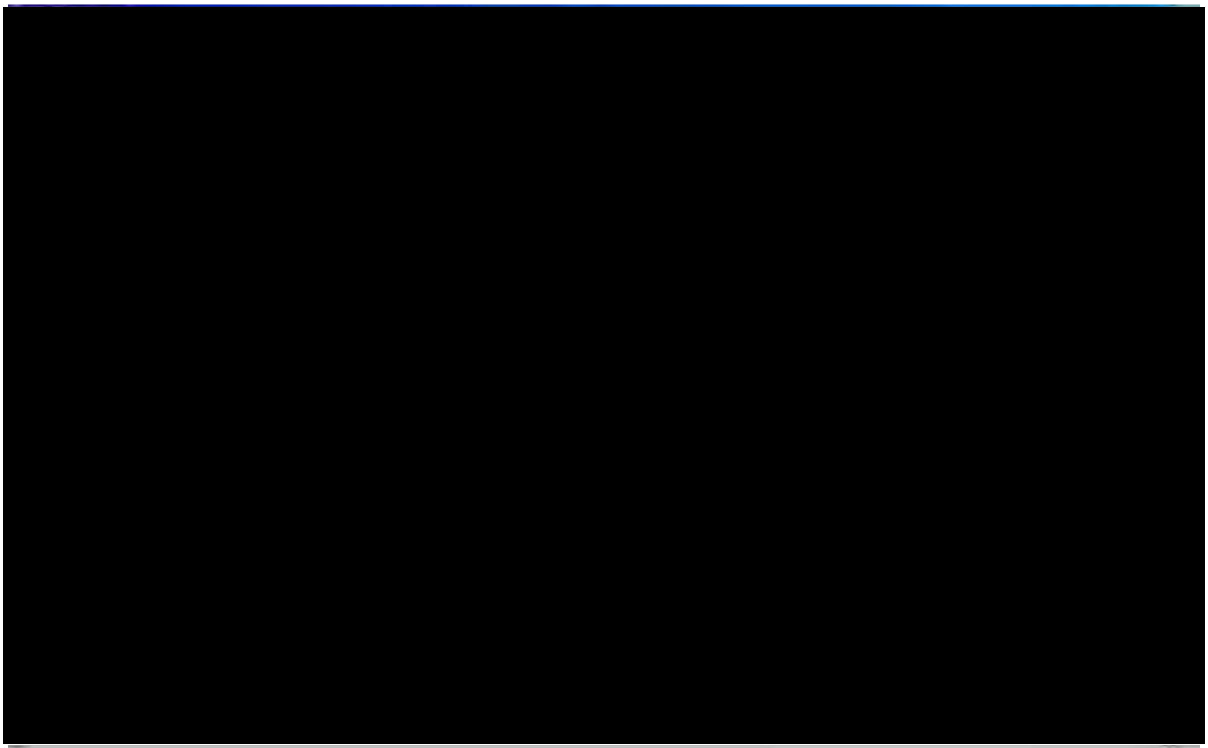

Fig. 3.2 Crytek Developer Kit

(Image from Crytek, Crysis Videogame, 2008)

Whilst Epic Games and Crytek can create detailed virtual environments, both of them charge a commercial development fee. Furthermore, both would require a complete design of virtual maps of an environment that may resemble an operational area alongside the need to develop all the game assets (those not included in the developer kit) using 3D tools such a Maya; 3DS Max or Blend to create these game assets. The Author therefore looked at tools which were free of charge and with more efficient development times (e.g. having pre-created game assets and maps).

3.6.3 America's Army

America's Army (AA) is free to download and develop for (AmericasArmy, 2011). AA was originally developed for the Office of the Assistant Secretary of the Army: Manpower and Reserve Affairs in the United States (US). This videogame was developed for recruitment purposes as it gave potential new recruits an insight into army life. As it became popular, it became freely downloadable along with its development kit. The development kit is based on the UDK. Fig. 3.3 shows examples of the types of environment that can be created using AA as well as an example of game assets provided. Game assets are known as 3D objects in the virtual environment. The game assets in AA have a military orientation and the environment maps in AA resemble areas in which the US Army was operating in. This military focus was of interest to the Author as it could reduce the time need to create a virtual environment that resembled an operation theatre. It also reduced the time to add in new military assets that would be needed.

Fig. 3.3 Images of America's Army Virtual Environment and Developer Kit

(Image from America's Army SDK (PC), 2008)

3.6.4 VBS2

VBS2 is a licensed tool developed by Bohemian Interactive Simulations (BIS). An interactive 3D training tool built on the company's Real Virtual 2 engine and exploited by a range of international Armed Forces, particularly (but not exclusively) for land warfare simulation.

VBS2 offers a rich virtual environment to work from which includes various Military organisations. A big advantage is that is has most of the British Army military included in the package. Fig. 3.4 and Fig. 3.5 are two example scenes developed by the Author and illustrate the level of detail that can be achieved easily using VBS2.

Fig. 3.4 British Army Lynx Helicopter

Fig. 3.5 British Army Patrol in Land Rover Wolf

To some extent VBS2 removes the need to be an expert 3D modeller because numerous assets (e.g. different types of civilians, vehicles, boxes, etc.) are already available in the SDK. This allows the developer to focus on the game play/scenario.

Another advantage of using VBS2 is that anything developed under the UK MOD license is freely available on request. For example, a full Afghanistan Green Zone developed for VBS2 was obtained from Roke Manor Research Limited.

The development aspect of VBS2 uses scripted language or commands to move or trigger events. Fig. 3.6 illustrates a camera script which follows a specific path, in this example the civilian in the red dress.

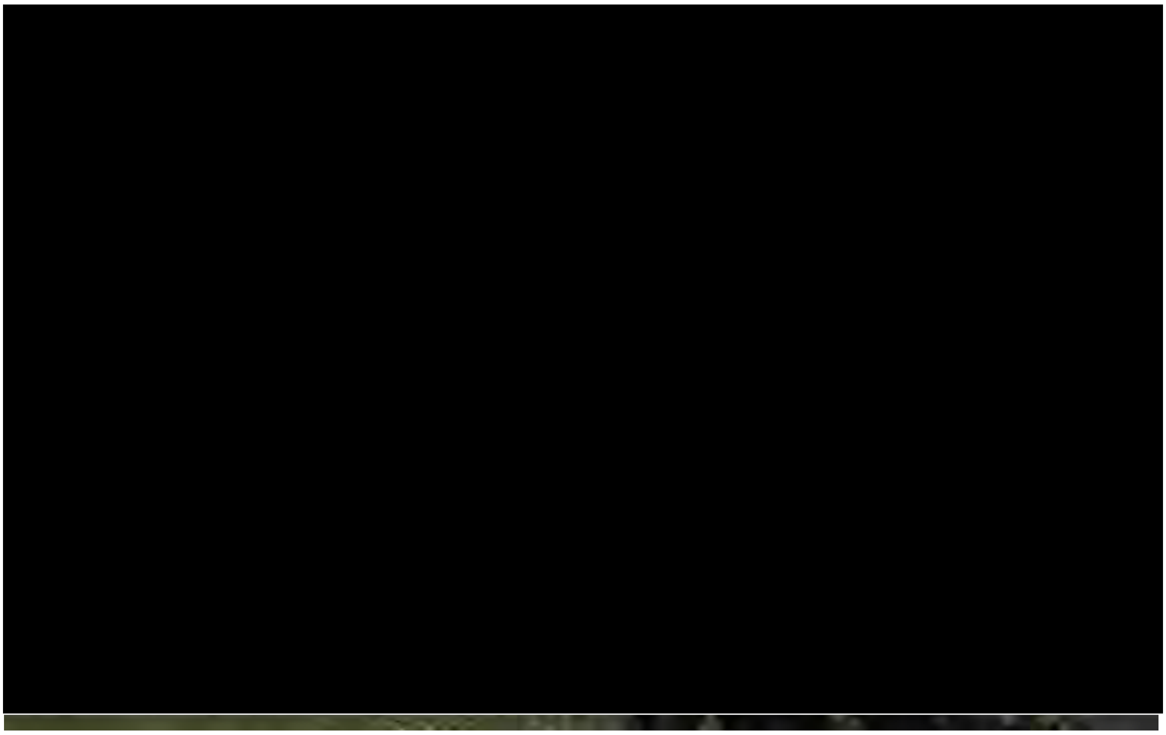

Fig. 3.6 Varied Civilians getting Supplies

Other scripts allow for actions, like smoking, running or talking on a mobile phone, etc. However there are limitations in the level of realism. For example, although the man in Fig. 3.7 looks as if he is looking at something (it is meant to be a mobile phone), the phone is missing even though he acts as if he is using one.

Fig. 3.7 Test Script Man using Mobile Phone

VBS2 also allows easy movement by connecting waypoints together and commanding a character to move along a path. Fig. 3.8 illustrates the developer's environment, the black line is the waypoint in which the character moves along (black dress lady) and the green cylinder is the command (MOVE).

However, there are limitations to this as well because although the line is drawn in a specific place, each character in the virtual world has a collision radius which cannot be seen/changed. This restriction is set by the game engine. Therefore sometime the character may do obscure thinks which the developer has no control over, such as walking over a barrier.

[redacted]

Fig. 3.8 Waypoint towards Supply Area

3.6.5 NetLogo

NetLogo is a multi-agent programming toolkit that delivers a flexible two-dimensional abstracted environment that helps to develop simulations of natural or social events. NetLogo allows the user to program simple agents to execute various rules and goals that can be viewed in real-, slow- or accelerated timeframes via a simple 2D (—plan view”) test environment. The 2D view enables researchers not only to observe micro-level behaviours of individual agents but also supports the visualisation of macro-level patterns of behaviours when individual agents interact with others.

NetLogo has been used to simulate population dynamics, negotiation strategies, even interactions between the civilians, peacekeepers and insurgents. An Australian Defence Science & Technology Organisation DSTO study conducted in 2005 concluded that NetLogo was a —valuable technique for conducting preliminary low fidelity studies, in minimal time and with minimal effort. It has been shown capable of modelling a subset of the physical, social, and behavioural interactions in guerrilla warfare and in that capacity is useful for developing conceptual models and providing insights for further study” (Wheeler, 2005b).

NetLogo uses its own software language to move ‗turtles‘ along an X Y coordinate. The ‗turtles‘ can take various shapes depending on what the developer wants them to be. NetLogo allows the developer to develop rules that the turtles must abide by. As such, NetLogo allows the developer to create different scenarios and test the outcome. (Wheeler, 2005a) used it to see how easily civilians could be persuade to assist guerrilla fighters.

Fig. 3.9 is an illustration of a scenario developed by the Author in NetLogo. The turtles are stickmen (white = civilian, blue = polices, green = soldiers and soldiers and grey = insurgents that detonate SIEDs and kill nearby people). The Author developed rules and used the simulation to see if by adding more polices and soldiers the level of insurgency decreased.

Fig. 3.9 Big Surge vs. SIED

The Main screen in the middle shows the turtle (people) randomly walking around. The sliders allow the user to increase or decrease the population numbers (i.e. 3 police). The graph shows the number of civilians (blue on graph); number of police (dark blue line) and number of soldiers (dark green line) remaining in the population as time elapses. The aim of the insurgents is to avoid the police and soldiers, if they are able to, then their purpose is to kill *n-X* of people within a specific radius.

NetLogo allows numerous iterations of the scenario to be conducted, which for this scenario was either increasing or decreasing the numbers of soldiers/polices until the insurgents could not kill any civilians or the number of civilians killed was reduced.

Fig 3.10 is another illustration of a scenario developed by the Author in NetLogo. This time people need to obtain food from the food aid distribution. Once the food runs out, people within a given radius of those who have food try to steal it from them. As this scenario is run the people change different colours to denote either if they were given food from the food aid, stole the food or have no food.

Fig. 3.10 Food Aid Distributions

3.7 Selection of SDKs for the experiments

The Author selected VBS2 and NetLogo as the SDKs used for, respectively, the higher and lower fidelity scenarios in the experiments.

The choice of tool for the higher fidelity scenarios came down to ease of development and speed for a single developer. VBS2 was chosen because of the links to the MoD and the fact that it gives a good level of fidelity in 3D compared to other software tools whist being free and relatively efficient for a single developer to use.

The other software tool NetLogo was chosen because it was free to use and had a visibly different (lower) level of fidelity to VBS2, allowing for obvious differences between the two to be noticed. NetLogo also allowed the freedom to create several ideas from scratch and test in a reasonably short time.

CHAPTER 5

EXPERIMENT I

5.1 Introduction and Objectives

The aim of the first experiment was to explore an observer's ability to identify the intent of a crowd and whether any individuals are deviating from the norm.

Observers were shown the same scene using two different levels of fidelity and four different crowd sizes. In alternative cases one individual's actions was not in line with the rest of the crowd. The observers were then provided with five possible explanations of what was happening and asked to record how strongly they felt that these were plausible answers. Each explanation was measured using a 10-point scale system, with 1 being not to confident and 10 being highly confident. This scale could then measure the confidence of the observers who when trying to accurately depict what was happening.

The objective of the experiment was to gain an insight into the following:

- Hypothesis linked with thesis question 1: do observers have strong opinion of what is happening in a particular situation?
- Hypothesis linked with question 1b: does the distribution of written and verbal information have a different effect on the observer's opinion?
- Hypothesis linked with question 1 & 2: can observers identify the person who is not acting in line with the crowd?
- Hypothesis in line with question 2: does fidelity make a difference?
- Does the size of the crowd have an impact? [Not linked to thesis question but necessary to tell if crowd size would influence observers]

As access to experts was limited, this experiment was conducted with all naïve observers. As such thesis question 3 was not touched in this experiment.

5.2 Scenario Description and Design

This scenario was based loosely around the earthquake that happened in Haiti in 2010, i.e. individuals collecting food aid. The design of the experiment began by developing two types of briefing: one written and one verbal. The reasoning behind the two brief types was to see if verbal or written briefings altered an observer's confidence or understanding of a scenario. This is based on the way the British Army distributes intelligence, either in a written form or a verbal briefing. The reason for using different types of briefing was derived from speaking to SMEs (see Section 4.2.3).

The purpose of the brief was to provide information about the scenario such as the place, what had happened, the state of the people and that the United Nations had started to bring in supplies to the region. Pictures where provide to show the different states of a food aid distribution both in the written and verbal briefings.

The information pointed towards two facts: that the individuals in the scenario where hungry and that a rule based systems was being set up to allow food aid to be distributed evenly and fairly. To keep bias to a minimum, no information was given about the individual's intent (i.e. that one individual may break the rules). The software tools used to design the computer based scenes were, NetLogo (Fig. 5.1), and VBS2 (Fig. 5.2).

The area in NetLogo consisted of a black background with coloured squares. VBS2 had a higher fidelity with a texture area for the background (consisting of grass, sand and walled areas) and it had a 3D perspective.

Fig. 5.1 NetLogo food aid scenario with 3 individuals

Fig. 5.2   VBS2 food aid scenario with 3 individuals

The design created a rule based system for the individuals in the virtual environment to follow. The first rule was to have the individual collect a food ration coupon in one area (white square in NetLogo (Fig. 5.1), far right walled area in VBS2 (Fig. 5.2)). The next rule was for the individual to proceed to another area where they can exchange the ticket for food/provisions (purple square in NetLogo (Fig. 5.1), and nearby walled area with food sacks in VBS2, (Fig. 5.2)) before finally returning to the ‗base‘ location where they started from.

Squares and walled areas replicate different state changes in the rule system, each time an individual enters these areas their colour change to denote a change in the state and that the individual is following the rules correctly. The colours are as follows:

- In NetLogo the individual were simple circles with a unique identifier number. Colour changes were used to indicate the different states of the individual (Fig. 5.3).
  - Blue = empty handed
  - Red = has food ration coupon
  - Green = has food
- In VBS2 the individuals were 3D. The colour of their clothes changes used to indicate the different rule changes of the individual (Fig. 5.4)
  - Black = empty handed
  - Blue = has food ration coupon
  - Red (carrying parcel) = has food

The design also considered variations in crowd sizes for both NetLogo and VBS2. The different crowd sizes consisted of 3, 6, 9 and 12 individuals.

Another four crowd sizes where produced using the same numbers, but including one individual who did not participate in the food collection. This individual does not follow any rules (breaking the ‗norm‘) but waits at the ‗base‘ until another individual has obtained the food/provisions and returned to ‗base‘. The individual breaking the ‗norm‘ then steals the food/provision from the other individual. The theft is shown through changes in status in both the thief and victim. The Individual who steals in NetLogo is depicted as a circle. In VBS2 it is depicted as a man wearing a blue trouser and red top. Fig. 5.3 and Fig. 5.4 show the differences between NetLogo and VBS2.

Fig. 5.3 NetLogo with 4 individual 1 who steals (top left corner)

[redacted]

Fig. 5.4 VBS2 with 4 individual 1 who steals (top left corner)

Squares and walled areas replicate different state changes in the rule system, each time an individual enters these areas their colour changes to denote a change in the state and that the individual is following the rules correctly.

The virtual environments were designed using VBS2 tool (high fidelity) and NetLogo tool (lower fidelity) to create the differences in fidelity. The virtual environments were then made into a series of static images. The static images were put together into slides using Microsoft Power Point for VBS2 and Microsoft Image Viewer for NetLogo. The static images on the slides were then presented as images moving from one state to another. The whole slide show showed a complete process from getting the food ration coupon, getting food/provision, and returning safely to the ‗base' (originally starting point). Fig.5.5 and Fig. 5.6 show an example of the different slide shows. Only one visual perspective was used to keep the process simple (Projected slides in front of a class using the projector and laptop running Windows Vista)

[redacted]

Fig. 5.5 Example slide sequence NetLogo

[redacted]

Fig. 5.6 Example slide sequence VBS2

Along with the virtual environment static images, observers were given five different options. Observers were given different types of briefings and were required to rate their confidence in each option if they felt it was a viable explanation using a scale of 1 to 10. The observers could choose more than one option. Table 5.1listls the options. These are referred to as “text options” in the analysis of experiment results.

Table5.1 Options Choices

Options	Description
1	Individuals are moving from the city towards the UN run vaccination centre. The white square is the first stop to get the vaccination and then towards the magenta square for proof of vaccination. Once they have been confirmed they move back to the city. The colour change shows the process.
2	Individuals are moving from the city to a safe haven, the white square set up by the UN. They have then been forced from the white square because of a potential aftershock so the individuals move to the magenta square where they are told that it is not safe so they move back. The colour change shows emotional state.
3	Individuals are moving from the city towards the UN white square where ration coupons are being provided. Once the coupon has been received the individual then moves to the magenta square to get some provisions. Once provisions have been obtained, they return back to the city. The colour change shows the process.
4	Individuals are moving from the city in search of provisions. The UN white square is where the individuals storm first. Realising there is no provision there; the individuals move to the magenta square and storm again. On realising that there are no provisions they return to the city. The colour change shows emotional state.
5	Iindividuals are moving from the city towards the UN white square where ration coupons are being provided. Once the coupon has been received the individual then moves to the magenta square to get some provisions. But on leaving the magenta square an individual has not followed the correct rules and has mugged an individual with provisions. The colour change shows the process.

The only difference to the options for VBS2 is the referral to squares. The squares in VBS2 are actually walled areas. So in VBS2 the first square is referred to as the ‗first walled area' and the second square is referred to as the ‗second walled area'. The purpose of the options was to measure observer's confidence when viewing the static images. The confidence measurements could then be used to answer the thesis question.

The correct option for when the crowd of individuals following the correct rule based system (‗norm')was option 3. The information provided in the brief give subtle queues which was hope to ted the observers responding and giving a high rating level of confidence in option 3. When an individual broke from the ‗norm' and stole from an individual option 5 was the correct answer. Option 5 was the only option given to the observers that mention an individual not conforming to the rules.

5.3 Experiment Setup and Running

The experiment was conducted with two separate groups, one group who read the written briefing and the other group who were given the briefing verbally. Each group were split into either the NetLogo scene or the VBS2 scene (not both). The total numbers of observers was 27. The groups spilt between written and verbally briefing and between NetLogo and VBS2 are as followed:

- NetLogo, written brief = 6 observers;
- VBS2, written brief = 7 observers;
- NetLogo, verbal brief = 8 observers; and
- VBS2, verbal brief = 6 observers.

The observers were given 10 minutes to read through the brief and were allowed to ask questions pertaining to the brief only. The observer where then given a further five minutes to read through the five different options provided. The observers could ask questions pertaining only to the five choices. The same times where given to those observers who got the briefing verbally.

The first run of the experiment was a dry run to get observers familiar to the scenario and the process of selecting from the five options. The dry run slide show ran for approximately 30 seconds. A further 120 seconds was given to the observers to select from the five options and a rate their confidence using a 10-point scale system. Questions could be asked again to ensure that the observers understood the experiment process.

Once it was clear that all observers understood the dry run, the experiment was conducted a further six times. Each run was randomly selected from the four types of crowd sizes. Three runs would include the ‗norm' where no individual trying to steal food/provision and three runs would include the deviation from the ‗norm' with an individual trying to steal food/provisions. Each run lasted approximately 150 seconds. The test runs are per Table 5.2 below.

Table 5.2 Experiment Runs for Written and Verbal Brief

Run Number	Test Number	NetLogo	Test Number	VBS2
One	Test 2	6 Individuals following the rules to get food/provisions	Test 6	6 Individuals following the rules to get food/provisions
Two	Test 4	12 Individuals following the rules to get food/provisions	Test 3a	3 Individuals behaving normal following the rule system. 1 individual stealing the food/provisions
Three	Test 5	3 Individuals behaving normal following the rule system. 1 individual stealing the food/provisions	Test 9a	9 Individuals behaving normal following the rule system. 1 individual stealing the food/provisions
Four	Test 8	12 Individuals behaving normal following the rule system. 1 individual stealing the food/provisions	Test 12	12 Individuals following the rules to get food/provisions
Five	Test 6	6 Individuals behaving normal following the rule system. 1 individual stealing the food/provisions	Test 12a	12 Individuals behaving normal following the rule system. 1 individual stealing the food/provisions
Six	Test 3	3 Individuals following the rules to get food/provisions	Test 9	9 Individuals following the rules to get food/provisions

5.4 Experiment Results

The data was analysed using Excel and then statistically analysed using IBM SPSS Statistics Data Editor Version 19.

5.4.1 Observers opinions

The data gathered from experiment I was first analysed as a whole to develop a general baseline and see what confidence levels observers had overall in the five options. Taking an overall approach provides an understanding of whether any option(s) had a dominant presence.

The results described below are for the combined data from the written and verbal briefings, as well as the data with and without an individual breaking the rules (i.e. all six runs). The graphs below cover an overall understanding of what the level of confidence observers had when rating the options (means confidence) and how strong the consensus was of the observers (stand deviation error bars).

5.4.1.1 Combined data

Fig. 5.7 shows the total confidence ratings given by all observers to all five options over the six scenario runs, combining the totals from NetLogo & VBS2. The maximum confidence level possible for any option is 1620. To achieve a value of 1620 all 27 observers (fourteen observers in NetLogo and thirteen observers for VBS2) would have to have given a confidence rating of 10 for each option, in each of the six runs. The mean average of all five options is a confidence value of 586.2, which is equivalent to 36.1 percent of the maximum confidence level possible.

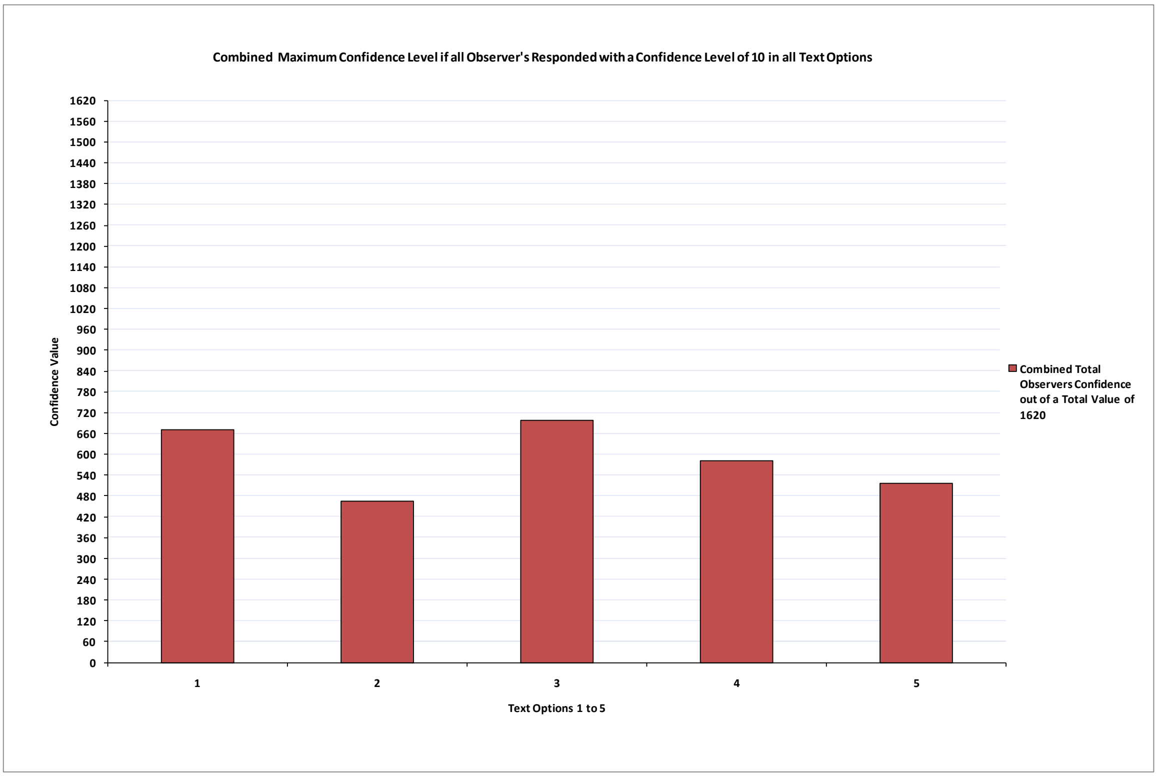

Fig. 5.7 Combined (NetLogo & VBS2) maximum confidence rating value of all options

Fig. 58 illustrates the mean confidence of the combined data from both NetLogo and VBS2 alongside the standard deviation for each of the options. This includes all 27 observers and all six runs of the experiment. 60 is the total maximum value that each option can be given when combining a maximum score of 10 through all of the six runs. The total mean confidence for all five options is 21.7 which is 36.1 percent confidence. The standard deviation is shown as error bar in Fig. 5.8. The error shows that the observer's confidence had a similar level of deviation when rating all five of the options. However the deviation overall mean average value is 12.5 which is high and indicates a low level of consensus amongst the observers.

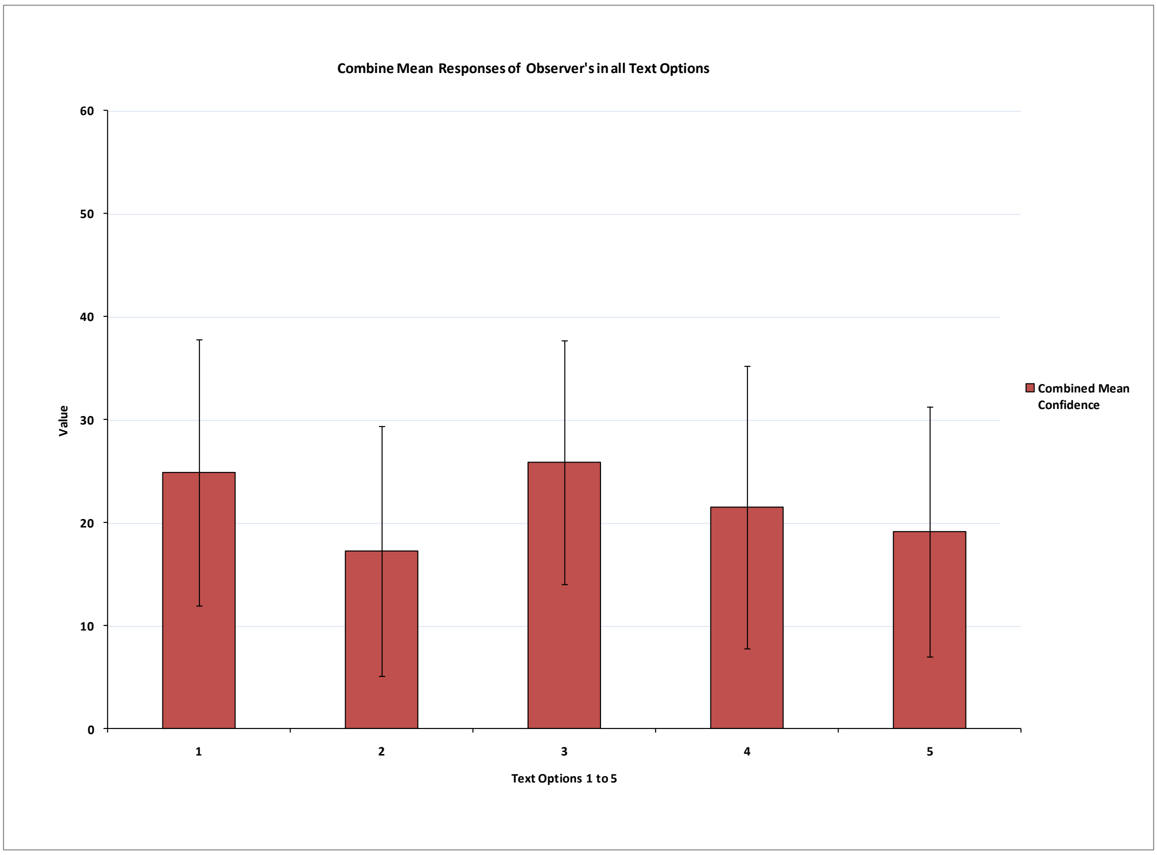

Fig. 5.8 Mean confidence for all five options with SD error bars (Combined)

5.4.1.2 NetLogo Data

Having analysed the results overall, the second step was to look at results for NetLogo and VBS2 separately. The results below are for the NetLogo software tool only and included the combined data of the written and verbal briefing. These graphs display the overall result of observer's confidence.

Fig. 5.9 shows the total response to all five options over the six scenario runs for NetLogo. The maximum confidence value for each option in NetLogo is 840. This is calculated by taking all 14 observers multiplied by six runs each with a maximum rating of 10. The mean average of all five options for NetLogo gives a mean confidence value of 293.5, which is equivalent to a 34.9 percent confidence rating for NetLogo.

This rating was lower than the 31.6 percent confidence rating of NetLogo and VBS2 combined which suggests that overall NetLogo produced less confident responses than VBS2.

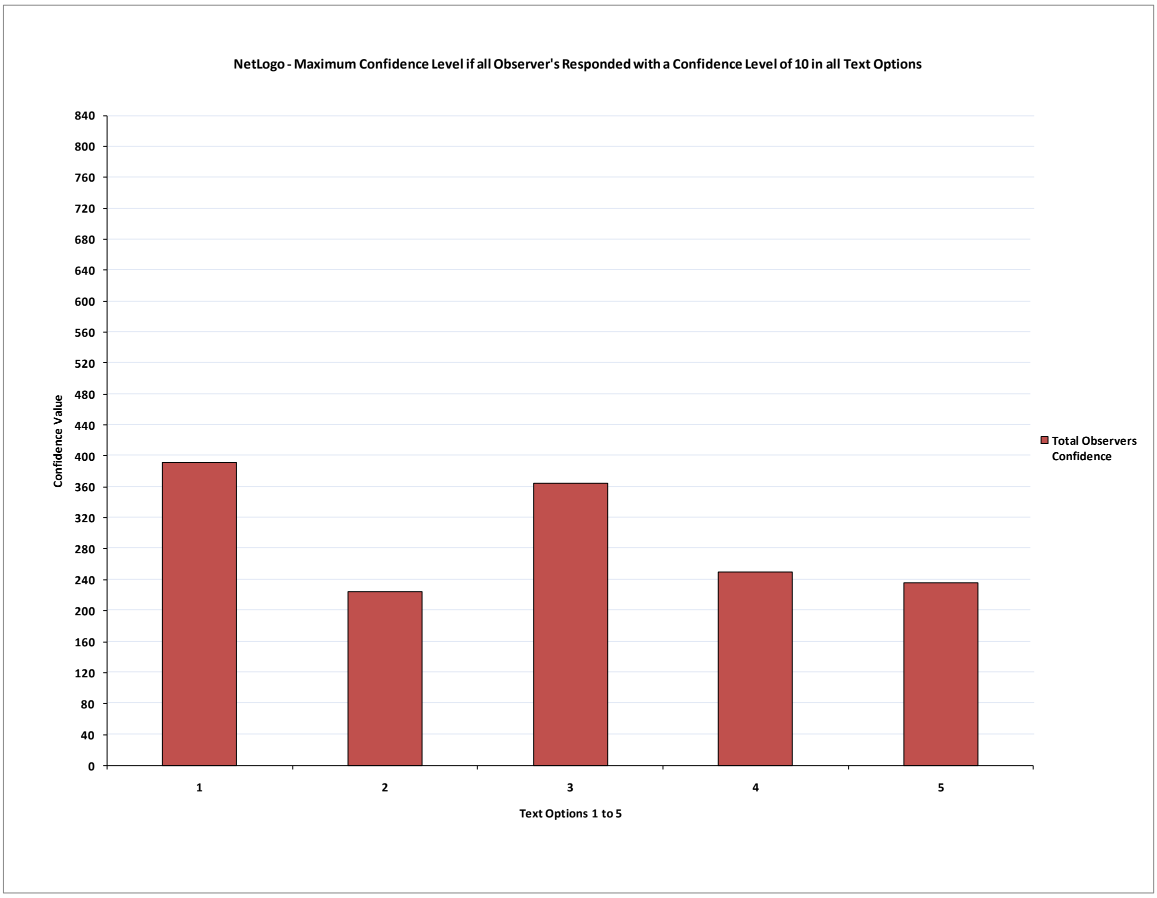

Fig. 5.9 Maximum confidence rating value of all five options for NetLogo

Fig. 5.10 is the mean confidence of NetLogo alongside the standard deviation (shown as the error bar). 60 is the total maximum value that each option can be given through all of the six runs. The mean confidence for all five options is 20.7 which is only 34.9 percent confidence. The standard deviation varied between the options. For example, option 4 showing a higher deviation compared to option 1. However, regardless of the variation, the overall mean standard deviation value is 11.9 which, whilst lower that the combined NetLogo and VBS2 value, is still high.

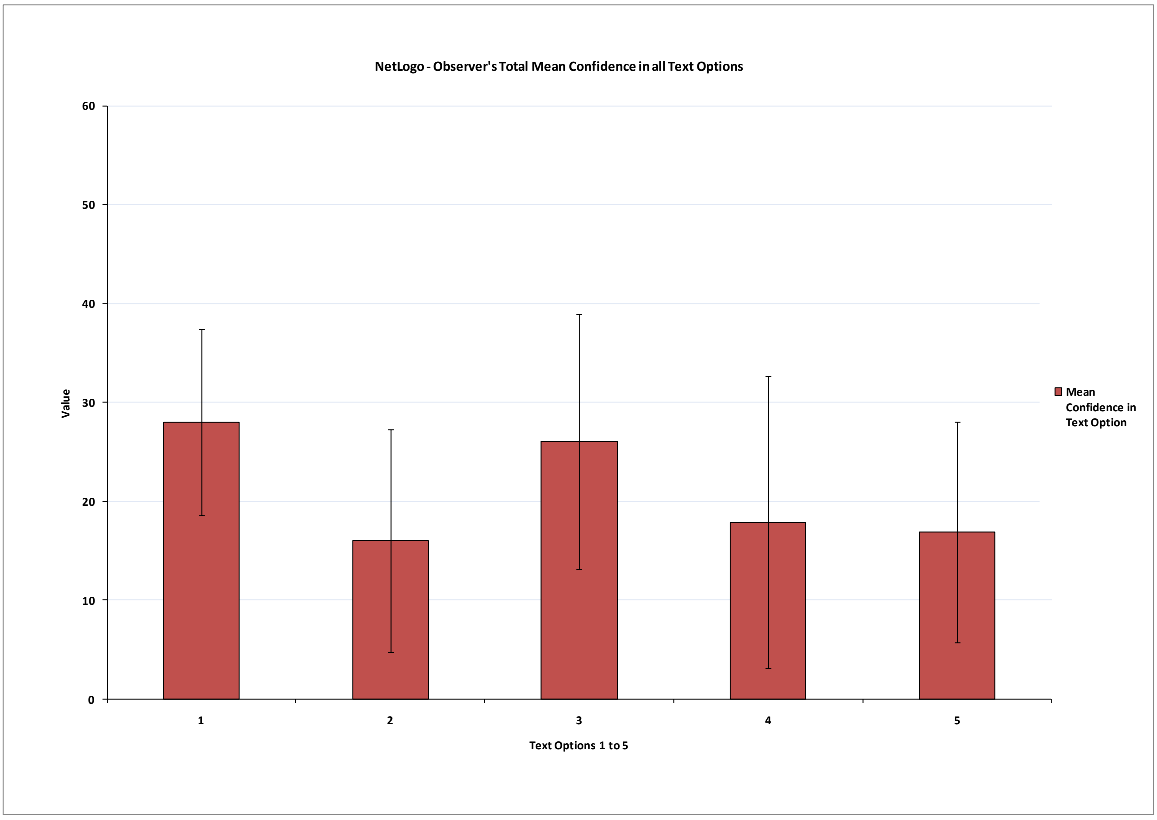

Fig. 5.10 Mean confidence rating for all options with SD error bars (NetLogo)

5.4.1.3 VBS2 Data

The results below are for the VBS2 software tool only and included the combined data of the written and verbal briefing. These graphs display the overall result of observer's confidence.

Fig. 5.11 shows the total response to all five options over the six scenario runs for VBS2. The total confidence value is 780, which was calculated by multiplying the total number of observers (13) with the number of runs (6) and maximum confidence rating (10). The mean average of all five options for VBS2 gives a mean confidence value of 292.7, which is equivalent to a 37.5 percent confidence rating overall. The VBS2 confidence rating is higher than that achieved for NetLogo.

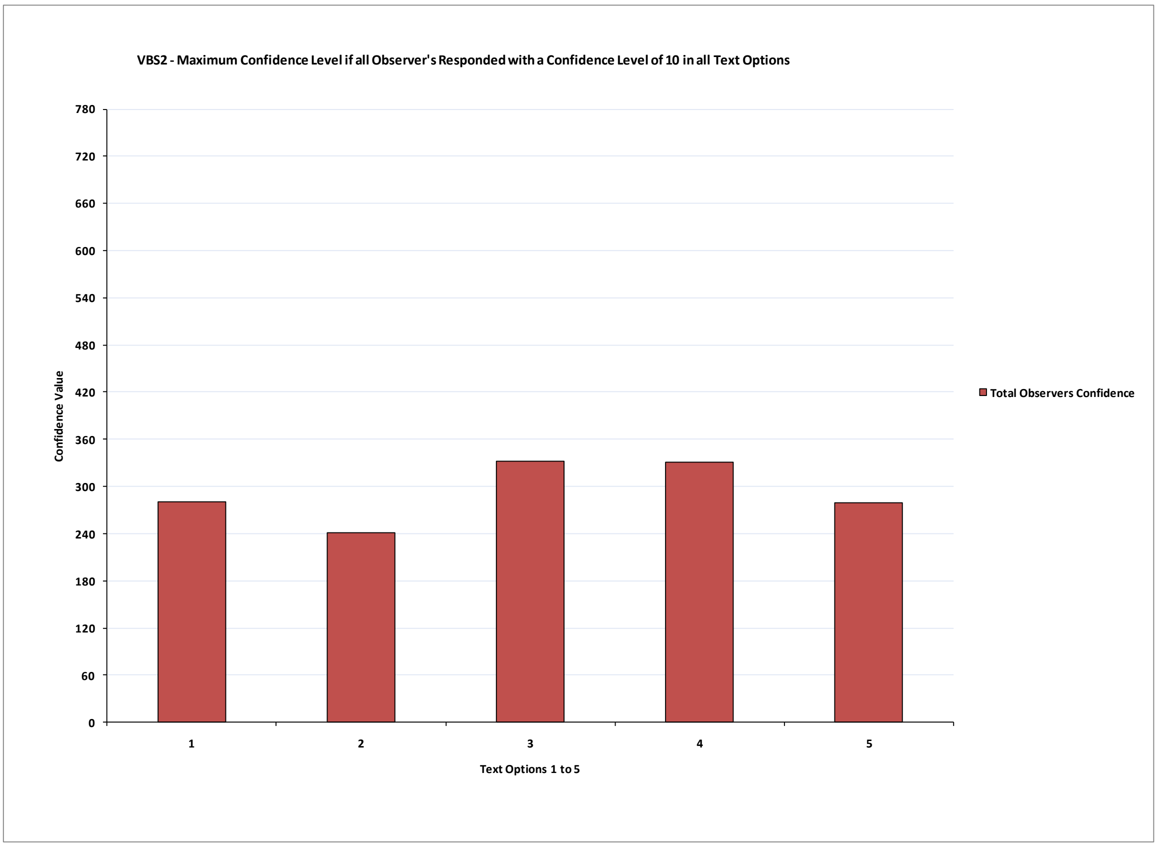

Fig. 5.11 Maximum confidence rating value of all options for VBS2

Fig. 5.12 is the mean confidence and standard deviation for VBS2. 60 is the total maximum value that each option can be given through all of the six runs. The total mean confidence for all five options is 22.5 which is only 37.5 percent confidence. Standard deviation is shown as error bar in Fig. 5.12. The deviation overall mean average value is 13.0 which is higher than for NetLogo. The lowest deviation is for option 3, which is one of the two correct options.

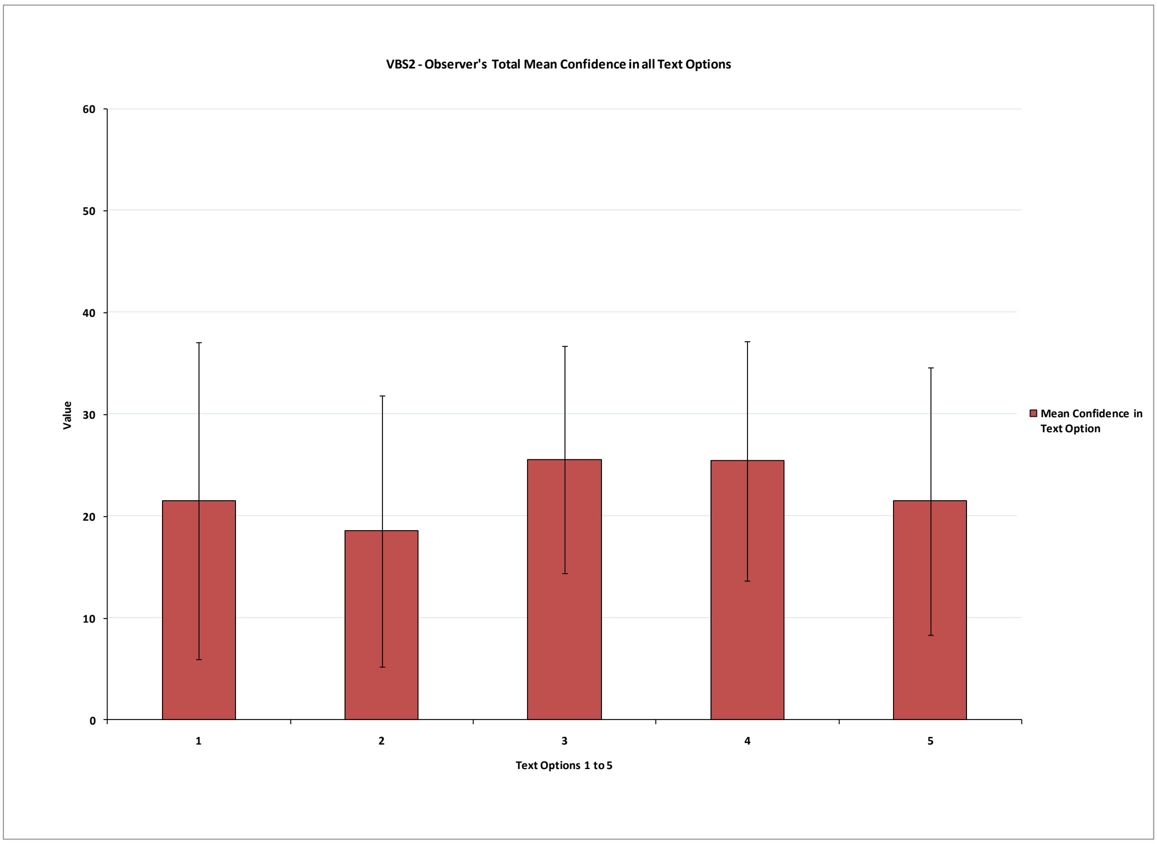

Fig. 5.12 Mean confidence for all options with SD error bars (VBS2)

5.4.1.4 Comparison of data (NetLogo & VBS2)

The following result below display a comparison between the NetLogo and VBS2 results described above.

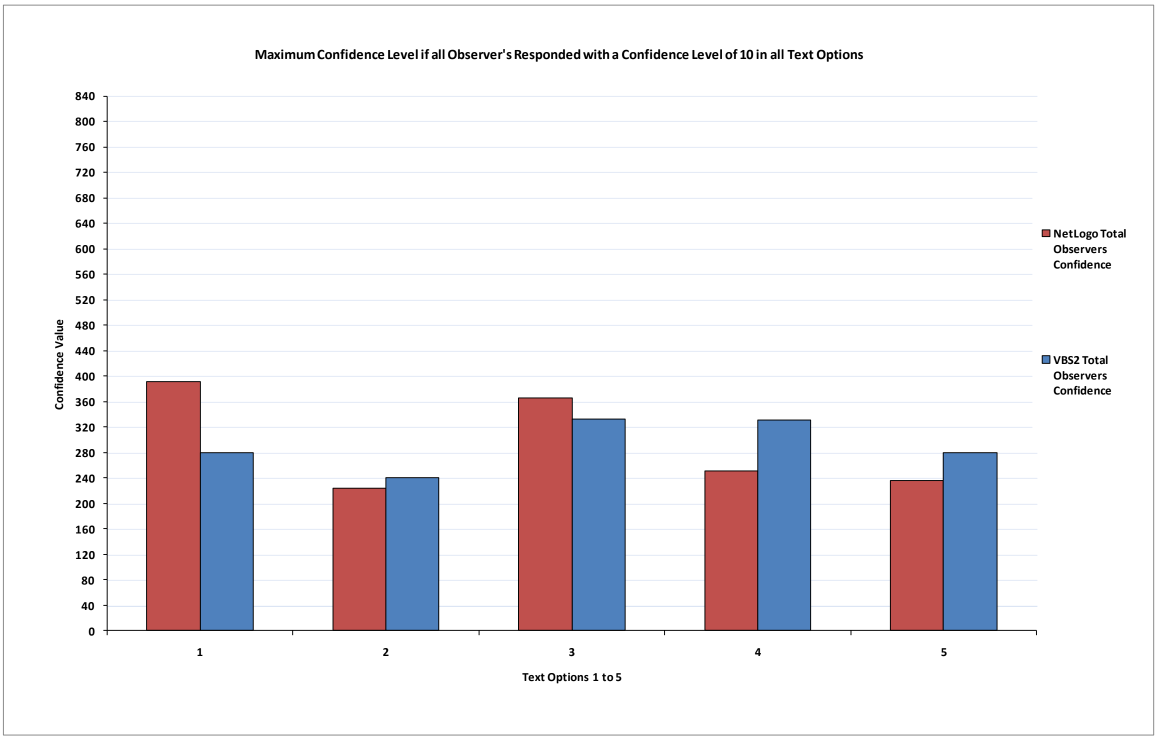

Fig. 5.13 Maximum confidence rating value of all five options (Compared)

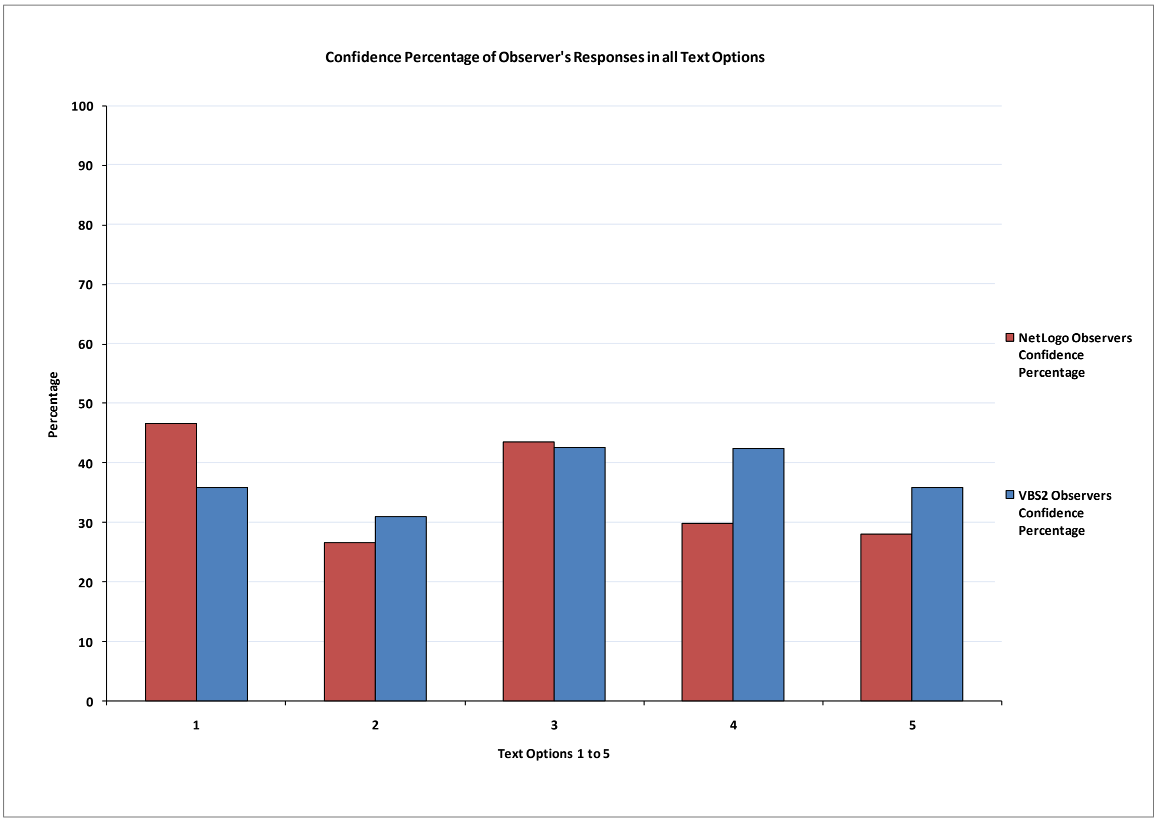

Fig. 5.14 Confidence percentage for each option (Compared)

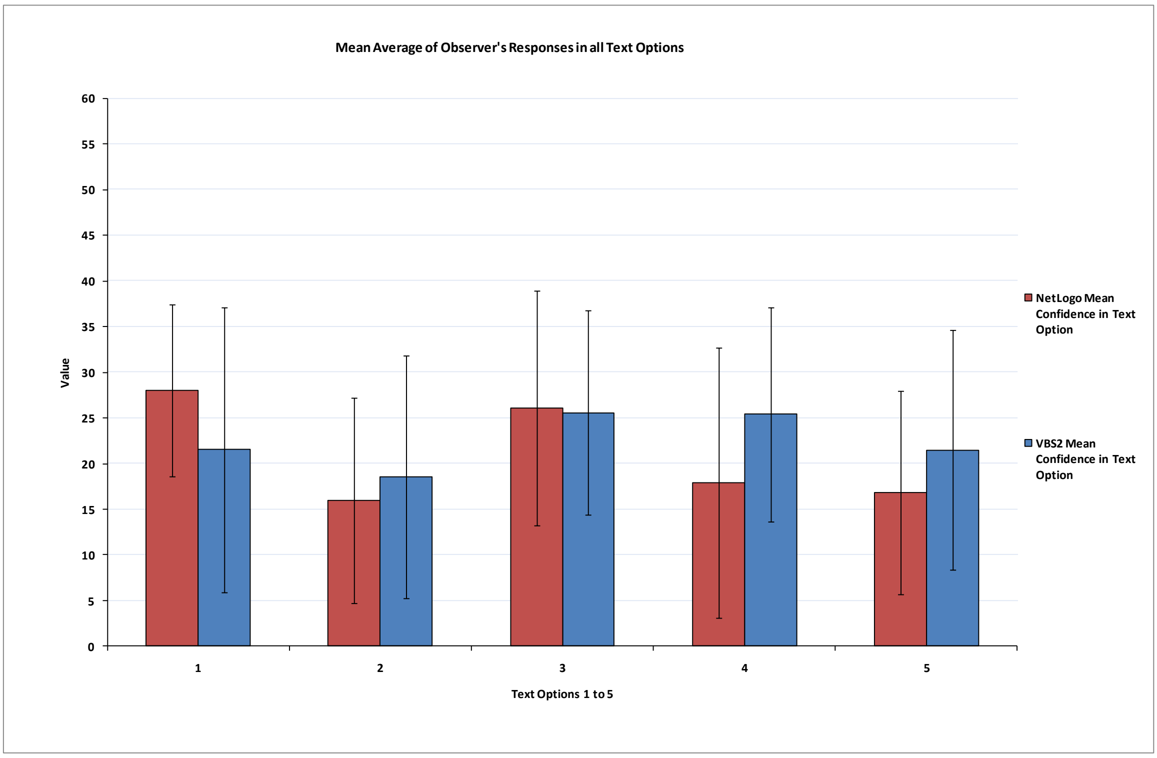

Fig. 5.15 Mean Confidence for all five options with SD error bars (Compared)

5.4.1.5 Observations and ANOVA significant main effect

Looking at the results described above in the context of thesis question 1, the Author feels that observers could not confidently distinguish what was happening in the virtual environment and they did not display a strong consensus.

The data show that confidence levels in all five options were considerably low and that, whilst there was some variation, no one option clearly stood out from the others. It also shows that the consensus between the observers was also low; as such the standard deviation was high in all five options. This would suggest that the observers did not have a clear consensus on the five options.

The comparison graphs also indicate that for some of the options there are differences in the NetLogo and VBS2 confidence ratings. Fig. 5.15 shows that options 1, 2, 4 and 5 show a visible variation in confidence levels between the two difference software tools (fidelity). Despite the fact that low confidence levels were reported in all of the options, these differences between the two software tools could suggest that fidelity may play a part in an observer‘s perception and confidence levels.

To obtain further evidence of a possible relationship the Author analysed the data using a two-way mixed ANOVA design between-subjects and within-subjects. This analysis would show if fidelity had any significant main effect between the two software tools.

The two-way measurements between-subjects would compare NetLogo five options with VBS2 five options similar to Fig. 5.15. However, the results showed there is a non-significant main effect between NetLogo and VBS2 [F (1, 25) < 0.669, ns].

In the within–subjects measurement there is a significant main effect when comparing between all five options [F (4, 100) = 4.013, p < 0.005] which suggest that there is little consensus in confidence ratings between all five options.

There is non-significant main effect on the interaction between options and fidelity (NetLogo and VBS2) [F (4, 100) < 0.075, ns]. This could suggest that there is a similar level of confidence between the options in both fidelities. In other words there little difference between NetLogo‘s five options and VBS2‘s five options (as shown in Fig.5.15).

5.4.2 Distribution of information

Having looked at the overall data, and different software tools, the next step was to determine if the distribution of information increased an observer‘s confidence. Did the provision of written briefings result in meaningful differences in results compared to observers who were given the information through verbal presentation?

This is to hypothesise that giving information by verbal or written means has a different effect on the observer‘s confidence rating for the five options during. For example, could the information increase the observer‘s confidence and understanding of the virtual environment? It is hoped that it would also lead the observers to selecting the correct option (3 or 5). Option 3 was the correct option when individuals in the crowd followed the rules to obtain food aid, option 5 was the correct one where an individual broke the rules and stole the food aid from another individual.

5.4.2.1 Comparison of different brief types

Fig. 5.16 compares the differences in confidence between the verbal briefing and written briefing using NetLogo. The confidence rating across all six runs is compared to the two brief and the five options. The mean confidence across all five options for verbal briefing was 20.7, which is equivalent to 34.5 percent. The written briefing mean confidence rating across all five options was 21.3, which is 35.5 percent. Fig. 5.16 shows considerable difference between option 1 and 5. The standard deviation error bars are high across all five options, giving an average standard deviation of 11.9 for verbal and 10.3 for written. This would suggest that the observer‘s confidence across the two briefing types had a similar level of deviation, which means that the spread of confidence is very large.

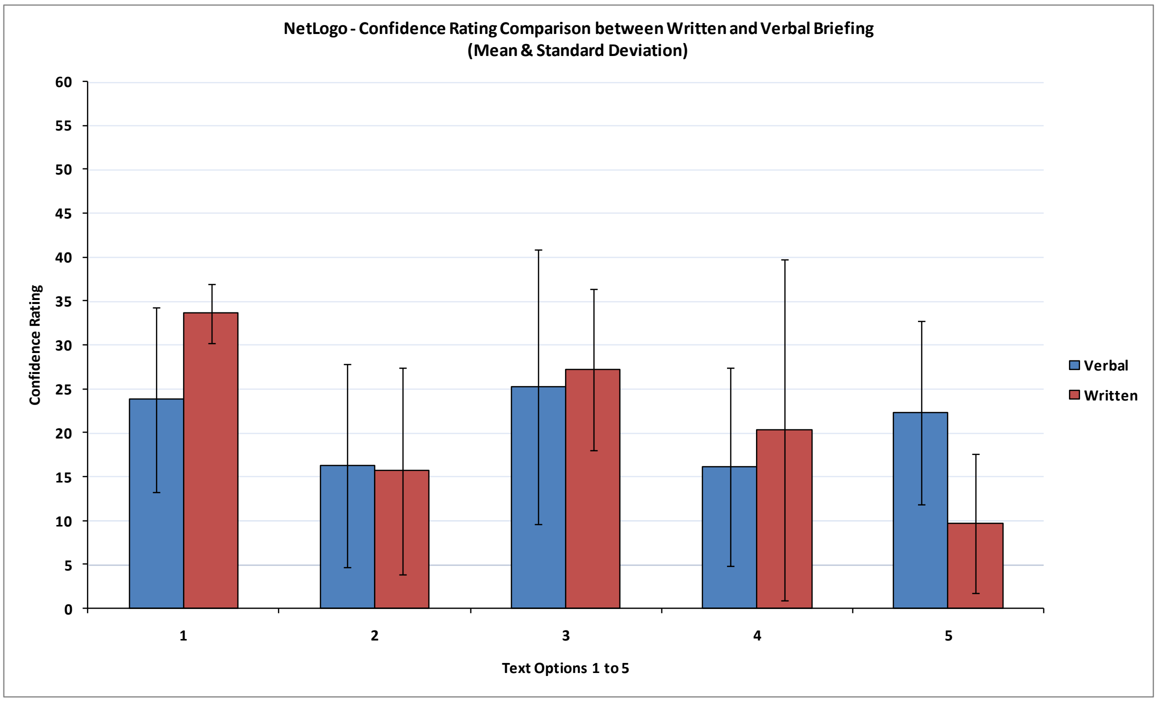

Fig. 5.16 NetLogo comparison between written & verbal briefing

Fig. 5.17 compares the results of the verbal briefing and written briefing using VBS2. The mean confidence across all five options for verbal briefing was 21.4, which is equivalent to 35.6 percent. The written briefing mean confidence rating across all five options was 23.5, or 39.1 percent.

The Fig. 5.17 shows considerable difference between verbal and written briefings for option 1 and 5, which is similar to NetLogo Fig. 5.16. The standard deviation error bars are high across all options. The average standard deviation across five all option was 11.9 for verbal and 13.5 for written. Again this is very similar to NetLogo and would suggest that the observer‘s confidence had a similar level of deviation across both briefings.

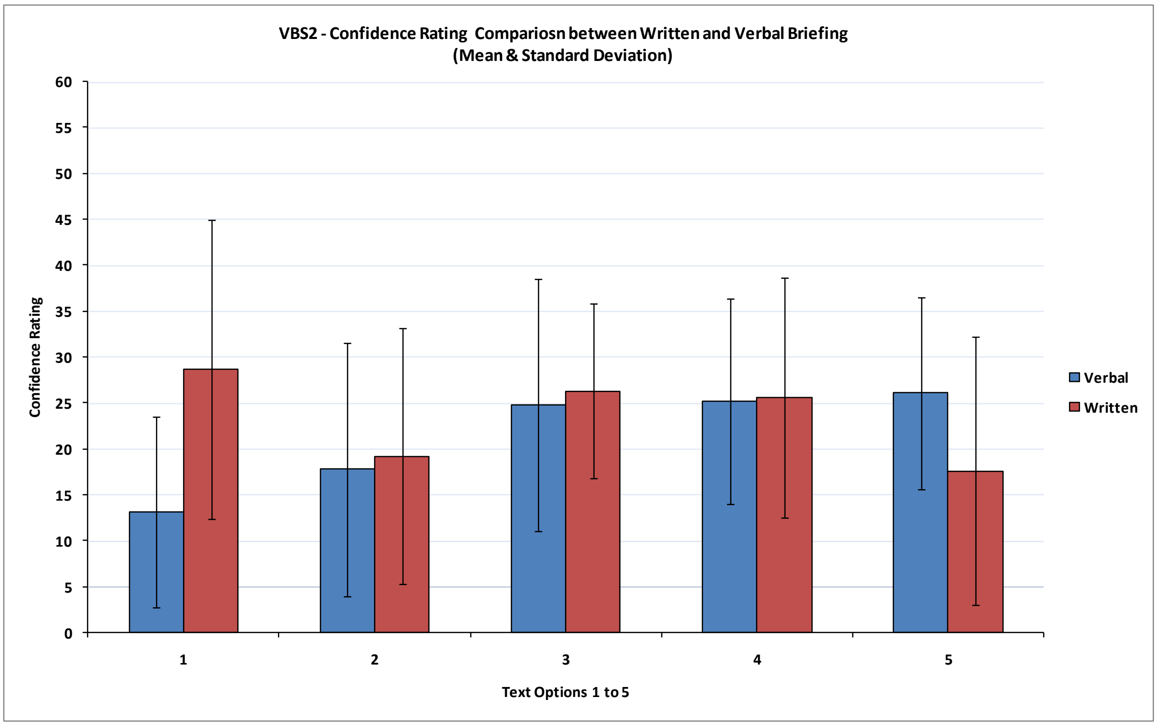

Fig. 5.17 VBS2 comparison between written & verbal briefing

Fig. 5.18 is a comparison between the two software tools (NetLogo and VBS2) and between written and verbal briefings. The written brief for both software tools suggest that when averaging all the means the high overall mean confidence is in the written brief for both VBS2 and NetLogo. The written briefing for VBS2 had an overall mean confidence rating of 39.1 percent whereby NetLogo an overall mean confidence written brief 35.5 percent for the written brief.

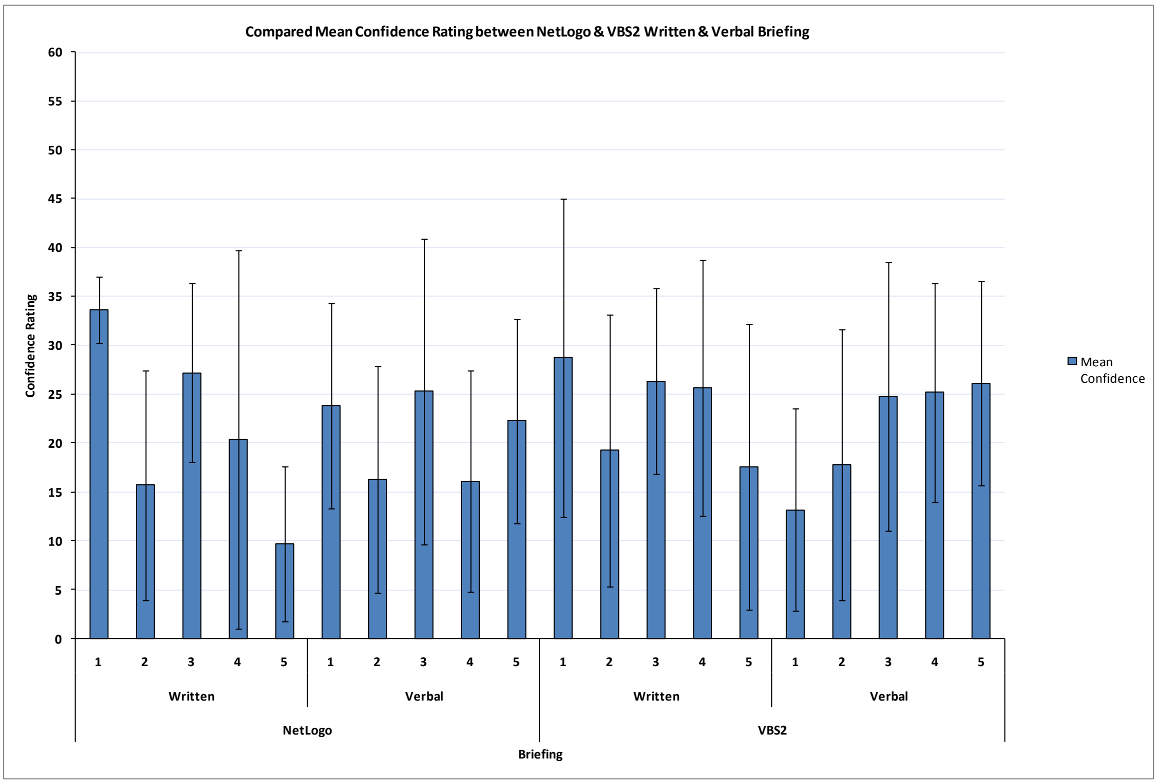

Fig. 5.18 Comparison between NetLogo and VBS2

5.4.2.2 Observations and ANOVA significant main effect

To find any significance in the way information was given to the observers prior to running the experiment, the data was analysed using three-way mixed ANOVA design measuring the data between subjects NetLogo with VBS2 (fidelity), written and verbal briefing (brief types) and the five options.

The first measurement result was for between-subjects when comparing the two fidelities NetLogo with VBS, this is the combination of all the data for NetLogo and VBS2 returns a non-significant effect between the two fidelities [F (1, 23) < 0.707, ns].

The second measurement result was between verbal and written brief not including fidelity. The comparison between the total data of written and verbal briefings returned a non-significant effect between the two briefs [F (1, 23) < 0.726, ns].

The final measurement measures to see if there is any significant interaction between the two fidelities and the types of briefings. The results suggest that there is a non-significant effect between fidelity and the brief types [F (1, 23) < 0.839, ns].

When measuring within-subject this measurement compares the factors, fidelity, brief types and the five options. The first measurement is the total data of NetLogo and VBS2 five options which suggests that when combining the data of both fidelities there is differences between the option (as a total) [F (4, 92) = 4.878, p < 0.001]. This finding is backed up with a Pairwise comparison of the five options. Pairwise measures one option with the other four and does this through all five (see Appendix A)

The measurement is then broken down into comparisons with other factors. The five options are compared with the fidelities. The comparison of the interaction between the options and fidelities does return a significant main effect between them. [F (4, 92) = 6.972, p < 0.017]. What this suggests is that by removing the type of briefing from the equation, there is a difference between VBS2 and NetLogo's five options.

Removing the fidelity from the measurement and measuring between the five options and the two briefing types then there is a significant main effect between them [F (4, 92) = 6.017, p < 0.000].

The final measurement was the interaction between the fidelities (NetLogo & VBS2), the briefing types (written/verbal) and the five options. This is now comparing the three factors to see if there is a significant effect between them all. This measurement returned a non-significant result [F (4, 92) < 0.868, ns]. This suggest that, as with the illustration of Fig. 5.19, there is no difference between NetLogo, VBS2, written, verbal, briefing type and the five options.

5.4.2.3 Conclusion

Analysing Fig. 5.18 shows that the two highest confidence ratings for each brief type in Table 5.3 are.

Table 5.3 Highest Mean Confidence for NetLogo and VBS2

Fidelity	Rank	Briefing Type	Option	Mean Confidence
NetLogo	1	Written	1	33.6
NetLogo	2	Written	1	27.7
VBS2	1	Written	1	28.7
VBS2	2	Written	3	26.7
NetLogo	1	Verbal	3	25.3
NetLogo	2	Verbal	1	23.8
VBS2	1	Verbal	5	26.1
VBS2	2	Verbal	4	25.2

What can be taken from this is that there are differences when removing the factor of the briefing types out of the equation. The above bullet points show that the highest two options for both NetLogo and VBS2 are options 1 and 3. Therefore it is clear that the written briefing had the same effect on those observers who viewed the NetLogo to those who view the VBS2. Also there must be something in option 1 and 3 to this cause similarity. Furthermore, the Author noted that option 5 did not make it into the top two options.

Verbal briefing does in some way have an effect on observer confidence as this alters slightly, although not significantly enough as the measurement have shown. In NetLogo a shift was observed from option 1 to option 3 as being the highest. This could suggest that the verbal briefing offered clarity to observers who now saw option 3 as the correct choice. However, if this suggestion is correct, then option 5 should also have been in the highest two, and this does not appear in the NetLogo results. The big difference appears with VBS2 which has the two high confidences rating in the correct options 5 and the second highest being option 4. This could suggest two things: that the fidelity helped the observers notice someone breaking the rules and or that the brief type helped clarify the five options.

The next analysis will break down the data further to build on the above findings. The test run will be broken into two groups of three: three runs with an individual breaking the rule, and three runs without.

5.4.3 Identifying correctly someone who is breaking the rules

During the running of the experiment cases where an individual was breaking the rules were inserted in half the runs. As such the data is split into two groups: *‗Individual breaking the rules',* which was three runs, and *‗Individual not breaking the rules',* which was also three runs, making a total of six runs.

The data gathered were analysed to try and answer thesis questions 1 and 2 i.e. can observers distinguish when someone is breaking the rules, and if so, does the fidelity of the software being use have an effect? To try to answer this, the mean confidence rating of each option was analysed to see if any significant main effects could be found or whether a strong consensus could be seen in the graphs. The data analysis first looks generally at NetLogo vs. VBS2, and then separates the brief types (written & verbal) and fidelities types (NetLogo & VBS2) to compare if any difference can be found between them.

5.4.3.1 Individual breaking the rules NetLogo vs. VBS2

Results in the graph below display the three runs which had an individual breaking the rules for both NetLogo and VBS2. The Author hoped to find option 5 for both NetLogo and VBS2 displaying the highest confidence level as this was the ―correct" answer. However, Fig 5.19 did not display this.

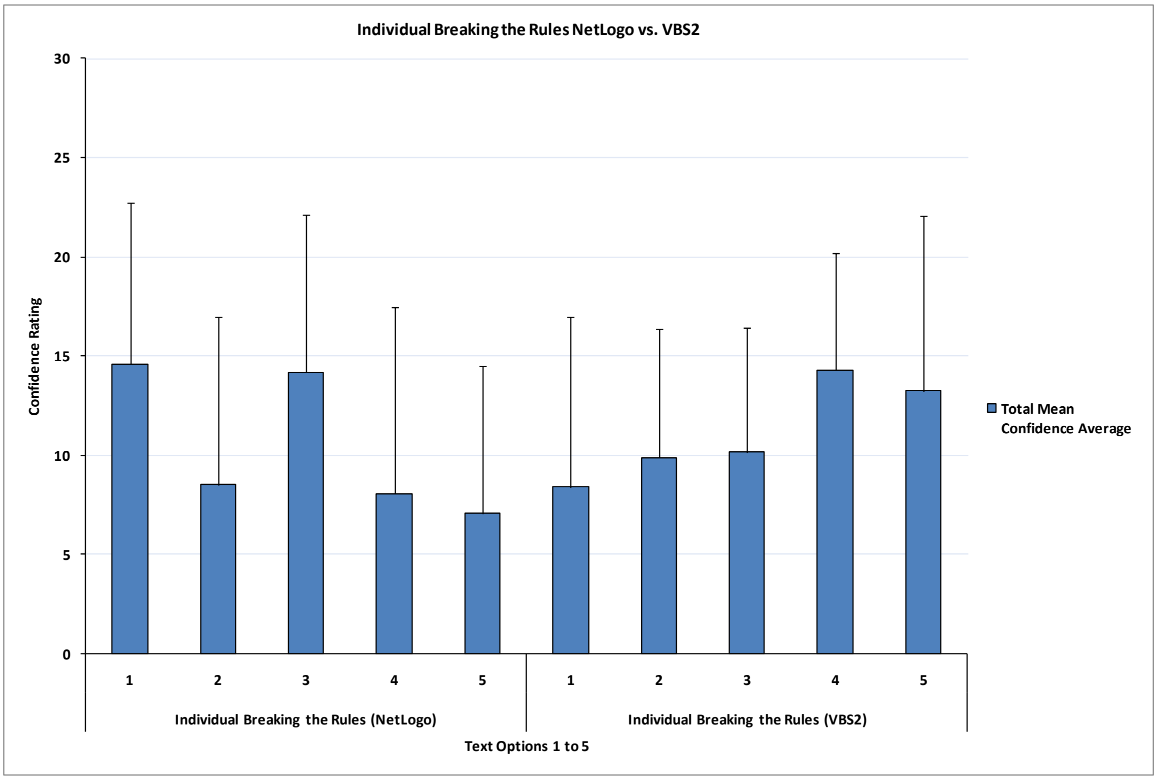

Fig. 5.19 Mean Confidence for Individual breaking the rules NetLogo vs. VBS2

The data was analysed conducting using a two-way mixed ANOVA to measure between two different groups of VBS and NetLogo observers and compare within-subject effects. The between-subjects are NetLogo and VBS2. This measures the total data of NetLogo between VBS2 total to compare if there is any significant effect between the two fidelities. This is the total data of the three runs. The result returned a non-significant main effect [F (1, 25) < 0.710, ns] which is expected as the results before also returned the same finding. It also shows that there is no difference between the observers mean confidence of NetLogo and VBS2. This can suggest that the fidelity at this point of measurement does not have an effect on the observer.

The next measurement is within-subjects to see if there are any significant effects between subjects (NetLogo, VBS2 and the five options).

The first measurement compared all the five options data, combining the data of NetLogo and VSB2. The result for this is measurement was a non-significant effect [F (4, 100) < 0.538, ns]. This result suggests that when analysing the data of the five options were an individual has broken the rules then there is no difference between 1 to 5 options. Table 5.4 shows that there is little difference to the combined mean confidence of NetLogo and VBS2 five options. This also suggests that the confidence in noticing whether that an individual has broken the rules could not be detected (overall). The highest combined confidence is option 3.

Table 5.4 Combined Total Mean Confidence for Five Options (Breaking Rules)

Options	Option 1	Option 2	Option 3	Option 4	Option 5
Combine Total Mean	22.4	17.7	23.6	21.4	19.4

The final measurement is the interaction between NetLogo, VBS2 and their five options. The result suggest that there is a significant main effect between them [F (4, 100) = 4.678, $p < 0.002$]. This finding concurs with Fig. 5.19 which shows that there is clear difference between NetLogo and VBS2 five options. It can also suggest that in some way fidelity does alter the observer's confidence.

However neither of the two fidelities led the observers to confidently identifying when an individual broke the rules. For NetLogo the highest confidence was in option 1 and for VBS2 option 4.

In the next section the data will be broken into the two brief types.

5.4.3.2 Individual breaking the rules NetLogo, VBS2 compared with brief type

The data are broken down to compared if the way in which information was given assisted the observer in noticing if an individual had broken the rules. Like the previous analysis which analysed the data at the fidelity level, the data now is separated into the two briefing types. It will also use the data from the three run that had an individual breaking the rules. Fig. 5.20 illustrates the differences between NetLogo, VBS2 and the two types of briefings written and verbal - to see if the information helped observers make the correct choice.

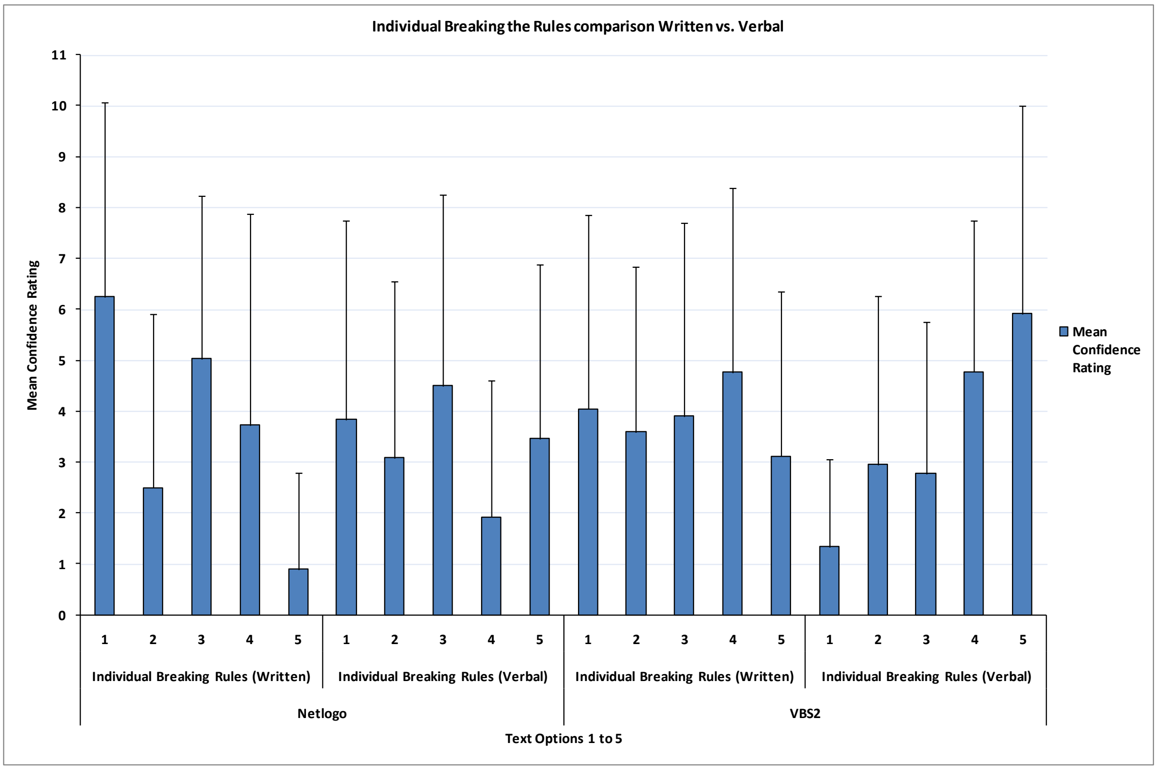

Fig. 5.20 Individual breaking the rules comparison between brief types

Fig. 5.20 illustrates that, for VBS2, option 5 in verbal briefing has the highest mean confidence and concurs with the findings in section 5.4.2.1 when analysing the whole data of written and verbal briefing data (all six runs). This suggests that observers either became aware of what was happening through the verbal briefing or that the fidelity helped (VBS2) them notice an individual breaking the rules.

The highest mean confidence rating for verbal briefing in VBS2 was for the correct option 5 with a mean confidence of 5.92. In comparison, the highest mean confidence for verbal briefing in NetLogo was for option 3 with a mean confidence of 4.50. Test option 5 verbal briefing in NetLogo had a mean confidence of 3.46.

The written briefs also had different outcomes for VBS2. The highest mean confidence was for option 4 with a mean of 4.76 and for NetLogo the highest mean confidence was for option 1 with a mean of 6.25.

As option 5 which is the "correct" option. Fig. 5.20 shows that it wasn't easy in any fidelity to spot an individual breaking the rules. However, although NetLogo highest confidence was not in the correct option, there is a significant difference between the confidence in option 5 verbal briefing and written briefing. For NetLogo verbal briefing the confidence mean for option 5 was 3.46 compared with the written brief confidence mean of 0.89.

VBS2 verbal briefing did have the highest confidence in the correct option, but not in the written and like NetLogo option 5 was the lowest mean out of the five options. Therefore it is possible to suggest that verbal briefing does help and improves the observer confidence but the increase in confidence does not ensure that they will get the correct options.

It is possible that the higher fidelity in VBS2 coupled with the verbal brief did raise awareness towards someone breaking the rules and led to increase confidence in the option 5.

The data will now be further analysed using a three-way mixed ANOVA design within-subject effects and between-subjects, which measures between NetLogo and VBS2 (fidelity), two briefing types, the runs which only had an individual breaking the rules and the five options. Due to a slight complexity Fig 5.21 helps illustrates this. Note that each run has five options and NetLogo is identical to VBS2.

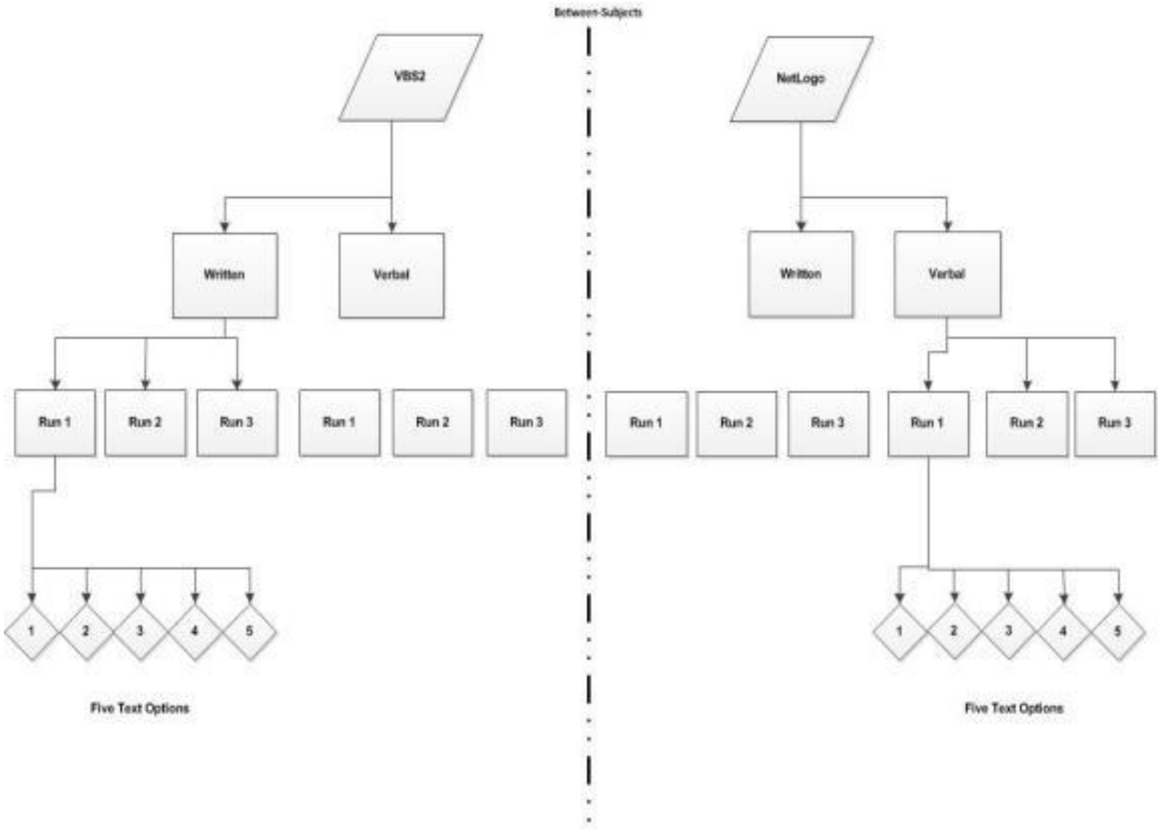

Fig. 5.21 Illustration of Measurements

The first group of measurements was to measure if there were any significant effect between-subjects. The first was to compare between the two fidelities to see if any effect between the two could be found. The measurement takes the total confidence data in NetLogo and VBS2 and compares. The data measures to find if there is any significant effect between the two fidelities.

This however returned a non-significant main effect between NetLogo and VBS2 [F (1, 23) < 0.763, ns].

The next measurement was to compare between the two brief types, written and verbal. This measurement does not compare between the NetLogo & VBS2 but takes the data from both fidelities and splits them into written and verbal briefing and compares to see if there is a significant effect between the two. The comparison between the two briefs also revealed a non-significant main effect between the two briefing types [F (1, 23) < 0.621, ns].

The final measurement in the between-subjects is the interaction between NetLogo, VBS2 and the two briefing. This now measures to see if there is a significant effect on the brief types and the fidelity. A non-significant main effect was found when comparing the interaction between fidelity and briefing [F (1, 23) < 0.990, ns].

The next of measurement measures the data within-subject to see if there is any significant effect between the different factors. The factors are the fidelity, the brief type and the five options.

Measuring the data within-subject effect there was non-significant main effect between the interaction of fidelity, brief type and the three runs of the experiment [F (2, 46) < 0.288, ns].

There is a significant main effect on the different scene runs (three in total for individual breaking the rules) [F (2, 46) = 7.638, $p < 0.001$]. This suggests that when combined all the data from NetLogo‘s three runs and VBS2 three runs and comparing between the three runs (Pairwise comparison) then there is a significant difference between them. Table 5.5 illustrates the Pairwise comparison of the total three runs. The difference between the runs is that the size of the crowd changes. The crowd sizes will be analysed further on in the thesis.

Table 5.5 Pairwise Comparison of the Three Runs (Individual Breaking Rules)

Measure:MEASURE_1

(I) Runs	(J) Runs	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.775*	.266	.024	-1.462	-.088
	3	-.830*	.231	.005	-1.425	-.234
2	1	.775*	.266	.024	.088	1.462
	3	-.055	.212	1.000	-.603	.493
3	1	.830*	.231	.005	.234	1.425
	2	.055	.212	1.000	-.493	.603

Based on estimated marginal means

Measure:MEASURE_1

(I) Runs	(J) Runs	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.775*	.266	.024	-1.462	-.088
	3	-.830*	.231	.005	-1.425	-.234
2	1	.775*	.266	.024	.088	1.462
	3	-.055	.212	1.000	-.603	.493
3	1	.830*	.231	.005	.234	1.425
	2	.055	.212	1.000	-.493	.603

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

There is a significant main effect on the interaction between the five options and fidelity (NetLogo and VBS2) [F (4, 92) = 6.179, $p < 0.000$]. So far the measurement findings have suggested that there are non-significant differences between the fidelities. However this changes when measuring the interaction between NetLogo, VBS2 and their five options as there is a significant difference between the interactions.

When fidelity is removed from the equation, there is also a significant main effect on the interaction between the five options and brief types (written and verbal) [F (4, 92) = 5.393, $p < 0.001$]. This suggests that when comparing the two briefings types and their five options by combing the data of NetLogo and VBS2 (removing the fidelity) there is differences between the five options of written and verbal. Further analysis of this measurement puts the verbal briefing option 5 as highest mean confidence of 4.68 mean overall.

For the written brief the highest mean confidence was option 1 with 5.15 mean overall. This suggests that the verbal briefing somehow assisted the observer's confidence towards selecting the correct option. However the same cannot be said for the written brief.

5.4.3.3 Conclusion

What can be concluded from the data analysed is that, when comparing the briefing types through different measurements by analysing one or more factors, it is suggested that the verbal briefing produces the highest confidence in the ability to detect if an individual had broken the rules. This was evident in the mean confidence rating for the correct option, option 5. It was also clear that VBS2 verbal briefing mean confidence in options was the highest. This suggests that fidelity did assist the observer visual in detect if an individual was breaking the rules. Even when the fidelity was removed the verbal briefing had the highest mean confidence for option 5.

This is also highlighted in Fig. 5.20 which illustrates that the highest confidence rating for VBS2 verbal briefing is option 5. It is worth highlighting an important fact that although NetLogo confidence mean in detecting an individual were low, there was an increase in confidence between the written and the verbal briefing for option 5.

5.4.4 Individuals not breaking the rules

Similar to 5.4.3, the data that will be analysed is the three runs were individuals did not break the rules. The first analysis compares the two fidelities (NetLogo and VBS) and their five options. This is to see where the confidence is in the options, which is hoped is option 3. The next analysis compares between the briefing types and fidelity. Is there a change from the first analysis and if so did the briefing type assist the observers confidence rating and towards answering the correct option.

Did the observer of the experiment easily identify that individuals in all three runs where abiding by the rules and getting food from the food aid station? The data was first analysed to see if the confidence rating of the five options altered between NetLogo and VBS2.

5.4.4.1 Individuals not breaking the rules NetLogo vs. VBS2

Fig. 5.22 compares the total results (combined data of the –three runs and including both brief types) of the observer‘s mean confidence rating between NetLogo and VBS2.

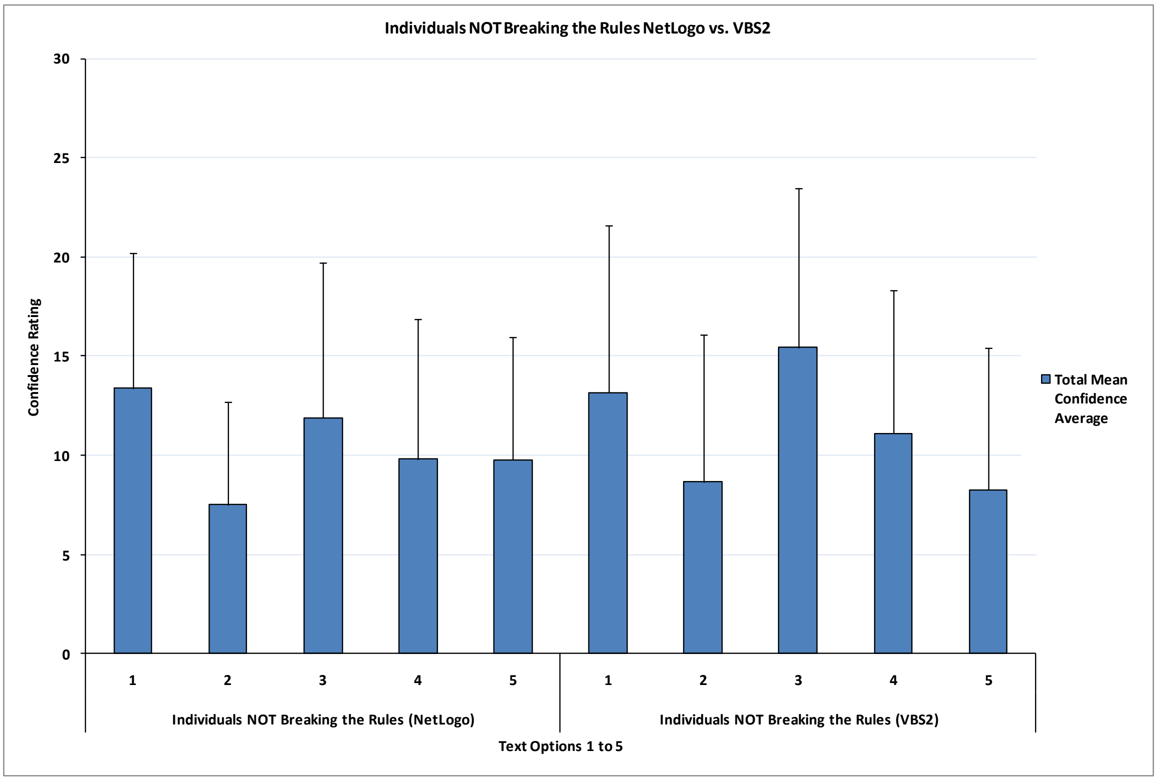

Fig. 5.22 Mean Confidence for Individual not breaking the rules NetLogo vs. VBS2

What can be seen is that for NetLogo the observer‘s strongest confidence was with option 1 (a mean confidence rating of 13.39) compared with VBS2 which was option 3 (a mean confidence rating of 15.42). What can be suggests from the Fig 5.22 is at this point is that observer‘s mean confidence in selecting the correct option was higher in VBS2 than in NetLogo. However the means are low for both fidelities considering the maximum confidence for each option is 30 and there are large standard deviations throughout all of the options (spread of confidence).

Analysing the data of the observer‘s confidence rating using a two-way mixed ANOVA design measurement to compare between NetLogo and VBS2 (between-subjects) to see if there is a significant difference between the fidelities. This resulted in a non-significant main effect between-subjects NetLogo and VBS2 (fidelity) [F (1, 25) < 0.648, ns]. This result is the same as previous measurements; and again at this level of measurement it would suggest that fidelity does not have an effect on the observers.

The next measurements are within-subject. The first looks at the total data of all five options. The result for not breaking the rules suggests that there is a significant main effect between the five options [F (4, 100) = 4.646, p < 0.002]. This result is the same for runs with an individual breaking rule. Table 5.6 below shows the total mean confidence of the five options and illustrates that there is little difference between the five options as the data analysis suggests. However it is worth noting that option 3 has the highest overall mean which is the correct option.

Table 5.6 Combined Total Mean Confidence of NetLogo and VBS2 Five Options

Options	Option 1	Option 2	Option 3	Option 4	Option 5
Combine Total Mean	25.6	15.5	26.2	20.1	17.5

The next measurement was to measure within-subject to see if there are any differences between the fidelities and their five options. The result was a non-significant main effect [F (1, 25) < 0.772, ns]. Fig. 5.22 illustrates the results which show that there are little differences between the mean confidence of NetLogo and VBS2. This suggests that there is no significant difference between the observer‘s confidence of NetLogo and VBS2.

The next analyse broke down the data into the two briefing types to see if this had any effect on the observers confidence, Fig. 5.23 below illustrates the differences between the written and verbal briefing NetLogo and in VBS2. This is to compare the briefing types with the five options mean confidence to see if the briefing helped observer‘s make the correct choice by displaying a higher mean confidence rating in option 3.

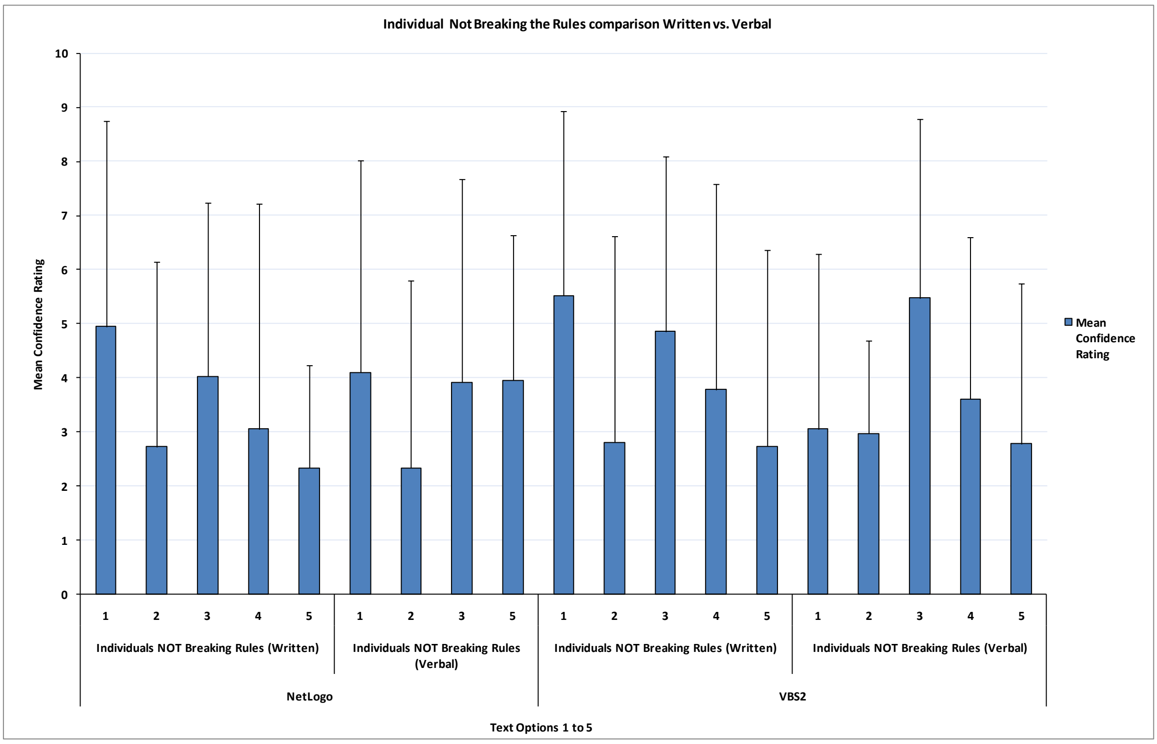

Fig. 5.23 Individuals not breaking the rules comparison between brief types

When comparing the two brief types, VBS2 highest mean confidence rating was for option 1 in the written briefing with mean confidence of 5.52. When comparing the written brief to the verbal brief, VBS2 highest confidence has altered from option1 to the correct option, option 3 giving a slightly lower mean confidence rating of 5.47. As in section 5.4.3, it could be suggested that the verbal briefing or the VBS2 fidelity helped to inform the observers of the correct option.

The NetLogo results show a different outcome. For the written briefing, option 1 had the highest mean confidence rating of 4.94 and for the verbal briefing option 1 again had the highest mean confidence rating of 4.10.

Analysing the data just comparing the two briefing types and the five options resulted in the verbal briefing showing the highest mean confidence for the correct option 3 (4.69 mean). The written briefing however showed the highest confidence in option 1 (5.23 mean) which was not the correct option. This finding concurs with that of when an individual broke the rules and could suggest that the verbal briefing type does help towards informing the observers.

The data was then analysed using a three-way mixed ANOVA design within-subject effects and between-subjects. This analysis between-subjects suggested that there is a non-significant main effect between NetLogo with VBS2 (fidelity) [F (1, 23) < 0.667, ns].

The difference between brief types also had a non-significant main effect [F (1, 23) < 0.857, ns] and the interaction between fidelity and brief type also had a non-significant main effect [F (1, 23) < 0.699, ns].

The measurements within-subjects effect took the data of the three runs, the five options (1 to 5), fidelity and brief type (written or verbal) and the interaction between them. The only significant main effect is for the five options [F (4, 92 = 4.656, p = 0.002].

The measurement showed non-significant main effect for any of the interaction with brief types which suggests that the brief type make had no impact on observer‘s confidence rating, nor does fidelity in which was being used.

5.4.4.2 Conclusion

What can be concluded from the analysing of the data for individuals who followed the rules is that the mean confidence rating when trying to find significant main effect between the fidelity and briefing type was inconclusive. That is fidelity type (NetLogo or VBS2) and the different briefing types (written or verbal) had no impact on the observer‘s confidence.

Although significant main effects were found in the options (1 to 5), Fig. 5.22 illustrates that the consensus of the observers was not present which created large standard deviations in the options. Although for verbal briefing and VBS2 option 3 was the highest, when analysed with all the data it prove to have no significant main effect.

Therefore there is no clear consensus that the observers could chose the correct options and that the different brief types had no significant effect. Based only on graph (Fig. 5.22) it could be suggested that fidelity briefing type does have some effect in VBS2, however this is not the same for NetLogo.

5.4.5 Impact of crowd size

Although not a directly related to any of the three thesis questions, Experiment I was conducted using different crowd sizes and therefore it is useful to understand if an alteration in the crowd size had an effect on the observer‘s choice of options and the confidence rating given.

The focus of this analysis is the number of individuals in the virtual environment (i.e. the crowd size). The measurements have been split into NetLogo and VBS2 with individuals not breaking the rules, individuals breaking the rules combining the data from both briefing types. The Graphs measure the observers mean confidence rating of the five options that they were given during each run of the food aid scenario.

The analysis then compares crowd sizes without individuals breaking the rules to the best match of crowd sizes with the individual breaking the rules. For example, a crowd size of 3 (three individuals not breaking rules) is compared to a crowds size of 3 plus 1 individual breaking the rule. The analysis also looks at whether the correct option has been selected (option 3 for no breaking rule, option 5 for breaking the rules).

The analysis above is conducted first for NetLogo, then for VBS2 and then concludes with a comparison between VBS2 and NetLogo to see if fidelity has an influence on confidence.

5.4.5.1 NetLogo crowd size comparison

Fig 5.24 and Fig. 5.25 illustrate the difference between observer‘s confidence rating and the different crowd sizes shown. The numbers in the graph index depicts the number of individuals in the crowd, the ‗+1‘ denotes that an individual was breaking the rules. The sometimes large variation in confidence ratings illustrated in both graphs could suggest that the crowd size does affect the confidence mean on some of the options, however again the standard deviations are very high, due to the spread of confidence.

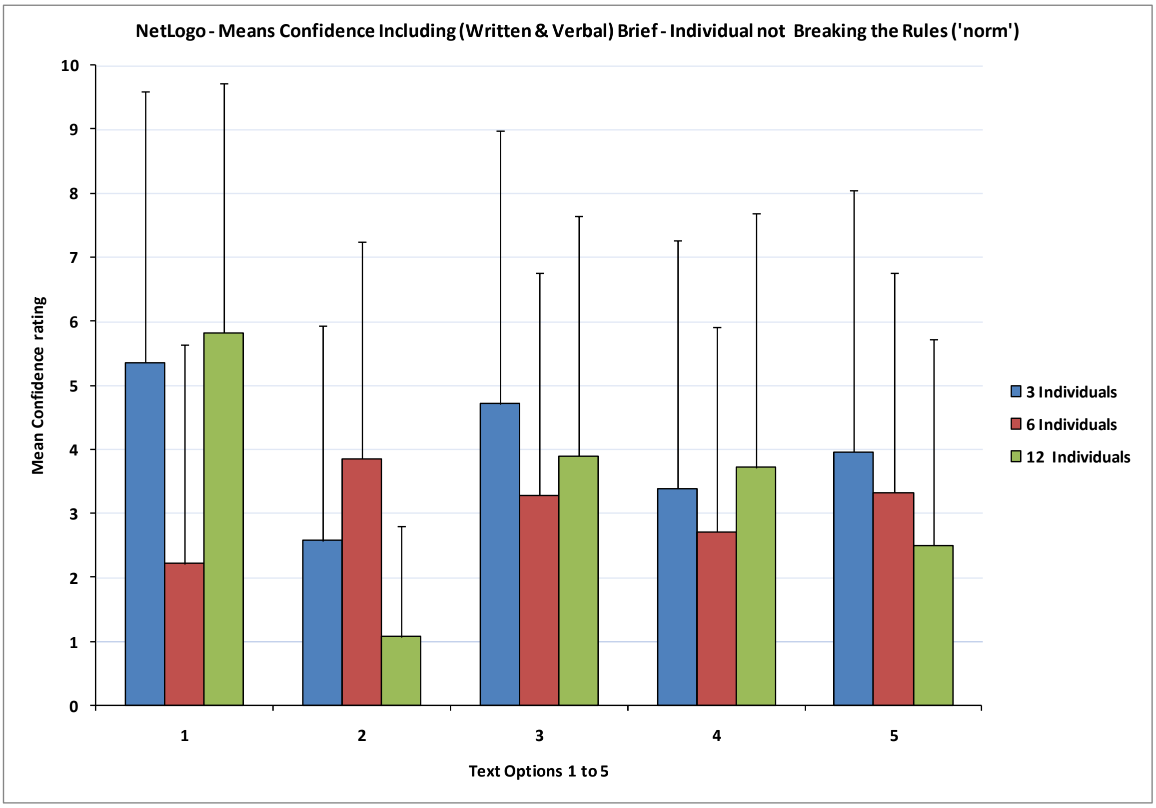

Fig. 5.24 Different crowd types and individuals not breaking rules (NetLogo)

Table 5.7 shows the highest mean confidence rating for Fig. 5.24 crowd sizes.

Table 5.7 Highest Mean Confidence Rating per Crowd Size (Not Breaking Rules) NetLogo

Crowd Size	Option	Mean Confidence	Standard Deviation (Error Bar)
3 Individuals	1	5.36	4.24
6 Individuals	2	3.86	3.39
12 Individuals	1	5.82	3.89

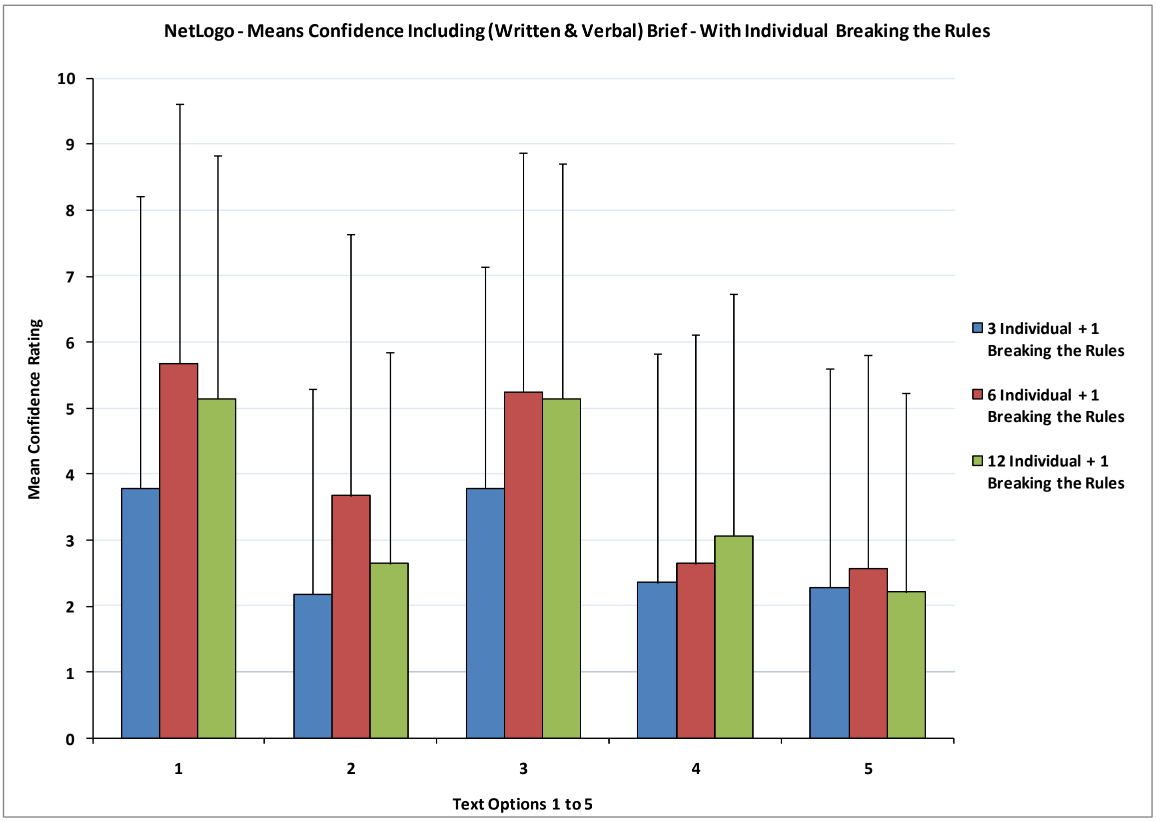

Fig. 5.25 Different crowd types individual breaking the rules (NetLogo)

Table 5.8 shows the highest mean confidence rating for Fig. 5.25 crowd sizes.

Table 5.8 Highest Mean Confidence Rating per Crowd Size (Breaking Rules) NetLogo

Crowd Size	Option	Mean Confidence	Standard Deviation (Error Bar)
3 Individuals + 1 Breaking the Rules	1 & 2	3.79, 3.79	4.42, 3.36
6 Individuals + 1 Breaking the Rules	1	5.68	3.94
12 Individuals + 1 Breaking the Rules	1 & 3	5.14, 5.14	3.69, 3.56

The next graphs are broken down by crowd size and compare the mean confidence for individuals breaking and not breaking the rules to see if size increased the observer‘s awareness of an individual breaking the rules or not breaking the rules.

Fig 5.26 below shows mean confidence in a crowd size of 3. It is suggested that when an individual was added who broke the rules the mean confidence decreased it also suggests that as the crowd size grew, then so did the confidence in some of the options. Confidence levels in option1 & 3 are higher when the crowd included an individual breaking the rules. However they were both incorrect options as option 5 was the correct choice for when an individual was breaking the rules.

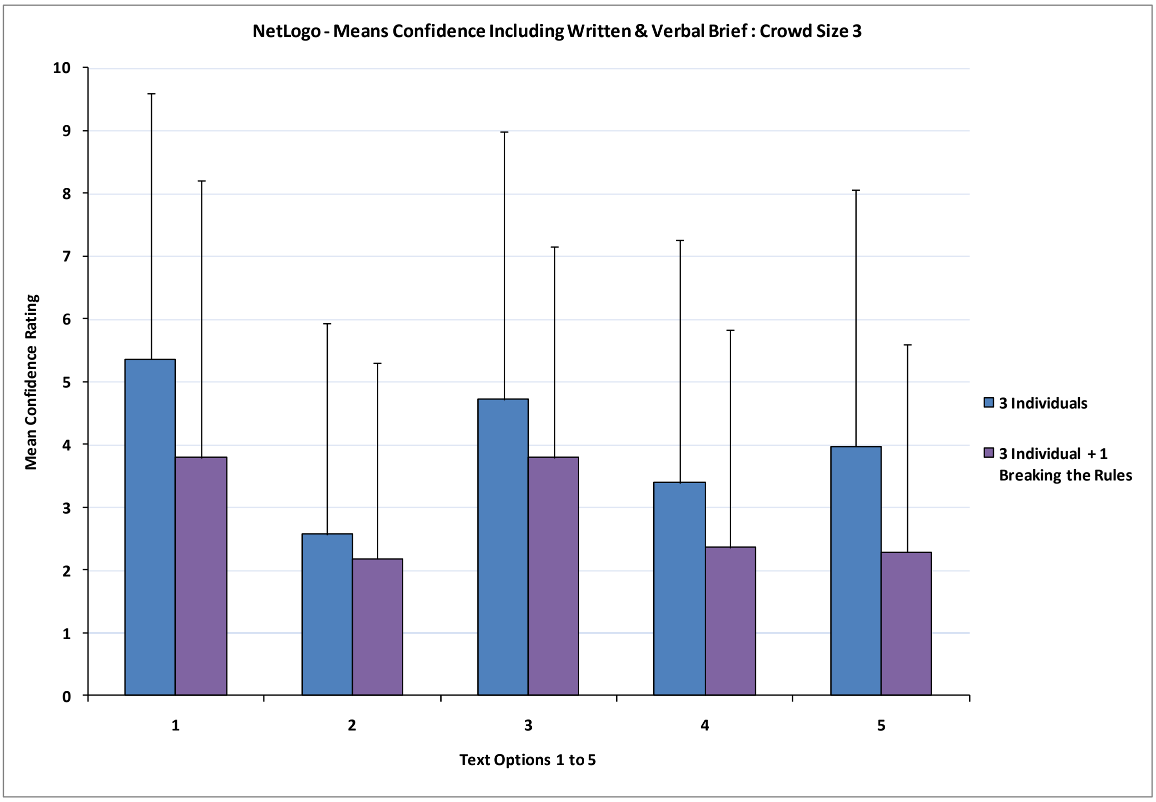

Fig. 5.26 Crowd size 3 NetLogo

Fig. 5.27 and Fig. 5.28 below shows the results for a crowd size of 6 and 12 respectively.

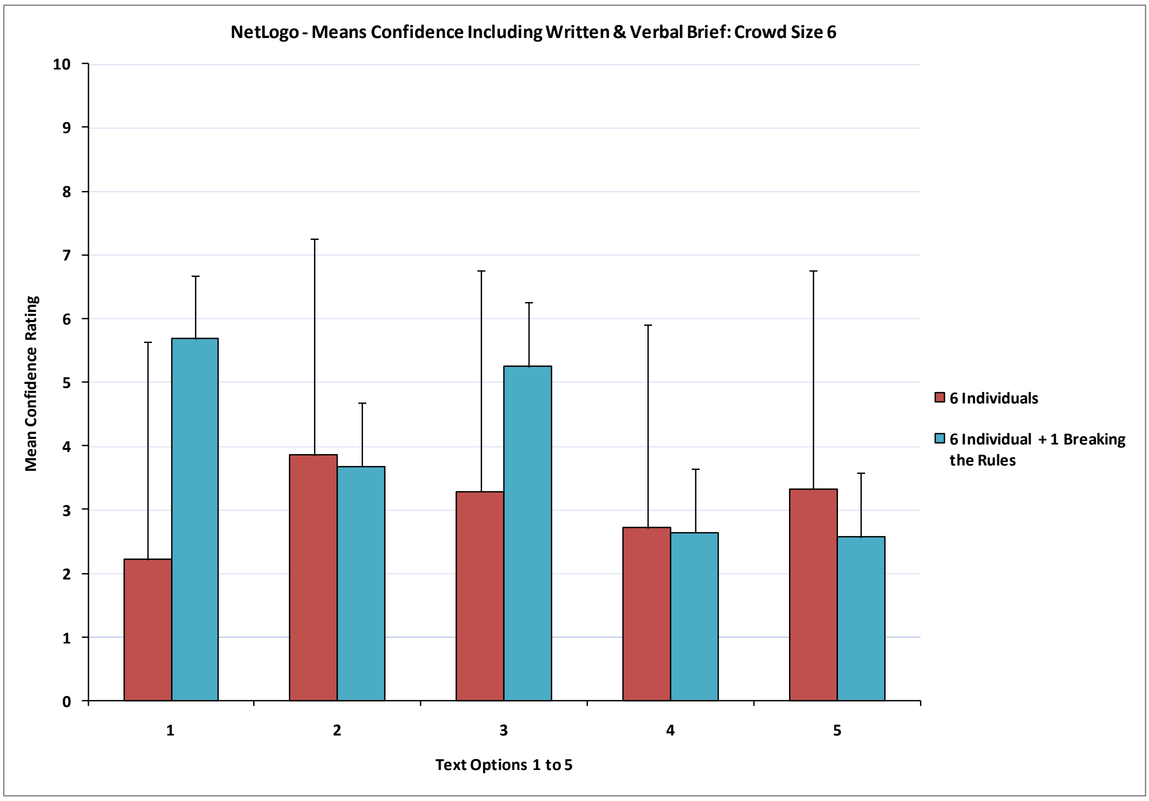

Fig. 5.27 Crowd size 6 NetLogo

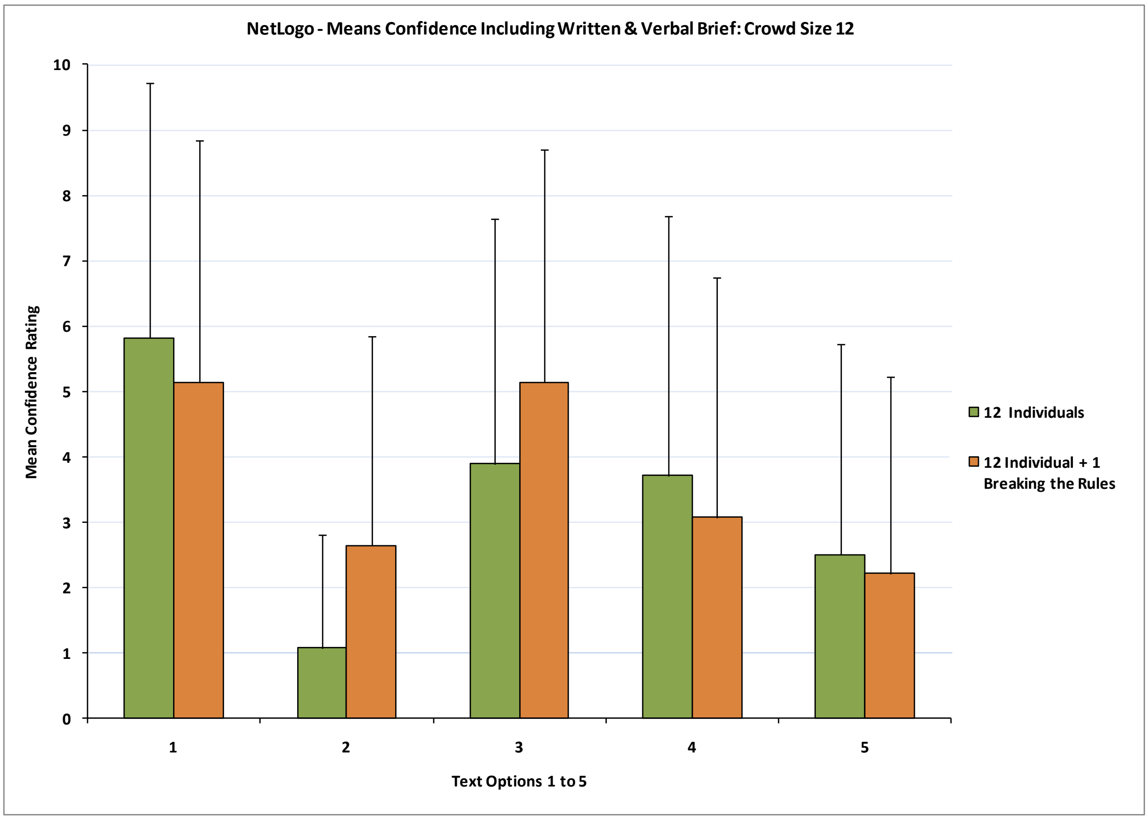

Fig. 5.28 Crowd size 12 NetLogo

When comparing the overall mean confidence between all the crowd sizes were the individuals followed the rules, crowd size 3 had the highest confidence with a mean of 4.00, followed by crowd size 12 (mean 3.40) and crowd size 6 (mean 3.08). Where an individual was breaking the rules the, highest mean was for crowd size 6 (mean of 3.96) followed by crowd size 12 (mean of 3.64) and last was crowd size 3 (mean 2.88).

A three-way repeated-measure ANOVA within-subject effect was performed to measure if there is any significant main effect differences between individuals not breaking the rules, individual who has broken the rules, crowd sizes and the five options.

When comparing between the six test runs, three runs where individuals followed the rules and three runs where an individual broke the rules then it is suggested that there is a non-significant main effect when comparing individuals who are not breaking the rules with individual breaking the rules [F (1, 13) < 0.993, ns].

Furthermore there is a non-significant main effect with crowd sizes [F (2, 26) < 0.925, ns]. Crowd size and interaction with the five options also had a non-significant main effect [F (8, 104) < 0.430, ns].

However, there is significant main effect with the five option [F (4, 52) = 5.118, p < 0.001] and there is also a significant main effect on the interaction of crowd size and those breaking rules and not breaking rules [F (2, 26) = 6.611, p < 0.005]. These findings correlate with what the graphs show, i.e. that there is a significant different between the five options and significant difference between the crowd sizes and comparing with an individual breaking the rules and those individual who followed the rules.

5.4.5.2 Conclusion NetLogo size comparison

It is suggested that there is a significant effect when comparing the difference between the individual not breaking the rules crowd sizes and that where an individual is breaking the rules crowds sizes. This could suggest that crowd size does have an effect on the observer‘s confidence rating.

For a crowd size of 3, the data shows highest mean for individual not breaking the rules was for the correct option: option 3. However, the same cannot be said for an individual breaking the rules was as the highest means were for option 1 & 3 (same mean of 3.29) which are both incorrect.

As the crowds grow so does the highest mean options change however this remains the incorrect option. In the runs were the individuals are not breaking the rules the highest confidence mean for crowd size 6 was for option 2 and for crowd size 12 option 1. Similarly _incorect‘ options resulted in runs where there was an individual who is breaking the rules. Crowd size 6 had the highest confidence mean rating for option 1 and crowd size 12 had option 1 &3 (same mean).

This would suggest that for NetLogo the crowd size does not led the observer to making the correct options nor is it clear that the crowd size has any effect on the observer's awareness of what is happening in the virtual world.

5.4.5.3 VBS2 crowd size comparison

Fig 5.29 and Fig. 5.30 illustrates the difference between observer's confidence rating and the different crowd sizes shown for VBS2 respectively with and without an individual breaking the rules. The graphs suggest that the crowd size does somehow affect the confidence mean on some of the options in VBS2.

In Fig 5.30 below the mean confidence rating for crowd size 9 stands out from the rest of the crowd sizes. Considering that option 3 is the correct option for when individuals are following the rules this is a good sign that VBS2 and crowd size 9 may potentially have a significant mean.

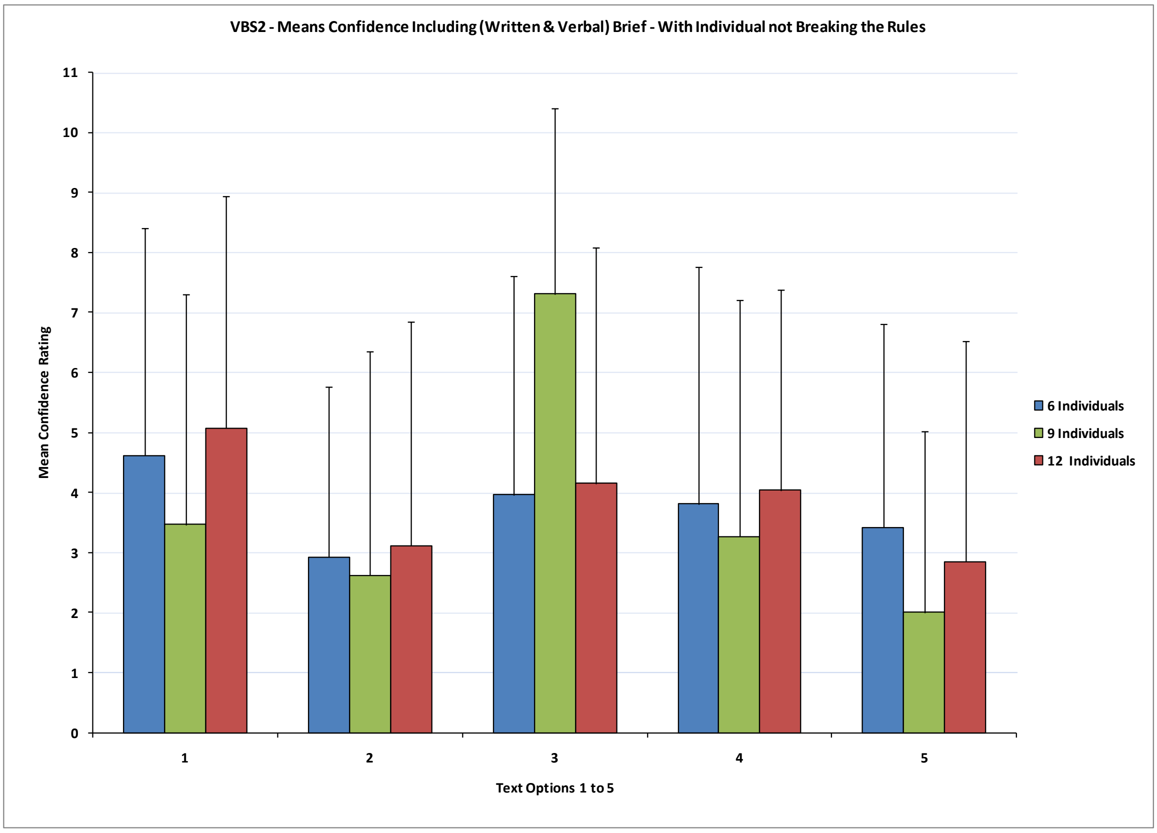

Fig. 5.29 Different crowd types and individuals not breaking rules (VBS2)

Table 5.9 Highest Mean Confidence Rating per Crowd Size (Not Breaking Rules) VBS2

Crowd Size	Option	Mean Confidence	Standard Deviation (Error Bar)
6 Individuals	1	4.62	3.80
9 Individuals	3	7.31	3.09
12 Individuals	1	5.82	3.86

Fig. 5.30 suggests that when adding an individual into the crowd who breaks the rules the mean confidence sways towards option 4 and option 2, neither of which are the correct option. However the variance between the highest option mean confidence and the correct option 5 is minimal as it differed from 0.50 to 1.00 mean. When comparing the total mean confidence rating for all the options, crowd size 12 had the highest mean for both and individual not breaking the rules (mean of 3.85) and the for an individual breaking the rules (a mean of 4.20).

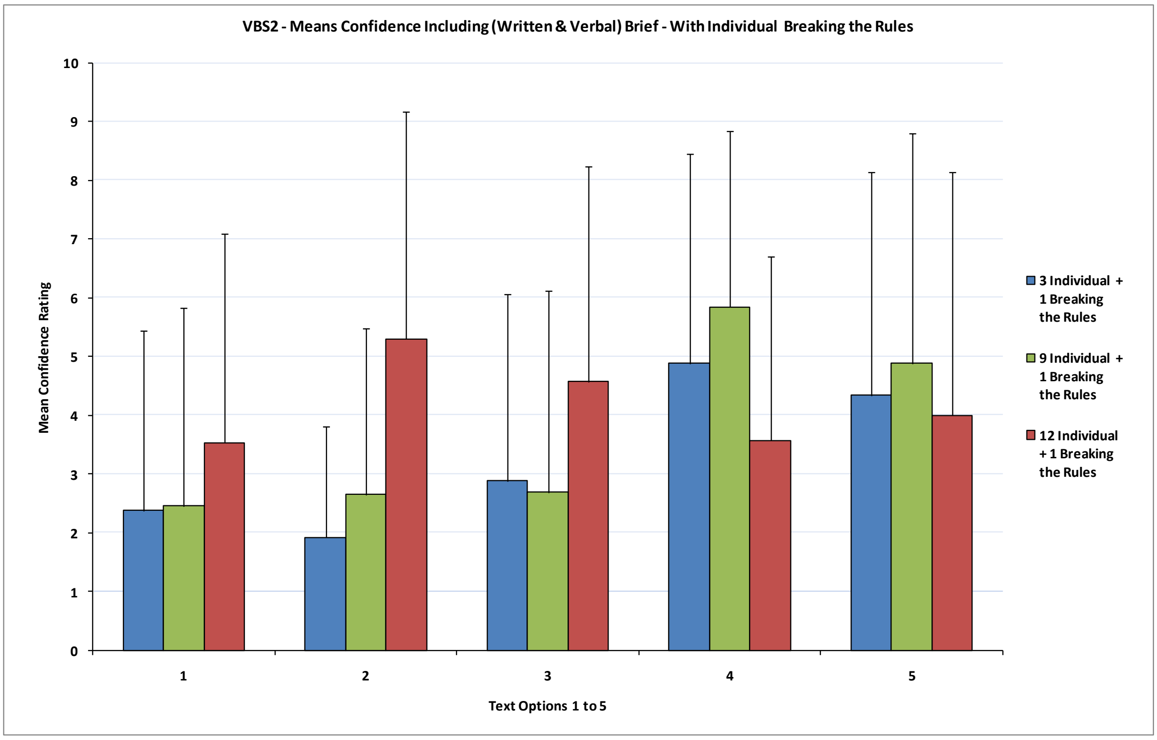

Fig. 5.30 Different crowd types individual breaking the rules (VBS2)

Table 5.10 Highest Mean Confidence Rating per Crowd Size (Breaking Rules) VBS2

Crowd Size	Option	Mean Confidence	Standard Deviation (Error Bar)
3 Individuals + 1 Breaking the Rules	4	4.88	3.57
6 Individuals + 1 Breaking the Rules	4	5.85	3.00
12 Individuals + 1 Breaking the Rules	2	5.31	3.86

A crowd size of 6 was randomly chosen for the crowd size of individuals not breaking the rules and as such Fig 5.31 compares the crowd size of 3 plus 1 individual breaking the rules to the crowd size of 6. The graph shows - the mean confidence across all five options.

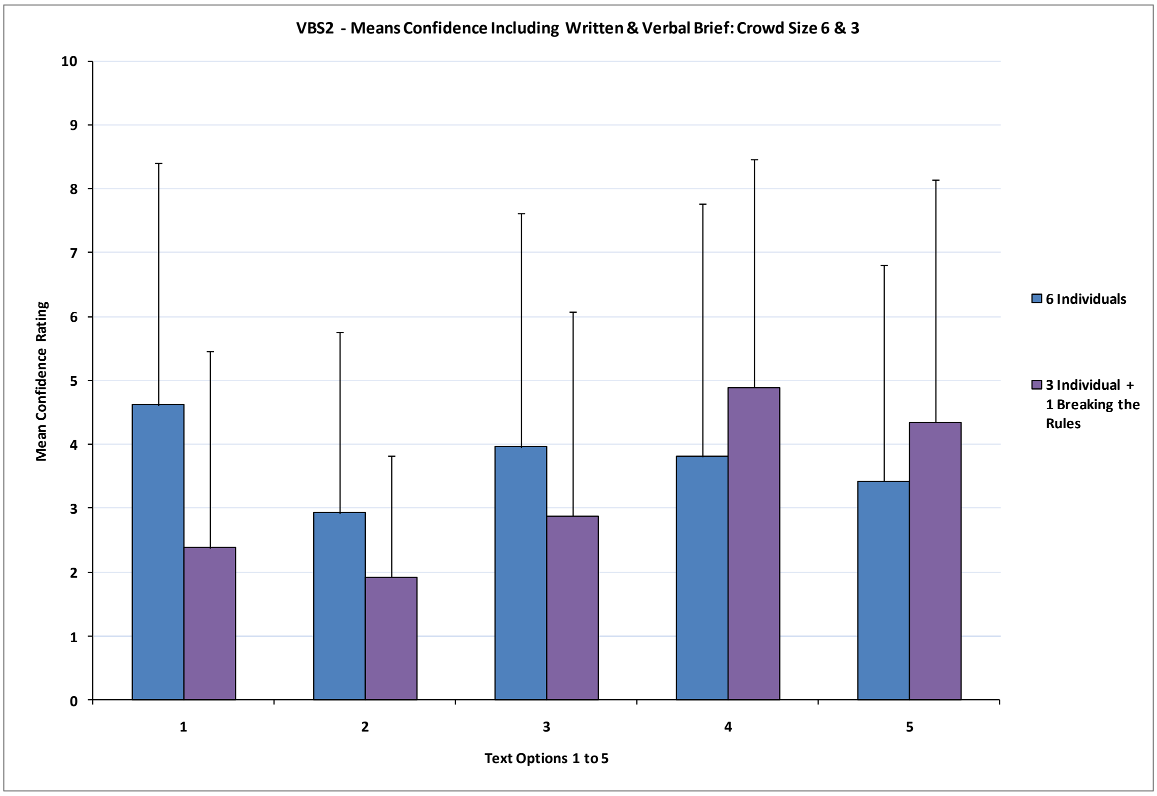

Fig. 5.31 Crowd size 6 & 3 VBS2

Fig 5.32 below shows that, for VBS2 crowd size 9 without an individual breaking the rules, observer‘s confidence had a high confidence rating mean of 7.31 for option 3 which is the correct option. What is also interesting is that when adding an individual who broke the rules the confidence in option 4 rose and second was option 5. Albeit option 5 is the correct option, the difference between the mean confidence in option 4 and 5 was only 0.96.

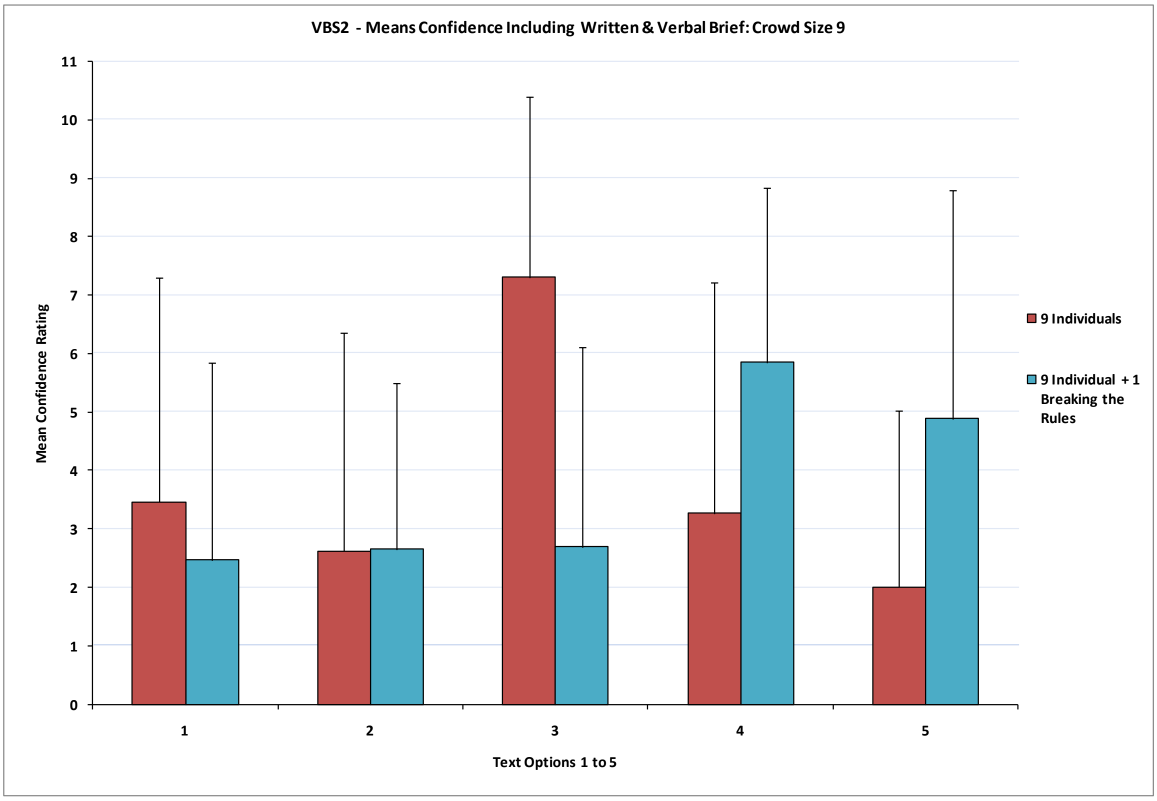

Fig. 5.32 Crowd size 9 VBS2

What is suggested in Fig. 5.33 is that in an increased crowd size of 12 for VBS2 the confidence seems to level out across the different options compared to the smaller crowd sizes. The observers mean confidence rating was highest in option 1 when all individuals are following the rules. This changed to option 2 when adding in an individual breaking the rules. This is in contrast to crowd size 3, 6 and 9.

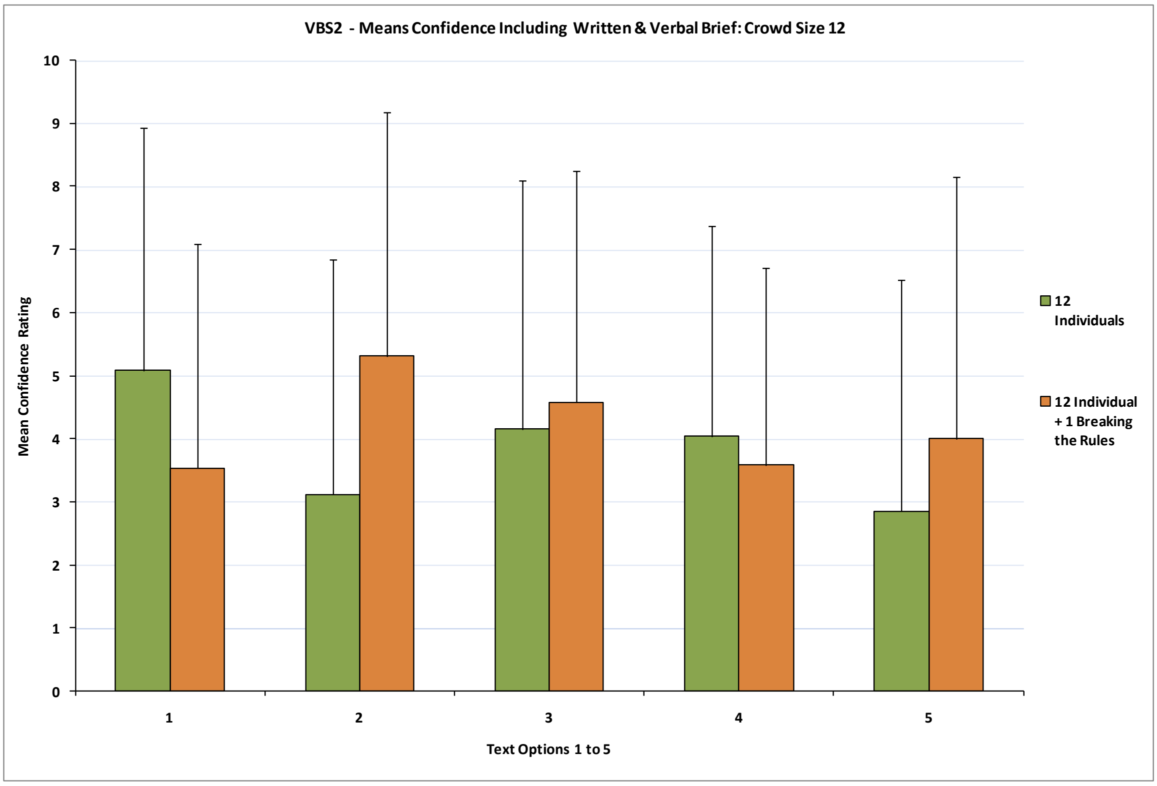

Fig. 5.33 Crowd size 12 VBS2

Similarly to NetLogo, the data was analysed to see if any significant main effects differences could be found towards crowd sizes. A three-way repeated-measure ANOVA within-subject effects tested if there were any significant main effect between, individuals not breaking rules, individual who has broken the rules, crowd sizes and the five options.

The measured data suggest that there is a non-significant main effect on individuals who did not break the rules compared with individual who broke the rules [F (1, 12) < 0.804, ns].

There is also a non-significant main effect between the crowd sizes [F (2, 24) < 0.187, ns].

Furthermore, there is a non-significant main effect between the interaction of crowd sizes and the five options [F (8, 96) < 0.173, ns].

However, there is a significant main effect on the interaction between individuals who did not break the rules compared with individual who broke the rules and the five options [F (4, 48) = 4.434, p < 0.004] which is similar to NetLogo.

5.4.5.4 Conclusion VBS2 crowd sizes

The only suggested significant main effect existed when comparing between when an individual does not break the rules with that where an individual does break the rules and the interaction of the five options. However the crowd sizes don't have any significant effect towards giving the observer a great confidence mean rating towards the correct option.

The means confidence did change depending on the crowd size. In the food aid scene were individuals did not break the rules for crowd size 6 the highest mean confidence was in option 1 (4.62 mean), this changed with crowd size 9 with the highest confidence mean in option 3 (correct option) with a mean of 7.31. And when the crowd size increased to crowd size 12 the highest option, was option 1 (5.17 mean). These alterations were apparent also happened in the runs were an individual broke the rules however none of the options with the highest mean were for the correct option 5. This would suggest that for VBS2, like NetLogo, the crowd size do not help the observer select the correct options nor is it clear that the crowd size has any effect on the observer's awareness of what is happening in the virtual world.

5.4.5.5 Comparison between NetLogo and VBS2 crowd sizes

The final comparison is to see if there are any differences between NetLogo and VBS2. Fig. 5.34 and Fig. 5.35 illustrate the confidence mean rating for the five options and compare VBS2 and NetLogo side by side.

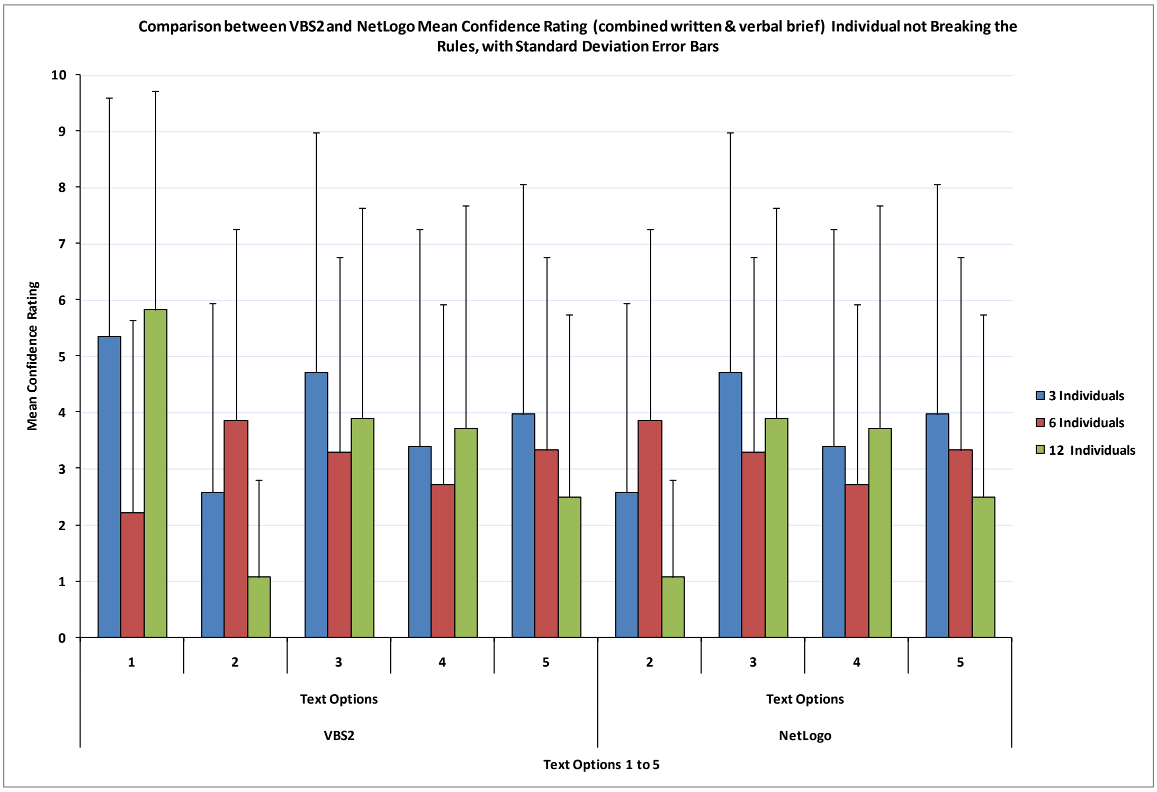

Fig. 5.34 Comparison between NetLogo and VBS2 crowd sizes (not breaking rules)

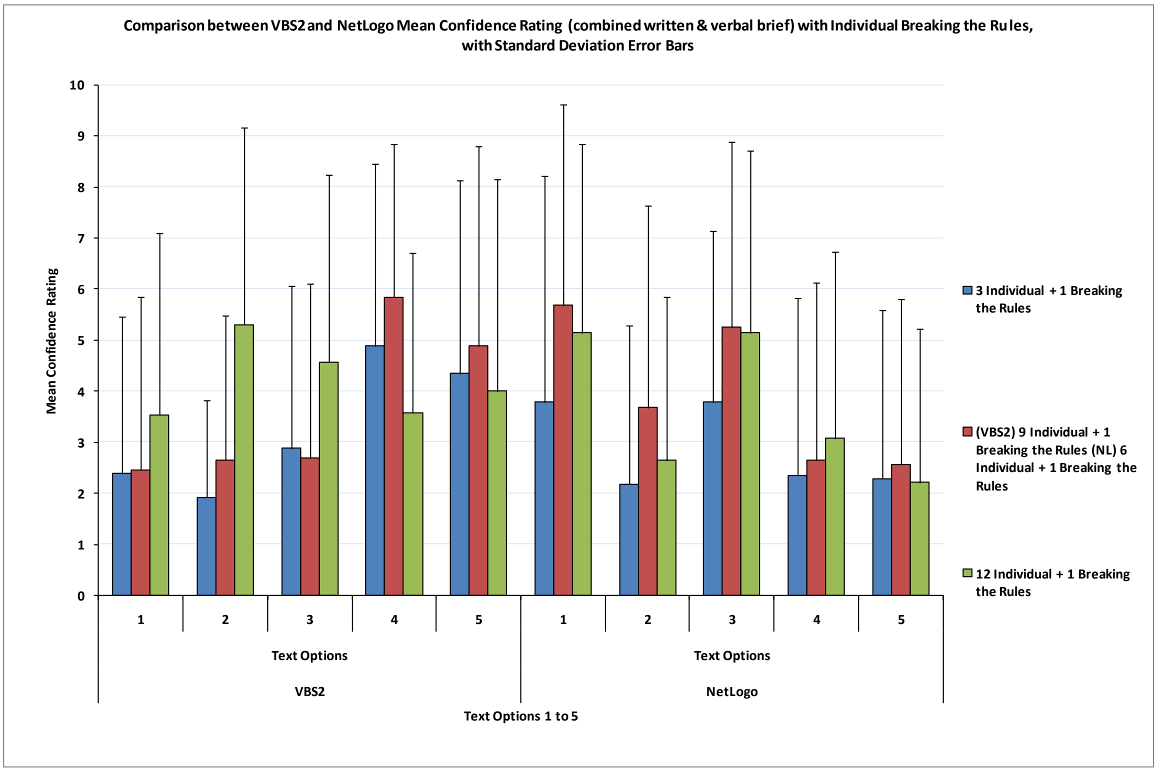

Fig. 5.35 Comparison between NetLogo and VBS2 crowd sizes (breaking rules)

A three-way mixed measured ANOVA within-subjects and between subjects, was performed to compare NetLogo with VBS2. The with-in subjects are breaking rules, not breaking rules, crowd size and option and between subjects suggests NetLogo and VBS2 (fidelity).

For the between subject NetLogo and VBS2 there is a non-significant main effect [$F.$ $(1, 25) < 0.669$, ns]

There is also a non-significant main effect between the breaking of rules and not breaking rule interaction between fidelity [$F.$ $(1, 25) < 0.882$, ns].

Furthermore there is a non-significant main effect with crowd size and the interaction with fidelity [$F.$ $(2, 50) < 0.480$, ns].

There is also a non-significant main effect with options and the interaction with fidelity [$F.$ $(4, 100) < 0.075$, ns]

However, there is a significant main effect within-subject not breaking rules, breaking rules and it interaction between crowd sizes [F $(2, 50) = 5.811$, $p < 0.005$] which suggest that there is some relationship between the different crowd sizes between when individuals that followed the rules and those who did not.

There is also a significant main effect with-subject, not breaking rules, breaking rules and interaction between options and fidelity [F (4, 100) = 3.772, $p < 0.007$] which suggest there is an influence between NetLogo, VBS2, when comparing between the individuals who did not break the rules with the individual that broke the rules and the five options.

5.5 Conclusions from Experiment I

As the results from the data analysed are numerous, this section summarises them and refers them back to the objectives of the experiment which were defined in section 5.1 as follows:

- Do observers have strong opinion of what is happening in a particular situation?
- Does the distribution of written and verbal information have a different effect on the observer‘s opinion?
- Can observers identify the person who is not acting in line with the crowd?
- Does fidelity make a difference?
- Does the size of the crowd have an impact?

The ANOVA results have been summarised in tables showing the measurement and corresponding result. There are two main result columns. ―Comparison” is used when the NetLogo and VBS2 data were compared with each other. ―Combined” is used when the measurement used the combined data of both NetLogo and VBS2. Table 5.11 and 5.15 also include columns only for NetLogo only and only for VBS2.

Do observers have strong opinion of what is happening in a particular situation? And, does fidelity make a difference? Table 5.11 shows that ignoring the different fidelities and comparing the five options there is a significant difference which suggest observers do have an opinion of what is happening. However to suggest it was strong would not be accurate as the mean confidences were low and the spread in the standard deviation was high which suggest that there was little consensus. Furthermore, at the highest level, there is no significant difference between the two fidelities (i.e. when comparing total VBS2 with total NetLogo data).

Table 5.11 Summary of Findings - Five Options

Mean Confidence Level in all Five Options				
Measurement	**NetLogo Only**	**VBS2 Only**	**Comparison**	**Combined Data**
Total VBS2 compared with NetLogo			Non-Significant	
Five Options (1…5)	[*F* (4, 52) = 5.118, p < 0.001]	Non-Significant		[*F* (4, 100) = 4.013, p < 0.005]

Did the type of briefing have an effect on the observers? Overall the results appear to be inconclusive. Table 5.12 summarises the experiment results when comparing brief types. At the highest level when (i.e. combining the data of VBS2 & NetLogo) the results suggest briefings did not have an effect on the observer. However when comparing between briefing types (written, verbal) and fidelity (VBS2, NetLogo), then there is a difference, which suggests that at this level briefing and fidelity did have an effect on the observer‘s confidence. But these results seem to change as the data is analysed in further detail, e.g. by adding the five options, ultimately returning a non-significant difference between them.

Table 5.12 Summary of Findings - Brief Types

Different Brief Types		
Measurement	**Comparison**	**Combined Data**
Fidelity removed - Verbal Briefing compared with Written Briefing		Non-Significant
Fidelity - VBS2 compared with NetLogo	Non-Significant	
Fidelity - VBS2, NetLogo, Written and Verbal Briefing	Non-Significant	
Fidelity removed - Five Options (1…5)		[*F* (4, 92) = 4.878, p < 0.001]
Fidelity - VBS2, NetLogo and the Five Options for each (1…5)	[*F* (4, 92) = 6.972, p < 0.017]	
Fidelity removed - Verbal, Written Briefing (Non- fidelity) and the Five Options for each (1…5)	[*F* (4, 92) = 6.017, p < 0.000]	

Fidelity - VBS2, NetLogo, Written, Verbal briefing and five options for each (1…5)	Non-Significant	

Can observers identify the person who is not acting in line with the crowd? The observers did not detect if an individual broke the rules. The data showed low mean confidence in the options and the spread of deviation was high between the options.

Table 5.13 summarises the results with an individual breaking the rules and shows that neither fidelity nor the type of brief led the observer‘s to be confidently aware of someone acting out of the ‗nom‘ (i.e. breaking the rules). However a significant effect is also found looking solely at the VBS2 & NetLogo‘s five options, briefing types and five options, and the five options.

Interestingly, in the case of an individual breaking the rules, briefing type has an effect on the observer‘s confidence of choice between the five options (again combining the total data of NetLogo & VBS2).

Table 5.13 Summary of Findings - Individual Breaking the Rules

Individual Breaking the Rules		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo	Non-Significant	
Fidelity - VBS2, NetLogo and the Five Options for each (1…5)	[F (4, 100) = 4.678, p < 0.002]	
Fidelity removed - Five Options (1…5)		Non-Significant
Individual Breaking the Rules & Briefing Types		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo	Non-Significant	
Fidelity removed - Verbal Briefing and Written Briefing		Non-Significant
Fidelity - VBS2, NetLogo and Verbal Briefing, Written Briefing	Non-Significant	

Fidelity removed - Five Options (1…5)		Non-Significant
Fidelity removed - Three runs with individuals breaking rules (1...3)		[*F* (2, 46) = 7.638, p < 0.001]
Fidelity - VBS2, NetLogo and the Five Options for each (1…5)	[*F* (4, 92) = 6.179, p < 0.000]	
Fidelity removed - Verbal Briefing compared with Written Briefing and Five Options for each (1…5)		[*F* (4, 92) = 5.393, p < 0.001]

Table 5.14 summarises the results when only the data without individuals breaking the rules. The results suggest that observers where more in unison in their confidence between the two fidelities and similarly for briefing types as both resulted in non-significant differences. However, removing fidelity and briefing types from the measurement and comparing the results only between the five options shows there is a difference of opinion/confidence between them (i.e. indicating a positive answer to the first question – there is a strong opinion of what is happening in a particular situation). It is clear that similarities in the option made it difficult for the observer to confidently understand what was happening in the scenarios, when the individuals were conducting the ‗nom‘.

Table 5.14 Summary of Findings - Individuals Not Breaking the Rules

Individuals Not Breaking the Rules		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo	Non-Significant	
Fidelity - VBS2, NetLogo and the Five Options for each (1…5)	Non-Significant	
Fidelity removed - Five Options (1…5)		[*F* (4, 100) = 4.646, p < 0.002].
Individual Not Breaking the Rules & Briefing Types		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo	Non-Significant	

Fidelity removed - Verbal Briefing compared with Written Briefing		Non-Significant
Fidelity - VBS2, NetLogo and Verbal Briefing, Written Briefing	Non-Significant	
Fidelity removed - Five Options (1…5)		[*F* (4, 92 = 4.656, p = 0.002].
Fidelity removed - Three runs with individuals not breaking the rules (1...3)		Non-Significant
Fidelity - VBS2, NetLogo and the Five Options for each (1…5)	Non-Significant	
Fidelity removed - Verbal Briefing compared with Written Briefing and Five Options for each (1…5)		Non-Significant

Does the size of the crowd have an impact? Although not directly linked to any of the three hypothesis questions in Chapter 1, this was analysed out of interest and because it linked to the literature reviewed on crowds. Table 5.15 first separates the data into individual fidelities, and then compares/combines the two fidelities.

The results for NetLogo alone suggest that the confidence over all the crowd sizes was similar hence having a non-significant effect. However if results for crowd sizes are split between those with individual breaking rules and individuals not breaking rules then the observers confidence does show a significant difference. A significant effect was also found when analysing all five options.

VBS2 had similar results to NetLogo with the exception that there was a non-significant effect in the confidence for all five options.

When comparing/combining NetLogo and VBS2, the only noticeable change to the observer's confidence is when comparing the three measurements, fidelity, crowd sizes (breaking rules & not breaking rules) and the five options which shows the amongst this level of measurement there is a difference towards the observers confidence

Table 5.15 Summary of Findings - Impact of Crowd Size

Impact of Crowd Size (NetLogo only and VBS2 only)		
Measurement	**NetLogo Only**	**VBS2 Only**
Crowd Sizes	Non-Significant	Non-Significant
Crowd sizes (combined) and Five Options (1…5) for each	Non-Significant	Non-Significant
Five Options (1…5)	[*F* (4, 52) = 5.118, p < 0.001]	Non-Significant
Crowd sizes Breaking rules compared with Crowds sizes not breaking rules	[*F* (2, 26) = 6.611, p < 0.005].	[*F* (4, 48) = 4.434, p < 0.004]
Three runs (Total) individuals not breaking the rules, compared with Three runs (Total) individual breaking the rules	Non-Significant	Non-Significant
Crowd Size Comparison of NetLogo and VBS2		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo	Non-Significant	
Fidelity removed - Six runs individuals not breaking the rules, compared with Six runs individual breaking the rules		Non-Significant
Fidelity - VBS2, NetLogo and Crowd Size	Non-Significant	
Fidelity - VBS2, NetLogo and the Five Options for each (1…5)	Non-Significant	
Fidelity removed - Crowd sizes Breaking rules compared with Crowds sizes not breaking rules		[*F* (2, 50) = 5.811, p < 0.005]
Fidelity - VBS2, NetLogo, Crowd sizes Breaking rules compared with Crowds sizes not breaking rules and Five Options (1….5) for each	[*F* (4, 100) = 3.772, p < 0.007]	

CHAPTER 6

EXPERIMENT II

6.1 Introduction and Objectives

The concept for experiment II was to design a simulation of a CCTV operator‘s display area. Building on the lessons learnt from the first experiment, experiment II the Author altered the way in which the information was to be delivered from a static slide presentation to a fully animated virtual environment showing on 12 different cameras playing 3 different scenarios over two minutes. The 10-point scale system used to measure observer‘s confidence was replaced by an eye tracking equipment which collects measurements of dwell time and fixation times of the observer eyes on the screen. This data would enable a more accurate way of detecting were the observer was looking and how long for during each of the experiment runs.

Information would also be provided verbally to the observers in the form of an intelligence report. This was to measure if information altered the observers understanding of the virtual world.

Unlike the first experiment, this experiment had some expert observers in the form of an intelligence officer who is currently serving in the British Army and a student who had experience in CCTV operations. This enabled the Author to analyse how experts see a virtual world considering their training is based on real world events and compare this with naive observers.

The objective of the experiment was to gain an insight into answering the following:

- Can the observer detect what is happening in the virtual world without any information provided? Does the observer‘s understanding or focus change when information is provided in the form of a verbal intelligence report? [Thesis question 1, 1a, 1b].
- Does fidelity have an effect on the observer‘s response? Does it become more accurate with VBS2 or NetLogo? [Thesis question 2,2a, 2b]
- Can a difference can be found between a naïve observe and an expert observer? [Thesis question 3]

The following sections will describe the design of the experiment and the reason behind CCTV camera design and then look at the results from the experiment to see if the thesis questions can be answered.

6.2 Scenario Description and Design

The design for experiment II allowed the setup to have different camera angles in the virtual world. The first experiment covered only one vantage point which was almost top down where the observer was looking down onto the crowds. This was slightly different for VBS2 due to the 3D environments that can be created in VBS2. Experiment II created four different vantage points, one per camera, with a total of four per scenes. This reflects the input from the literature review and SME interview which suggest observations are conducted at different vantage points.

The design was built using three software tools: CS4 Adobe Premier Pro, NetLogo and VBS2. The use of each software tool is described below.

CS4 Adobe Premier Pro was used to develop a way of allowing three scenarios each with four cameras (12 cameras in total) to play simultaneously. Fig 6.1 below illustrates the template designed and shows the similarities between the Met Police visit to an operations control room (left image) and the Adobe CS4 design to simulate a CCTV operator display (right image).

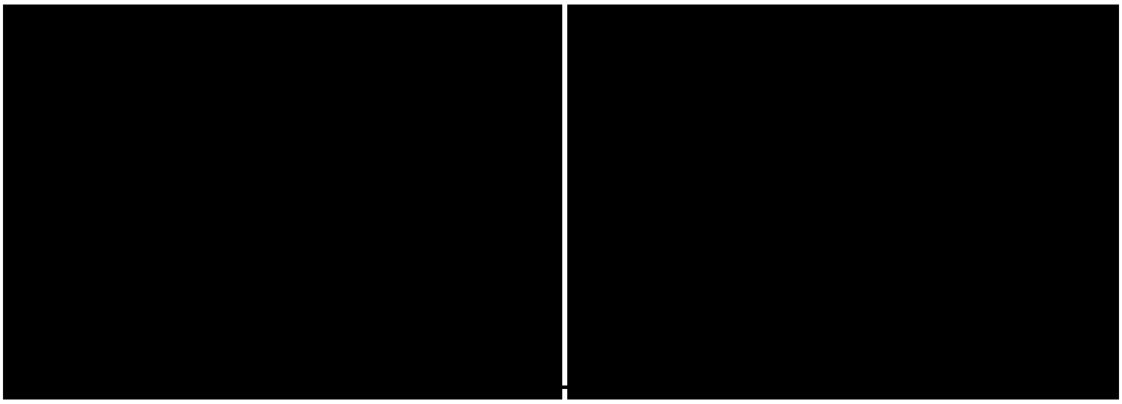

Fig. 6.1 CCTV Operator's display

The VBS2 software development kit (SDK) was used to design three different scenarios for a 3D (higher fidelity) virtual environment. Each scenario had four cameras at different vantage points. The scenarios for the 3D virtual environment were designed using VBS2 assets which come pre-built in the SDK. Assets are object created for the 3D virtual environment these assets can for example be building, roads, people etc. All three scenarios were designed by the Author.

The first scenario was based around people getting food from a food aid area. People in different outfits move forward to a collection point to collect the food then turn around to return to home (or safe place). The top rows of cameras in Fig. 6.2 (i.e. cameras 01 to camera 04) display the four different vantage points. The vantage points in VBS2 used waypoints (waypoints follow a design path in the virtual world) with virtual cameras attached to them (virtual camera record like real cameras accept they record what is happening in the virtual world). The different vantage points are different waypoints setup in the virtual environment to record at different angles of the virtual world. An example is camera 01 in Fig. 6.2 which was positioned on top of a pole above the food aid station. Likewise, camera 04 was positioned as an unmanned aerial vehicle (UAV) flying over the food aid area. The purpose behind this is to collect data to see if vantage points have an effect on the observer‘s choice of cameras.

The second scenario (cameras 05 to 08) depicted a gated area controlled by military sentries. This scenario shows British military personnel patrolling the gated area along with a group of military vehicles (one Jackal and three Mastiffs) driving along the road outside the gated area.

The third scenario (cameras 09 to 12) depicted a helicopter landing zone where four military personnel (three British and one American) are waiting for a Chinook helicopter to land. The helicopter comes into land and then the four military personnel board the helicopter. The Chinook then takes off with the four military personnel onboard.

An example of all three scenarios and 12 cameras designed using VBS2 is shown in Fig 6.2 below.

Fig. 6.2 VBS2 3D design

NetLogo was used for the 2D (lower fidelity) virtual environment design but keeping the same scenarios as VBS2. The design tries to copy the scenarios as closely as the software would allow to that of VBS2. The purpose was of creating two identical virtual environments was so that the observers could identify one with the other and make an informed judgment on fidelity.

The virtual environment in NetLogo had to be designed completely from scratch by the Author. The different coloured areas in the virtual world are X and Y patch coordinates. For example, the gate area in the middle row of Fig.6.3 was created using X Y coordinates and different colours in NetLogo. Some predefined assets like a stick person are available in NetLogo. These ‗stick persons‘ where modified to look like military personnel in the gated area and helicopter landing zone. The Author intentionally made the people in the food aid scenario the same colour as the VBS2 food aid people. The helicopter was built using a design tool in NetLogo.

It was difficult to set up the vantage points identically to VBS2, primarily due to the limited level of detail that a 2D SDK can provide. The work around for this was to change the level of detail of the people in the virtual world. To create a top down vantage point, the Author made the people in the virtual world circles that moved round. Changing the angle slightly in NetLogo, and inserting stick people, allowed to create a feeling of depth to the world and make it look as if the camera vantage point had changed. NetLogo also did not have waypoint or camera. This people in the world are programmed to follow a path of X and Y coordinates to make it look as if they are walking from one point to another.

An example of all three scenarios and 12 cameras designed using NetLogo is shown in Fig 6.3 below.

Fig. 6.3 NetLogo 2D design

NetLogo does not have a virtual camera. Therefore the video for each camera was captured using a specialist software called FRAPS and then edited in Adobe CS4.

The three scenarios (twelve cameras in total) run for an approximately 120 seconds. Each contextual scenario (virtual environment) had people and or vehicles moving in them over a period of time. Each contextual scenario had four vantage points (4 cameras per scenario). The contextual scenario time sequence is approximately 30-40 seconds and are compressed and built into cameras template using Adobe CS4, making the whole scenario run last 120 seconds. This creates an impression that each camera is monitoring different areas in the virtual environment. Repeating the 30 - 40 second scenario clips over 120 seconds made it feel as if the cameras were refreshing.

6.3 Experiment Setup and Running

The observer's in this experiment wore a head mounted eye tracker as shown in Fig. 6.4 below. This equipment allows tracking of where the observer's eyes are focusing on the screen and for how long.

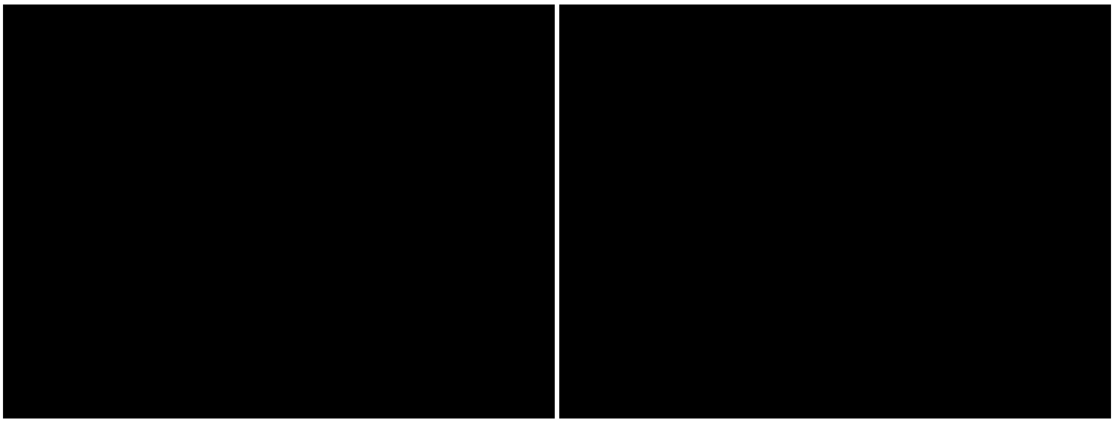

Fig. 6.4 Headset worn by the observer's

Each eye needs to be calibrated so that it known exactly what the observer is looking at Fig. 6.5 shows the setup procedures. The observer is required to look at the one black dot in the centre of the screen. This black dot appears every second, first in the centre, then in each four corners (one at a time) of the screen. Once the eyes have been calibrated, the system needs to carry out a validation which can only commence if calibration has been successful.

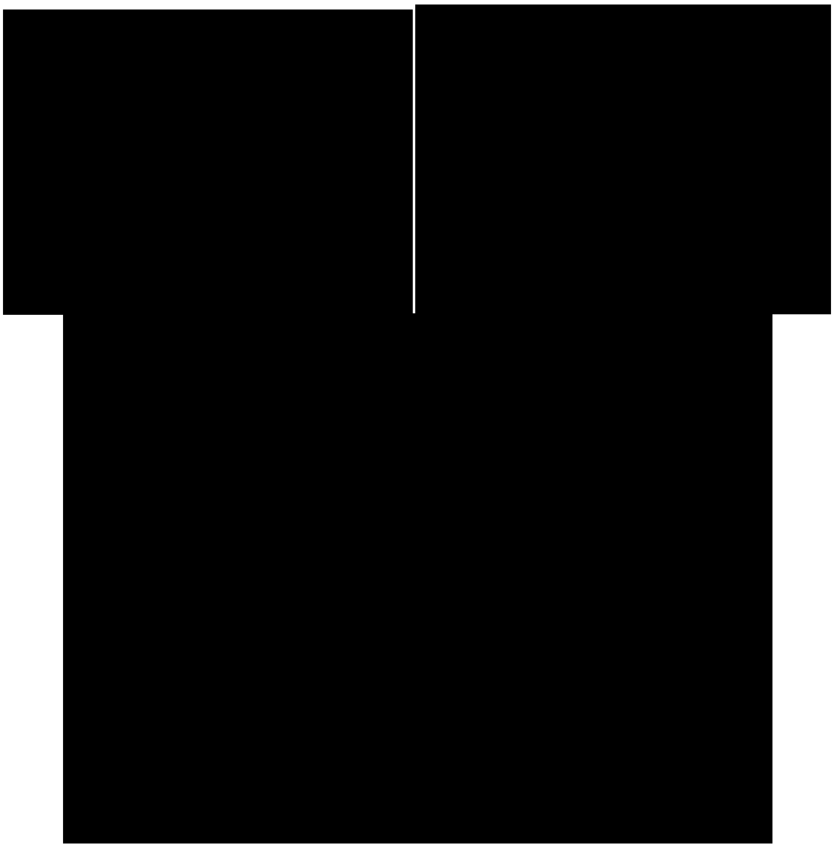

Fig. 6.5 Setup calibration & validation for eye tracking equipment

Validation is similar to the calibration. It requires the observer will be to look at the black dots on the screen. The purpose of the validation is to measure the difference between the computed fixation and the fixation position for the target obtained in calibration. The lines on the screen show the error of the gaze. If the error is too great this will cause the validation to fail (shown in red as —poor‖eye gaze). Fig 6.5 (bottom image) shows the validation green and blue ‗+‘symbols.

Once the calibration and validation is completed the experiment can commence. The data collected by the eye tracking equipment is stored in files until the user defines what data is needed. Once this data is defined the eye tracker software produces a comma separated values (CSV) file. The data collected for experiment II were interest areas defined by the experimenter using the eye tracker software. The interest areas were as follows:

- Twelve cameras, 12 separate interest areas;
- Three scenarios (all four cameras together) 3 interest areas, top, middle bottom area; and
- Total area incorporating all 12 cameras, 1 interest area.

This was a total of 16 different interest areas. These are depicted in Fig. 6.6 below.

Once the interest areas were defined, they could be saved and used for each of the observers. This maintains continuity to the measurements. The data recorded for these interest areas was dwell time percentage, fixation percentage and fixation count.

Total time is taken from the total interest area (yellow are in Fig 6.6 below). This is the most accurate start time to the experiment, which was approximately 120 seconds from start to finish. The reason for taking total time from the total interest area was that experiment was only interested in the camera areas. The dwell time percentage and fixation percentage is a percentage of time from the total interest area.

Fig. 6.6 Setup Areas of Interest (IA)

Once all the calibration and validation was completed the observer was taken through a trial run (trial one) similar to Fig 6.1 just showing red information for the cameras. This was to get observers use to looking at the screen and the 12 cameras.

Once the observers were comfortable with the trial run, they were taken through a total of six trials, with one VBS2 and one NetLogo sequence similar to Fig. 6.2 & Fig. 6.3.

For trials two (VBS2) and three (NetLogo), the observers were given no intelligence and were allowed to look freely around the screen. After trial two VBS2 had finished they were ask to describe what he/she thought was happening and which cameras they felt were import, three in total. The process is repeated for NetLogo trial three.

For trial four (VBS2) and five (NetLogo), the experiment leader (and Author) provided intelligence verbally. The observer was informed that certain people in the area wanted to disrupt the food aid area and that cameras had been setup to see if any suspicious people could be seen. After trial four (VBS2) had finished the observer was asked to describe what he/she thought was happening and which cameras they felt were import, three in total. The process was repeated for trial five (NetLogo).

For trial six (VBS2) and seven (NetLogo), the experiment leader (and Author) also provided verbal intelligence. The observer was informed that a possible unauthorised vehicle was in the area and that cameras had positioned to see if any suspicious vehicles could be identified in the area. After trial six (VBS2) had finished the observer was asked to describe what he/she thought was happening and which cameras they felt were import, three in total. The process was repeated for trial seven (NetLogo).

Although the intelligence informed the observers that a suspicious person could be present in the food aid area and a suspicious vehicle was in the area, no such person or vehicle was put into the scenarios. As experiment I suggested little influence of adding a person into the virtual environment, the omission of a suspicious person or vehicle in experiment II was intentional to see if it influenced observers. Furthermore the intelligence was provided during the experiment to see if observers became more aware of the scenario and concentrated on one area more than another (potentially missing something in another camera). In addition, this allows a comparison between how experts and naïve observers digest the information and process it into a search.

For each trial, the scenarios positions would change between the top, middle and bottom interest areas. For example, if VBS2 had the helicopter landing zone in the top interest area, NetLogo may have the helicopter landing zone in the middle interest area instead. This was done to stop observers becoming accustomed to the position of a scenario and also to see more clearly whether a particular scenario/camera attracted the most interest from the observer.

Although the scenarios rotated, the camera perspective positions remained in the same order. For example, as shown in Fig. 6.2, camera 01 shows the food aid vantage point looking at the food aid distributors. When the scenarios rotate, the position of camera 01 can only be swapped with in either camera 05 (middle area) or camera 09 (bottom area), i.e. it will always remain on the far left of the screen. Table 6.1 shows the cameras and the possible rotation combinations.

Table 6.1 Camera Rotations

Scenario Positions	Camera 1	Camera 2	Camera 3	Camera 4
Top Area	Camera 01	Camera 02	Camera 03	Camera 04
Middle Area	Camera 05	Camera 06	Camera 07	Camera 08
Bottom Area	Camera 09	Camera 10	Camera 11	Camera 12

The number of observers for this experiment was 12, with two having experience in intelligence and CCTV operations. Therefore the data has been split between 10 naïve observers and two experienced observers. Section 6.4 will try to answer some or all of the hypothesis questions by analysing the data collected from the eye tracker and the observers responses to the questions asked.

6.4 Experiment Results

Similarly to experiment I, the data were analysed using Excel and IBM SPSS Statistics Data Editor Version 19.

6.4.1 Camera preferences measuring mean dwell time

6.4.1.1 Mean dwell time per camera based on no context

The starting point for the analysis was to establish a baseline by analysing which cameras in the experiment produced the highest dwell times. This was to firstly understand which cameras the observers were concentrating on the most regardless of content. Fig 6.7 illustrates the total combined (NetLogo & VBS2) mean dwell time for each camera with standard deviation shown as error bars. Due to the low number of expert observers their data is represented as circles and a dashed line to distinguish it clearly from the results of naïve observers.

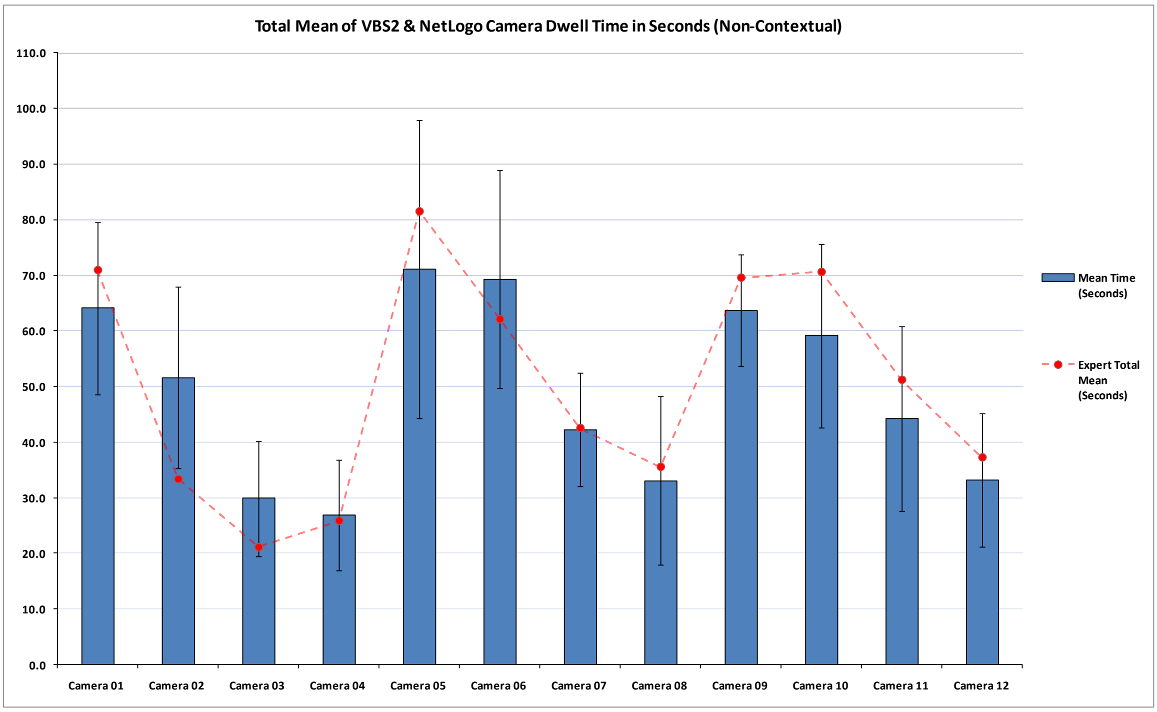

Fig. 6.7 Total mean dwell time per camera

Table 6.2 shows the mean dwell times data for each camera (Fig. 6.7 illustrates the data).

Table 6.2 Total Mean Dwell Time (Seconds)

	Naïve Total Mean Dwell Time (Seconds)	Standard Deviation (Seconds)	Expert Total Mean Dwell Time (Seconds)	Standard Deviation (Seconds)
Camera 01	64.08	15.46	70.94	17.75
Camera 02	51.64	16.30	33.37	0.55
Camera 03	29.87	10.32	21.10	0.20
Camera 04	26.91	10.00	25.79	9.20
Camera 05	71.07	26.77	81.55	28.34
Camera 06	69.33	19.51	62.15	1.34
Camera 07	42.26	10.22	42.47	1.15
Camera 08	33.08	15.14	35.52	16.79
Camera 09	63.67	10.05	69.58	8.58
Camera 10	59.12	16.57	70.64	15.82
Camera 11	44.22	16.63	51.15	9.70
Camera 12	33.22	11.99	37.19	3.24

The combined data and graph (Fig 6.7) suggests that the mean dwell time patterns of expert observers and naïve observers for each camera are similar. In both cases there is a trend to focus more on the right hand cameras, (Cameras 01, 02, 05, 06, 09 and 10 refer to Fig. 6.6 for visual) and less on the left. The standard deviation (similar to the first experiment) is high. Also there is a difference in the deviation of the naïve observers compared to the experts who seem to be in a greater consensus of which cameras to dwell on. However, with only two experts this is hard to verify. This suggests that although there is different dwell time for each camera there are also large deviations attached to them, suggesting a low level of consensus.

Fig 6.8 and Fig. 6.9 show the mean dwell times of both software tool NetLogo and VBS2. This enables a comparison to be made between the two levels of fidelity. The red circle and dash line shows mean dwell time of the expert observers.

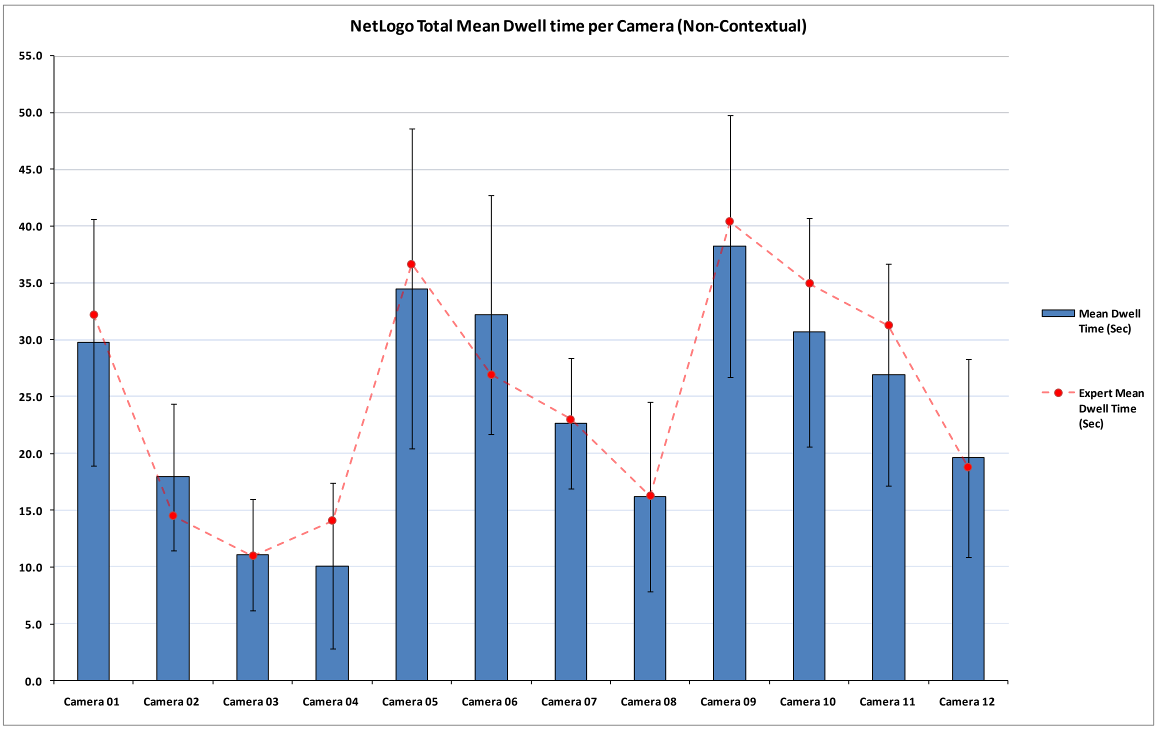

Fig. 6.8 NetLogo's mean dwell time per camera

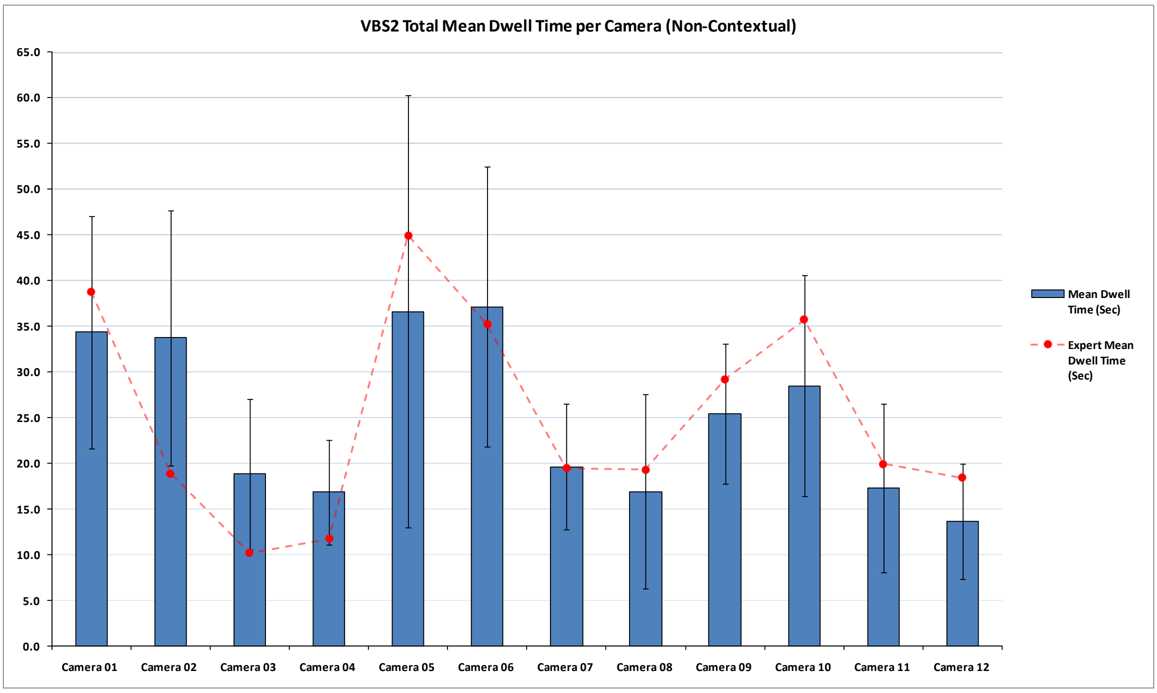

Fig. 6.9 VBS2 mean dwell time per camera

Table 6.3 is mean dwell time per camera for both NetLogo and VBS2 for both naïve and expert observers. This data is illustrated in Fig. 6.8, 6.9 and the comparison in Fig 6.10. Fig 6.10 also includes the expert mean dwell times, the black circle and dash line donates VBS2 and the yellow circles and dash line donates NetLogo.

Table 6.3 NetLogo and VBS Mean Dwell Times

	VBS2 Mn Dwell Time (Sec)	SD (Sec)	Expert VBS2 Mn Dwell Time (Sec)	SD (Sec)	NetLogo Mn Dwell Time (Sec)	SD (Sec)	Expert NetLogo Mn Dwell Time (Sec)	SD (Sec)
Camera 01	34.32	12.75	38.72	32.53	29.76	10.88	32.23	14.78
Camera 02	33.72	13.96	18.89	4.56	17.92	6.48	14.48	4.01
Camera 03	18.80	8.26	10.14	0.79	11.07	4.90	10.96	1.00
Camera 04	16.81	5.78	11.73	0.62	10.10	7.29	14.06	8.58
Camera 05	36.57	23.65	44.90	18.10	34.49	14.08	36.65	10.24
Camera 06	37.12	15.35	35.19	2.54	32.21	10.51	26.97	3.88
Camera 07	19.59	6.86	19.45	0.36	22.67	5.74	23.02	0.79
Camera 08	16.90	10.61	19.27	5.77	16.19	8.32	16.25	11.02
Camera 09	25.43	7.66	29.17	15.88	38.24	11.51	40.41	7.29
Camera 10	28.46	12.07	35.69	0.94	30.66	10.04	34.96	14.89
Camera 11	17.27	9.23	19.89	3.22	26.95	9.76	31.26	6.47
Camera 12	13.63	6.31	18.40	9.13	19.59	8.70	18.79	5.89

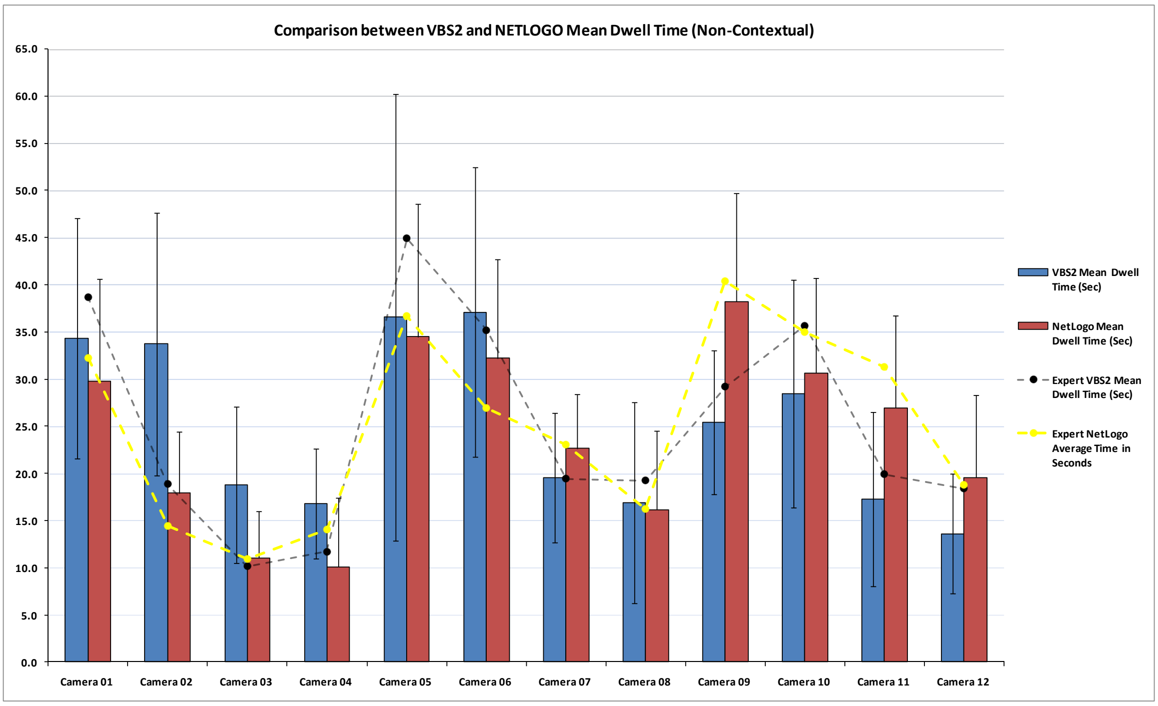

Fig. 6.10 NetLogo & VBS2 comparison of mean dwell time per camera

Differences become apparent when comparing between the two fidelities (NetLogo and VBS2). The pattern for naïve observers seem to differ between the mean dwell time of NetLogo (red) and VBS2 (blue). This also appears seem to appear as well in the expert. Camera's 01 to 06 and Camera 08 have longer mean dwell times in VBS2, whereas Camera's 07, 09, 10, 11 and 12 are have a longer mean dwell times in NetLogo.

The expert's black and yellow dashed line suggests similar patterns to that of the naïve observers. The mean dwell time for VBS2 longer than NetLogo up to Camera 08 and this is then reversed with NetLogo having the longer mean dwell time for the later cameras.

What can be suggested thus far is that the patterns of dwell time are similar between the naïve observes and the expert observers.

Another interesting factor is the difference between the NetLogo and VBS2 mean dwell times. The data were analysed using the dwell times of 10 naïve observers to find if there was any significant main effect between NetLogo, VBS2 and the 12 cameras.

A two-way repeated-measure ANOVA within-subject effects measured the differences between the two fidelities (NetLogo and VBS2) and 12 cameras. The measurement suggests there is a significant main effect when comparing NetLogo and VBS2 [$F\ (1, 9) = 12.003$, $p < 0.007$]. Fig 6.11 illustrates that by adding the total mean dwell times for NetLogo and VBS2, there is a difference between the two fidelities. Although the times for each run where the same, this suggests that within the total area of interest VBS2 had the higher dwell times. This means that potentially in NetLogo the observers had a tendency to dwell outside of the interest area creating a significant effect between VBS2 and NetLogo. However, this difference is small. VBS2 only has 8.79 seconds more dwell time compared to NetLogo. The deviation difference between the VBS2 and NetLogo is also low with only a 4.20 second difference between them.

There is a significant main effect on all the 12 cameras [F (11, 99) = 10.151, p < 0.000]. This concurs with the Fig. 6.7 which illustrates the difference in dwell times of each of the cameras. To analyse the data further and see where the differences are, a Pairwise comparisons analysis was produced. This analysis compares each camera with one another, comparing their dwell times. Table 6.4 is small sample taken which shows Camera 01 compared to the other 11 cameras. This data suggest that there is significant difference between camera 04, 08 and 12 (full results in Appendix A).

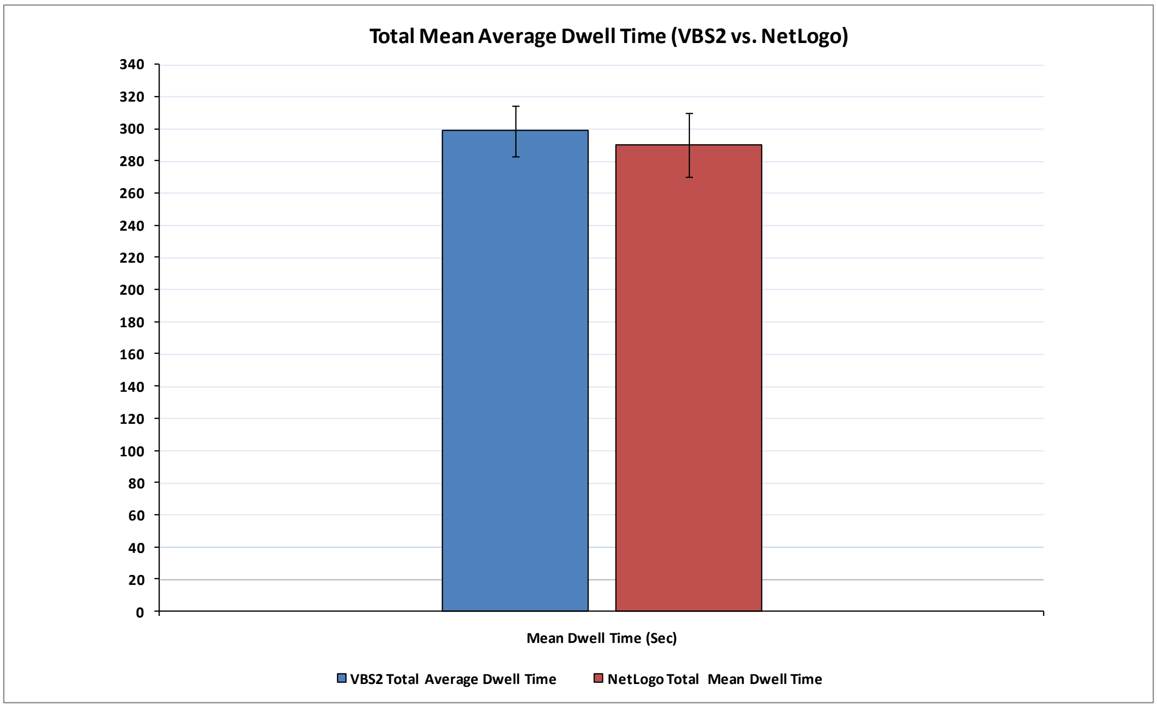

Fig. 6.11 Total mean dwell time NetLogo vs. VBS2

Table 6.4 Pairwise Comparison Table

(I) Cameras	(J) Cameras	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	6.223	3.029	1.000	-8.869	21.315
	3	17.108*	2.101	.001	6.643	27.574
	4	18.589*	2.754	.006	4.869	32.310
	5	-3.491	5.423	1.000	-30.509	23.528
	6	-2.622	4.822	1.000	-26.644	21.400
	7	10.910	3.112	.440	-4.595	26.415
	8	15.500*	2.466	.009	3.216	27.784
	9	.209	3.550	1.000	-17.476	17.894
	10	2.479	4.059	1.000	-17.744	22.702
	11	9.934	4.106	1.000	-10.522	30.390
	12	15.429*	2.755	.022	1.705	29.154

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni

The final measurement compares the interaction between NetLogo, VBS2 and their 12 cameras to see if there are any differences. The result is a significant main effect VBS) [F (11, 99) = 2.664, $p < 0.005$]. This correlates with the first two measurements which both had a significant difference. By comparing NetLogo's 12 cameras with VBS2's 12 cameras a difference is found, this presumably being the dwell times. This would concur with Fig 6.10 which illustrates the dwell times of NetLogo and VBS2 and their 12 cameras.

6.4.1.2 Conclusion on dwell times non-contextual

Table 6.5 Summary of Non-Contextual Findings

Comparison of Mean Dwell Time per Camera (12) between NetLogo & VBS2 (Non-Contextual)		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo (Total Dwell Time)	[F (1, 9) = 12.003, p < 0.007]	
Fidelity removed - 12 Cameras		[F (11, 99) = 10.151, p < 0.000]
Fidelity - VBS2, NetLogo and the 12 Cameras each	[F (11, 99) = 2.664, p < 0.005]	

Without considering the context (i.e. scenarios) it can be suggested that there are significant differences between software tools that display different levels of physical fidelity. Furthermore the data suggests that cameras on the right hand side of the screen received more attention from the observers than cameras on the left.

What is also interesting is that there appear to be similarities in the mean dwell time trends between the naïve observers and the expert observers suggesting they were both focusing on similar scenes. Table 6.5 shows the hypothesis question 2 and 2b have been answered by showing observers do perceive differently in different fidelities and that there is a significant difference to demonstrate this.

6.4.1.3 Mean dwell time based on three contextual scenarios

This section breaks down the data from the non-contextual information above, to the contextual understanding of what the observers were dwelling on the most for all of the trial. The contextual environment is the virtual environments that have been created for the cameras. The three virtual environments (scenarios) are the food aid (FA), gate and sentries (GS), and helicopter landing zone (HP).

The measurements taken from all observers was organised into the three scenarios stated above. For each scenario there are 4 cameras (total of 12) three scenarios times by four cameras. Each camera for each scenario is labelled camera 1 to 4 for simplicity. As the scenarios were rotated this does not mean that camera 1 to 4 for each of the scenarios necessarily equals the top row of cameras 01 to 04.

The analysis for this is not to see where exactly the observer was dwelling on (i.e. Top Area, Camera 01) but at what contextual scenarios and their four vantage point cameras the observer was dwell on (i.e. helicopter landing camera 1) Fig. 6.12 represents the mean dwell time per camera per scenario regardless of fidelity (i.e. taking the combined data from NetLogo and VBS2). Each contextual scenario is represented by a different colour. Red is the food aid distribution environment, green the gate and sentry environment and blue is the helicopter land zone environment. The black circles and dash line represent the experts mean dwell times.

Table 6.6 Dwell Time Data for Fig. 6.12

Contextual Scenario	Camera per Scenario (1 to 4)	Mean Dwell Time	Standard Deviation (Sec)	Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)
Food Aid	Camera 1	102.92	30.33	123.91	5.64
	Camera 2	76.90	25.43	75.15	9.68
	Camera 3	55.14	17.05	58.39	7.88
	Camera 4	43.80	12.44	44.71	3.26
Gate & Sentry	Camera 1	55.38	12.71	54.19	8.33
	Camera 2	50.30	16.19	41.52	12.55
	Camera 3	30.21	8.49	22.42	7.92
	Camera 4	24.37	11.79	24.82	6.97
Helicopter Pad	Camera 1	40.51	10.23	43.98	16.49
	Camera 2	52.90	14.07	49.50	4.51
	Camera 3	31.00	7.24	33.91	7.46
	Camera 4	25.05	9.51	28.96	18.99

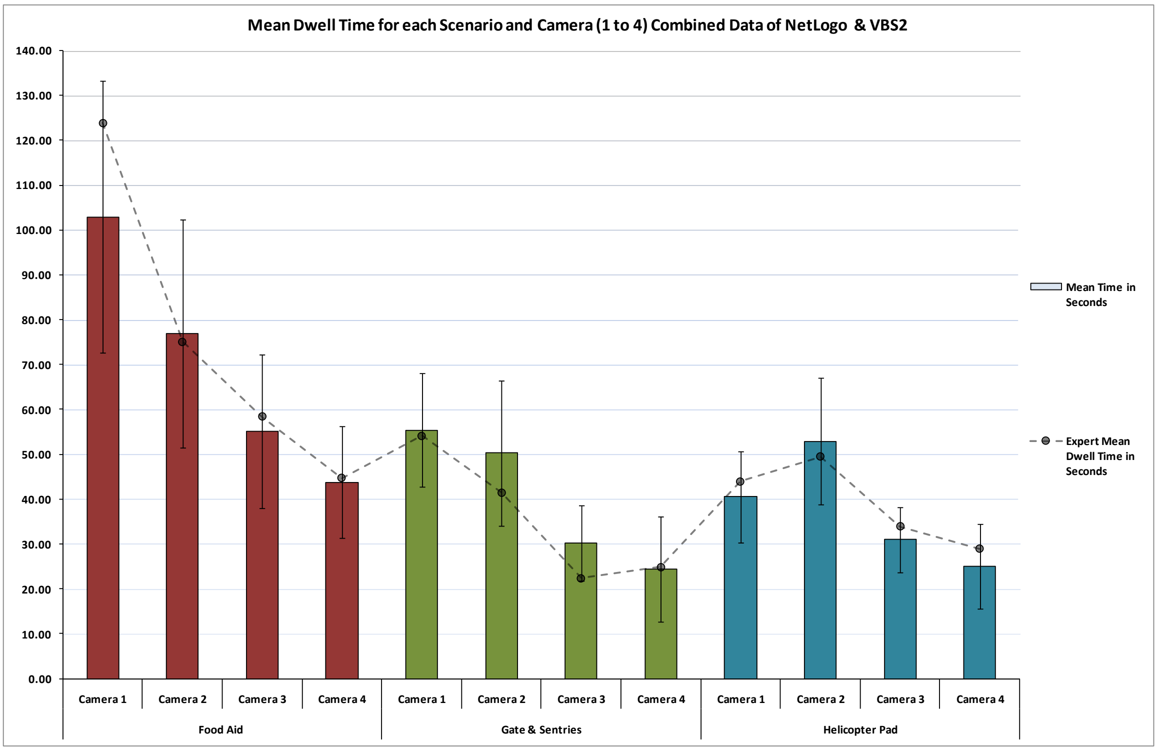

Fig. 6.12 Mean Dwell Time for each Scenario

The data at this point suggests that the observer's interest is primarily on camera 1 of the food aid scenario. Both experts spent most dwell time in camera 1. As previously shown in Fig. 6.2, camera 1 of the food aid scenario shows the vantage point were the camera is focused on the food aid distribution area. What is interesting from this data is that even though the food aid scenario was rotated randomly throughout the exercise it still generated the highest dwell time of both types of observers.

To analyse the data a two-way repeated-measure ANOVA within-subjects was conducted. The data that was to be analysed was the combined dwell times of both NetLogo and VBS2. The purpose was to measure the three contextual scenarios and their four vantage point cameras to find if there was any significant difference between them.

The first analysis was to measure the differences between the three contextual scenarios. The result suggests that there is a significant main effect between the three contextual scenarios [F (2, 18) = 21.422, $p < 0.000$]. This can be seen clearly in Fig. 6.12 whereby the dwell times are considerably different across the three scenarios. Why the food aid scenario has the highest dell time is unclear at this time. It is suggested that people read from left to right which would then mean observers commenced there search at Camera 01, however this is unlikely to be the case as the data focuses on how much time the observers spent on the three contextual scenarios and the four vantage points cameras for each scenario, not on the sequence of looking from one point to the next.

The following measurement is a comparison between the total dwell times of the three contextual scenarios and the four vantage point cameras. The result suggests that there is a significant main effect between the cameras [F (3, 27) = 26.782, $p < 0.000$]. A Pairwise Comparison Table 6.7 of all the camera is also produced which can be further analysed to see which shows how each camera differs for each other.

Table 6.7 Comparison of the Four Vantage Point Cameras

Measure:MEASURE_1

(I) VPCameras	(J) VPCameras	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	6.240	4.933	1.000	-10.355	22.836
	3	27.489*	4.500	.001	12.349	42.628
	4	35.200*	4.007	.000	21.721	48.679
2	1	-6.240	4.933	1.000	-22.836	10.355
	3	21.248*	4.064	.003	7.576	34.921

	4	28.960*	5.926	.005	9.023	48.896
3	1	-27.489*	4.500	.001	-42.628	-12.349
	2	-21.248*	4.064	.003	-34.921	-7.576
	4	7.711	3.748	.419	-4.898	20.321
4	1	-35.200*	4.007	.000	-48.679	-21.721
	2	-28.960*	5.926	.005	-48.896	-9.023
	3	-7.711	3.748	.419	-20.321	4.898

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

The final analysis is to compare the three contextual scenarios and their four vantage point cameras with one another. The results suggest that there is a significant main effect between the interaction of three contextual scenarios and their four vantage point cameras - [F (6, 54) = 7.473, $p < 0.000$]. Fig.6 .12 clear shows that there is large difference between the dwell times in each of the three scenarios and their four vantage point cameras this correlates with the analysis.

6.4.1.4 Mean dwell time based on three contextual scenarios – VBS2 only

Table 6.8 and Fig 6.13 show the mean dwell time for VBS2 only, for each of the contextual scenario. The data in the coloured columns is for the 10 naïve observers mean dwell time for each of the contextual scenarios and the back circle and dash line is the mean dwell time of the experts for VBS2.

Table 6.8 Dwell Time Data for VBS2 Fig. 6.13

Contextual Scenario	Cameras	VBS2 Mean Dwell Time (Sec)	Standard Deviation	VBS2 Expert Mean Dwell Time (Sec)	Standard Deviation
Food Aid	Camera 1	55.17	21.60	70.23	2.10
	Camera 2	35.83	17.69	32.30	3.99
	Camera 3	17.00	5.29	18.32	3.36
	Camera 4	18.08	5.98	19.46	4.00
Gate & Sentry	Camera 1	26.11	12.31	24.62	0.01
	Camera 2	30.13	12.59	29.92	12.33
	Camera 3	18.42	5.08	17.21	6.79
	Camera 4	13.13	6.78	14.40	3.26
Helicopter Pad	Camera 1	15.04	6.84	27.82	10.43
	Camera 2	33.34	10.62	26.34	7.07
	Camera 3	20.24	5.43	19.59	1.48
	Camera 4	16.13	8.22	21.19	0.24

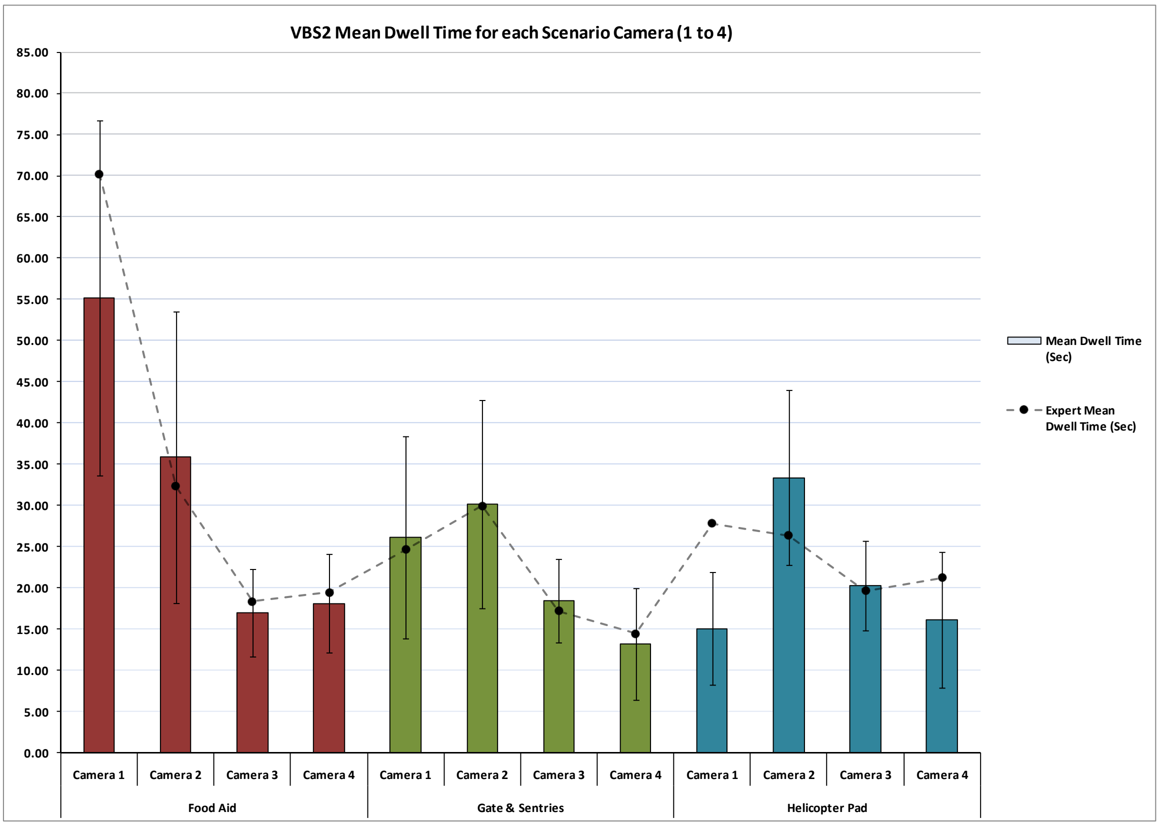

Fig. 6.13 VBS2 Mean Dwell Time for each Scenario

The data shows camera 1 in the food aid scenario has the highest dwell mean time for both the naïve and expert observers. The second highest dwell time is for cameras 2.

For the food aid and gate & sentry contextual scenarios, the first two vantage point cameras had the highest dwell time. However in the helicopter pad contextual scenario, vantage point cameras are 2 and 3 have the highest mean dwell times. This is suggesting that it was not always the left hand camera that had the greatest dwell time for each scenario and that the rotating of the position did not alter this.

6.4.1.5 Mean dwell time based on three contextual scenarios – NetLogo only

Like VBS2 the data has been analysed in Excel Table 6.9 and Fig.6.14 show the mean dwell time data base on the three contextual scenarios for NetLogo. The data in the coloured columns is for the 10 naïve observers mean dwell time for each of the contextual scenarios and there four vantage point cameras. The black circles and dash line are experts mean dwell time.

Table 6.9 Dwell Time Data for NetLogo Fig. 6.14

Contextual Scenario	Cameras	NetLogo Mean Dwell Time(Sec)	Standard Deviation	NetLogo Expert Mean Dwell Time (Sec)	Standard Deviation
Food Aid	Camera 1	47.75	13.74	53.68	3.54
	Camera 2	41.07	11.04	42.85	13.66
	Camera 3	38.14	13.90	40.07	4.51
	Camera 4	25.72	8.80	25.25	0.74
Gate & Sentry	Camera 1	29.27	6.63	29.56	8.34
	Camera 2	20.17	6.58	11.59	0.22
	Camera 3	11.79	5.95	5.22	1.13
	Camera 4	11.24	8.05	10.42	3.71
Helicopter Pad	Camera 1	25.47	8.30	26.05	12.93
	Camera 2	19.55	7.20	21.96	0.87
	Camera 3	10.76	4.67	19.95	0.96
	Camera 4	8.92	3.24	13.43	10.73

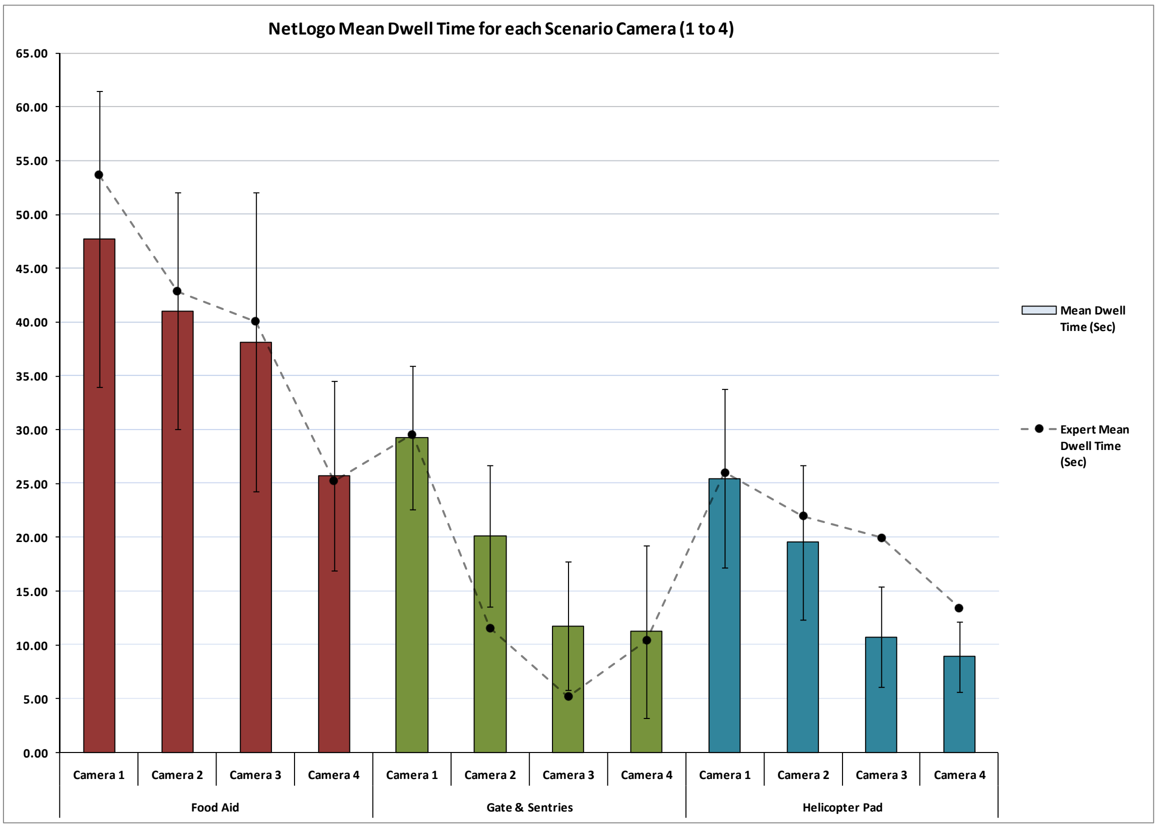

Fig. 6.14 NetLogo Mean Dwell Time for each Scenario

The results for NetLogo are similar to VBS2 as camera 1 in the food aid scenario shows the highest mean dwell time for both naïve and expert observers. Considering that there is a significant difference in the level of fidelity between VBS2 and NetLogo, it is interesting to find such similarities in the camera dwell times. However, whilst there are clear similarities in the trends, the mean dwell time in absolute terms does show some differences when it comes to expert observers. As shown in Table 6.8 and 6.9, expert observers had a mean dwell time of 70.23 seconds for camera 1 food aid in VBS2 compared to only 53.68 seconds for the same camera in NetLogo.

What is interesting is that by adding all the mean dwell times of the entire three contextual scenarios four vantage point camera it is clear that both naive and experts spent more time in VBS2 then NetLogo. This concurs with the first analyse when comparing the total dwell times between VBS and NetLogo, whereby there was a significant difference between the two suggesting that the observers may have tendency to dwell outside the total area of interest in NetLogo.

Looking at the data, some of the NetLogo‘s mean dwell times are higher in certain cameras. The Author would have presumed that observers would have spent more time dwelling on NetLogo scenarios because of the lower fidelity. However this is not the case and VBS2 had the highest total dwell time.

Furthermore, the Author observed that the same preference for cameras on the left hand side is also evident in the NetLogo data, i.e. higher dwell times for camera 1 and camera 2 for each of the contextual scenario.

6.4.1.6 Mean dwell time based on three contextual scenarios – compared analysis

Fig 6.15 combines Fig 6.13 and Fig. 6.14 data to illustrate the difference between VBS and NetLogo. The data is the same as in Tables 6.8 and 6.9.

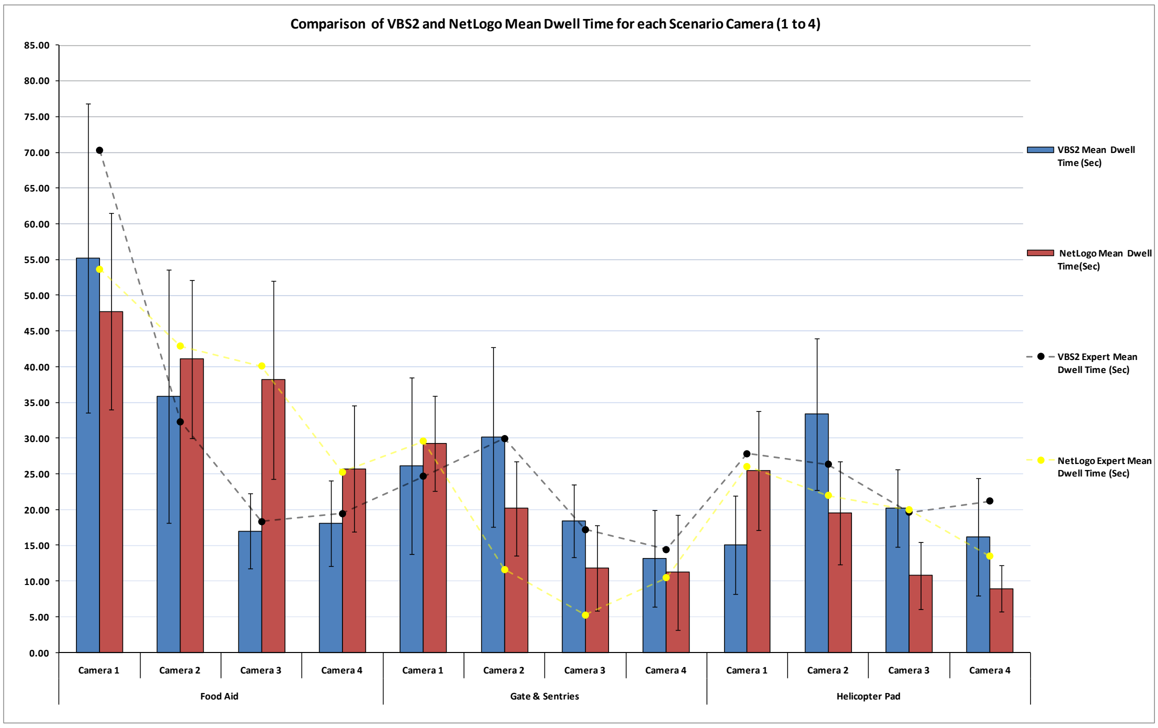

Fig. 6.15 Comparison of Mean Dwell Time for each Scenario

This comparison highlights more clearly that for some cameras and scenarios there are stronger differences between NetLogo and VBS2 mean dwell times. For example, in Fig. 6.15 the NetLogo dwell time is markedly higher in camera 3 for the food aid scenario compared with VBS2. Fig. 6.16 shows what the observer would have been looking at in Camera 03 and Camera 11.

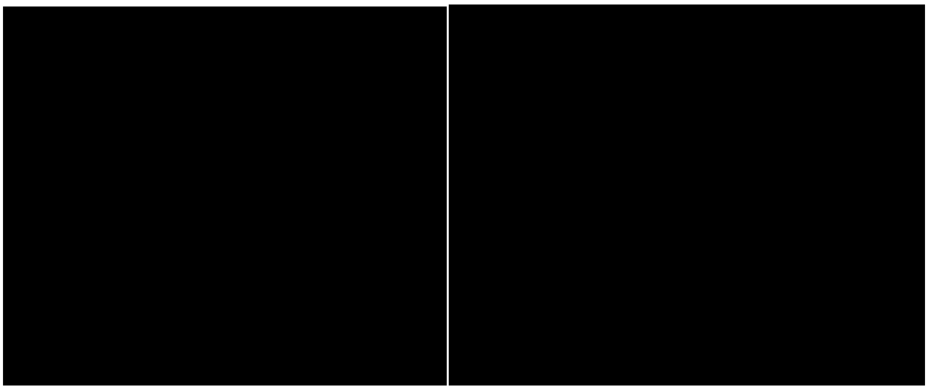

Fig. 6.16 Extracts from Fig. 6.2 (VBS2) and Fig 6.3 (NetLogo) respectively

It is clear that a higher interest was in NetLogo's food aid camera 3 compared with VBS2 camera 3 despite the lower level in fidelity. It is not obvious why this is the case, given that for other cameras the difference in fidelity did not play such a big part, but one suggestion could be greater the clarity in perspective for the NetLogo vantage point compared to the VBS2 vantage point and lighting.

The data was analysed using a three-way repeated-measure ANOVA within-subjects. This would identify if there are any differences between VBS2 and NetLogo, the three contextual scenarios and the four vantage point cameras (camera 1 to 4).

When comparing the difference between VBS and NetLogo it is suggested that there is a significant main effect between them [F (1, 9) = 12.002, p < 0.007]. This result is similar to the results found in section 6.4.1.1 again this could suggest that the observer had a tendency to dwell outside of the total area of interest, this seem to happen when observing in NetLogo.

Analysing between the three contextual scenarios showed that there is a significant main effect between them [F (2, 9) = 21.419, p < 0.000]. This measurement doesn't factor in fidelity but combines the data of NetLogo and VBS2 three contextual scenarios and analysis only the three contextual scenarios. What this analysis suggests is that there is difference between the food aid, gate & sentry, and helicopter pad dwell times overall. This correlates with the Fig. 6.12 graph.

A further analysis of the three contextual scenarios Table 6.10 analyses the data into a Pairwise comparison which show the measurements between the three contextual scenarios. The number 1 (food aid), 2 (gate & sentry) and 3 (helicopter pad). The table demonstrates that there difference between the scenarios.

Table 6.10 Pairwise Comparison between Contextual Scenarios

Measure:MEASURE_1

(I) ContextualScenarios	(J) ContextualScenarios	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	14.812*	3.479	.006	4.607	25.017
	3	16.163*	2.960	.001	7.481	24.845
2	1	-14.812*	3.479	.006	-25.017	-4.607
	3	1.351	1.289	.966	-2.431	5.133
3	1	-16.163*	2.960	.001	-24.845	-7.481
	2	-1.351	1.289	.966	-5.133	2.431

Based on estimated marginal means

Measure:MEASURE_1

(I) ContextualScenarios	(J) ContextualScenarios	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	14.812*	3.479	.006	4.607	25.017
	3	16.163*	2.960	.001	7.481	24.845
2	1	-14.812*	3.479	.006	-25.017	-4.607
	3	1.351	1.289	.966	-2.431	5.133
3	1	-16.163*	2.960	.001	-24.845	-7.481
	2	-1.351	1.289	.966	-5.133	2.431

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni

When analysing the four vantage point cameras there is a significant main effect between the four vantage point cameras [F (3, 27) = 26.783, p < 0.000]. Similar to the three contextual scenarios, this data is the combine data of both VBS2 and NetLogo three scenarios vantage point camera total. The analyse is the same as the three contextual scenarios whereby a Pairwise comparison like (Table 6.10) shows were the significant difference is when comparing each vantage point camera with one another. (See Appendix A for results)

Thus far the analysed data concurs with the graphs in which there are differences between NetLogo and VBS2, the scenarios and the four cameras vantage points. A further analysis also shows that there is difference between NetLogo‘s three contextual scenarios and VBS2 three contextual scenarios. The results show that there is a significant main effect when comparing the interaction between VBS2, NetLogo and the three contextual scenarios [F (2, 18) = 8.078, p < 0.003]. Fig. 6.15 illustrates this by showing the differences in dwell time between the three contextual scenarios and the two fidelities. This means that observers dwell time differed depending on which fidelity they were using.

Analysing the data further by removing the contextual scenarios from the equation and only measuring the four vantage point cameras for each of the fidelity then there is still a significant main effect when comparing the interaction between VBS2, NetLogo and the cameras [F (3, 27) =

5.361, $p < 0.005$]. Mean that also observer dwelling in difference cameras depending on the fidelity.

The final analyses factors in all three measurements, fidelity (NetLogo, VBS), three contextual scenarios and the four vantage points per scenario to see if there is a significant main effect between them. The results return a significant main effect when comparing the interaction between VBS2, NetLogo, three contextual scenarios and the four vantage point cameras [F (6, 54) = 6.792, $p < 0.000$].

6.4.1.6 Conclusion on contextual scenarios

Table 6.11 below summarises the findings from contextual scenarios.

Table 6.11 Summary of Contextual Findings

Comparison Between NetLogo & VBS2 - Mean Dwell Times (Three Contextual Scenarios)		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo (Total Dwell Time)	[F (1, 9) = 12.002, p < 0.007]	
Fidelity - VBS2, NetLogo and Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) – Comparison between the 3	[F (2, 18) = 8.078, p < 0.003]	
Fidelity - VBS2, NetLogo and Four Vantage Point Cameras (n…1)	[F (3, 27) = 5.361, p < 0.005]	
Fidelity removed - Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad)		[F (2, 9) = 21.419, p < 0.000]
Total Dwell Time - Four Vantage Point Cameras (n…1)		[F (3, 27) = 26.783, p < 0.000]
Fidelity - VBS2, NetLogo, Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) and Four Vantage Point Cameras (1…4) per Scenario	[F (6, 54) = 6.792, p < 0.000]	
Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) – Comparison between the 3		[F (2, 18) = 21.422, p < 0.000]
Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) and the Four Vantage Point Cameras per Scenario		[F (6, 54) = 7.473, p < 0.000]
Total Dwell Time - Four Vantage Point Cameras		[F (3, 27) = 26.782, p < 0.000]

Comparing the total dwell times of NetLogo and VBS2 at the highest level resulted in a significant effect between the two levels of fidelity VBS2 and NetLogo, meaning that there is a difference in dwell time occurring between the two fidelities. This is relevant to thesis question 2 as it suggests fidelity does have an effect on how observers perceive.

There is also a significant effect between the three contextual scenarios: food aid distribution area, gate & sentry area and the helicopter landing zone. Although not directly linked to a thesis question, this is in line with what was gleaned from SMEs, i.e. that context is important, as it had an effect on the dwell time of the observers. Furthermore these times differed between the experts and naïve observers. This impacts the answer to thesis question 3 as it implies that there are differences between experts and naïve observers.

There is also a significant effect between the vantage points that the cameras are angled at in the scenarios. This is an interesting find for the design/development in a virtual environment as it suggests that different angles of visualisation have an effect on an observer‘s dwell time.

The interaction measurement between all three factors (fidelity, context and cameras) also showed a significant effect. This suggests that fidelity, contextual scenarios and vantage points all have an effect on the observers dwell time. This correlates with the significant effects found when the data analysed separately. These findings are unlike those for experiment I.

Cameras 1 and 2 for all three contextual scenarios generated the highest dwell times amongst the naïve and expert observers, suggesting that the right hand of the screen had a predominant effect on both the expert‘s and naïve observers. It could also suggest that what is contextual happening in these cameras created greatest interest to the observers no matter what the display position was (top, middle or bottom).

The following sections in this chapter will analyse the data further by breaking down the data into the different trial runs. Because this analysis concentrated on the whole data which combined all seven trials what is not clear is what effect did the intelligence reports given in trial 4, 5, 6, and 7 affect the observers (did/didn‘t). Using the same dwell times and analysing the data in the same three factors of fidelity, context and cameras. It is hope that differences will appear between the trials suggesting that information alters the observer dwell times. But also clarify if fidelity does also alter the way in which the observers dwelled.

6.4.1.7 Experiment trial analysis between VBS2 and NetLogo

The data analysed for the analysis by trial is the same data i.e. dwell times gathered from the 10 naïve observers. The first trial was a dry run and as such is ignored in the analysis. Trial 2 (VBS2) and trial 3 (NetLogo) allowed the observer to look around freely without the provision of any intelligence. Trial 4 (VBS2) and trial 5 (NetLogo) informed the observers of a suspicious person in the food aid area. Trial 6 (VBS2) and trial 7 (NetLogo) informed the observers of a suspicious vehicle in the area.

The Author presumed that the results from trial 2 & 3 would show an even distribution of dwell time per scenario and per camera as no intelligence was provided and the observer was allowed to look around freely. The Author also had an expectation that trial 4 & 5 would show an increase in dwell time for the food aid scenario as the observer looks for a suspicious person. The final trials, 6 & 7, could have mixed results as there are vehicles in the food aid scenario, a military convoy in the gate sentry scenario and the helicopter landing and taking off in the helicopter landing zone.

<u>Trial 2 & 3</u>

Fig 6.16 and Fig 6.17 below illustrate the mean of the dwell time for trial 2 and trial 3 respectively. The black circle and dash line shows the mean dwell time of the experts.

Fig 6.18 compares the dwell time data between VBS2 with NetLogo. Data for the naïve observers is shown in blue for VBS2 and red for NetLogo. Expert data is shown in black for VBS2 and yellow for NetLogo. The data for all the trial 2 & 3 graphs are in Table 6.10.

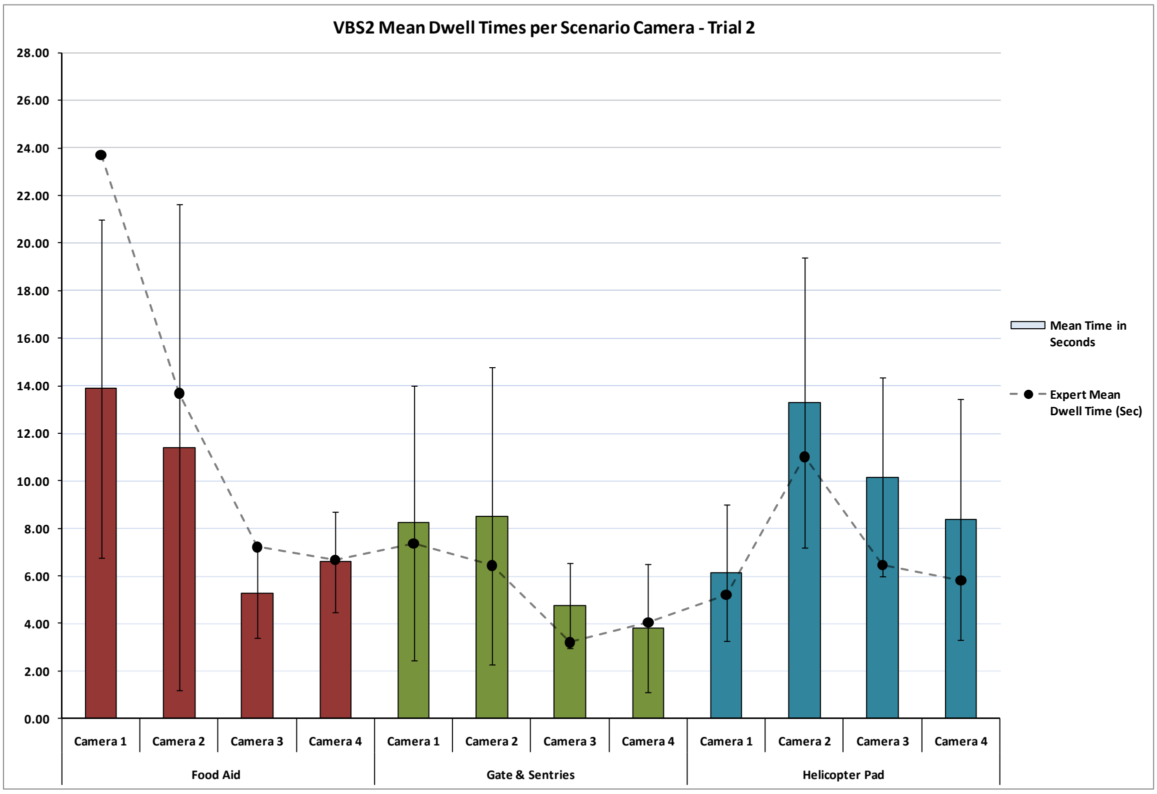

Fig. 6.17 VBS2 Trial 2

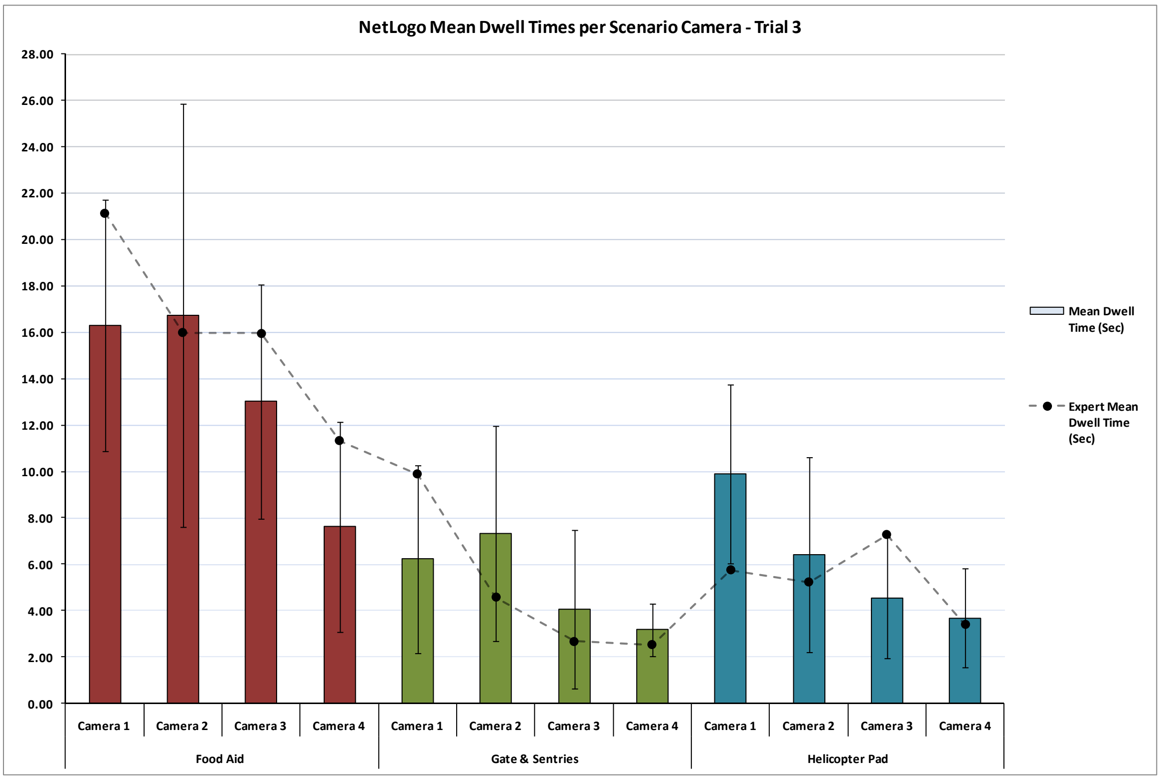

Fig. 6.18 NetLogo Trial 3

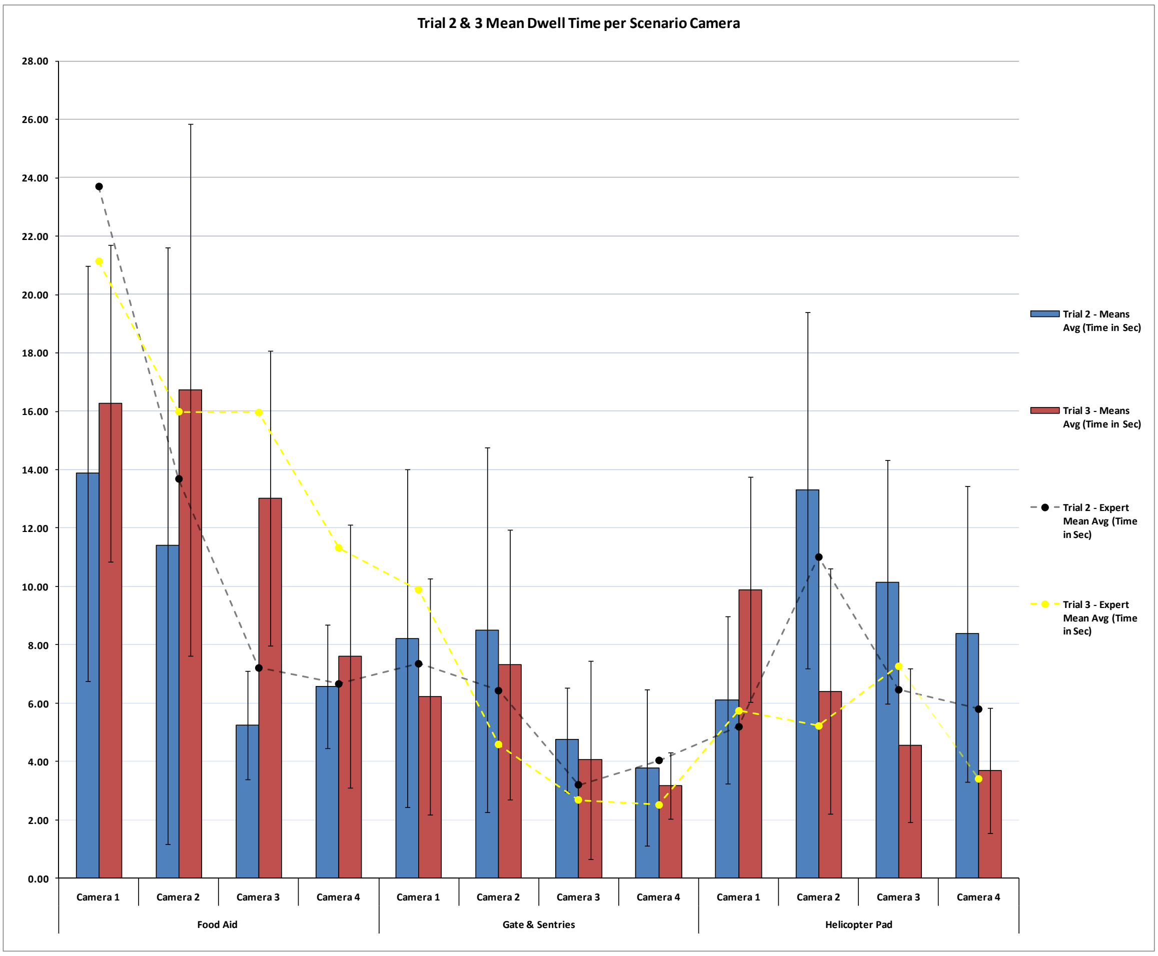

Fig. 6.19 Comparison of trial 2 & 3 VBS2 and NetLogo dwell times

Table 6.12 Mean Dwell Times for VBS2 and NetLogo (Trial 2 & 3)

		Trial 2 - Means Dwell Time (Sec)	Standard Deviation (Sec)	Trial 2 - Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)	Trial 3 - Means Dwell Time (Sec)	Standard Deviation (Sec)	Trial 3 - Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)
Food Aid	Camera 1	13.88	7.11	23.69	16.84	16.28	5.42	21.14	0.12
	Camera 2	11.40	10.22	13.68	3.74	16.74	9.12	15.99	1.60
	Camera 3	5.25	1.86	7.21	1.95	13.02	5.05	15.97	1.37
	Camera 4	6.58	2.12	6.66	3.80	7.61	4.52	11.32	3.99
Gate & Sentries	Camera 1	8.23	5.79	7.36	3.70	6.23	4.06	9.88	2.55
	Camera 2	8.51	6.25	6.44	0.17	7.32	4.63	4.58	0.14
	Camera 3	4.75	1.79	3.20	1.38	4.05	3.41	2.67	1.25
	Camera 4	3.79	2.69	4.04	1.79	3.18	1.13	2.52	0.06
Helicopter Pad	Camera 1	6.12	2.87	5.19	3.28	9.89	3.86	5.74	0.63
	Camera 2	13.30	6.11	11.01	7.40	6.41	4.19	5.23	1.22
	Camera 3	10.16	4.18	6.47	0.58	4.56	2.63	7.27	0.74
	Camera 4	8.37	5.08	5.80	1.26	3.68	2.15	3.40	0.90

As can be seen from the graphs above, the dwell times of the naïve observers display a similar pattern to that of the experts. Table 6.12 shows that camera 1 and 2 in the food aid scenario and gate & sentry scenario have the highest dwell times for both naïve and expert. This alters slightly in the helicopter pad scenario which for the naïve observer is cameras 2 and 3, whilst for the expert observers is camera 2 and 4, suggesting that for the experts camera 4 held greater interest.

Fig. 6.19 shows similarities between VBS2 and NetLogo, for the food aid and gate & sentry for both naïve and expert cameras 1 and 2 had the highest dwell times. This alters again when dwelling on the helicopter pad. The naïve highest dwell time where in camera 1 and 2 and for the experts cameras 1 and 3. This is interesting considering that the both groups of observers were freely allowed to dwell where they wanted.

A further difference between VBS2 and NetLogo is found in the total dwell times for trial 2 and trial 3. Trial 2 (VBS2) had the highest total dwell time compared with trial 3 (NetLogo) for the naïve observers. However, the opposite was true for the experts whose highest total dwell was for NetLogo. This could suggest that the expert spent more time in NetLogo to understand what was happening in the scenarios whereas the naïve preferred to dwell on the more graphically enhanced virtual world of VBS2.

The data was analysed further using three-way repeated-measure ANOVA within-subjects. Three-way ANOVA compares between the fidelities (NetLogo & VBS2), the three contextual scenarios and their four vantage point cameras.

The measurement should return a significant difference because no rules or information was given to the observers. The observers were allowed freedom to dwell at what interested them.

The analysis looked at whether there is a difference between the fidelities as is suggested in Fig. 6.19 which illustrates difference between dwell times of VBS2 and NetLogo. Furthermore it looked at whether a significant difference exists between the total interaction between VBS2, NetLogo, three contextual scenarios and the four vantage point cameras per scenario. If so, this would suggest that there is a difference in how the observer spends their dwell time between the fidelities, the contextual information and the vantage point of the camera.

When comparing the total data of VBS2 and NetLogo a non-significant main effect was found between the two occurs [$F.$ $(1, 9) < 0.372$, ns], which is unlike the analysis of the total dwell times (section 6.4.1.6), whereby there was a difference between VBS2 and NetLogo. Trial 2 & 3 there is not a difference which shows that the observer were less likely to wonder outside the total area of interest.

When analysing the total data (combined NetLogo and VBS) which remove the fidelity factor of from the measurement then the three contextual scenarios have a significant main effect on between the scenarios [F $(2, 18) = 9.115$, $p < 0.002$]. Like previous results, this suggests that there is difference in the observers dwell time over the total dwell time of the three contextual scenarios. The same effect happens when removing the factors of the three contextual scenarios, fidelity by combining the data of NetLogo and VBS2. By analysing that data of the four vantage point camera (overall) then between them there is a significant main effect [F (3, 27) = 13.866, p < 0.000]. This can also be further analyse in a Pairwise comparison (similar to Table 6.10) which shows the each comparison between the different scenario and cameras (see Appendix A)

This analyse is important, by measuring the interaction between the two factors, VBS, NetLogo, three scenarios. This will see if fidelity and context have a different effect on the observers dwelling. The result for this measurement returns a significant main effect when comparing the interaction between VBS2, NetLogo and the three scenarios, [F (3, 27) = 13.866, p < 0.012].

The final analysis measures the interaction of all three factors, fidelity, context and their four vantage point cameras. When comparing the interaction between VBS2, NetLogo, three contextual scenarios and the four vantage point cameras per scenario then this also suggest that there is a significant main effect between them [F (6, 54) = 3.410, p < 0.006].

Trial 2 & 3, findings show that the observers for both NetLogo and VBS2 spent equal time dwelling in the total area of interest which is interesting as this could suggest that as the trials went on so does the lost of interest in the area.

By analysing the data further into differences factors it is clear that there is differences between the three factors (fidelity, scenario and vantage point) which all had a significant effect on observers dwelling time.

What can be concluded for trial 2 & 3 is that the observers spent equal time dwelling in NetLogo and VBS2 on a whole. However what changes is where the observers dwell and how long for. This could be either due to the level of detail (fidelity), context of what is happening in scenario or the vantage point of the camera. The graphs clear illustrate that the observers find the highest interest in the food aid area before any information is provided.

Trial 4 & 5

Trial 4 & 5 introduce intelligence report informing the observers that a suspicious person may want to disrupt the food aid area. The Author expected that the highest dwell times should be on the four vantage point cameras of the food aid scenario. However what might also be found is that naïve observers get distracted by the other scenarios such as a helicopter landing or a vehicle passing. This could suggest that even thought important information is provided about a possible threat this could be easily over written by something more interesting happening. If such distraction occurs then it should be evident that the experts dwell times differ to those of the naïve observers. What is also important to bear in mind, is that trial 2 & 3 already showed a high dwell time in the food aid scenario despite no intelligence having been provided.

Fig 6.20, Fig 6.21 illustrates the mean dwell time for both naïve and expert observers. The expert data is depicted as black circles and dash line.

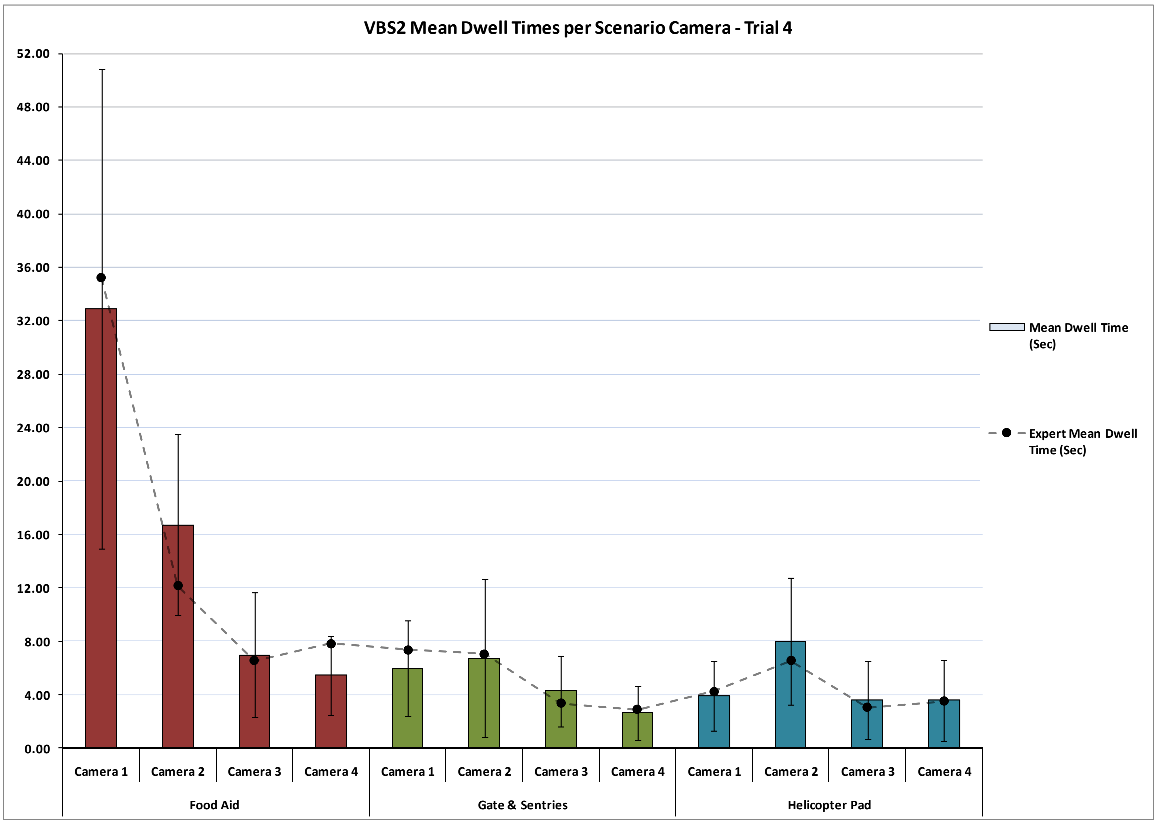

Fig. 6.20 VBS2 Trial 4

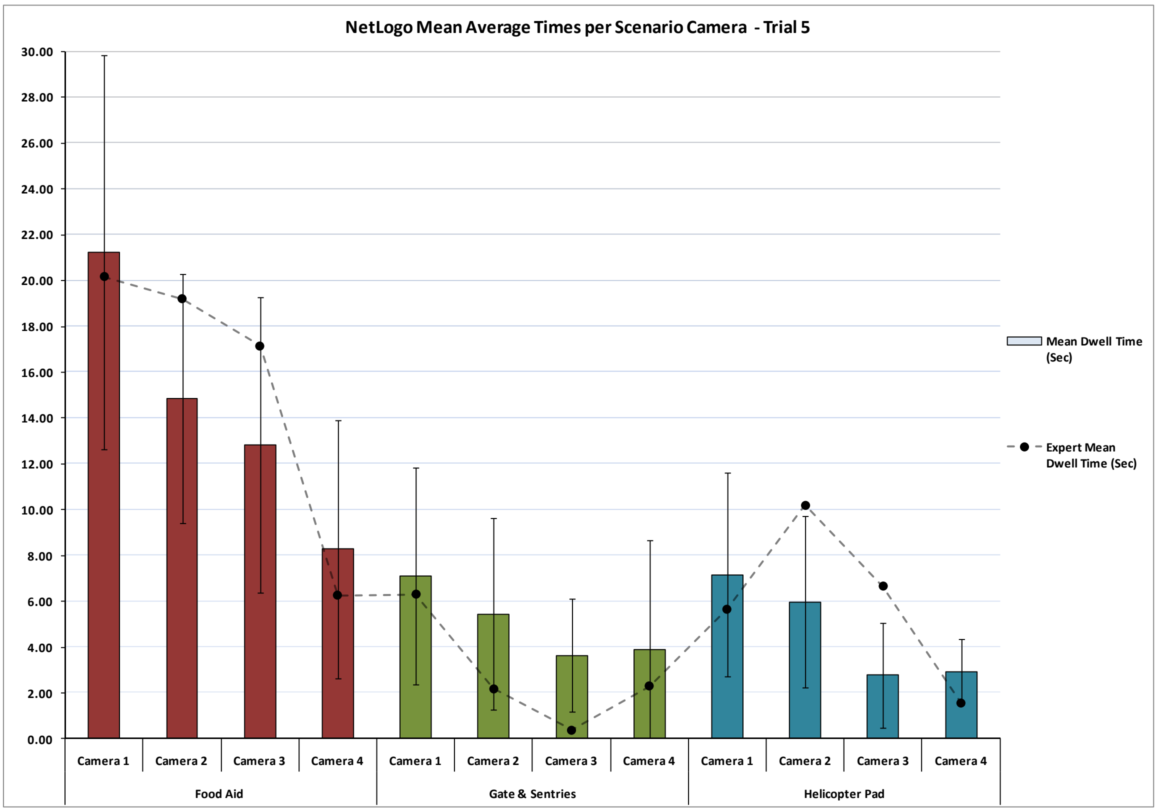

Fig. 6.21 NetLogo Trial 5

What can be seen in the graphs is a clear response to the provision of information. Both the VBS2 and NetLogo graphs show a predominantly high dwell time for food aid scenario.

When comparing the data from naïve and expert observers, Fig. 6.20 shows that the dwell times of both groups are similar for each camera. This suggests that there is not clear distinction between the observation dwell times of naïve and expert for VBS2. The data gathered for NetLogo, however, shows a different pattern in the dwell times between naïve and expert observers. It is clear that the experts spent more time on camera 2 & 3 of the food aid than the naïve. Also the experts had interest not only in the food aid are but also in the helicopter pads camera 2 & 3. This suggests that the experts spent more time in the area of interest (food aid) and that something in NetLogo raised awareness for the experts to dwell not only in the food aid area but also in the helicopter pad area. Fig. 6.21 shows a difference in naïve and expert dwelling patterns.

It had been suggested by an observer that he felt the two scenarios could be linked in the fact that he thought that helicopter pad area could be used to transport transporting food to the food aid area. There is therefore the possibility that the expert may have felt a similar link was plausible and therefore that this was enough to warrant observing other scenarios as well. Both naïve and expert observers spent more time in VBS2 which may suggest that the higher visual quality of VBS2 compare to NetLogo was more stimulating to the observers.

Table 6.13 Mean Dwell Times for VBS2 and NetLogo (Trial 4 & 5)

		Trial 4 - Means Dwell Time (Sec)	Standard Deviation (Sec)	Trial 4 - Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)	Trial 5 - Means Dwell Time (Sec)	Standard Deviation (Sec)	Trial 5 - Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)
Food Aid	Camera 1	32.92	17.96	35.19	13.34	21.24	8.60	20.17	7.63
	Camera 2	16.72	6.79	12.16	1.40	14.85	5.44	19.20	5.31
	Camera 3	6.98	4.67	6.53	3.10	12.82	6.46	17.15	1.39
	Camera 4	5.46	2.96	7.82	2.00	8.27	5.65	6.26	3.21
Gate & Sentries	Camera 1	5.98	3.58	7.37	2.71	7.10	4.75	6.30	2.77
	Camera 2	6.75	5.90	7.05	3.98	5.44	4.17	2.18	0.45
	Camera 3	4.28	2.64	3.35	2.86	3.63	2.47	0.39	0.21
	Camera 4	2.66	2.03	2.89	0.93	3.89	4.78	2.30	0.43
Helicopter Pad	Camera 1	3.91	2.62	4.23	4.86	7.16	4.44	5.67	2.12
	Camera 2	7.99	4.77	6.55	5.22	5.98	3.76	10.20	8.09
	Camera 3	3.59	2.93	3.06	2.75	2.78	2.28	6.65	2.97
	Camera 4	3.58	3.05	3.51	3.72	2.91	1.43	1.56	0.54

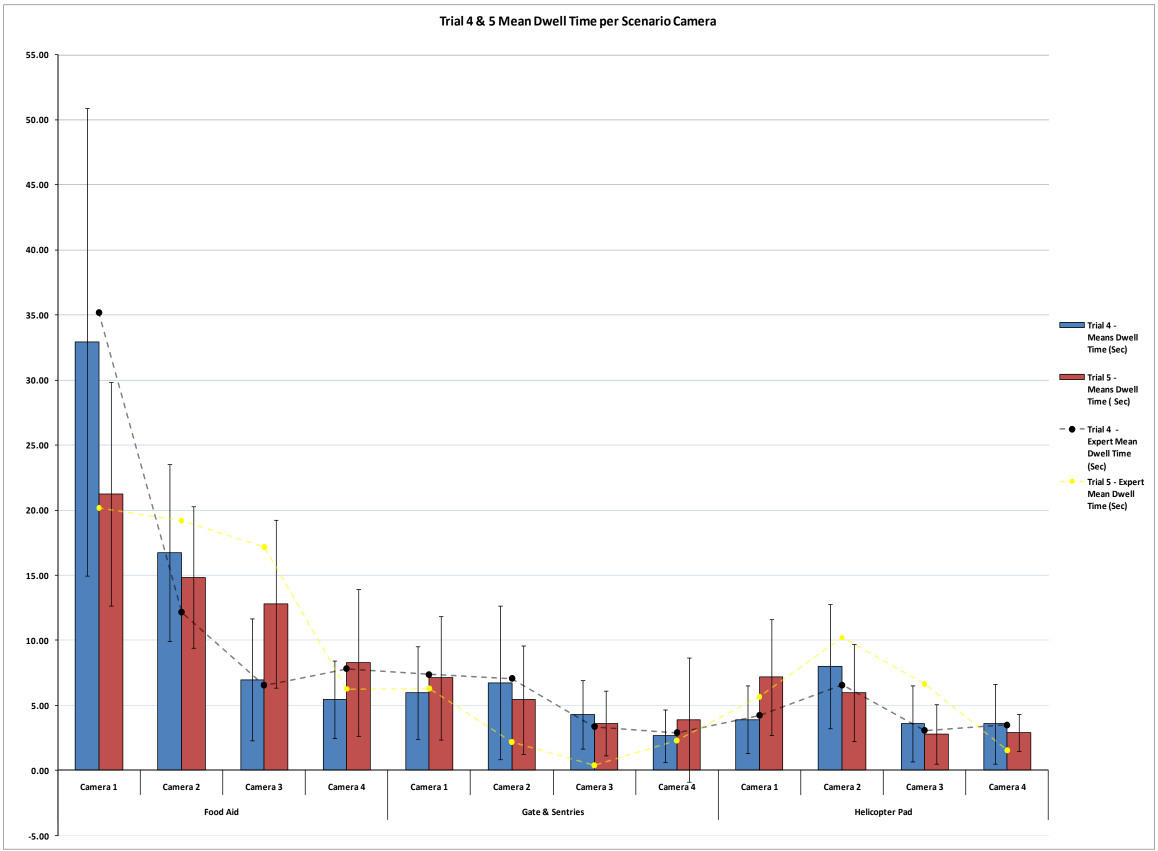

Fig. 6.22 Comparison of trial 4 & 5 VBS2 and NetLogo dwell times

The data was analysed using the three-way repeated-measure ANOVA within-subjects. The measurement is the same as trial 2 & 3 and will look at the interaction between the fidelity (VBS2 & NetLogo), three contextual scenarios and the four vantage point cameras.

A comparison between VBS2 and NetLogo for trial 4 & 5 resulted in a non-significant mean effect [*F. (*1, 9) < 0.073, ns]. Which suggests, like in trial 2 & 3, the observers did not lose interest due to the fidelity and that an equal concentration was within both fidelities' total interest areas.

There is also a significant main effect for the three contextual scenarios [*F* (2, 18) = 23.653, p < 0.000] and four the cameras [*F* (3, 27) = 26.142, p < 0.002]. These results are similar to the results for trial 2 & 3 and produce the same Pairwise comparison analysis graph (similar to Fig. 6.9) which compares the difference between three contextual scenarios and the same for the four vantage point cameras.

However, interestingly and unlike trial 2 and 3, there is a non-significant main effect [*F. (*2, 18) < 0.742, ns] between the interaction of VBS2, NetLogo and the three contextual scenarios. What is interesting about this find is the suggestion that there is no difference between the fidelities and the three scenarios. This means that the observers took in the information from the intelligence report and concentrated in the food aid area looking for a potential suspect and that it made no difference whether it was in VBS2 or NetLogo. This is important because it shows that fidelity did not have an impact on the observers, and that the task of searching for a potential suspect overrides the quality of the virtual world.

This finding is also supported by the measurement of the interaction between VBS2, NetLogo and the vantage point cameras, which is also resulted in a non-significant mean effect [*F. (*3, 27) < 0.056, ns].

What may cause confusion with the above finding is the fact that the measurement of the interaction between VBS2, NetLogo, three contextual scenarios and the four vantage point cameras this did result in a significant main effect [*F* (3, 27) = 26.142, p < 0.002]. However Fig 6.21 clear shows that there are differences between the scenarios and the cameras and this would be reflected when removing the fidelity aspect and combining the data for the three contextual scenarios. This would be the same for the four vantage point cameras as Fig.6.21 also illustrates that there are significant differences between the cameras. Combing the data in Excel to compare with - ANOVA analyses would suggest that there is differences between camera 1, camera 2, camera 3 and camera 4.

Trial 4 & 5 clearly suggests that the intelligence report given to the observers did led - the observers to spending the majority of their dwell time on the food aid scenario and that fidelity had no impact on this. This is seen to be a significant find. However the risk attached to this find, is that by neglecting the other two contextual scenarios then potential something could have happened in them and this could have lead to the observer missing this. This may/may not be a potential explanation for why the experts also looked at the helicopter pad scenario, as shown in Fig. 6.21. However it was not replicated in VBS2.

Trial 6 & 7

In trial 6 & 7 the observer had been informed, again verbally via an intelligence report, that there may be a suspicious vehicle in the area. This will increase the difficulty of the search as there are vehicles in two of the three contextual scenarios (gate & sentry and helicopter pad). As such, the Author expected more mixed results compared to trail 4 & 5.

The Author thought it likely that the gate and sentry scenario would have the highest dwell time, because there is a slow moving convoy of four vehicles. Albeit these vehicles are not suspicious it should raise awareness in the area. Another potential scenario that could experience an increase in dwell time was the helicopter pad scenario, as the helicopter is continually landing and taking off.

There are some differences between VBS2 and NetLogo camera vantage points for the three contextual scenarios this is because of the limitation of the NetLogo's software. The one difference is the food aid scenario of NetLogo the vehicles can be seen but not in VBS2. Also camera 1 was the only camera in the gate & sentry scenario that could display the convoy of vehicles moving, again due to limitations when angling the vantage points cameras (the vehicle would have driven sideways). This will create a difference between NetLogo and VBS2 however if in NetLogo gate & sentry scenario camera 1 has the highest dwell time then this would clearly suggest that the observer notice the vehicles and that the intelligence report was listened to. It could also suggest that even though the fidelity is lower in NetLogo that observes still associated the lower quality vehicles as vehicles.

The results from trial 6 & 7 are illustrated in the Fig. 6.23 and 6.24 below.

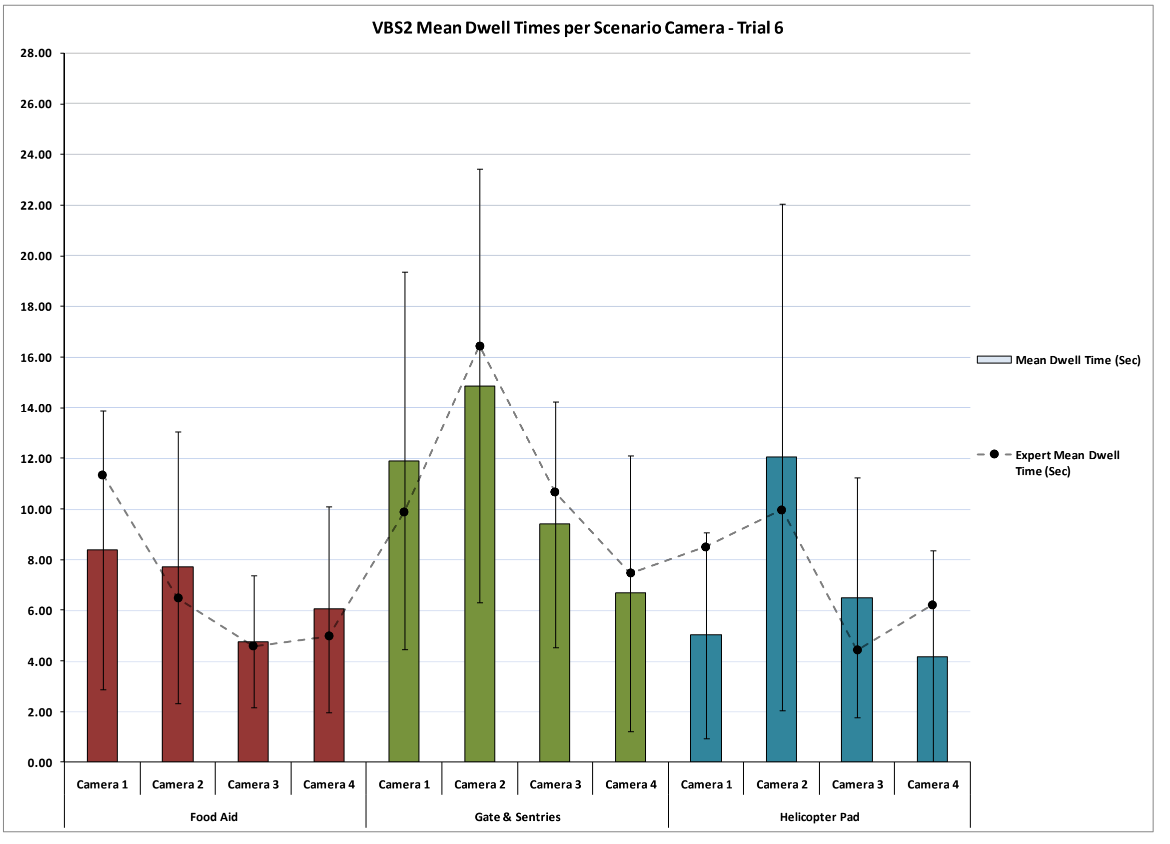

Fig. 6.23 VBS2 Trial 6

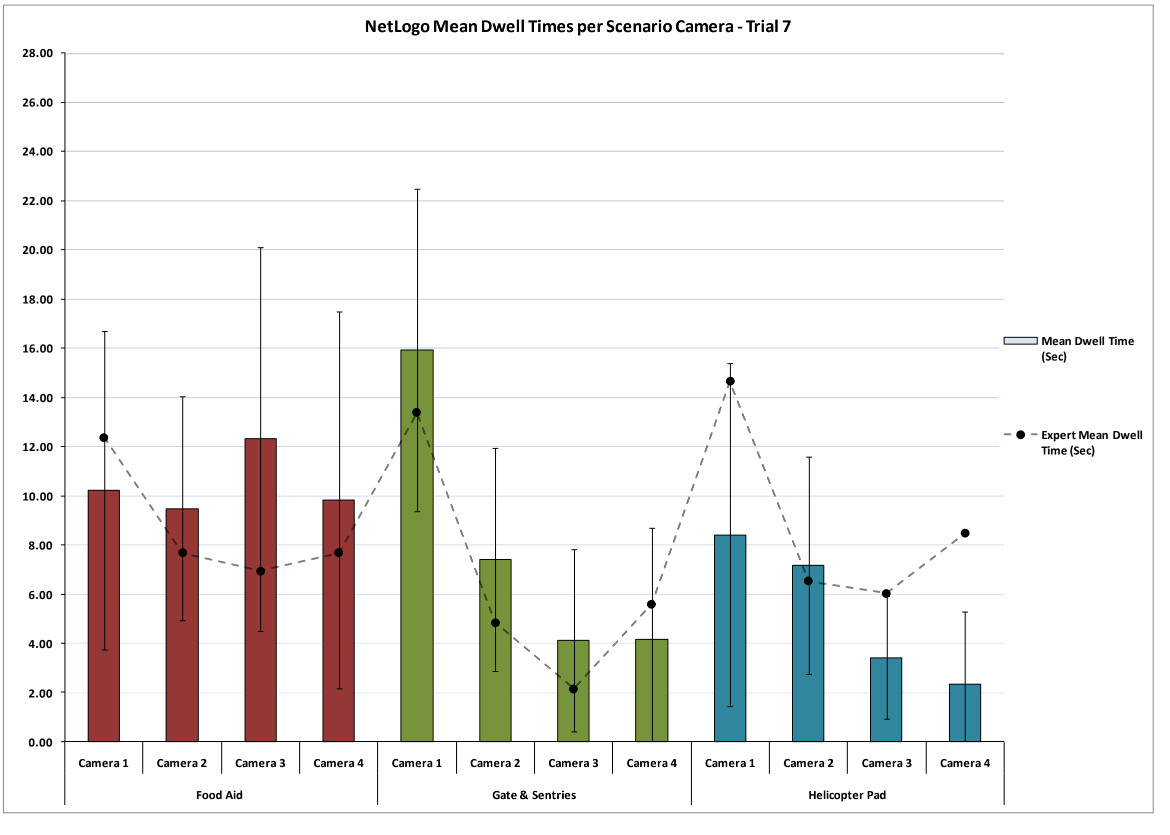

Fig. 6.24 NetLogo Trial 7

In Fig 6.23 it can be seen that the dwell time had shifted from the food aid scenario to the gate and sentry and the helicopter pad scenario, compared to trials 2, 3, 4 and 5. This suggests that the observers did notice the vehicles in these scenarios as there is a raised awareness (dwell time) towards these interest areas. The helicopter pad also had an increase in dwell time especially around camera 2. Again a similar pattern to dwell times between the naïve and experts observer appears.

NetLogo dwell times also seem to shift (see Fig. 6.24) and show that the gate & sentry scenario camera 1 (the camera that has the vehicle in it) has the highest dwell time, which is important considering the difference in fidelity been VBS2 and NetLogo. The vehicle in NetLogo is composed of simple moving blocks. Also important is the dwell time‘s increase in the food aid scenario compared with VBS2 which also shows that the observers notice the food aid trucks. There is a reduction in the helicopter pad scenario dwell time compared with VBS2. This could suggest that visual it is not as interesting as the VBS2 helicopter pad.

What is interesting is that Fig. 6.24 shows that there is a clear difference between the naïve and experts dwell time patterns. Camera1 in the helicopter pad scenario had a large increase of dwell time from the expert observers compared to the naïve observers. This suggests that the helicopter scenario attracted greater interest (or suspicion).

Table 6.14 Mean Dwell Times for VBS2 and NetLogo (Trial 6 & 7)

		Trial 6 - Means Dwell Time (Sec)	Standard Deviation (Sec)	Trial 6 - Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)	Trial 7 - Means Dwell Time (Sec)	Standard Deviation (Sec)	Trial 7 - Expert Mean Dwell Time (Sec)	Standard Deviation (Sec)
Food Aid	Camera 1	8.38	5.52	11.34	5.59	10.23	6.47	12.37	11.29
	Camera 2	7.70	5.36	6.47	1.16	9.48	4.55	7.67	6.75
	Camera 3	4.77	2.60	4.58	4.52	12.30	7.82	6.95	7.27
	Camera 4	6.04	4.06	4.98	1.80	9.84	7.66	7.67	7.93
Gate & Sentries	Camera 1	11.91	7.46	9.89	6.41	15.94	6.55	13.38	3.01
	Camera 2	14.87	8.55	16.43	8.52	7.41	4.53	4.84	0.08
	Camera 3	9.39	4.85	10.66	2.55	4.11	3.71	2.16	0.09
	Camera 4	6.68	5.44	7.47	0.54	4.17	4.53	5.60	3.34
Helicopter Pad	Camera 1	5.01	4.07	8.51	1.97	8.41	6.97	14.64	11.43
	Camera 2	12.06	9.99	9.98	7.56	7.17	4.43	6.53	6.01
	Camera 3	6.50	4.73	4.43	3.16	3.42	2.48	6.04	4.67
	Camera 4	4.17	4.17	6.22	5.80	2.33	2.97	8.48	9.29

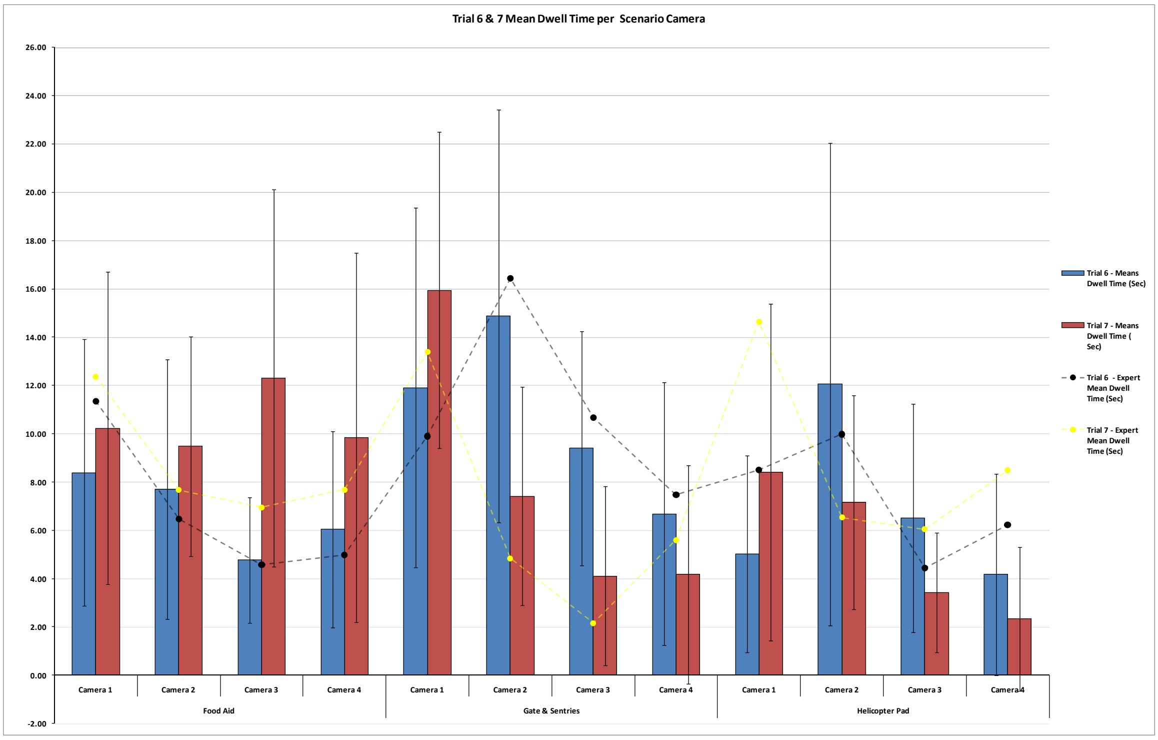

Fig. 6.25 Comparison of trial 6 & 7 VBS2 and NetLogo dwell times

Fig 6.25 illustrates the difference in dwell times between VBS2 and NetLogo. The graph clearly shows a change in the dwell times between the experts VBS2 (black line) and NetLogo (yellow line) dwell times, the same for the naïve VBS2 (blue) and NetLogo (red). However this is expected to happen because of the limitation that NetLogo had with some of the camera vantage points.

The intelligence given to the observers does seem to have an impact when looking at the dwell time data. Table 6.14 shows that observers were dwelling in the area that had vehicles. The alteration in dwell time patterns between VBS2 and NetLogo also shows that even when a decrease in fidelity occurs, the observers still found were the vehicles were.

VBS2 has the highest total dwell time for trial 6 & 7 for both naïve and expert observers. This could suggest that the higher visual appearance of the VBS2 virtual environments kept the focus of the observers for longer compared to NetLogo. This has been a consistent finding between all of the trials.

The data will be analysed further to find if there are any significant differences. It is suggested that, unlike trial 4 & 5, there will be an increased in the significant main effect in the area of interaction between VBS2 and NetLogo. This is because there was a limitation in the ability to replicate closely the VBS2 scenarios into NetLogo. However it is important to get a significant difference between NetLogo and VBS2 which then correlates to the graph data.

The data was analysed using a three-way repeated-measure ANOVA within-subjects. This compares VBS2, NetLogo, the three contextual scenarios and the four vantage point cameras for each of the scenarios.

At the highest factor level when comparing the total dwell times of VBS2 and NetLogo there is a non-significant main effect between the two fidelities [$F. (1, 9) < 0.205$, ns] (similar to trial 2, 3, 4 and 5). What is interest about this find is that by separating the trials for analysing each trial produced a non-significant effect when comparing the two fidelities which suggest that observers dwelled within the total interest area (suggest no differences).

When removing the fidelity factor out of the measurement and only analysing the three contextual scenarios then there is a non-significant main effect between the three contextual scenarios [$F. (2, 18) < 1.955$, ns]. This is the first trials which this has happened, which suggest that when combining total dwell times of NetLogo and VBS2 and comparing this to the three contextual scenarios there are no significant differences between the scenarios, meaning that the dwell times in each were similar. However when analysing the overall dwell times of the four vantage point cameras then there is a significant main effect [$F (3, 27) = 7.704$ $p < 0.001$]. By removing the factor of fidelity and context then there is a difference in the four camera vantage points.

This is interesting because this suggest when looking at fidelity there is no difference between the two, remove the fidelity and measure the context and this also returns no difference between the three scenarios. Remove context and we find a difference between cameras. Mean that the camera vantage point was more important to the observer.

When analysing the interaction VBS2, NetLogo and the three contextual scenarios, then there is a significant main effect between VBS2, NetLogo and the three contextual scenarios [$F (2, 18) = 4.235$, $p < 0.031$]. This suggests that there is a difference between the fidelities and the three contextual scenarios, concurring with Fig. 6.23 and 6.24, which shows that the observers were dwelling on different scenarios (because the vehicles were in different areas).

The other important significant main effect is the interaction between VBS2, NetLogo, three contextual scenarios and the four vantage point cameras [$F (6, 54) = 57.405$, $p < 0.037$]. This suggests what is illustrated in Fig 6.25 that there is difference between the 3 measured factors which is what was suggested needed to happen if there is to be a difference between were the observers dwell in on VBS2 and NetLogo.

The above results between the three factors are important because they are similar to trial 4 & 5. When the data is analysed on a whole by measuring the different factors, fidelity and contexts on their own a non-significant differences between them occurred. Suggest that neither the fidelity nor contextual scenario differed from each other. But the significant difference was in the combined data of the four vantage point cameras suggesting that the observer saw difference in the camera vantage points only,

However when the data is separated into the different factors and analysed between the interaction of NetLogo's, three contextual scenarios and their four vantage point cameras and VBS2, three contextual scenarios and their four vantage point cameras. Then like Fig. 6.25 clearly illustrates these differences.

What could be suggest is that although when all the factors (fidelity, context and cameras) are measured this did create a significant effect, however when broken down separately this did not happen which then leads to suggesting the difference is in the four vantage point cameras in each scenario. This then further suggests that the observer were aware of the intelligence report leading them to dwell in the cameras that had vehicles in them, not the scenario or fidelity.

This does suggest over the two trials that did have intelligence the observers were aware and did follow the information that the report provided. It also shows that overall neither NetLogo nor VBS2 had any impact on the observers, the task overrode the fidelity.

6.4.1.8 Conclusion on trial analysis between VBS2 and NetLogo

Table 6.15 Summary of Trial Findings

Trial 2 (VBS2) & 3 (NetLogo)		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo (Total Dwell Time)	Non-Significant	
Fidelity removed - Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) – Comparison between the 3		[F (2, 18) = 9.115, $p < 0.002$]
Total Dwell Time - Four Vantage Point Cameras (1…4)		[F (3, 27) = 13.866, $p < 0.000$]
Fidelity - VBS2, NetLogo and Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad)	[F (3, 27) = 13.866, $p < 0.012$]	
Fidelity - VBS2, NetLogo and Four Vantage Point Cameras (1…4)	Non-Significant	

Fidelity - VBS2, NetLogo, Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) and Four Vantage Point Cameras (n…1) per Scenario	[F (6, 54) = 3.410, p < 0.006]	
Trial 4 (VBS2) & 5 (NetLogo)		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo (Total Dwell Time)	Non-Significant	
Fidelity removed - Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) – Comparison between the 3		[F (2, 18) = 23.653, p < 0.000]
Total Dwell Time - Four Vantage Point Cameras (n…1)		[F (3, 27) = 26.142, p < 0.002]
Fidelity - VBS2, NetLogo and Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad)	Non-Significant	
Fidelity - VBS2, NetLogo and Four Vantage Point Cameras (1…4)	Non-Significant	
Fidelity - VBS2, NetLogo, Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) and Four Vantage Point Cameras (1…4) per Scenario	[F (3, 27) = 26.142, p < 0.002]	
Trial 6 (VBS2) & 7 (NetLogo)		
Measurement	**Comparison**	**Combined Data**
Fidelity - VBS2 compared with NetLogo (Total Dwell Time)	Non-Significant	
Fidelity removed - Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) – Comparison between the 3		Non-Significant
Total Dwell Time - Four Vantage Point Cameras (n…1)		[F (3, 27) = 7.704 p < 0.001]

Fidelity - VBS2, NetLogo and Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad)	[*F* (2, 18) = 4.235, p < 0.031]	
Fidelity - VBS2, NetLogo and Four Vantage Point Cameras (1…4)	[*F* (3, 27) = 4.235, p < 0.001]	
Fidelity - VBS2, NetLogo, Three Contextual Scenario (Food Aid, Gate & Sentry, Helicopter Pad) and Four Vantage Point Cameras (1…4) per Scenario	[*F* (6, 54) = 57.405, p < 0.037].	

Table 6.15 summarises the findings by trial groups. As in trial 2 & 3 the observers were allowed to look around freely it serves as the baseline. The baseline will help to answer the thesis question 1a, 2, 2b, and 3.

Comparing the total data between VBS2 and NetLogo returned a non-significant effect in all six trails. This suggests that there is no difference in dwelling times when the data is measured in trials. This finding then suggests that the difference between the two fidelities is not great.

When the removing fidelity from the measurement trials 2, 3, 4 and 5, all showed significant effects between the three contextual scenarios, suggesting that dwell times between the scenarios were still sufficiently different. However trial 6 & 7 returned a non-significant effect, suggesting that between the three contextual scenarios there was no significant difference. Intelligence was given in trial 6 & 7 which required the observers to look for suspicious vehicles which were in all three contextual scenarios. This could help explain the reason for the non-significant effect, as it could suggest observers spent an even level of time searching for the vehicles in all three contextual scenarios.

The six trials all returned a significant effect when comparing the total dwell time of all four vantage point cameras, which means that angles of the camera have an effect on the observers dwell time.

Trial 2, 3, 6 and 7 show a significant effect between the fidelities and the contextual scenarios which means the observers dwell differently on the contextual scenarios depending on the fidelity (answering question 2 – does the level of fidelity in the virtual environment have an effect on how observers perceive). However trial 4 & 5 were the intelligence was given regarding suspicion around the food aid, returned of a non-significant effect. This implies that the intelligence had effect which potentially was stronger than that between fidelities and contextual scenarios.

Trial 6 & 7 returned a different measurement compared to trial 4 & 5, however this does not contradict the suggestion that intelligence has a significant effect because unlike the food aid contextual scenario which was in one area, the vehicle intelligence encouraged observers to look at all three contextual scenarios. This is also shown in Fig. 6.23 and 6.24 which show the observers dwelling in the contextual scenarios that had vehicles. Therefore a difference in dwell time between fidelities and contextual scenarios should occur because vehicles were placed in different scenario for VBS2 and NetLogo which clearly shows that briefings have an effect on the observers (Thesis question 1a).

The measurement between fidelity and the four vantage point cameras (overall value of the three contextual scenarios vantage point cameras) for trial 2, 3, 4 and 5 returned a non-significant effect between VBS2 and NetLogo‘s four cameras. Although this analysis is an overall measurement, and removes the measurement of context, it has some significance as it shows that observers had similar dwell times in VBS2 and NetLogo regardless of whether intelligence was given or not. However trial 6 & 7 did return a significant effect implying that that intelligence did affect the observers somehow. The non-significant effect resulting from trail 4 & 5 could be because the intelligence informed the observer to concentrate in one contextual area (food aid) creating no differences in NetLogo or VBS2. This could be suggesting that ―physical” fidelity is being overshadowed by ‗psychological” fidelity through the briefings and that this has an effect on the observers. It was expected that trial 6 & 7 return a significant effect because the vehicles are in different cameras in VBS2 and NetLogo.

What is of interest which isn‘t so clearly definable is that trial 2 & 3 allowed observers freedom to dwell wherever, however this creates a non-effect in such that the was no difference in NetLogo four cameras to VBS2 four cameras.

The final measurement analysed the all three measurements fidelity, three contextual scenarios and their four vantage point cameras. All six runs produce a significant effect. This correlates with the graphs in section 6.4.1.7 which illustrates that there are difference between NetLogo, VBS2, three contextual scenario and the four vantage point cameras. This suggests that fidelity, combined with contextual scenarios and different angles of the cameras generates differences in the observers dwell times. The provision of a briefing also has an effect on the observers as it can still be seen in the graphs that the observer dwell times leaned towards what the intelligence was suggesting. Similar patterns appear when comparing the experts with naïve observers dwell time throughout the contextual scenarios and cameras, the differences in the dwell times suggest that experts are spending more time searching for suspicious activities.

6.4.2 Text analysis between VBS2 and NetLogo comments

What will also be analysed in this section is the text feedback from each observer, each observer was ask after each trial what they thought was happening. This data has been mined looking for strength in words. This builds a web connection of word that the observers kept using to describe each trial.

During the trials, the verbal feedback from observers was recorded. This data is analysed in this section of the thesis. The same question was asked to each observer every time a trial was conducted. The question was: —What do you think is happening?” This question would hopefully help gather information on what observes felt was happening in each trial. It could also help to see if there is a difference between the naïve and the expert responses.

The Author recorded the feedback provided by each observer and used firstly Excel and then the text mining software SPSS Text Analytical for Surveys 4 to analyse the comments made for consistencies and relationships between key works in each trial.

SPSS Text Analytical for Surveys 4 analyses key words or recurring words over the complete run of comments. It does this by analysing the words and then looking at the three words that follow to see if a correlation can be made. An example of this would be —helicopter space pad” would be analysed as —helicopter pad”. The SPSS Text Analytical for Surveys 4 software then highlights this as a keyword.

The Author built categories for these keywords. For example, if an observer suggested that a vehicle looked military and another observer suggested the vehicle was a military jeep then only one category was produces know as ‗Military Vehicles‘. The software produces a diagram of the categories of key words showing the strengths of the categories as a circle. The larger the circle, the more respondents used a word in that category. The circles are joined by lines which create a web diagram. The lines represent correlations between the categories, which are the number of shared respondents between those two categories. An example would be —landing zone” circle with a line drawn to —Helicopter” circle. Fig. 6.26 shows an example of the graphical representation of expert responses to trial 2 helicopter landing zone scenario.

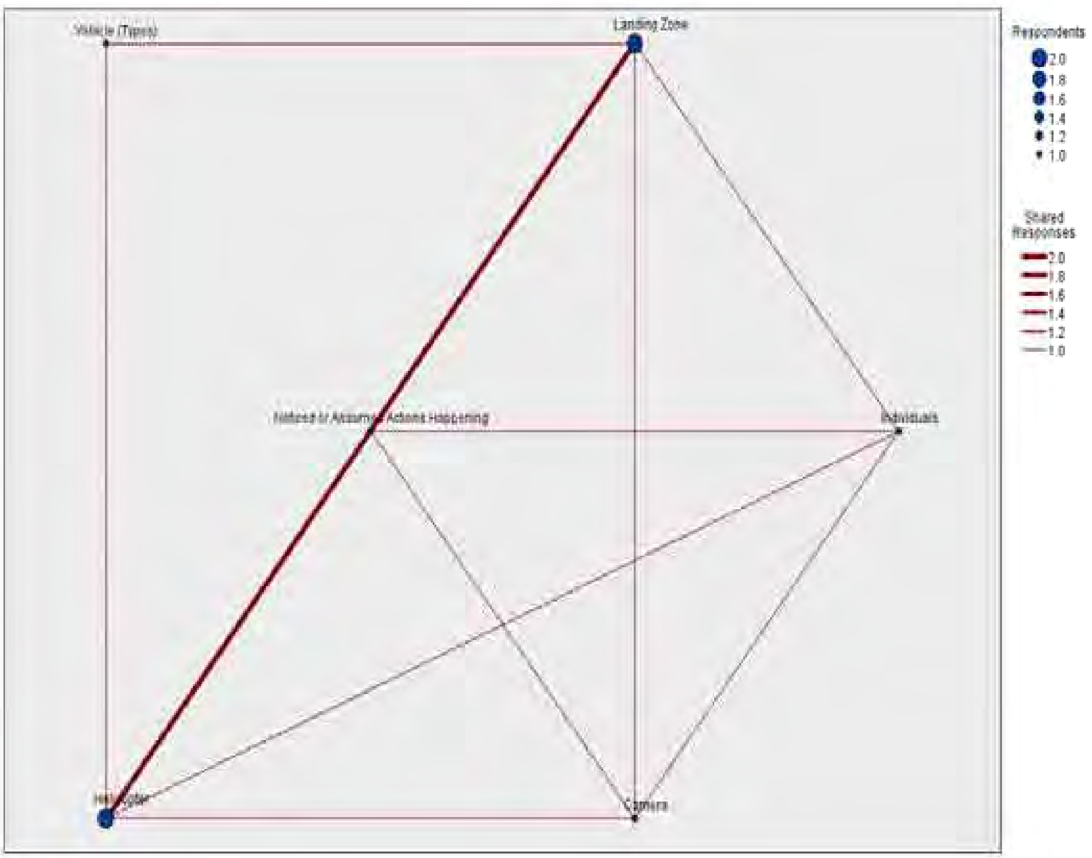

Fig. 6.26 Trial 2 expert responses to the helicopter landing zone scenario

All the web diagrams produced from the analysis of verbal responses can be found in Appendix B. In order to simplify the comparative analysis between VBS2/NetLogo and naïve/expert observers, the categories with the highest strengths have been summarised in a table, for each trial pair.

6.4.1.10 Trial 2 and 3 Text Analysis

Table 6.16 shows the three highest categories of keywords from the both the naïve and expert observers for trial 2 and 3.

Trial 2 and 3 allowed the observer to freely observer and it has brought with it responses of what the observers perceiving their world to be. This analysis is at its highest level, the tables are to comparing the categories of keywords between the expert and naïve observer. It also helps highlight difference between VBS2 and NetLogo. Table 6.16 shows how the observers saw the three contextual scenarios before intelligence was given to them.

Table 6.16 Text Analysis for Trial 2 & 3

Trial 2 & 3		Food Aid		Gate & Sentry		Helicopter Pad	
		Naïve	Expert	Naïve	Expert	Naïve	Expert
VBS2	1	People	People Moving/Walking	Area	Area	Helicopter Landing, Taking off or Loading	Helicopter
	2	Beach	Noticed/Assumed Actions Happening	Guards	Noticed/Assumed Actions Happening	Helicopter	Landing Zone
	3	Noticed/Assumed Actions Happening	People	Vehicle (Types)	Checkpoint	Buildings	Individuals
NetLogo	1	Types of People	Noticed/Assumed Actions Happening	Convoy/Moving Vehicles	Noticed/Assumed Actions Happening	Helicopter	Helicopter
	2	Noticed/Assumed Actions Happening	Supplies	Area	Area	Noticed/Assumed Actions Happening	People
	3	Groups	Types of People	Vehicle (Types)	People	People	Camera

Interestingly there are similarities between the naïve responses and the expert responses. An example of this is the VBS2 food aid which sees both naïve and expert observers referring to people in the virtual environment. However the expert responses had a tendency focus more on what was happening in the virtual environment. For example, in the food aid scenario, the number one keyword category was the referral to what the people where doing, which was that people where moving/walking, not the simple fact that they were there.

The keyword category ―Noticed or Assumed Actions Happening" is a group of keywords that the observers used to describe what he/she perceived was happening. For example an expert observer noticed that people were jumping over the barrier of the food aid, whilst the naïve observer noticed that people were walking along a beach front. Two completely different opinions were recorded possibly because both observers saw the virtual world differently from a contextual point of view. One observer was from Saudi Arabia. This was the one who felt the food aid scenario was a beach front and not as a food aid. The expert saw things differently and noticed it was a food aid area and that someone had been jumping over the barrier. However this is only one example and some naïve observers did notice greater detail than the experts.

Another interesting find is how close the naïve observer‘s categories are to the experts. In the lower fidelity NetLogo it suggests that the observers become more descriptive. For example in the food aid scenario, both naïve and expert observers suggest that not all people were the same by referring to the colour they were (the category for this was" Types of People").

6.4.1.11 Trial 4 and 5 Text Analysis

The text analyse of trial 4 & 5 should hopefully illustrate a difference between trial 2 & 3 and how experts differs from naïve observers when intelligence is given. Trial 4 and 5 provided the observer with information that there was a suspicious person in the food aid area. Table 6.17 illustrates that the observer‘s began to talk more about the food aid area and what was happening. The keyword categories ―Types of People" and ―Groups" seem to increase. For example, the category —Types of People" has keywords in them that observer suggest that colour was different types of people. These colours include blue people, white people etc.

This shows the naïve observers began to suggest that colour represented different type's people and that different people where in groups. Throughout both fidelities the reference towards people, supplies and the movement suggests that the fidelity played little or no part in assisting the task of identifying "if" there was a suspicious person in the food aid scenario.

It is unclear from Table 6.17 if the expert observers described anything differently compared to naïve observers. The only main difference is that experts saw groups.

Table 6.17 Text Analysis for Trial 4 & 5

Trial 4 & 5		Food Aid		Gate & Sentry		Helicopter Pad	
		Naïve	Expert	Naïve	Expert	Naïve	Expert
VBS2	1	People Moving/Walking	Groups	Area	Noticed/Assumed Actions Happening	Helicopter	Helicopter
	2	Supplies	Types of People	Vehicle (Types)	Convoy/Moving Vehicles	People	People
	3	Types of People	Noticed/Assumed Actions Happening	Guards	Area	Noticed/Assumed Actions Happening	Noticed/Assumed Actions Happening
NetLogo	1	People	People Moving/Walking	Area	Convoy/Moving Vehicles	Helicopter	Helicopter Landing, Taking off, or Loading
	2	Types of People	Groups	Guards	Area	Helicopter Landing, Taking off, or Loading	Helicopter
	3	Supplies	Noticed/Assumed Actions Happening	Vehicle (Types)	People	People	People

6.4.1.12 Trial 6 and 7 Text Analysis

This analysis of trial 6 & 7 should illustrate a shift in what observers described towards vehicle types, convoys and movement compared to in trial 4 & 5. This is because the intelligence suggested that there was a suspicious vehicle in the area. Trial 4 & 5 illustrated an increased in descriptiveness about the food aid area.

Table 6.18 Text Analysis for Trial 6 & 7

Trial 6 & 7		Food Aid		Gate & Sentry		Helicopter Pad	
		Naïve	Expert	Naïve	Expert	Naïve	Expert
VBS2	1	People	Vehicle (Types)	Vehicle (Types)	Convoy/Moving Vehicles	Helicopter	Landing Zone
	2	People Moving/Walking	Noticed/Assumed Actions Happening	Area	Vehicle (Types)	Helicopter Landing, Taking off, or Loading	People
	3	Vehicle (Types)	Supplies	Individuals	Area	People	Noticed/Assumed Actions Happening
NetLogo	1	People	Groups	Vehicle (Types)	People	Helicopter	Landing Zone
	2	People Moving/Walking	Vehicle (Types)	Convoy/Moving Vehicles	Vehicle (Types)	Helicopter Landing, Taking off, or Loading	People
	3	Vehicle (Types)	Noticed/Assumed Actions Happening	Noticed/Assumed Actions Happening	Area	People	Noticed/Assumed Actions Happening

Table 6.18 shows an increase in the observers responding to the intelligence report by describing more about vehicles in the three contextual scenarios. The category —Vehicle (Types) comprises different keywords that observers used to describe the vehicles they saw in the cameras (i.e. truck, car, jeep etc.). Other categories used describe if the vehicles where moving or if they were in convoy (which the expert used frequently).

Tables 6.16 and 6.17 show how the descriptive change in the observer‘s response in the food aid area now includes the vehicle category. This clearly suggests that the observer had listened to the information provided. The gate & sentry area also shows an increased response to the types of vehicles.

Table 6.17 does not clear distinguish between the expert and naïve observer responses. The only difference is in the number of responses. However, the experts noticed that there was a convoy moving in VBS2 gate & sentry scenario whilst the naïve observers did not. This could be because one expert was a serving British Army Office and could identify the actual types of vehicles in the virtual environment. This is something that the naïve observers could not do; however they did notice that the vehicles were not the same.

6.4.1.13 Trial text analysis conclusion

The text analysis provides an insight into to what observers saw in the virtual world. The Author noted that even thought the world was virtual, the observers still associated that there was people, types of people and groups even when the fidelity was lowered.

The only suggestion that can be gathered from how the experts analysed the world virtual to the naïve was by the ranking of the categories. The detail which is in the Excel spreadsheet shows that the British Army officer could identify the types of vehicles in the virtual world.

It can also be concluded that intelligence does alter how the observer responded to the three contextual scenarios and this is evidenced in the verbal responses to the trials. By analysing the verbal text from trial 2 & 3 which was a free to look anywhere it allows a baseline to be drawn on and a comparison can be made when looking at the rest of the trials. It is clear from this comparison that there is a clear alteration when intelligence is provided to how the perception (response) of the observer changes.

When looking to answer thesis question 2b (Will differences be sufficient to demonstrate that fidelity does make a difference) just using the text analysis, then it appears that fidelity doesn‘t alter the observer‘s awareness of what is happening. This is evidenced in the tables which indicate that the keyword categories are similar in both VBS2 and NetLogo.

The text analysis also demonstrates that both levels of observers had awareness of what was happening in the virtual environment meaning that thesis question 1 can be answered that overall the observers did understanding what was happening in each contextual scenario and that it did not drastically alter between the fidelities.

Thesis 1a asks whether briefings matter to an observer's ability to detect what is happening in a virtual environment. When intelligence (briefing) was provide in trial 4 & 5 there was increase awareness of people, group's types of people, movement and noticing something etc. The same happened when intelligence was provided for trial 6 & 7 evidenced by an increased awareness of vehicles. This suggests that intelligence does affect the observer's ability to detect what is happening in the virtual environment.

What was less clear was whether there is a definitive difference between the experts and naïve observer's text, almost suggesting there isn't any. What might affect the responses is if the task required a detail description of objects, then an expert might have produced a more detailed account of the objects in the virtual environment. There was some small evidence of this as the experts could identify vehicle types.

CHAPTER 7

CONCLUSION AND RECOMMENDATIONS

7.1 Conclusions

The thesis was originally sponsored by HS-C to conduct research towards understanding insurgents ‗perception of intent' using different levels of fidelity. This did not come about and the research developed more into a semi-military/civilian and focused on understanding how observers perceive whilst using different levels of physical fidelity (achieved through games technologies). Two experiments were designed and conducted.

In the first experiment differences in physical fidelity were achieved by using NetLogo and VBS. The contextual scenario which was needed and seem to reflect on SMEs experience was hypothesised around a food aid distribution, in which people need to get food and followed the rules in doing so. To add the perception of ‗intent' in the contextual scenario, an individual breaking the rule and stealing the food from other individuals was added. Two different briefs, verbal and written were given in this experiment.

Experiment II developed from experiment I, removing the briefing types with only a verbal brief given. The static slide show converted into a dynamic motion video to create a richer virtual environment based on SME and organisational visits (CCTV) and adding more contextual scenarios to the environment which included the food aid distribution, gated area around a military base and a helicopter landing zone.

The thesis question set out to answer question about fidelity, provision of information (from briefing) and naïve and expert differences.

7.1.1 Detecting what is happening in a virtual environment

1. Can observers detect what is happening in a virtual environment?

In Experiment I, it appears that observers have difficulty detecting what is happening correctly in a virtual environment when no direct briefing on what to look for was provided. In this experiment the observers were provided with verbal and written information about the scenario and what was happening, the information did not clearly state what the people in the virtual environment where doing, however hints were provided in the briefings. Therefore it can be said that the observers were using their own understand understanding to determine what was happening in the virtual world. The information provided in the five options was varied and they had to make a choice based on the briefing and what was in front of them. Whilst there were some options that had higher mean confidence than others, overall the mean confidences were low in all options (never above 50%) and there was high standard deviation indicating a low level of consensus.

All observers in Experiment I were novices and the experiments were timed. There is therefore the possibility that they needed greater clarification to gain an understanding of what was happening in the virtual environment and/or that they needed more time to get familiar with the scenarios. Parallels could be drawn with (G. Klein, 1999) *"Recognition-Primed Decision Model"* (see Fig. 2.1) as the observers had no experience in the situation and were provided with limited relevant cues and expectancies. As such it was difficult for them to identify anomalies, as evidenced by the broad spread in standard deviation. They also had limited opportunity to clarify the data given. The outcome may have been different if the information in the options provided was more precise.

What is also not known is what impact the time restrictions had on the observer‘s ability to learn. Repeating the experiment with same students again or not having a time restriction could have returned different results.

Informal comments relayed back from both experiments, and hence not included in the results, also provided interesting insights. Some observers assumed that the different colour of individuals in the virtual world meant that they belonged to different groups. Other observers thought that the colours represented emotional states. This implies that observers showed some indication of forming groups or creating mental rules without any intelligence having been provided.

Experiment I also considered whether or not observers were able to notice someone breaking the rules, which they did not. The data showed low mean confidence in the options and the spread of deviation was high between the options.

Experiment II had different effect on the observers suggest through the text analysis that there was an understanding of what was happening in the virtual environment if before any information was given. This could be because Experiment I did not have dynamic motion but used static slide shows to create movement potentially overriding the impact of confidence on what was happening. The removal of static slide shows and moving to dynamic motion in experiment II may have helped responders understand the virtual environment better. This draws some parallel to (Michotte, 1963) notion that motion is key not fidelity and that low fidelity can still produce results (perception of causality).

These last two observations (the instinct to create groups and the inability of observers to detect someone breaking the rules in Experiment I) both reinforce the importance of crowds and understanding their behaviour and context. Without the existence of a crowd it becomes harder to understand and detect when someone is doing something ‗out of the norm'. A good understanding of the ‗norm is important. This ties in closely to both fidelity and expertise conclusions below.

1a. Does the provision of a briefing help?

The provision of a briefing appears to have an impact on the observer's. Unlike Experiment I, observers in Experiment II were provided with direct intelligence (i.e. the briefings) and there was a visible shift in the focus of their attention, evidenced by the change in dwell times in trials 4 and 5. The shift was not visible for trials 6 and 7, but this was expected because the intelligence could have been applicable to all three scenarios.

What cannot be said is whether the provision of a briefing is necessarily a good thing. During the visit to REME, the Author was informed that the trainers had included a person on the rooftop in the training but had not informed the trainees. The trainees did not notice the presence of this individual presumably because they were focused on the task at hand. Hence their briefing redirected their attention but also eroded their awareness of other factors that were happening, even if these were potentially relevant to the task at hand. The challenge is to get the level of briefing right.

1b. Does the type of briefing matter?

During SME interviews the Author was informed that both verbal and written briefings were provided in military and policing contexts. Therefore, Experiment I included the provision of either verbal or written information during the experiment. The results overall were inconclusive.

Whilst verbal and written briefings were not re-examined in Experiment II, this experiment offered a more direct briefing which did have an effect.

7.1.2 Effect of fidelity on observer perception

2. Does the level of fidelity in the virtual environment have an effect on how observers perceive?

2a. Do responses become more accurate?

2b. Will differences be sufficient to demonstrate that fidelity does make a difference?

The results were mixed.

At the highest level, based on experiment II, physical fidelity does have an effect on observers. However what causes this is not clear because further analysis cannot conclude on whether it is the context, the fidelity or the vantage points of the cameras that are causing the significant effect to occur. Furthermore a significant effect was not shown in the high level analysis of experiment I.

Concentrating on the impact of physical fidelity in experiment II, VBS2 showed an overall higher dwell time compared to NetLogo. This implies that the higher fidelity was able to attract the attention of the observer for longer times. However, this is not always a good thing in terms of being able to achieve the task at hand.

When direct information to search in an area was provided in experiment II, the focus on the task overrode the physical fidelities of VBS2 and NetLogo and showed a non-significant difference for trials 4 & 5. This suggests that psychological fidelity (task) became important than the physical fidelity. Trial 6 & 7 showed a significant difference which could suggest that the assumption of trial 4 & 5 are incorrect. However trial 4 & 5 concentrated on one are (food aid), whereas the intelligence in trial 6 & 7 referred to vehicles which were in different areas. Therefore it is believed the information proved in all four trials drew the observer away from the physical fidelity and leaned more to conduct what the task required the observers to do (psychological fidelity).

Looking at experiment II trial 2 & 3, observers gave similar responses to what was happening in the virtual environment regardless of the level of fidelity (VBS2 vs. NetLogo) and without any information having been provided. In other words, responses did not become more accurate.

This would be similar to the Author‘s own experience in Radars systems. In his experience it isn‘t the level of fidelity that is need to make informed judgments. Why don‘t Radar systems have high fidelity, with ships and airplanes? This is because it is not necessary to clutter the screen with realistic information to achieve the desired task. Radar systems remove distractions as the operator needs to make fast decisions on potential threats and navigational safety. High fidelity could lead to distraction (unnecessary noise). The increased dwell times in VBS2 could be due to similar increased noise. Whilst observers may engage more with the environment this does not mean they do well at the task.

The visit REME School reflects this to a certain extent in that the change from the level of fidelity from OHP training to VBS2 created a greater interaction in the classroom participants. However it is not known whether the greater interaction also had a corresponding effect on how trainees perceived or how well they performed the task.

Furthermore experiment II showed clear preference to certain cameras and therefore suggested that the vantage points in these cameras provided a more interesting or clearer perspective.

7.1.3 Differences between naïve and expert observers

3. Is there a difference between a naïve observer and one who has had experiences in observation or intelligence gathering skills (expert)?

Even prior to running the experiments, SME interviews and literature reviewed implied that experience and a good understanding of context played an important part in an observer's ability to understand what is happening and identifying deviations from the norm. Furthermore, it was apparent that passing on cultural awareness and personal experience can be challenging. However, identifying the ‗norm' is important because without understanding ‗everyday things' you cannot identify deviations. The Author's attempts to define cues in a military context were not successful, but this would have been ambitions as it became apparent that the best providers of intelligence are the individuals on the ground, who are regularly on patrol.

Obtaining SMEs for the experiments was not easy and therefore only experiment II looked at this thesis question. Even in this case the number of experts was not the same as the number of novices which meant that meaningful statistical analyses could not be conducted.

However, the dwell time patterns of SMEs and novices were compared and in most cases were similar. Whilst the patterns were similar, the actual dwell times differed in some cases. This could be a potential indication that SMEs are better trained and therefore better at distinguishing between background noise and relevant information. Experts were also marginally more descriptive in the text analysis. Unfortunately these hypotheses cannot be proven due to the limitations described above.

7.2 Recommendations

The Author feels that there were a number of interesting insights from this thesis. Due to the breadth of the topic and the large volume of variables needed to be considered, it would be recommended that further analyses are conducted. Further investigation of existing experimenting experiment results could consider the following:

- Realistic (or improved) definition of context and cues to be tested;
- Secure a larger number of experts to perform additional experiments addressing thesis question 3;
- Re-running experiment I adding dynamic motion;
- Re-running experiment I without time limits or increasing the time;
- Re-running experiment I limiting the option selection to 1 option;
- Re-running experiment without

- Introducing someone ‗breaking the rules‘ in experiment II;
- Reducing/varying the level of information provided in experiment II;
- Varying brief types for experiment II; and
- Comparing VBS2 and Videos for experiment II.

LIST OF REFERENCES

Alexander, A.L., Brunye, T., Sidman, J., et al. (2005) From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-In and their Effects on Transfer in PC-based Simulation and Games. [Accessed 2010].

AmericasArmy (2011) **America's Army** [online]. http://www.americasarmy.com/ [Accessed 3 August 2008]

Berk, R. (ed.) (1972a) **The Controversy Surrounding Analyses of Collective Violence: Some Methodological Notes***,* Chicago: Aldine-Atherton.

Berk, R. (ed.) (1972b) **The Emergence of Muted Violence in Crowd Behaviour: A Case Study of an Almost Race Riot***,* Chicago: Aldine-Atherton.

Berk, R. (1974a) **Collective Behaviour**. Dubuque, IA: Brown. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Berk, R. (1974b) A gaming Approach to Crowd Behaviour. **American Sociological Review***,* 39: (3): 355-373.

Berlonghi, A.E. (1995) Understanding and Planning for Different Spectator Crowds. **Safety Science***,* 18: (4).

Bessel, R. and Emsley, a.C. (2000) **Patterns of Provocation Police and Public Disorder**. New York USA: Berghahn Books.

Boersema, T. and Zwaga, H.G. (1990) "Searching for Routing Signs in Public Buildings". In Brogan, D. (Ed.) **Visual Search**. London, Taylor and Francis.

Braun, A., Musse, S.R., de Oliveira, L.P.L., et al. (2003) "Modeling individual behaviors in crowd simulation". In Musse, S.R. (Ed.) **Computer Animation and Social Agents, 2003. 16th International Conference on***.*

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009a) "Understand Crowd Behaviour: Guidance and Lessons Identified". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Simulation Tools". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009c) "Understand Crowd Behaviour: Supporting Documentation". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009d) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown.

Choi, H. and Scholl, B.J. (2004) Effects of Grouping and Attention on the Perception of Causality. **Perception & Psychophysics***,* 66: (6): 926-942.

Cottrell, N., B. (ed.) (1972) **Social Facilitation***,* New York: Holt, Rinehart and Winston. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Courtney, M. (2011) "Public Eyes Get Smart". **Engineering and Technology (E&T)**. United kingdom, IET Publishing & Information Services.

CRYMOD (2011) **CRYMOD DISCUSS. CREATE. SHARE.** [online]. http://www.crymod.com/portal.php 2011 Crytek GmBH [Accessed 23 June 2008]

CRYTEK (2011) **CRYTEK** [online]. http://crytek.com/ CRYTEK [Accessed 23 June 2008]

Diener, E. (ed.) (1989) **Deindividuation: The Absence of Self-Awareness and Self-Regulation in Group Members**, Hillsdale, NJ: Lawrence Erlbaum.

Donald, I. and Canter, D. (1992) Intentionality and Fatality During the King's Cross Underground Fire. **European Journal of Social Psychology**, 22: (3): 203-218.

Drury, J. and Reicher, S.D. (1999) The Intergroup Dynamics of Collective Empowerment: Substantiating the Social Identity Model of Crowd Behavior. **Group Processes & Intergroup Relations**, 2: (4): 381-402.

Festinger, L., Pepitone, A. and Necombe, T. (1952) Some Consequences of De-individuation in a Group. **The Journal of Abnormal and Social Psychology**, 47: (2): 382-389.

Freud, S. (1951) "Group Psychology and the Analysis of the Ego". In Jones, E. (Ed.) 6th ed. New York, Liveright Publishing, Corporation.

Gaskins, R.C., Boone, C.M., Verna, T.M., et al. (2004) Modeling & Simulation Group Conference on Modeling and Simulation to Address NATO's New and Exisiting Military Requirement. (1 - 6).

Gaskins, R.C., Kruszewski, P.A., McKenzie, F.D., et al. (2008) Integrating Crowd-Behaviour Modeling into Military Simulation using Game Technology. **Simulation & Gaming**, 39: (1).

Hoggett, J. and Clifford, S. (2010) Crowd Psychology, Public Order Police Training and the Policing of Football Crowds. **Policing: An International Journal of Police Strategies and Management**, 33: (2).

Houghton, R.J. and Baber, C. (2009) **Detecting Deviants within Flocks**. Maastricht: Shaker.

Johnson-Laird, P.N. (1983) **Mental Models Towards a Cognitive Science of Languages, Inference, and Consciousness**. Cambridge: Cambridge University Press.

Klein, G. (1999) **Source of Power How People Make Decision**. United States of America: Massachusetts Institute of Technology.

Klein, O., Spears, R. and Reicher, S. (2007) Social Identity Performance: Extending the Strategic Side of SIDE. **Personality and Social Psychology Review**, 11: (1): 28-45.

Koenig, S.C., Liebhold, G.M.Y. and Gramopadhye, A.K. (1998) "Training for Systematic Search Using a Job Aid". **Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting**. Santa Monica, The Human Factors and ergonomic Society.

Latane, B., Williams, K. and Harkins, S. (1979) Many Hands make Light the Work: The Causes and Consequences of Social Loafing. **Journal of Personality and Social Psychology**, 37: (6): 822-832.

Le Bon, G. (1896) "The Crowd: A Study of the Popular Mind". New York, The Macmillan Co.

Michotte, A. (1963) **The Perception of Causality**. London: Methuen & Co Ltd.

Moray, N. (ed.) (1976) **Attention, Control and Sampling Behaviour**, New York: Plenum Press.

Pelachaud, C., Martin, J.-C., Andre, E., et al. (2007) "Intelligent Virtual Agents". **7th International Conference, IVA 2007**. Paris, France, Springer-Verlag.

Pelechano, N., O'Brien, K., Silverman, B., et al. (2005) Crowd simulation incorporating agent psychological models, roles and communication. **First International Workshop on Crowd Simulation**.

Prentice-Dunn, S. and Rogers, R., W. (eds.) (1989) **Deindividuation and the Self-Regulation of Behavior**, Hillsdale, NJ: Lawrence Erlbaum.

Reicher, S.D. (1984a) Social Influence in the Crowd - Attitudinal and Behavioural Effects of De-Individuation in Conditions of High and Low Group Salience. **British Journal of Social Psychology.**, 23: 341-350 (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. (1984b) The St Pauls' Riot: An Explanation of the Limits of Crowd Action in Terms of a Social Model. **European Journal of Social Psychology**, 14: (1): 1-21.

Reicher, S.D. (ed.) (1987) **Crowd behaviour as Social Action**, Oxford: Basil-Blackwell.

Reicher, S.D. (1996a) 'The Battle of Westminster': Developing the Social Identity Model of Crowd Behaviour in Order to Explain the Initiation and Development of Collective Conflict. **European Journal of Social Psychology**, 26: (1): 115-134.

Reicher, S.D. (1996b) 'The Crowd' Century: Reconciling Practical Success with Theoretical Failure. **British Journal of Social Psychology.**, (35): 535-553. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. (ed.) (1996c) **Social Identity and Social Change: Rethinking the Context of Social Psychology**, London: Butterworth. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. (1997) Collective Psychology and the Psychology of the Self. **British Psychological Society Social Section Newsletter**, 36: 3-15. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. and levine, M. (1994a) Deindividuation, Power Relations between Groups and the Expression of Social Identity. **British Journal of Social Psychology.**, 33: 145-163 (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. and Levine, M. (1994b) On the Consequences of Deindividuation Manipulation for the Strategic Communication of Self: Identifiability and the Presentation of Social Identity. **European Journal of Social Psychology**, 24: (4): 511-524.

Reicher, S.D., Spears, R. and Postmes, T. (1995) A Social Identity Model of Deindividuation Phenomena. **European Review of Social Psychology**, 6: 161-198. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reynolds, C.W. (1987) "Flocks, herds and schools: A distributed behavioural model". **Computer Animation: CG87. Proceedings of the Conference held at Computer Graphics 87**. London, UK, Online Publications.

Scholl, B.J. and Nakayama, K. (2002) Causal Capture: Contextual Effects on the Perception of Collision Events. **Psychological Science***,* 13: (6): 493-498.

Scholl, B.J. and Tremoulet, P.D. (2000) Perceptual Causality and Animacy. **Trends in Cognitive Sciences***,* 4: (8): 8.

Smyth, M.M., Collins, A.F., Morris, P.E., et al. (1995) **Cognition In Action 2nd Edition**. 2nd.Hove: Lawrence Erlbaum Associates Ltd.

Soraia Raupp, M., Branislav, U., Amaury, A., et al. (2005) "Groups and Crowd Simulation". **ACM SIGGRAPH 2005 Courses**. Los Angeles, California, ACM.

Stone, J.R. (2008) "Human Factors Guidelines for Interactive 3D and Game-Based Training Systems Design". In MoD (Ed.), Human Factors Integration Defence Technology Centre.

Stott, C. and Drury, J. (eds.) (1999) **The Intergroup Dynamics of Empowerment: A Social Identity Model***,* London: Macmillan. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Tajfel, H. (ed.) (1978) **Social Categorization, Social Identity, and Social Comparison***,* London: Academic Press (for) European Association of Experimental Social Psychology.

Tajfel, H., Billig, M., Bundy, R.P., et al. (1971) Social Categorization and Intergroup Behaviour. **European Journal of Social Psychology***,* 1: (2): 149-177.

Tajfel, H. and Turner, J.C. (eds.) (1979) **An Integrative Theory of Intergroup Conflict***,* Monterey, CA: Brooks/Cole. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Thalmann, D. and Musse, S. (2007) **Crowd Simulation**. Springer London.

Turner, J.C. (ed.) (1982) **Towards a Cognitive Redefinition of the Social Group***,* Cambridge: Cambridge University Press.

Turner, J.C. (ed.) (1985) **Social Categorization and the Self-Concept: A Social Cognitive Theory of Group Behaviour***,* Greenwich, CT: JAI Press. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Turner, J.C., Hogg, M.A., Oakes, P.J., et al. (1987) **Rediscovering the Social Group: A Self-Categorization Theory**. Oxford: Basil Blackwell.

Turner, R.H. (ed.) (1964) **Collective Behavior***,* Chicago: Randy McNally.

Turner, R.H. and Killian, L.M. (1957) **Collective behavior**. Englewood, NJ: Prentice-Hall.

UDK (2010) **UNREAL Development Kit** [online]. http://www.udk.com/ Epic Games Inc. [Accessed 3 August 2009]

Wheeler, S. (2005a) It Pays to Be Popular: a Study of Civilian Assistance and Guerilla Warfare. **Journal of Artificial Societies and Social Simulation***,* 8: (4): 9.

Wheeler, S. (2005b) "On the Suitability of NetLogo for the Modelling of Civilian Assistance and Guerrilla Warfare". In Laboratory, D.S.S. (Ed.) Edinburgh, South Australia.

Zanjoc, R.B. (1965) "Social Facilitation". **Science**. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Zimbardo, P.G. (ed.) (1970) **The Human Choice: Individuation, Reason, and Order Versus Deindividuation, Impulse and Chaos**, Lincoln: University of Nebraska Press. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

BIBLIOGRAPHY

Alexander, A.L., Brunye, T., Sidman, J., et al. (2005) From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-In and their Effects on Transfer in PC-based Simulation and Games. [Accessed 2010].

AmericasArmy (2011) **America's Army** [online]. http://www.americasarmy.com/ [Accessed 3 August 2008]

Andersen, J., vanDoorn, A.J. and Ridde, H.d. (2007) When and How Well do People See the Onset of Gestures? **Gesture**, 7: (3): 305-342.

Andrade, E.L., Blunsden, S. and Fisher, R.B. (2006) "Modelling Crowd Scenes for Event Detection". **Pattern Recognition, 2006. ICPR 2006. 18th International Conference on** Hong Kong, IEEE Xplore.

Andrew, C. (2007) "A Game AI Production Shell Framework: Generating AI Opponents for Geomorphic-Isometric Strategy Games via Modeling of Expert Player Intuition". **Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts**. Perth, Australia, ACM.

Arena, M.P. and Arrigo, B.A. (2006) **The Terrorist Identity: Explaining the Terrorist Threat**. New York: New York University Press.

Aubel, A., Boulic, R. and Thalmann, D. (2000) Real-time display of virtual humans: levels of details and impostors. **Circuits and Systems for Video Technology, IEEE Transactions on**, 10: (2): 207-217.

Baiget, P., Fernandez, C., Roca, X., et al. (2009) Generation of Augmented Video Sequences Combining Behavioral Animation and Mulit-Object Tracking. **Computer Animation and Virtual Worlds**, 20: (4): 473.

Bailey, P., Brodkin, D., Haisworth, J., et al. (2008) **The Virtual Building** [online]. http://www.arup.com/_assets/_download/D6E07D17-19BB-316E-40DD7479AA23FA6D.pdf Arup [Accessed 10 December 2010]

Baker, C.L., Goodman Noah D. and B., T.J. (2008) "Theory-Based Social Goal Inference". Massachusetts Institute of Technology.

Baldwin, D.A. and Baird, J.A. (2001) Discerning Intention in Dynamic Human Action. **Trends in Cognitive Sciences**, 5: (4): 171-178.

Bergesen, A.J. (2007) Three-Step Model of Terrorist Violence. **Mobilization**, 12: (2): 111-117.

Berk, R. (ed.) (1972) **The Controversy Surrounding Analyses of Collective Violence: Some Methodological Notes**, Chicago: Aldine-Atherton.

Berk, R. (ed.) (1972) **The Emergence of Muted Violence in Crowd Behaviour: A Case Study of an Almost Race Riot**, Chicago: Aldine-Atherton.

Berk, R. (1974) **Collective Behaviour**. Dubuque, IA: Brown. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Berk, R. (1974) A gaming Approach to Crowd Behaviour. **American Sociological Review**, 39: (3): 355-373.

Berlonghi, A.E. (1995) Understanding and Planning for Different Spectator Crowds. **Safety Science**, 18: (4).

Bessel, R. and Emsley, A.C. (2000) **Patterns of Provocation Police and Public Disorder**. New York USA: Berghahn Books.

BIS (2010) **Bohemia Interactive Simulations** [online]. http://www.bisimulations.com/products/vbs2 [Accessed 1 June 2009]

BISForum (2007) **BI Simulations Community Forums** [online]. https://forums.bisimulations.com/ucp.php?mode=login phpBB Group [Accessed 30 June 2009]

Bloom, M. and Horgan, J. (2008) Missing Their Mark: The IRA's Proxy Bomb Campaign. **Social Research**, 75: (2): 579-614.

Blythe, P.W., Todd, P.M. and Miller, G.F. (eds.) (1999) **How Motion Reveals Intention. Categorizing Social Interactions**, Oxford: Oxford University Press.

Boersema, T. and Zwaga, H.G. (1990) "Searching for Routing Signs in Public Buildings". In Brogan, D. (Ed.) **Visual Search**. London, Taylor and Francis.

Braun, A., Musse, S.R., de Oliveira, L.P.L., et al. (2003) "Modeling individual behaviors in crowd simulation". In Musse, S.R. (Ed.) **Computer Animation and Social Agents, 2003. 16th International Conference on**.

Bridger, R., S. (2003) **Introduction to Ergonomics**. 2nd.New York: Taylor & Francis.

Brill, L. (2003) Terrorism, Crowds and Power, and the Dogs of War. **Anthropological Quarterly**, 76: (1): 87-94.

Busby, J., Parrish, Z. and Van Eenwyk, J. (2005) **Mastering Unreal Technology: The Art of Level Design**. United States of America: SAMS.

Card, K., S., Moran, P., K. and Newell, A. (1983) **The Psychology of Human-Computer Interaction**. Hillsdale, New Jersey: Lawerence Erlbaum Associates.

Card, S.K., Moran, T.P. and Newell, A. (1983) **Pyschology of Human-Computer Interactiom**. Hillsdale: Lawrence Erlbaum Associates.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009) "Understand Crowd Behaviour: A Guide for Readers". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009) "Understand Crowd Behaviour: Guidance and Lessons Identified". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009) "Understand Crowd Behaviour: Simulation Tools". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009) "Understand Crowd Behaviour: Supporting Documentation". In Office, C. (Ed.) York, Crown.

Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown.

Choi, H. and Scholl, B.J. (2004) Effects of Grouping and Attention on the Perception of Causality. **Perception & Psychophysics**, 66: (6): 926-942.

Chung, D., deBuys, B.D. and Nma, C.S. (2007) "Influence of Avatar Creation on Attitude, Empathy, Presence, and Para-Social Interaction". In Jacko, J.A. (Ed.) **Human-Computer Interaction. Interaction Design and Usability 12th International Conference**. Beijing, China, Springer.

Clifford, S. and John, D. (2000) Crowds, Context and Identity: Dynamic categorization processes in the 'poll tax riot'. **Human Relations**, 53: (2): 247.

Committee, I.a.S. (2005) "Report into the London Terrorist Attacks on 7 July 2005". Intelligence and Security Committee.

Cooper, E.O. (2005) Crowd Control Computer Program Simulates Crowd Behaviour in Military Situations. **Quest**, 8: (2): 8 -10.

Cottrell, N., B. (ed.) (1972) **Social Facilitation**, New York: Holt, Rinehart and Winston. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Courtney, E.H. (2007) Training FOR THE WAR ON TERROR. **Military & Aerospace Electronics**, 18: (11): 18.

Courtney, M. (2011) "Public Eyes Get Smart". **Engineering and Technology (E&T)**. United Kingdom, IET Publishing & Information Services.

CrowdDynamics **Myriad II - Evacuation Module** [online]. http://www.crowddynamics.com/evacuation-demo.php [Accessed 10 December 2010]

CRYMOD (2011) **CRYMOD DISCUSS. CREATE. SHARE.** [online]. http://www.crymod.com/portal.php 2011 Crytek GmBH [Accessed 23 June 2008]

CRYTEK (2011) **CRYTEK** [online]. http://crytek.com/ CRYTEK [Accessed 23 June 2008]

Daniel, T., Laurent, K., William, O., et al. (2005) "Crowd and group animation". **ACM SIGGRAPH 2005 Courses**. Los Angeles, California, ACM.

David, F. (2004) Creating emotion in games: the craft and art of Emotioneering (TM). **Comput. Entertain.**, 2: (3): 15-15.

Dey, P. and Roberts, D. (2007) "A Conceptual Framework for Modelling Crowd Behaviour". In Roberts, D. (Ed.) **Distributed Simulation and Real-Time Applications, 2007. DS-RT 2007. 11th IEEE International Symposium**.

Diener, E. (ed.) (1989) **Deindividuation: The Absence of Self-Awareness and Self-Regulation in Group Members**, Hillsdale, NJ: Lawrence Erlbaum.

Dietrich, A., Gobbetti, E. and Yoon, S.E. (2007) Massive-Model Rendering Techniques: A Tutorial. **IEEE Computer Graphics and Applications**, 27: 20-34.

Donald, F.M. (2008) "Vigilance and the Implication of using Threat Image Projection (TIP) for CCTV Surveillance Operators". In Valli, C.A. & Brooks, D. (Eds.) **Proceeding of the 1st Security and Intelligence Conference**. Edith Cowan University Mount Lawely Campus Perth Western Australia, SECAU-Security Research Centre.

Donald, I. and Canter, D. (1992) Intentionality and Fatality During the King's Cross Underground Fire. **European Journal of Social Psychology**, 22: (3): 203-218.

Drury, J. and Reicher, S.D. (1999) The Intergroup Dynamics of Collective Empowerment: Substantiating the Social Identity Model of Crowd Behavior. **Group Processes & Intergroup Relations**, 2: (4): 381-402.

Drury, J. and Stott, C. (2001) Bias as a Research Strategy in Participant Observation: The Case of Intergroup Conflict. **Field Methods**, 13: (1): 47-67.

Drury, J. and Stott, C. (2003) The Role of Police Perception and Practices in the Development of "Public Disorder". **Journal of Applied Social Psychology**, 33: (7): 1480-1500.

Earnshaw, R. and Vince, J. (2002) **Intelligent Agents for Mobile and Virtual Media**. London: Springer-Verlag.

Eberly, D.H. (2004) **Game Physics**. San Francisco: Morgan Kaufmann.

Eberly, D.H. (2007) **3D Game Engine Design A Practical Approach to Real-Time Computer Graphics**. 2nd.San Francisco: Morgan Kaufmann

Elsa, S., Soraja Raupp, M., Fabien, G., et al. (1999) "An Architecture to Guide Crowds Using a Rule-Based Behavior System". **Proceedings of the third annual conference on Autonomous Agents**. Seattle, Washington, United States, ACM.

Erik, L.W. (2007) 3D video games: no programming required. **J. Comput. Small Coll.**, 22: (3): 105-111.

Evertsz, R., Pedrotti, M., Busetta, P., et al. (2009) "Populating VBS2 with Realistic Virtual Actors". **Behavior Representation in Modeling & Simulation (BRIMS)**. Sundance Utah, BRIMS.

Evertsz, R., Ritter, F.E., Busetta, P., et al. (2008) "Realistic Behaviour Variation in a BDI-based Cognitive Architecture". **SimTect 2008**. Melbourne, AOS Group.

Eysenck, M.W. and Keane, M.T. (2010) **Cognitive Psychology A Student's Handbook**. Sixth.Hove: Psychology Press.

Falmier, O. and Young, M.E. (2008) The Impact of Object Animacy on the Appraisal of Causality. **American Journal of Psychology**, 121: (3): 473-500.

Festinger, L., Pepitone, A. and Necombe, T. (1952) Some Consequences of De-individuation in a Group. **The Journal of Abnormal and Social Psychology**, 47: (2): 382-389.

Francois, G., Pattanaik, S., Bouatouch, K., et al. (2008) Subsurface Texture Mapping. **IEEE Computer Graphics and Applications**, 28: 34-42.

Freud, S. (1951) "Group Psychology and the Analysis of the Ego". In Jones, E. (Ed.) 6th ed. New York, Liveright Publishing, Corporation.

Fridman, N. and Kaminka, G. (2006) Modeling Crowd Behavior Based on Social Comparison Theory: Extended Abstract. **Cellular Automata: Lecture Notes in Computer Science**, Volume 4173/2006: (7the International Conference on Cellular Automata, for Research and Industry, ACRI 2006): 694–698.

FSEG (2009) **Fire Safety Engineering Group** [online]. http://fseg.gre.ac.uk/exodus/ University of Greenwich [Accessed 5 March 2008]

Garson, D.G. (2009) Computerized Simulation in the Social Science. **Simulation & Gaming**, 40: (2): 267-279.

Gaskins, R.C., Boone, C.M., Verna, T.M., et al. (2004) Modeling & Simulation Group Conference on Modeling and Simulation to Address NATO's New and Exisiting Military Requirement. (1 - 6).

Gaskins, R.C., Kruszewski, P.A., McKenzie, F.D., et al. (2008) Integrating Crowd-Behaviour Modeling into Military Simulation using Game Technology. **Simulation & Gaming**, 39: (1).

Gaskins, R.C., McKenzie, F.D., Weisel, E.W., et al. (2004) Developing a Crowd Federate for Military Simulation. **Proceedings of the Spring 2004 Simulation Interoperability Workshop**, 18-23.

Gerwehr, S. and Glenn, R.W. (2000) **The Art of Darkness: Deception and Urban Operations**. Santa Monica: RAND.

Gerwehr, S. and Glenn, R.W. (2002) **Unweaving the Web: Deception and Adaptation in Future Urban Operations**. Santa Monica: RAND.

Gigerenzer, G., Todd, P.M. and Group, A.R. (1999) **Simple Heuristics That Make Us Smart**. New York: Oxford University Press Inc.

Giustozzi, A. (2007) **Koran, Kalashnikov and Laptop: The Neo-Taliban Insurgency in Afghanistan**. London: HURST Publisher Ltd.

Google (2011) **Android Developers** [online]. http://developer.android.com/sdk/index.html [Accessed 11 November 2010]

Harris, S., Stedmon, A., Sharples, S., et al. (2008) "Keeping Track of a Suspect on CCTV: How Hard Can IT Be?!". In Bust, P.D. (Ed.) **Proceedings of the International Conference on Contemporary Ergonomics**. Nottingham, Taylor & Francis.

Havok (2011) **Havok** [online]. http://www.havok.com/ [Accessed 5 April 2010]

Hawes, N. (2000) Real-time goal-orientated behaviour for computer game agents. **Game-On 2000, 1st International Conference on Intelligent Games and Simulation**, 71–75.

Hawes, N. (2001) Anytime planning for agent behaviour. **Proceedings of the Twelth Workshop of the UK Planning and Scheduling Special Interest Group**, 157–166.

Hawes, N.A. and University of, B. (2003) **Anytime Deliberation for Computer Game Agents**. University of Birmingham.

Helbing, D., Johansson, A. and Al-Abideen, H.Z. (2007) Dynamics of crowd disasters: An empirical study. **Physical Review E**, 75: (4): 46109.

Hills, A. (2001) The Inherent Limits of Military Forces in Policing Peace Operations. **International Peacekeeping - London**, 00008: (00003): 79-99.

Hoggett, J. and Clifford, S. (2010) Crowd Psychology, Public Order Police Training and the Policing of Football Crowds. **Policing: An International Journal of Police Strategies and Management**, 33: (2).

Hopmeier, M., Ganor, B., Goodwin, T., et al. (2003) "There Are No Dangerous Weapons...": Suicide Attacks and Potential Responses. **Journal of Homeland Security**.

Horgan, J. (2005) **The Psychology of Terrorism**. Cornwall: Routledge.

Horgan, J. (2006) "Understanding Terrorism: Old Assuptions, New Assertions, And Challenges for Research". In Victoroff, J. (Ed.) **Tangled Roots: Social and Psychological Factors in the Genesis of Terrorism (NATO Security Through Science Series: Human and Societal Dynamics**. IOS Press 496.

Horgan, J. (2008) From Profiles to Pathways and Roots to Routes: Perspectives from Psychology on Radicalization into Terrorism. **The Annals of the American Academy of Political and Social Science**, 618: (1): 80-94.

Houghton, R.J. and Baber, C. (2009) **Detecting Deviants within Flocks**. Maastricht: Shaker.

IES (2011) **Integrated Environment Solutions** [online]. http://www.iesve.com/Software/VE-Pro/Simulex [Accessed 5 February 2008]

Ilya, B. and Jovan, P. (2007) "Automatic rigging and animation of 3D characters". **ACM SIGGRAPH 2007 papers**. San Diego, California, ACM.

Jennett, C., Cox, L., A., Cairns, P., et al. (2008) Measuring and Defining the Experience of Immersion in Games. **International Journal of Human-Computer Studies**, 66: (9): 641-661.

Johnson-Laird, P.N. (1983) **Mental Models Towards a Cognitive Science of Languages, Inference, and Consciousness**. Cambridge: Cambridge University Press.

Jur van den, B., Sachin, P., Jason, S., et al. (2008) "Interactive navigation of multiple agents in crowded environments". **Proceedings of the 2008 symposium on Interactive 3D graphics and games**. Redwood City, California, ACM.

Kasap, Z., Ben Moussa, M., Chaudhuri, P., et al. (2009) Making Them Remember-Emotional Virtual Characters with Memory. **Computer Graphics and Applications, IEEE**, 29: (1): 20.

Kasik, D.J. (2008) Tools and Products. **IEEE Computer Graphics and Applications**, 28: 79-80.

Keval, H. and Sasse, A.M. (2008) "To Catch a Thief -You need at least 8 Frames Per Second: The Impact of Frame Rates on User Performance in a CCTV Detection Task". **Proceeding of the 16th ACM international conference on Multimedia**. Vancouver, British Columbia, Canada, ACM.

Keval, H.U. and Sasse, M.A. (2006) "Man or Gorilla? Performance Issue with CCTV Technology in Security Control Rooms". **16th World Congress on Ergonomics**. Maastricht, International Ergonomics Association.

Klein, G. (1999) **Source of Power How People Make Decision**. United States of America: Massachusetts Institute of Technology.

Klein, O., Spears, R. and Reicher, S. (2007) Social Identity Performance: Extending the Strategic Side of SIDE. **Personality and Social Psychology Review***,* 11: (1): 28-45.

Klupfel, H., Schreckenberg, M. and Meyer-Konig, T. (2003) "Models for Crowd Movement and Egress Simulation". In Hoogendoorn, S.P. (Ed.) **Workshop on Traffic and Granular Flow***.* Delft, Netherlands, Berlin.

Koenig, S.C., Liebhold, G.M.Y. and Gramopadhye, A.K. (1998) "Training for Systematic Search Using a Job Aid". **Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting***.* Santa Monica, The Human Factors and ergonomic Society.

Latane, B., Williams, K. and Harkins, S. (1979) Many Hands make Light the Work: The Causes and Consequences of Social Loafing. **Journal of Personality and Social Psychology***,* 37: (6): 822-832.

Le Bon, G. (1896) "The Crowd: A Study of the Popular Mind". New York, The Macmillan Co.

Legion (2007) **Legion #1 in Pedestrian Simulation Solutions** [online]. http://www.legion.com/ 2008 [Accessed 5 February]

Lipton, A.J.D. "Intelligent Video Surveillance in Crowds". Reston, VA, Object Video Ltd.

Lorenz, F.M. (1996) Non-Lethal Force: The Slippery Slope to War? **Parameters***,* 26: (3): 52-62.

Luck, M. and Aylett, R. (2000) Applying Artificial Intelligence to Virtual Reality: Intelligent Virtual Environments. **Applied Artificial Intelligence***,* 14: (1): 3 - 32.

Machover, C. and Kasik, D.J. (2007) Tools and Products. **IEEE Computer Graphics and Applications***,* 27: 100-199.

Major, J.A. (2002) Advanced Techniques for Modeling Terrorism Risk. **The Journal of Risk Finance***,* 4: (1).

Mandun, Z., Jian, Y., Bin, D., et al. (2005) "Fast individual face modeling and animation". **Proceedings of the second Australasian conference on Interactive entertainment***.* Sydney, Australia, Creativity \& Cognition Studios Press.

Mark, A.G. (1999) Civil disturbance operations for mechanized engineer battalions. **Engineer***,* 29: (1): 32.

Masaki, O. and Akifumi, M. (2002) "A dynamic motion control middleware for computer games". **ACM SIGGRAPH 2002 conference abstracts and applications***.* San Antonio, Texas, ACM.

Mayer, I. and Bekebrede, G. (2004) "Serious Games and `Simulation Based E-Learning' for Infrastructure Management". In Pivec, M. (Ed.) **Affective and emotional aspects of human-computer interaction: game-based and innovative learning approaches***.* Portschach, Austria, Amsterdam.

McAleer, P. and Pollick, F.E. (2008) Understanding Intention from Minimal Displays of Human Activity. **Behavior Research Methods***,* 40: (3): 830-839.

McKenzie, F.D., Xu, Q. ,Nguyen, Q. A. H. ,Petty, M. D. (2004) Crowd Federate Architecture and API Design. **Proceedings of the Fall 2004 Simulation Interoperability Workshop***,* 19-24.

Meador, D.P., McClure, R.S. and Brett, B.E. (2006) Development of a Virtual Terrorist. **Human Factors and Ergonomics Society Annual Meeting Proceedings***,* 50: 2255-2258.

Michotte, A. (1963) **The Perception of Causality**. London: Methuen & Co Ltd.

Microsoft (2011) **App Hub** [online]. http://create.msdn.com/en-US/ [Accessed 20 January 2008]

Moray, N. (ed.) (1976) **Attention, Control and Sampling Behaviour***,* New York: Plenum Press.

Nguyen, Q., McKenzie, F.R.D. and Petty, M.D. (2005) Crowd Behavior Cognitive Model Architecture Design. **Proceedings of the 2005 Conference on Behavior Representation in Modeling and Simulation (BRIMS)***,* 16-19.

Nitschke, B. (2007) **XnaProjects.Net** [online]. http://xnaprojects.exdream.com/ [Accessed 5 March 2008]

Norman A., D. (1988) **The Design of Everyday Things**. United States of America: MIT Press edition.

OED (2010) **Oxford English Dictionary** [online]. http://www.oed.com/ Oxford University Press [Accessed 22 March 2008]

Pattie, M. (1995) Artificial life meets entertainment: lifelike autonomous agents. **Commun. ACM***,* 38: (11): 108-114.

Pelachaud, C., Martin, J.-C., Andre, E., et al. (2007) "Intelligent Virtual Agents". **7th International Conference, IVA 2007***.* Paris, France, Springer-Verlag.

Pelechano, N., O'Brien, K., Silverman, B., et al. (2005) Crowd simulation incorporating agent psychological models, roles and communication. **First International Workshop on Crowd Simulation**.

Philips, Z. (2007) Criminal Cues. **Government Executive***,* 39: (6): 17.

Preece, J., Rogers, Y., Sharp, H., et al. (1994) **Human-Computer Interaction**. Woking, England: Open University (Addison-Wesley).

Prentice-Dunn, S. and Rogers, R., W. (eds.) (1989) **Deindividuation and the Self-Regulation of Behavior***,* Hillsdale, NJ: Lawrence Erlbaum.

Rabin, S. (2006) **AI Game Programming Wisdom 3**. Boston: Charles River Media.

Rahmalan, H., Nixon M.S. and Carter, J.N. (2007) "On Crowd Density Estimation for Surveillance". **Crime and Security, 2006. The Institution of Engineering and Technology Conference on** London, IEEE Xplore.

Reicher, S. (2004) **Self and Social Identity (Perspectives on Social Psychology)**. Cornwall UK: WileyBlackwell

Reicher, S., Stott, C., Cronin, P., et al. (2004) An Integrated Approach to Crowd Psychology and Public order Policing. **An International Journal of Police Strategies & Management***,* 27: (4):

558-572.

Reicher, S.D. (1984) Social Influence in the Crowd - Attitudinal and Behavioural Effects of De-Individuation in Conditions of High and Low Group Salience. **British Journal of Social Psychology.**, 23: 341-350 (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. (1984) The St Pauls' Riot: An Explanation of the Limits of Crowd Action in Terms of a Social Model. **European Journal of Social Psychology**, 14: (1): 1-21.

Reicher, S.D. (ed.) (1987) **Crowd behaviour as Social Action**, Oxford: Basil-Blackwell.

Reicher, S.D. (1996) 'The Battle of Westminster': Developing the Social Identity Model of Crowd Behaviour in Order to Explain the Initiation and Development of Collective Conflict. **European Journal of Social Psychology**, 26: (1): 115-134.

Reicher, S.D. (1996) 'The Crowd' Century: Reconciling Practical Success with Theoretical Failure. **British Journal of Social Psychology.**, (35): 535-553. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. (ed.) (1996) **Social Identity and Social Change: Rethinking the Context of Social Psychology**, London: Butterworth. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. (1997) Collective Psychology and the Psychology of the Self. **British Psychological Society Social Section Newsletter**, 36: 3-15. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. and levine, M. (1994) Deindividuation, Power Relations between Groups and the Expression of Social Identity. **British Journal of Social Psychology.**, 33: 145-163 (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reicher, S.D. and Levine, M. (1994) On the Consequences of Deindividuation Manipulation for the Strategic Communication of Self: Identifiability and the Presentation of Social Identity. **European Journal of Social Psychology**, 24: (4): 511-524.

Reicher, S.D., Spears, R. and Postmes, T. (1995) A Social Identity Model of Deindividuation Phenomena. **European Review of Social Psychology**, 6: 161-198. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Reynolds, C.W. (1987) "Flocks, herds and schools: A distributed behavioural model". **Computer Animation: CG87. Proceedings of the Conference held at Computer Graphics 87**. London, UK, Online Publications.

Rittenbrunch, M. (2002) Atmosphere: A Framework for Contextual Awareness. **International Journal of Human-Computer Studies**, 14: (2): 159-180.

Rob, S. and John, S.G. (2004) Curious Agents and Situated Design Evaluations. **Artif. Intell. Eng. Des. Anal. Manuf.**, 18: (2): 153-161.

Robert, P., Thomas, A., Ted, S., et al. (2004) Countering Terrorism Through Information Technology. **Commun. ACM**, 47: (3): 36-43.

Royan, J., Gioia, P., Cavagna, R., et al. (2007) Network-Based Visualization of 3D Landscapes and City Models. **IEEE Computer Graphics and Applications**, 27: 70-79.

Schbley, A. and McCauley, C. (2005) Political, Religious, and Psychological Characteristics of Muslim Protest Marchers in Eight European Cities: Jerusalem Day 2002. **Terrorism and Political Violence**, 17: (4): 551 - 572.

Schlottman, A., Ray, E.D., Mitchell, A., et al. (2005) Perceived Physical and Social Causality in Animated Motions: Spontaneous Reports and Ratings. **Acta Psychologica**, 123: (1-2): 112-143.

Scholl, B.J. and Nakayama, K. (2002) Causal Capture: Contextual Effects on the Perception of Collision Events. **Psychological Science**, 13: (6): 493-498.

Scholl, B.J. and Nakayama, K. (2004) Illusory Causal Crescents: Misperceived Spatial Relations Due to Perceived Causality. **Perception**, 33: (4): 455-469.

Scholl, B.J. and Tremoulet, P.D. (2000) Perceptual Causality and Animacy. **Trends in Cognitive Sciences**, 4: (8): 8.

Schweingruber, D. and Wohlstein, R.T. (2005) The Madding Crowd Goes to School: Myths about Crowds in Introductory Sociology Textbooks. **Teaching Sociology**, 33: (2): 136-153.

Sheng, L. (2007) Real-Time Rendering of Natural Light Shafts for Indoor and Outdoor Scenes. **Journal of Computer Aided Design and Computer Graphics**, 19: 1580-1586.

Silverman, B., Johns, M., Cornwell, J., et al. (2006) Human Behavior Models for Agents in Simulators and Games: Part I: Enabling Science with PMFserv **Presence: Teleoperators and Virtual Environments**, 15: (2): 139-162.

Silverman, B., O'Brien, K., Cornwell, J., et al. (2006) Human Behavior Models for Agents in Simulators and Games: Part II Gamebot Engineering with PMFserv **Presence: Teleoperators and Virtual Environments**, 15: (2): 163-185.

Silverman, B., Shin, H., Dubois, R., et al. (2001) "Modeling and Simulating Terrorist Decision-making: A 'Performance Moderator Function' Approach to Generating Virtual Opponents". **Proceedings of the 10th Conference On Computer Generated Forces and Behavioral Representation**.

Sime, J.D. (1995) Crowd psychology and engineering. **Safety Science**, 21: (1): 1-14.

Smith, E., E, Nolen-Hoeksema, S., Fredrickson, L., B., et al. (2003) **Atkinson & Higard's Introduction to Psychology**. 14th.United States of America: Thomson Wadsworth.

Smyth, M.M., Collins, A.F., Morris, P.E., et al. (1995) **Cognition In Action 2nd Edition**. 2nd.Hove: Lawrence Erlbaum Associates Ltd.

Smyth, M.M., Morris, P.E., Levy, P., et al. (1987) **Cognition In Action**. London: Lawrence Erlbaum Associates Ltd.

Soraia Raupp, M., Branislav, U., Amaury, A., et al. (2005) "Groups and Crowd Simulation". **ACM SIGGRAPH 2005 Courses**. Los Angeles, California, ACM.

Soraia Raupp, M. and Daniel, T. (2000) "From One Virtual Actor to Virtual Crowds: Requirements and Constraints". **Proceedings of the fourth international conference on Autonomous agents**. Barcelona, Spain, ACM.

Soraia, R.M., Christian, B., Tolga, C., et al. (1998) "Crowd Modelling in Collaborative Virtual Environments". **Proceedings of the ACM symposium on Virtual reality software and technology**. Taipei, Taiwan, ACM.

Sredmon, A., Harris, S., Carse, A., et al. (2008) "Tracking a Suspicious Person using CCTV: But What Do We Mean By Suspicious Behaviour?". In Bust, P.D. (Ed.) **Proceeding of the International Conference of Contemporary Ergonomics**. Nottingham UK, Taylor & Francis.

Stone, J.R. (2008) "Human Factors Guidelines for Interactive 3D and Game-Based Training Systems Design". In MoD (Ed.), Human Factors Integration Defence Technology Centre.

Stott, C. (2009) **Crowd Psychology & Public Order Policing: An Overview of Scientific Theory and Evidence, Submission to the HMIC Policing of Public Protest Review Team** [online]. http://www.liv.ac.uk/psychology/cpd/Stott_(2009).pdf Liverpool University [Accessed 5 January 2011

Stott, C. and Drury, J. (eds.) (1999) **The Intergroup Dynamics of Empowerment: A Social Identity Model**, London: Macmillan. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Stott, C. and Reicher, S. (1998) Crowd Action as Intergroup Process: Introducing the Police Perspective. **European Journal of Social Psychology**, 00028: (00004): 509-531.

Tajfel, H. (ed.) (1978) **Social Categorization, Social Identity, and Social Comparison**, London: Academic Press (for) European Association of Experimental Social Psychology.

Tajfel, H., Billig, M., Bundy, R.P., et al. (1971) Social Categorization and Intergroup Behaviour. **European Journal of Social Psychology**, 1: (2): 149-177.

Tajfel, H. and Turner, J.C. (eds.) (1979) **An Integrative Theory of Intergroup Conflict**, Monterey, CA: Brooks/Cole. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Tajfel, H. and Turner, J.C. (eds.) (1986) **The Social Identity Theory of Intergroup Behaviours**, Chicago, IL: Nelson-Hall. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Taylor, M. and Horgan, J. (2006) A Conceptual Framework for Addressing Psychological Process in the Development of the Terrorist. **Terrorism and Political Violence**, 18: (4): 585-601.

Taylor, S.E. and Fiske, S.T. (1975) Point of View and Perception of Causality. **Journal of Personality and Social Psychology**, 32: (3): 439-445.

Thalmann, D. and Musse, S. (2007) **Crowd Simulation**. Springer London.

Ting, S.-P. and Zhou, S. (2009) Dealing wih Dynamic Changes in Time Critical Decision-Making for MOUT Simulations. **Computer Animation and Virtual Worlds**, 20: (2/3): 427.

Trenholme, D. and Smith, S. (2008) Computer Game Engines for Developing First-Person Virtual Environments. **Virtual Reality**, 12: (3): 181-187.

Troscianko, T., Holmes, A., Stillman, J., et al. (2004) What Happens Next? The Predictability of Natural Behaviour Viewd Through CCTV Cameras. **Perception***,* 33: (1): 87-101.

Turner, J.C. (ed.) (1982) **Towards a Cognitive Redefinition of the Social Group***,* Cambridge: Cambridge University Press.

Turner, J.C. (ed.) (1985) **Social Categorization and the Self-Concept: A Social Cognitive Theory of Group Behaviour***,* Greenwich, CT: JAI Press. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Turner, J.C., Hogg, M.A., Oakes, P.J., et al. (1987) **Rediscovering the Social Group: A Self-Categorization Theory**. Oxford: Basil Blackwell.

Turner, J.C. and Oakes, P.J. (1986) The Significance of the Social Identity Concept for the Social Psychology with Reference to Individualism, Interactionism and Social Influence. **British Journal of Social Psychology.***,* 25: (3): 237-252. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Turner, R.H. (ed.) (1964) **Collective Behavior***,* Chicago: Randy McNally.

Turner, R.H. and Killian, L.M. (1957) **Collective behavior**. Englewood, NJ: Prentice-Hall.

UDK (2010) **UNREAL Development Kit** [online]. http://www.udk.com/ Epic Games Inc. [Accessed 3 August 2009]

Ulicny, B. and , T., D. (2001) Crowd Simulation for Interactive Virtual Environments and VR Training Systems. **Computer Animation and Simulation 2001: Proceedings of the Eurographics Workshop in Manchester, UK, September 2-3, 2001***,* 163-171.

Velastin, S.A. (2009) "CCTV Video Analytics: Recent Advances and Limitations". In Zaman, H.B.;Robinson, P.;Petrou, M.;Olivier, P.;Schroder, H. & Shih, T.K. (Eds.) **First International Visual Informatics Conference***.* Kuala Lumpur, Malaysia, Springer.

Vider, S. (2004) Rethinking Crowd Violence: Self-Categorization Theory and the Woodstock 1999 Riot. **Journal for the Theory of Social Behaviour***,* 00034: (00002): 141-167.

Waddington, D. and King, M. (2005) The Disorderly Crowd: From Classical Psychological Reductionism to Socio-Contextual Theory - The Impact on Public Order Policing Strategies. **Howard Journal of Criminal Justice***,* 00044: (00005): 490-504.

Warren, R. (2002) Situating the city and September 11th: military urban doctrine, 'pop-up' armies and spatial chess. **International Journal of Urban and Regional Research***,* 26: (3): 614-619.

Weaver, R., Silverman, B.G., Shin, H., et al. (2001) Modeling and Simulating Terrorist Decision-making: A ―Performance Moderator Function" Approach to Generating Virtual Opponents. **Proceedings of the Tenth Conference on Computer Generated Forces and Behavioral Representation***,* 39–44.

Wee Ling, W., Cuihua, S., Luciano, N., et al. (2007) "Serious video game effectiveness". **Proceedings of the international conference on Advances in computer entertainment technology***.* Salzburg, Austria, ACM.

Werner, B., Helmut, P. and Mitsuru, I. (2007) "Automated generation of non-verbal behavior for virtual embodied characters". **Proceedings of the 9th international conference on Multimodal interfaces**. Nagoya, Aichi, Japan, ACM.

Wheeler, S. (2005) It Pays to Be Popular: a Study of Civilian Assistance and Guerilla Warfare. **Journal of Artificial Societies and Social Simulation**, 8: (4): 9.

Wheeler, S. (2005) "On the Suitability of NetLogo for the Modelling of Civilian Assistance and Guerrilla Warfare". In Laboratory, D.S.S. (Ed.) Edinburgh, South Australia.

Wilensky, U. (2011) **NetLogo** [online]. http://ccl.northwestern.edu/netlogo/ [Accessed 1 April 2008]

Wood, J. (ed.) (2007) **CCTV Ergonomics: Case Studies and Practical Guidance**, Amsterdam: Elsevier Ltd.

Xuelong, L., Maybank, S.J., Shuicheng, Y., et al. (2008) Gait Components and Their Application to Gender Recognition. **IEEE Transactions on Systems, Man and Cybernetics. Part C**, 38: (2): 145-155.

Yu, W. and Johansson, A. (2007) Modeling Crowd Turbulence by Many-Particle Simulations. **American Physical Society**, 76: (4): 046105-046101-046105.

Zanjoc, R.B. (1965) "Social Facilitation". **Science**. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

Zhigeng, P. (2007) Overview of Emotion for Virtual Agent. **Journal of Computer Aided Design and Computer Graphics**, 19: 1509-1516.

Zimbardo, P.G. (ed.) (1970) **The Human Choice: Individuation, Reason, and Order versus Deindividuation, Impulse and Chaos**, Lincoln: University of Nebraska Press. (Cited in Challenger, R., Clegg, C.W., Robinson, M.A., et al. (2009b) "Understand Crowd Behaviour: Supporting Evidence". In Office, C. (Ed.) York, Crown).

APPENDECIES

APPENDIX A

1.0 Pairwise Comparison

Table 1.1 Pairwise Comparison (Distribution of Information)

Measure:MEASURE_1

(I) TextOptions	(J) TextOptions	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	7.599*	1.802	.003	2.006	13.192
	3	-1.044	2.253	1.000	-8.038	5.950
	4	3.018	2.625	1.000	-5.129	11.165
	5	5.926	1.967	.062	-.180	12.033
2	1	-7.599*	1.802	.003	-13.192	-2.006
	3	-8.643*	2.626	.032	-16.795	-.491
	4	-4.581	2.619	.936	-12.710	3.547
	5	-1.673	2.081	1.000	-8.132	4.787
3	1	1.044	2.253	1.000	-5.950	8.038
	2	8.643*	2.626	.032	.491	16.795
	4	4.062	3.004	1.000	-5.262	13.385
	5	6.970*	2.173	.039	.226	13.715
4	1	-3.018	2.625	1.000	-11.165	5.129
	2	4.581	2.619	.936	-3.547	12.710
	3	-4.062	3.004	1.000	-13.385	5.262
	5	2.908	2.342	1.000	-4.360	10.177
5	1	-5.926	1.967	.062	-12.033	.180
	2	1.673	2.081	1.000	-4.787	8.132
	3	-6.970*	2.173	.039	-13.715	-.226
	4	-2.908	2.342	1.000	-10.177	4.360

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Table 1.2 Pairwise Comparison (12 Cameras)

Measure:MEASURE_1

(I) Cameras	(J) Cameras	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	6.223	3.029	1.000	-8.869	21.315
	3	17.108*	2.101	.001	6.643	27.574
	4	18.589*	2.754	.006	4.869	32.310
	5	-3.491	5.423	1.000	-30.509	23.528
	6	-2.622	4.822	1.000	-26.644	21.400
	7	10.910	3.112	.440	-4.595	26.415
	8	15.500*	2.466	.009	3.216	27.784
	9	.209	3.550	1.000	-17.476	17.894
	10	2.479	4.059	1.000	-17.744	22.702
	11	9.934	4.106	1.000	-10.522	30.390
	12	15.429*	2.755	.022	1.705	29.154
2	1	-6.223	3.029	1.000	-21.315	8.869
	3	10.885*	1.926	.021	1.292	20.479
	4	12.366	2.983	.165	-2.496	27.229
	5	-9.714	5.908	1.000	-39.148	19.721
	6	-8.845	4.652	1.000	-32.021	14.331
	7	4.687	3.065	1.000	-10.584	19.958
	8	9.277	3.618	1.000	-8.747	27.301
	9	-6.014	3.408	1.000	-22.993	10.965
	10	-3.744	4.005	1.000	-23.697	16.209
	11	3.711	4.524	1.000	-18.825	26.247
	12	9.206	3.927	1.000	-10.356	28.769
3	1	-17.108*	2.101	.001	-27.574	-6.643
	2	-10.885*	1.926	.021	-20.479	-1.292
	4	1.481	1.802	1.000	-7.495	10.457
	5	-20.599	5.273	.237	-46.870	5.672
	6	-19.731	4.402	.101	-41.662	2.201
	7	-6.198	2.377	1.000	-18.043	5.646
	8	-1.608	2.489	1.000	-14.006	10.789
	9	-16.899*	2.742	.011	-30.559	-3.240
	10	-14.630	3.572	.178	-32.423	3.164
	11	-7.174	3.116	1.000	-22.700	8.351
	12	-1.679	2.766	1.000	-15.458	12.100
4	1	-18.589*	2.754	.006	-32.310	-4.869
	2	-12.366	2.983	.165	-27.229	2.496
	3	-1.481	1.802	1.000	-10.457	7.495
	5	-22.080	4.509	.056	-44.545	.385
	6	-21.212*	4.245	.049	-42.357	-.066
	7	-7.679	2.182	.430	-18.549	3.190
	8	-3.089	1.744	1.000	-11.776	5.597
	9	-18.380*	2.288	.001	-29.778	-6.983
	10	-16.111	3.948	.182	-35.781	3.560
	11	-8.656	3.259	1.000	-24.892	7.581
	12	-3.160	1.918	1.000	-12.717	6.397
5	1	3.491	5.423	1.000	-23.528	30.509

	2	9.714	5.908	1.000	-19.721	39.148
	3	20.599	5.273	.237	-5.672	46.870
	4	22.080	4.509	.056	-.385	44.545
	6	.868	3.355	1.000	-15.844	17.581
	7	14.401	3.832	.297	-4.689	33.490
	8	18.990	4.662	.184	-4.236	42.217
	9	3.700	3.276	1.000	-12.620	20.019
	10	5.969	5.796	1.000	-22.905	34.844
	11	13.424	5.345	1.000	-13.202	40.051
	12	18.920	4.378	.127	-2.893	40.733
6	1	2.622	4.822	1.000	-21.400	26.644
	2	8.845	4.652	1.000	-14.331	32.021
	3	19.731	4.402	.101	-2.201	41.662
	4	21.212*	4.245	.049	.066	42.357
	5	-.868	3.355	1.000	-17.581	15.844
	7	13.532	3.216	.150	-2.488	29.552
	8	18.122	4.882	.319	-6.198	42.442
	9	2.831	2.374	1.000	-8.995	14.657
	10	5.101	3.308	1.000	-11.378	21.580
	11	12.556	4.067	.857	-7.704	32.816
	12	18.052	3.823	.072	-.995	37.098
7	1	-10.910	3.112	.440	-26.415	4.595
	2	-4.687	3.065	1.000	-19.958	10.584
	3	6.198	2.377	1.000	-5.646	18.043
	4	7.679	2.182	.430	-3.190	18.549
	5	-14.401	3.832	.297	-33.490	4.689
	6	-13.532	3.216	.150	-29.552	2.488
	8	4.590	2.230	1.000	-6.518	15.698
	9	-10.701*	1.996	.030	-20.643	-.759
	10	-8.431	3.329	1.000	-25.017	8.155
	11	-.976	3.189	1.000	-16.862	14.910
	12	4.519	2.980	1.000	-10.327	19.366
8	1	-15.500*	2.466	.009	-27.784	-3.216
	2	-9.277	3.618	1.000	-27.301	8.747
	3	1.608	2.489	1.000	-10.789	14.006
	4	3.089	1.744	1.000	-5.597	11.776
	5	-18.990	4.662	.184	-42.217	4.236
	6	-18.122	4.882	.319	-42.442	6.198
	7	-4.590	2.230	1.000	-15.698	6.518
	9	-15.291	3.319	.084	-31.828	1.246
	10	-13.021	4.544	1.000	-35.658	9.616
	11	-5.566	3.835	1.000	-24.670	13.538
	12	-.071	2.805	1.000	-14.042	13.901
9	1	-.209	3.550	1.000	-17.894	17.476
	2	6.014	3.408	1.000	-10.965	22.993
	3	16.899*	2.742	.011	3.240	30.559
	4	18.380*	2.288	.001	6.983	29.778
	5	-3.700	3.276	1.000	-20.019	12.620
	6	-2.831	2.374	1.000	-14.657	8.995
	7	10.701*	1.996	.030	.759	20.643
	8	15.291	3.319	.084	-1.246	31.828
	10	2.270	3.196	1.000	-13.650	18.190

11	9.725	3.212	.944	-6.277	25.727
12	15.220*	2.169	.004	4.417	26.024

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Table 1.3 Pairwise Comparison (Four Advantage Point Camera n…1)

Measure:MEASURE_1

(I) Cameras	(J) Cameras	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	3.120	2.466	1.000	-5.177	11.418
	3	13.745*	2.250	.001	6.176	21.313
	4	17.600*	2.004	.000	10.860	24.340
2	1	-3.120	2.466	1.000	-11.418	5.177
	3	10.625*	2.032	.003	3.789	17.460
	4	14.480*	2.963	.005	4.511	24.448
3	1	-13.745*	2.250	.001	-21.313	-6.176
	2	-10.625*	2.032	.003	-17.460	-3.789
	4	3.855	1.874	.419	-2.450	10.160
4	1	-17.600*	2.004	.000	-24.340	-10.860
	2	-14.480*	2.963	.005	-24.448	-4.511
	3	-3.855	1.874	.419	-10.160	2.450

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Table 1.4 Pairwise Comparison Trial 2 & 3 (Four Advantage Point Camera n…1)

Measure:MEASURE_1

(I) Camera	(J) Camera	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.511	.921	1.000	-3.608	2.587
	3	3.141*	.683	.008	.844	5.437
	4	4.567*	.827	.002	1.785	7.349
2	1	.511	.921	1.000	-2.587	3.608
	3	3.651*	1.079	.048	.021	7.281
	4	5.077*	1.305	.022	.688	9.467
3	1	-3.141*	.683	.008	-5.437	-.844
	2	-3.651*	1.079	.048	-7.281	-.021
	4	1.426	.581	.219	-.528	3.380
4	1	-4.567*	.827	.002	-7.349	-1.785
	2	-5.077*	1.305	.022	-9.467	-.688
	3	-1.426	.581	.219	-3.380	.528

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Table 1.5 Pairwise Comparison Trial 4 & 5 (Four Advantage Point Camera n…1)

Measure:MEASURE_1

(I) Cameras	(J) Cameras	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	3.430	1.294	.159	-.925	7.785
	3	7.373*	1.311	.002	2.961	11.784
	4	8.590*	1.329	.001	4.120	13.060
2	1	-3.430	1.294	.159	-7.785	.925
	3	3.942*	.733	.003	1.476	6.408
	4	5.160*	.988	.003	1.836	8.483
3	1	-7.373*	1.311	.002	-11.784	-2.961
	2	-3.942*	.733	.003	-6.408	-1.476
	4	1.218	.590	.414	-.766	3.201
4	1	-8.590*	1.329	.001	-13.060	-4.120
	2	-5.160*	.988	.003	-8.483	-1.836
	3	-1.218	.590	.414	-3.201	.766

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Table 1.6 Pairwise Comparison Trial 6 & 7 (Four Advantage Point Camera n…1)

Measure:MEASURE_1

(I) Cameras	(J) Cameras	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	.199	1.335	1.000	-4.292	4.690
	3	3.230	1.205	.151	-.825	7.285
	4	4.443*	.800	.002	1.753	7.133
2	1	-.199	1.335	1.000	-4.690	4.292
	3	3.031*	.699	.011	.679	5.383
	4	4.244	1.420	.091	-.532	9.019
3	1	-3.230	1.205	.151	-7.285	.825
	2	-3.031*	.699	.011	-5.383	-.679
	4	1.213	1.124	1.000	-2.570	4.995
4	1	-4.443*	.800	.002	-7.133	-1.753
	2	-4.244	1.420	.091	-9.019	.532
	3	-1.213	1.124	1.000	-4.995	2.570

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

APPENDIX B

1.0 Text Analysis Naïve Observers (Food Aid)

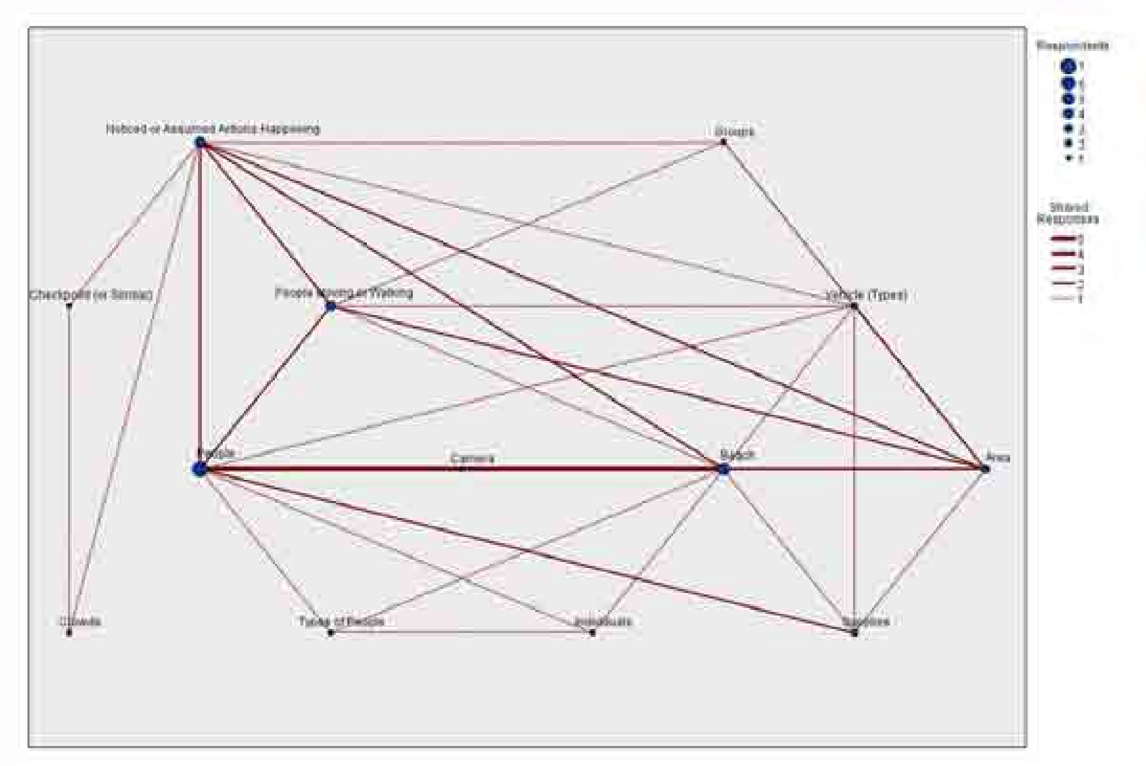

Fig 1.1 Trial 2

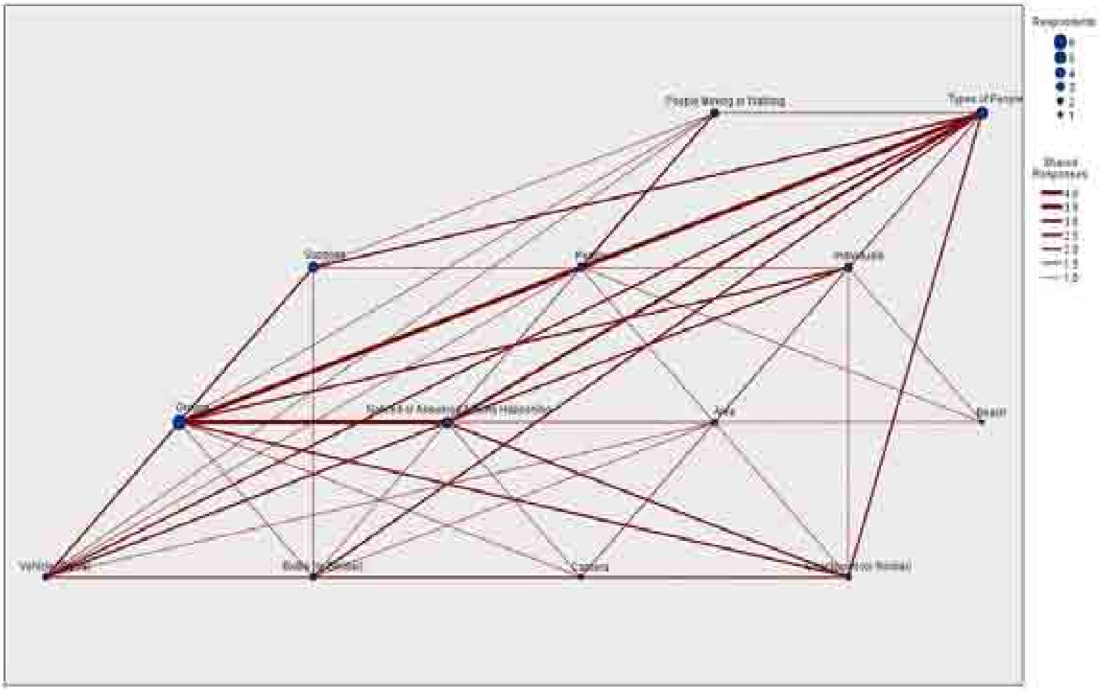

Fig 1.2 Trial 3

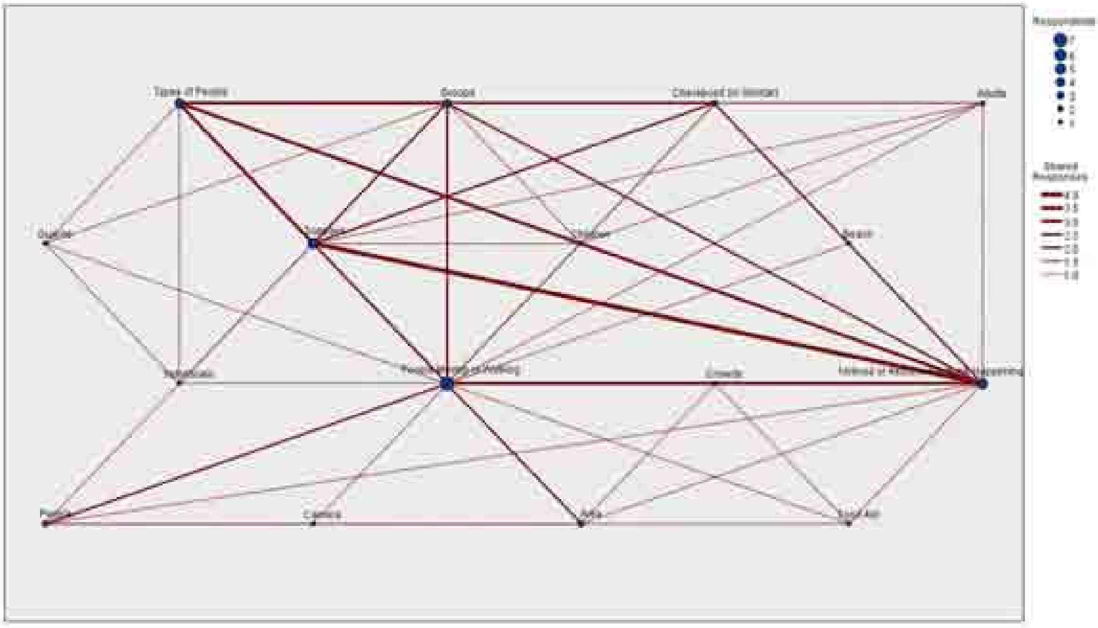

Fig. 1.3 Trial 4

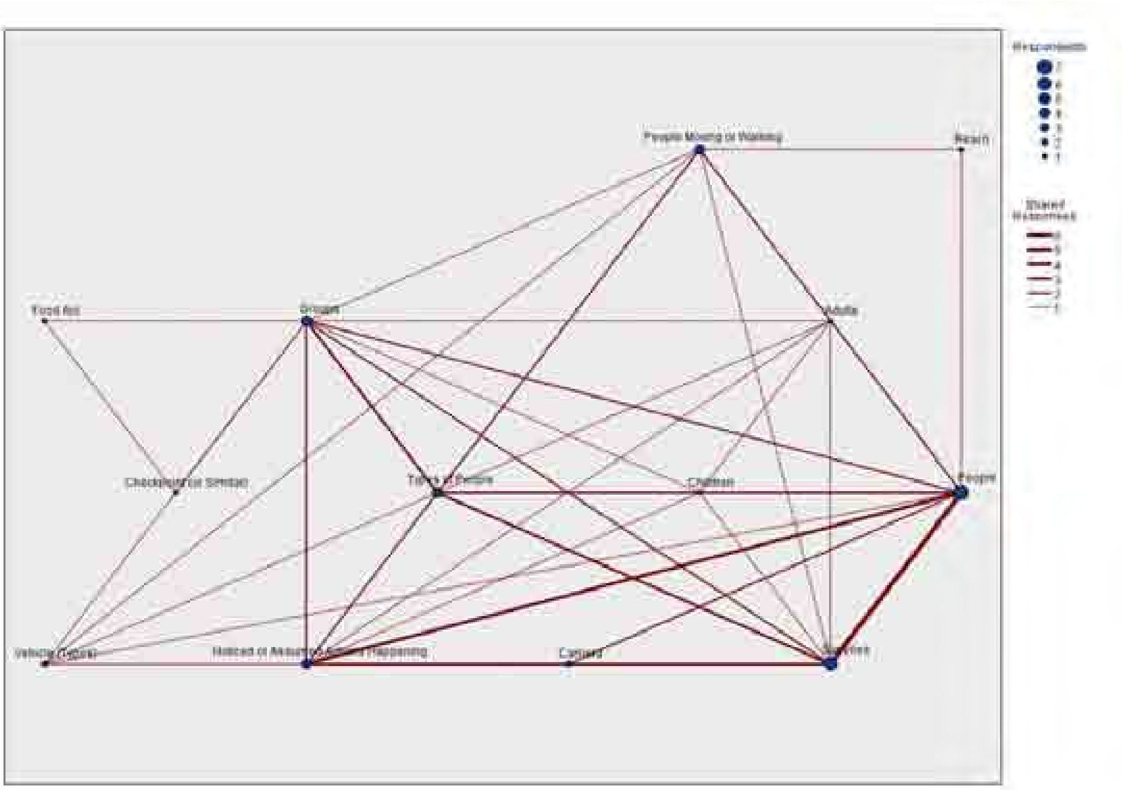

Fig. 1.4 Trial 5

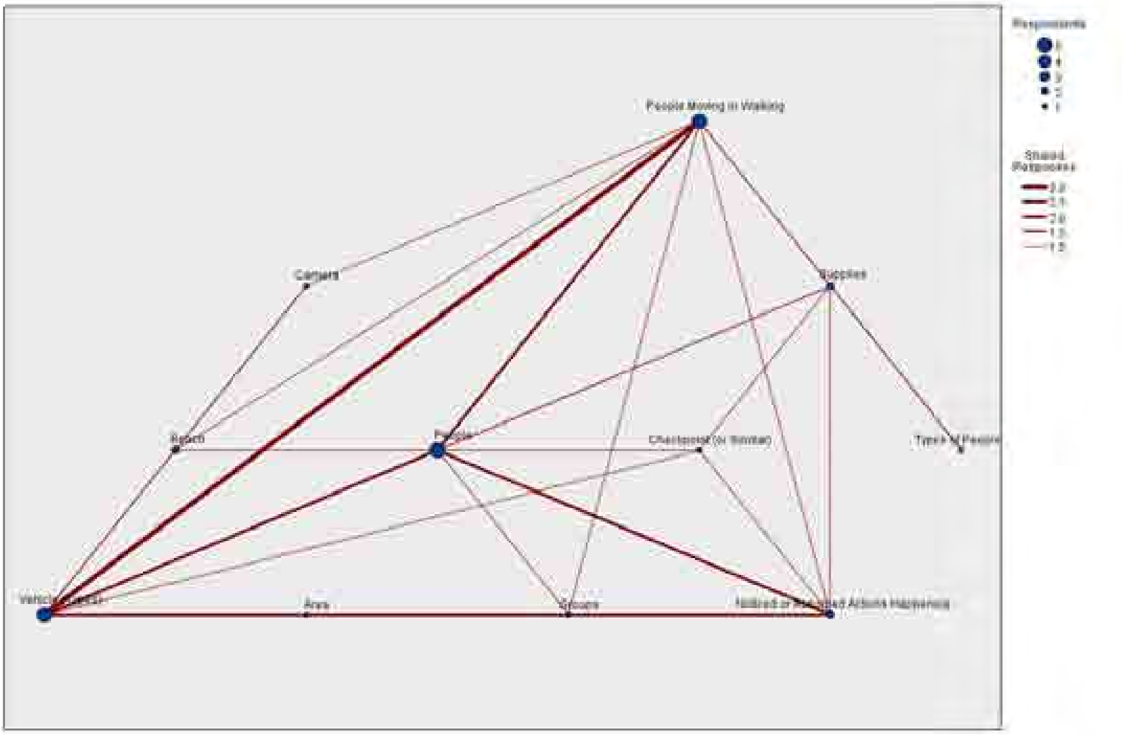

Fig 1.5 Trial 6

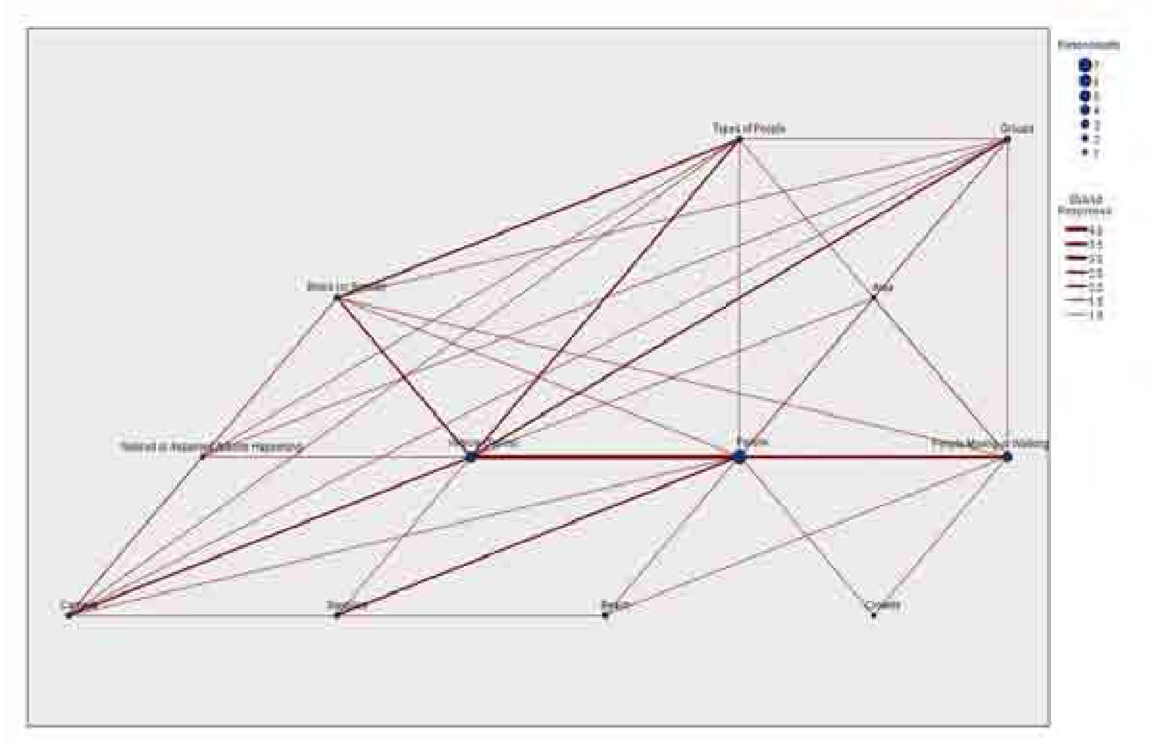

Fig. 1.6 Trial 7

2.0 Text Analysis Naïve Observers (Gate & Sentry)

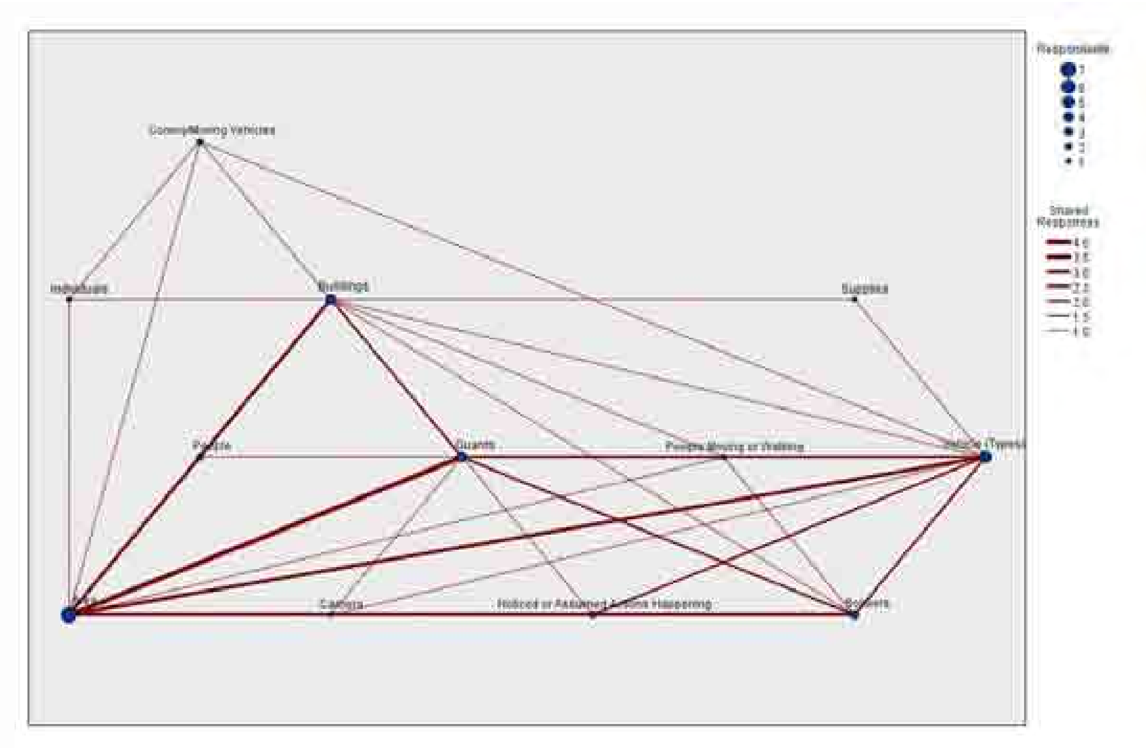

Fig. 2.1 Trial 2

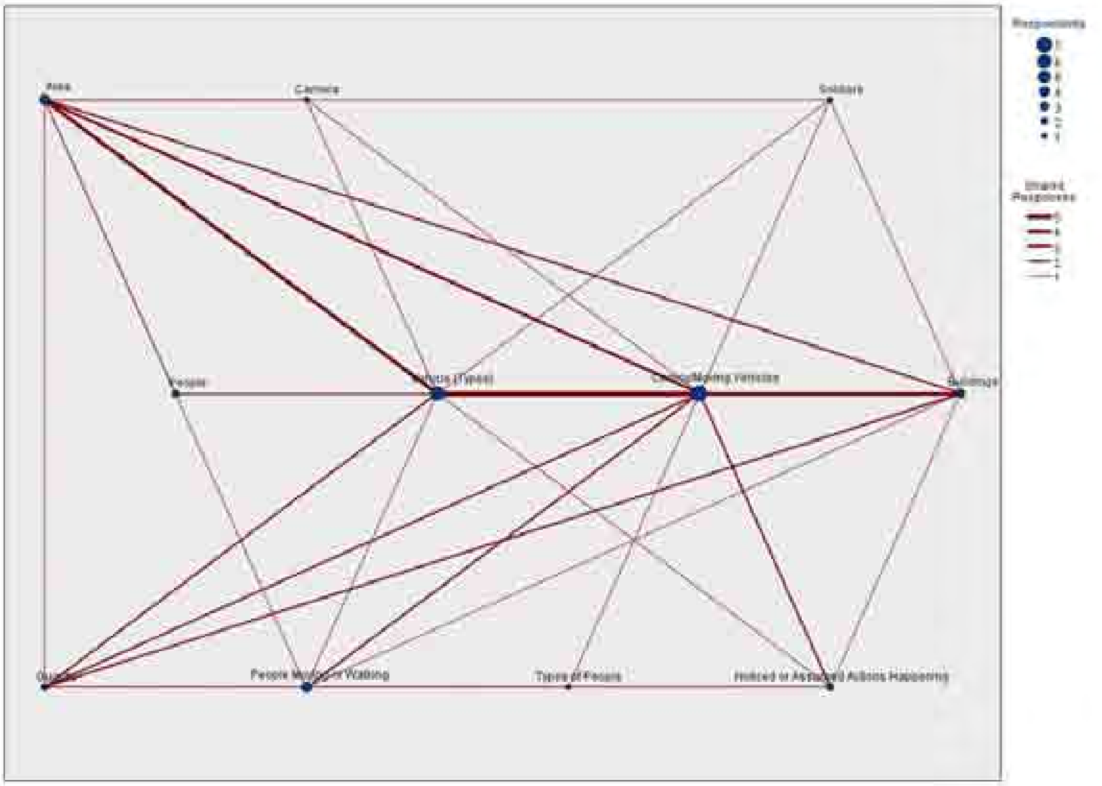

Fig. 2.2 Trial 3

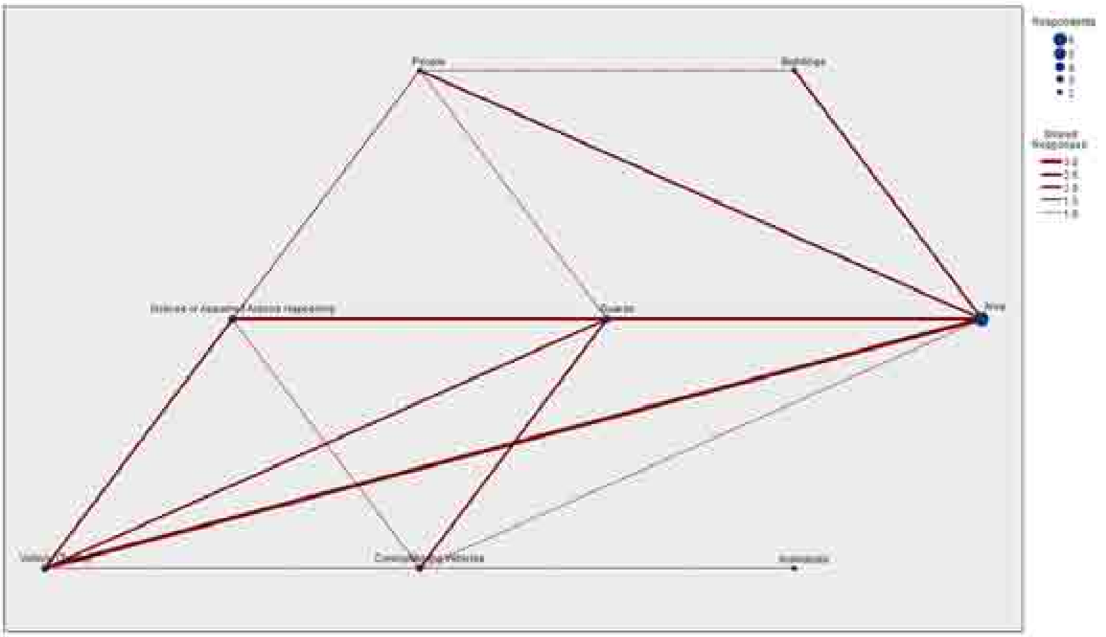

Fig. 2.3 Trial 4

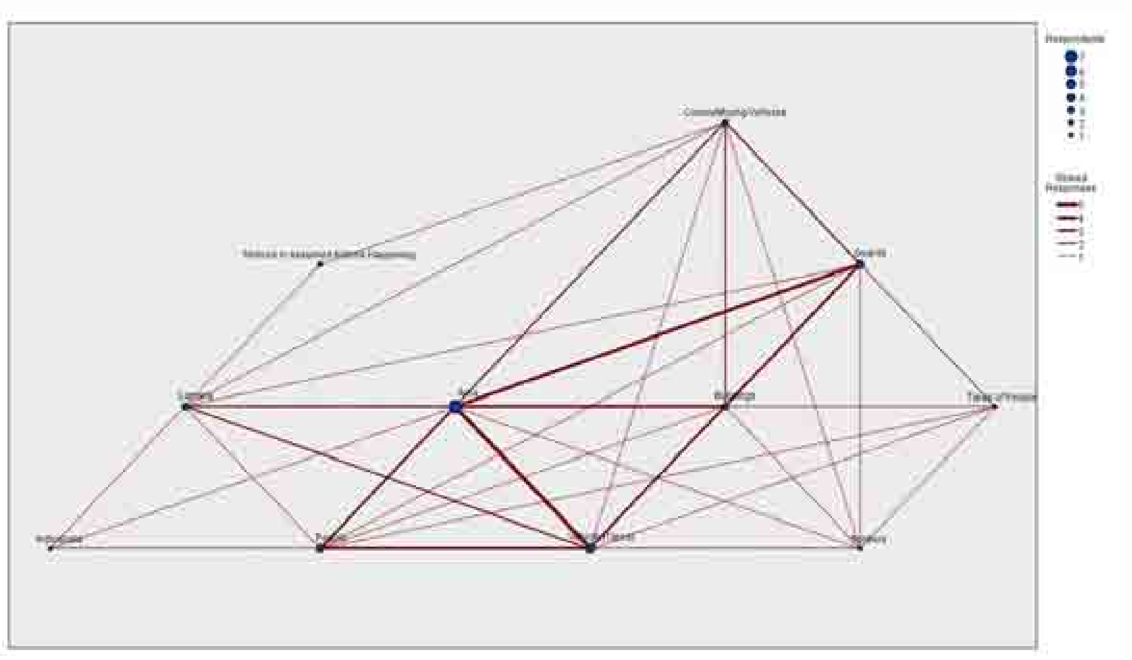

Fig. 2.4 Trial 5

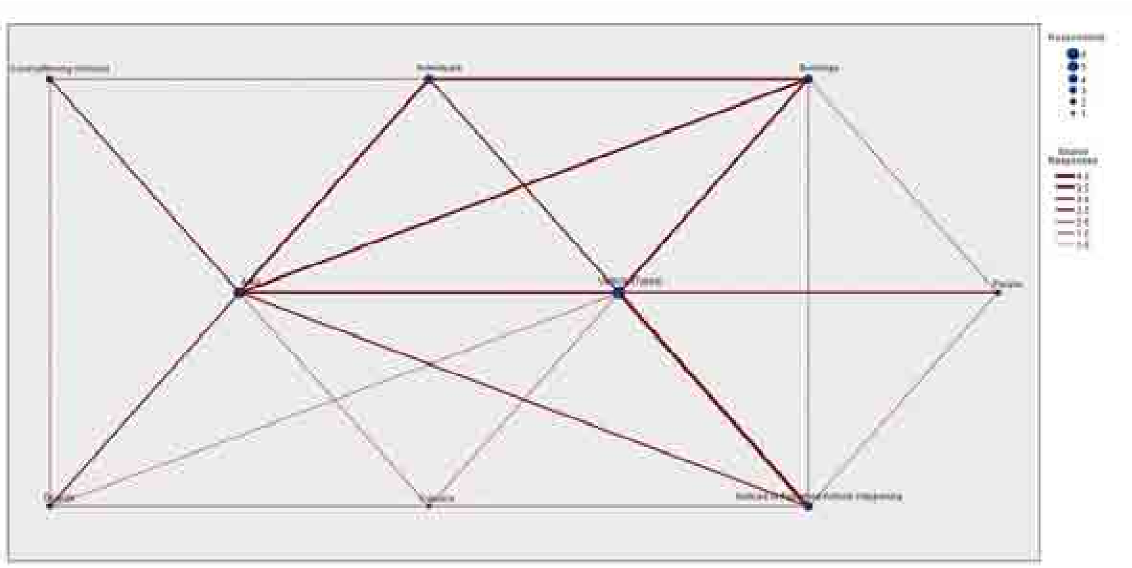

Fig. 2.5 Trial 6

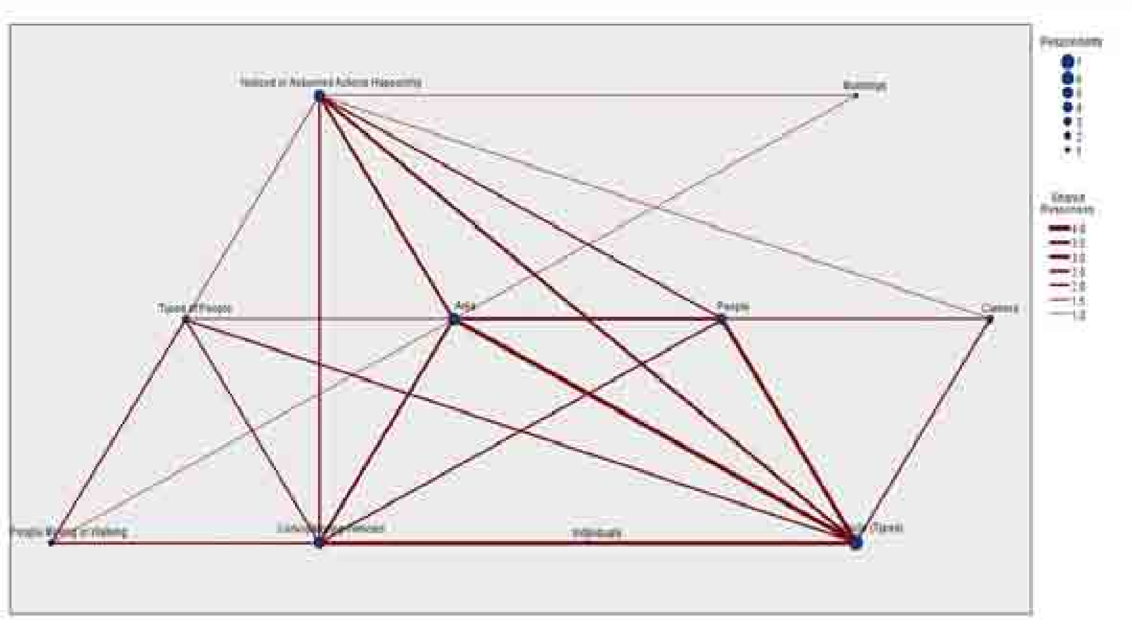

Fig. 2.6 Trial 7

3.0 Text Analysis Naïve Observers (Helicopter Pad)

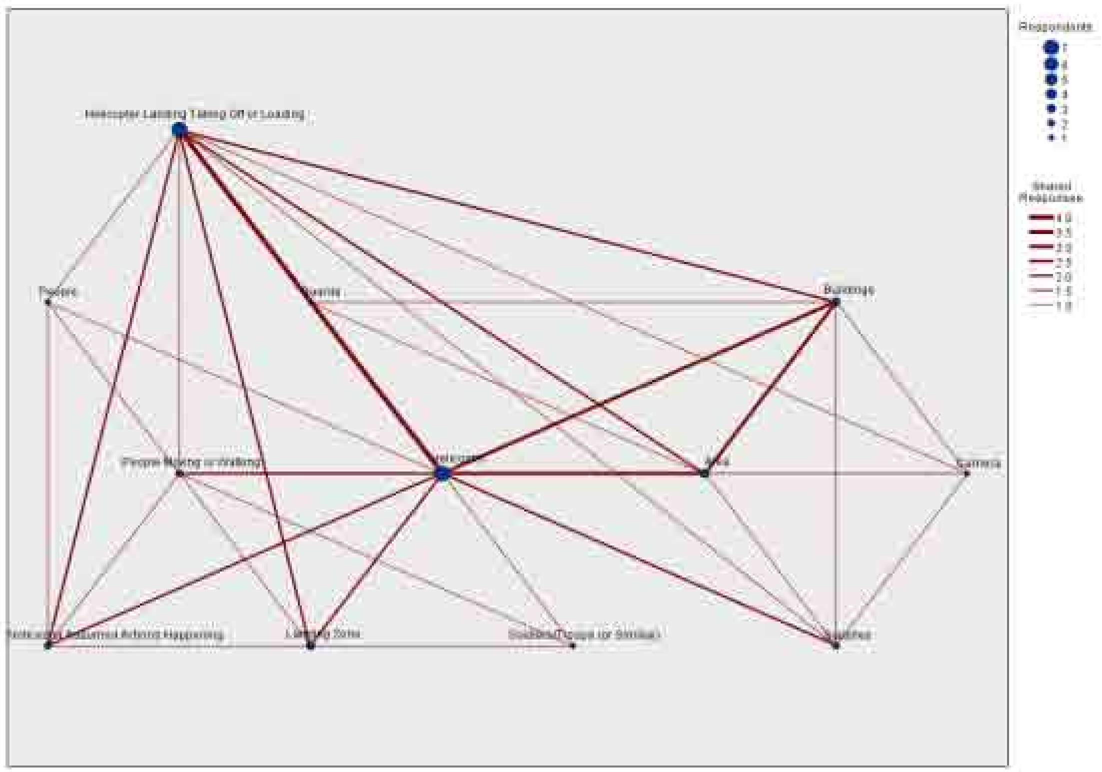

Fig. 3.1 Trial 2

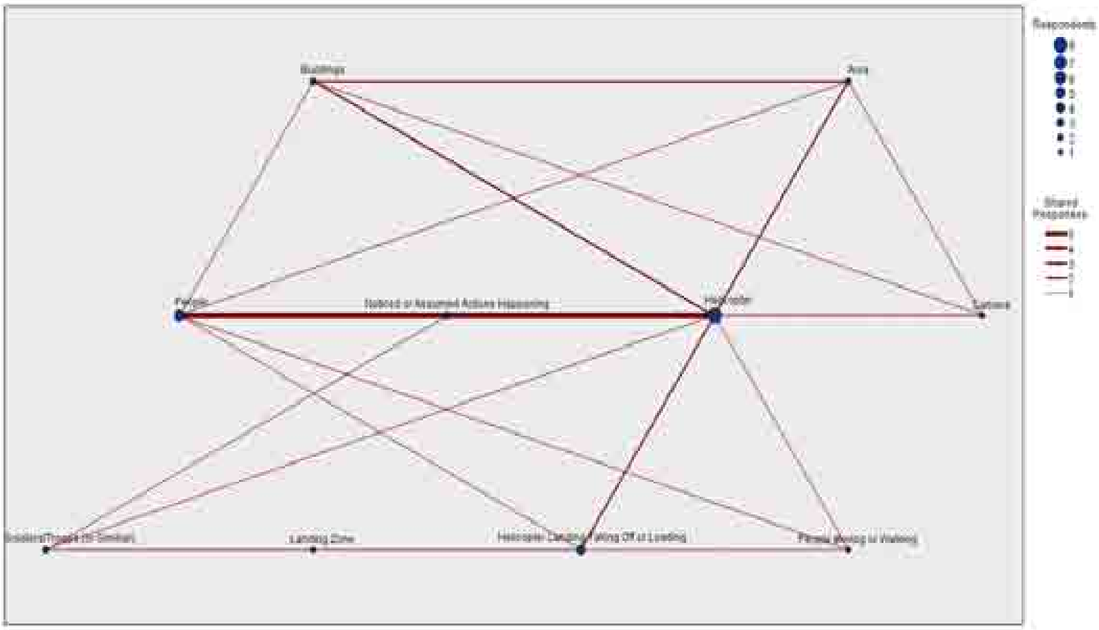

Fig. 3.2 Trial 3

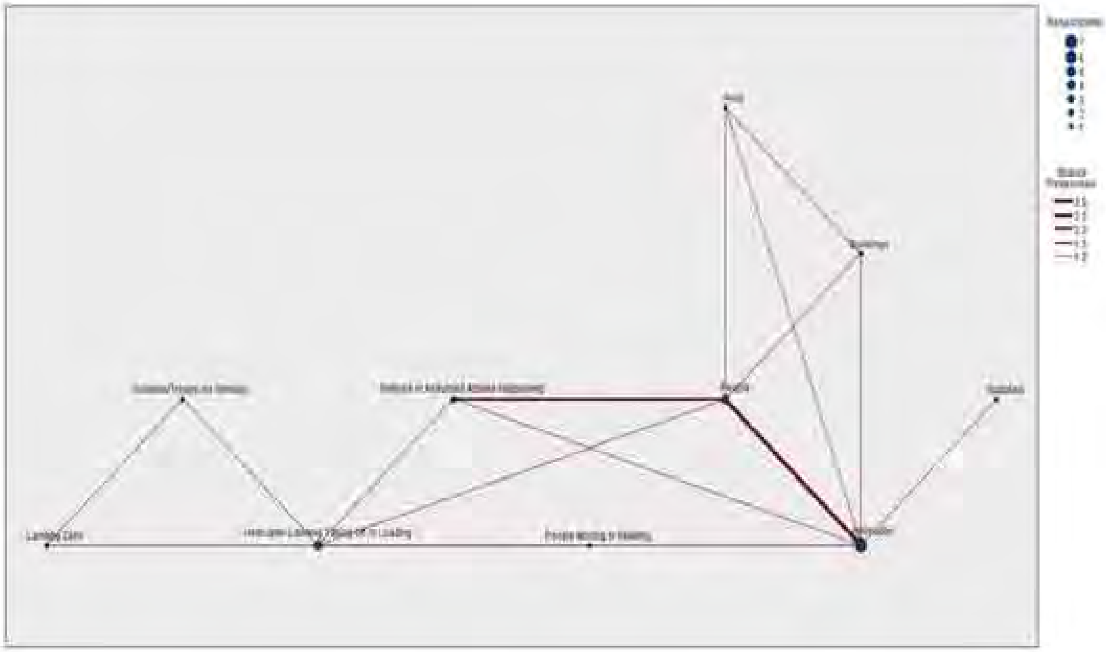

Fig. 3.3 Trial 4

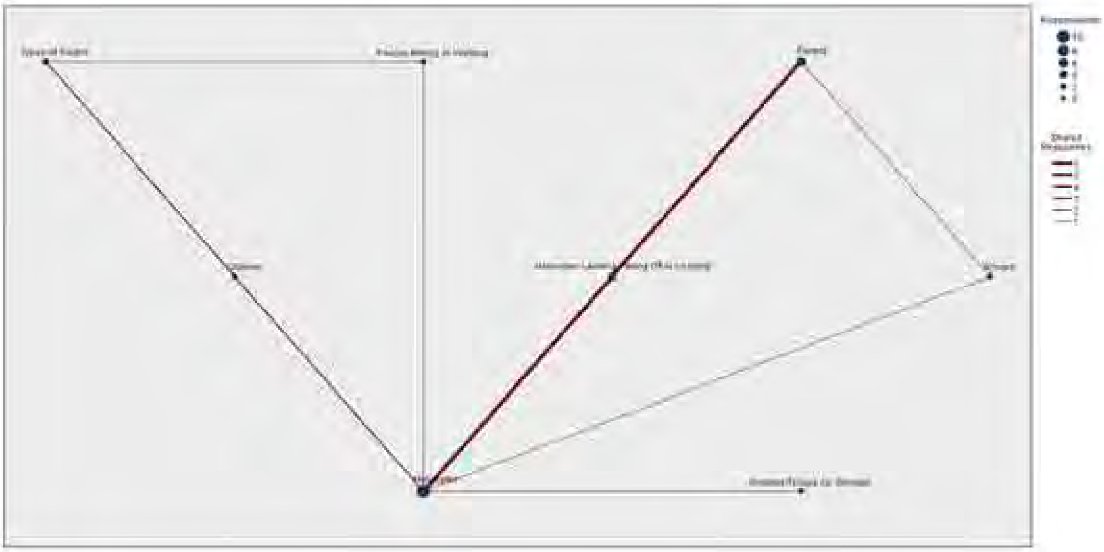

Fig. 3.4 Trial 5

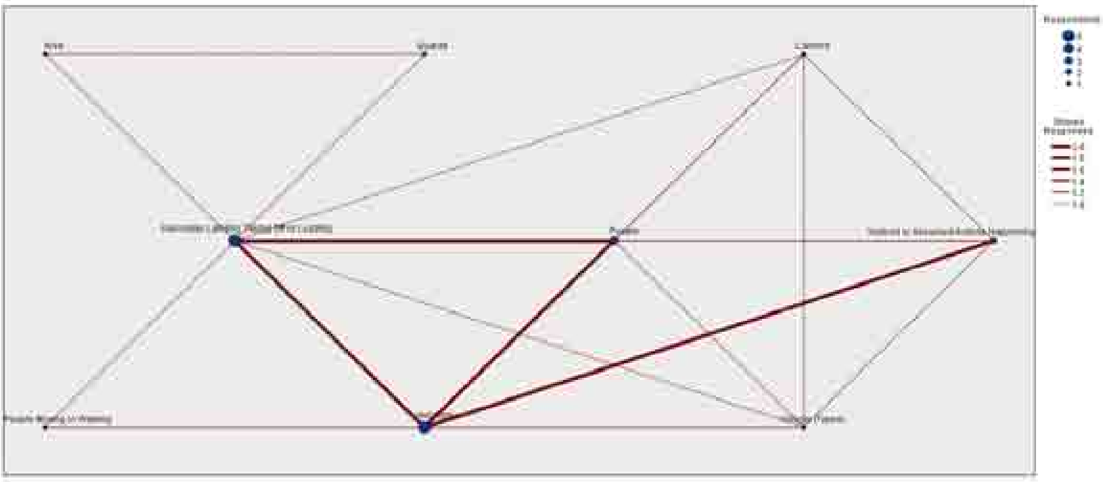

Fig. 3.5 Trial 6

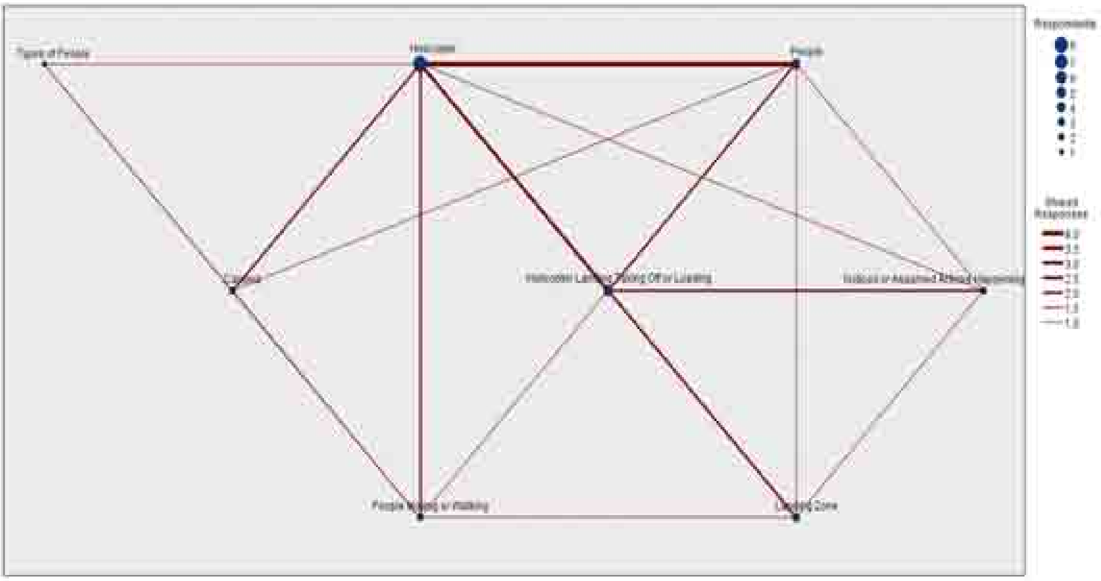

Fig. 3.6 Trial 7

4.0 Text Analysis Expert Observers (Food Aid)

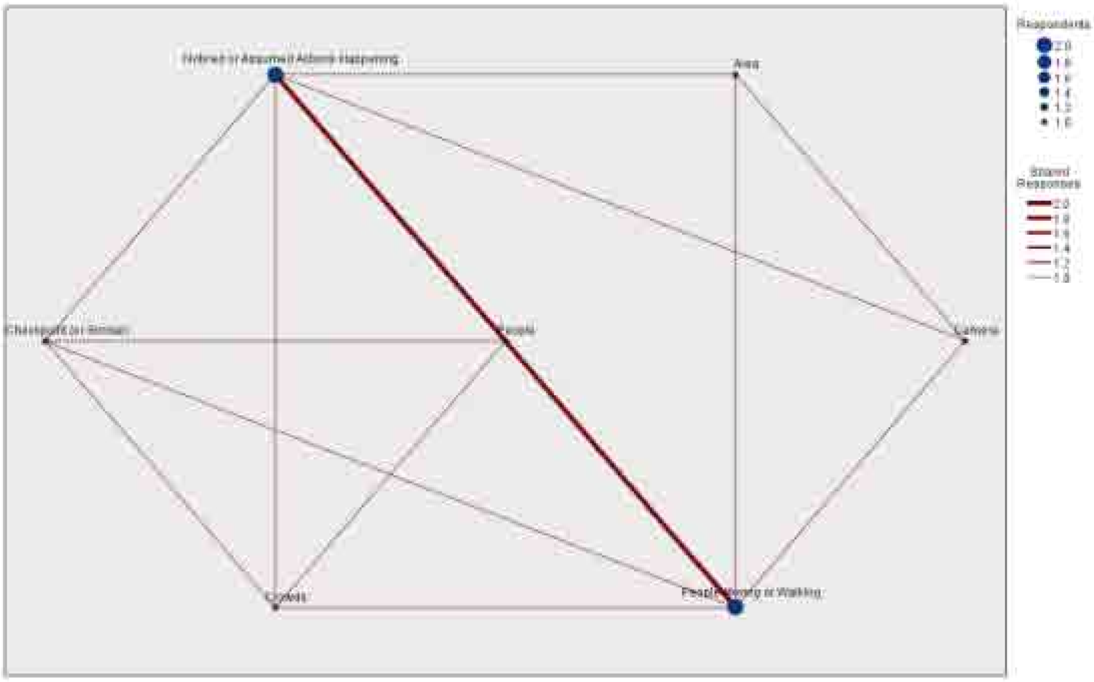

Fig. 4.1 Trial 2

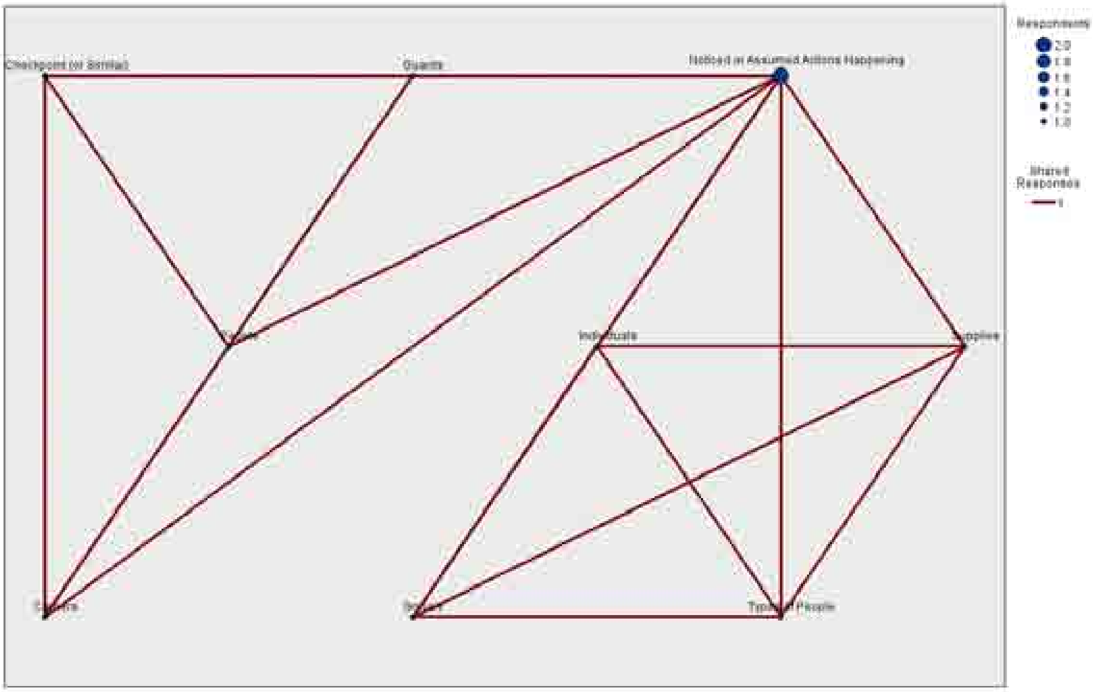

Fig. 4.2 Trial 3

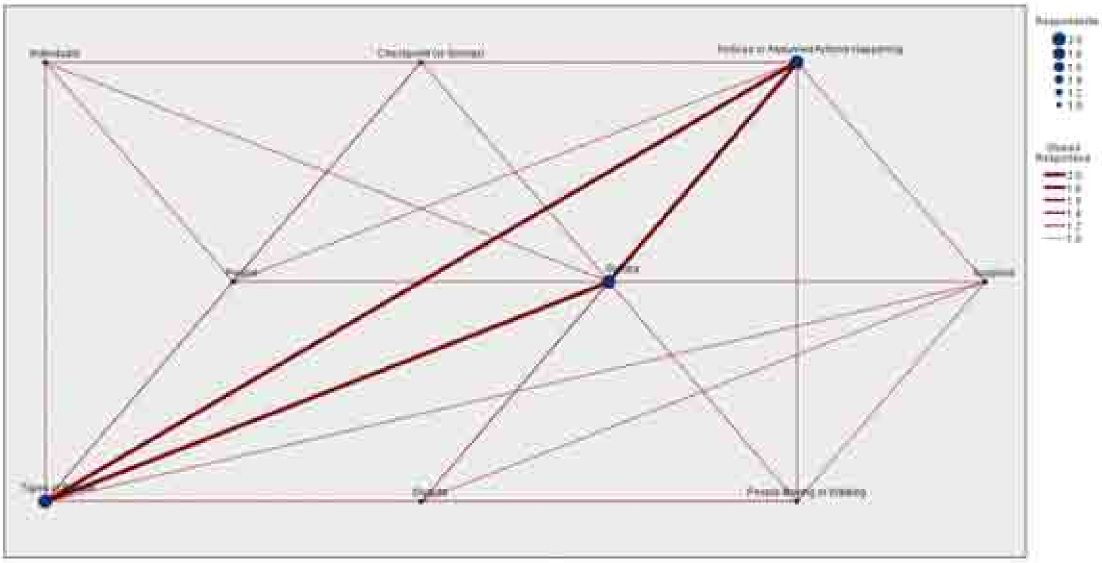

Fig. 4.3 Trial 4

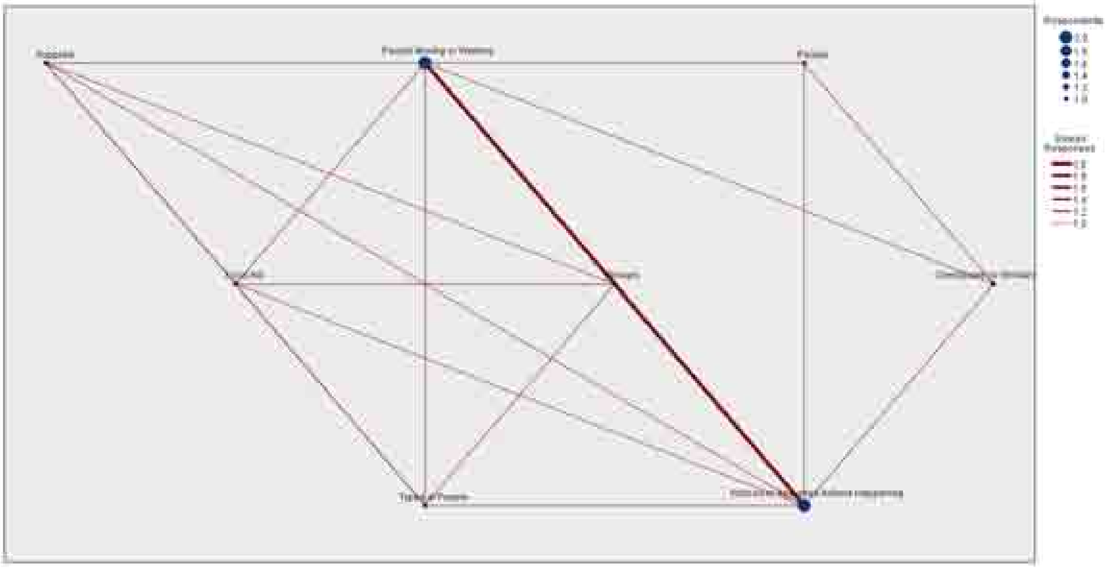

Fig 4.4 Trial 5

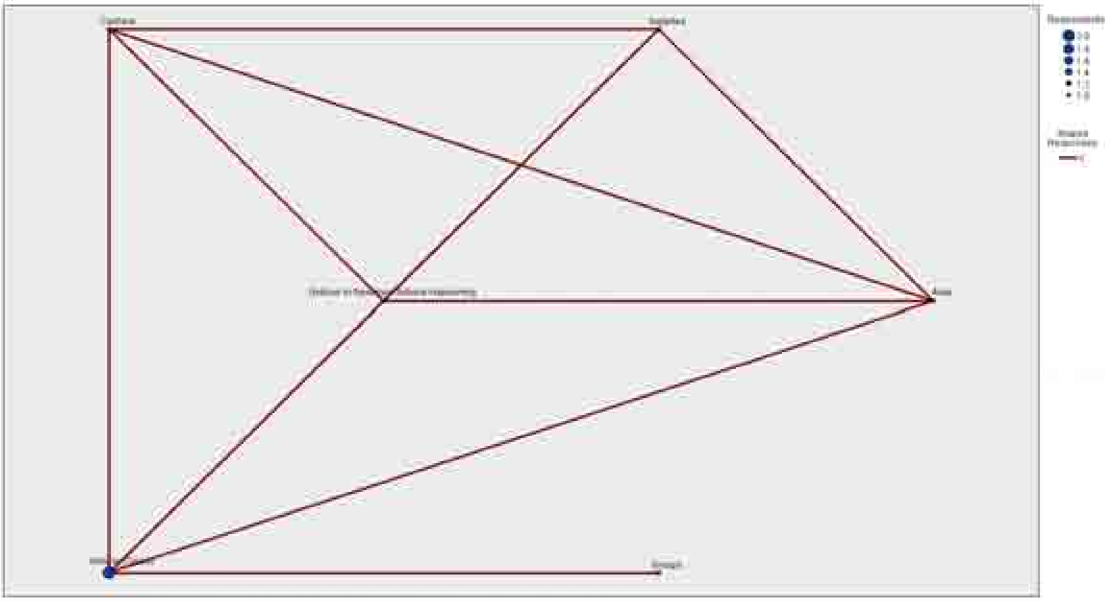

Fig. 4.5 Trial 6

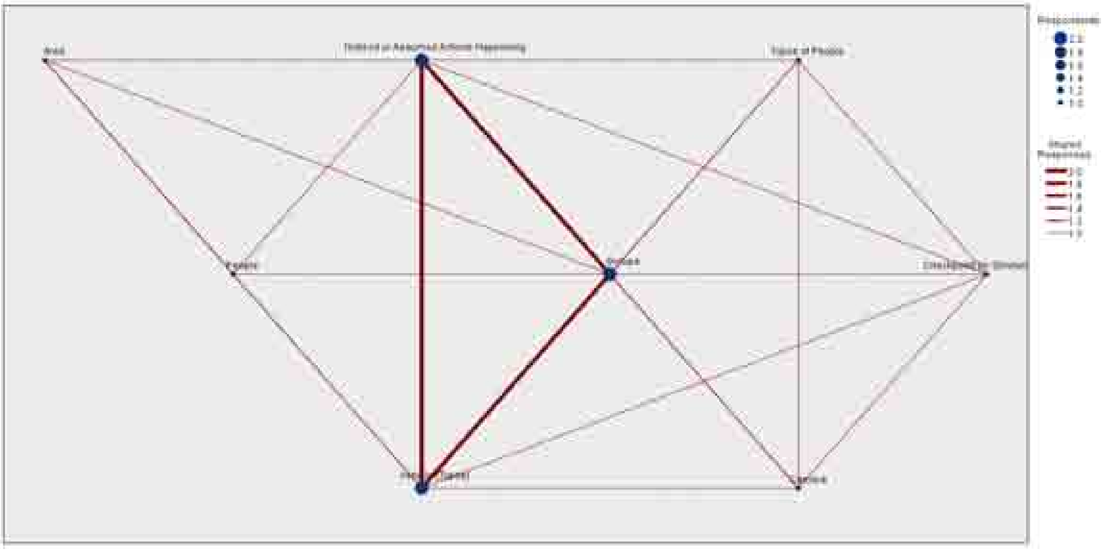

Fig. 4.6 Trial 7

5.0 Text Analysis Expert Observers (Gate & Sentry)

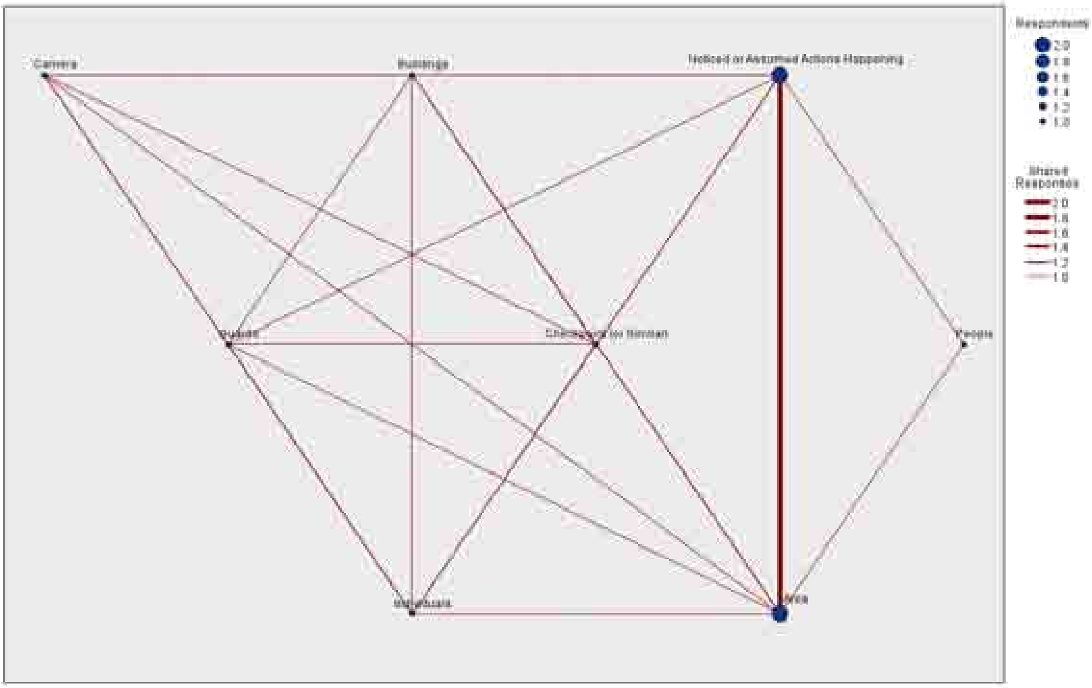

Fig. 5.1 Trial 2

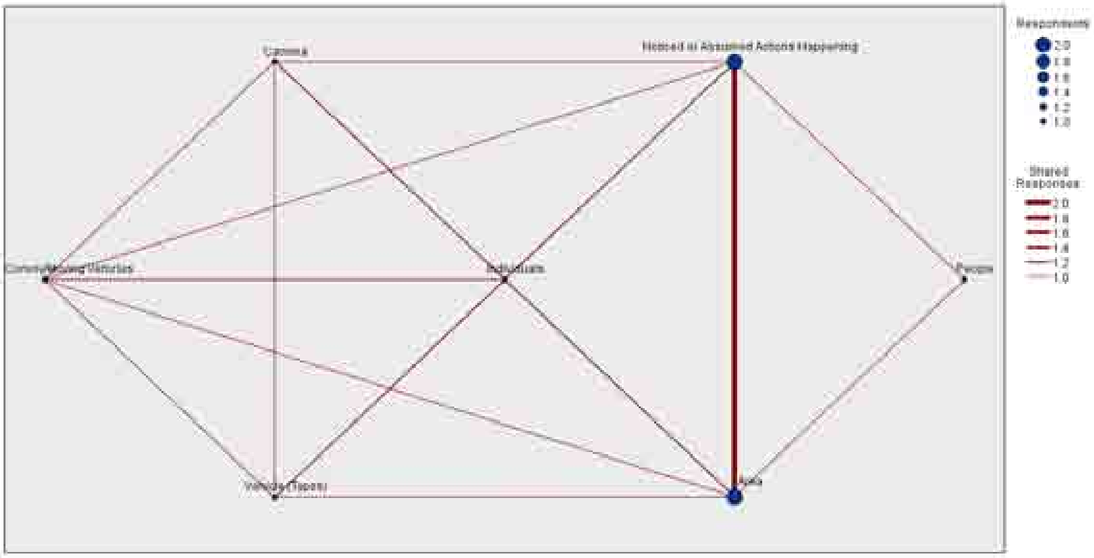

Fig. 5.2 Trial 3

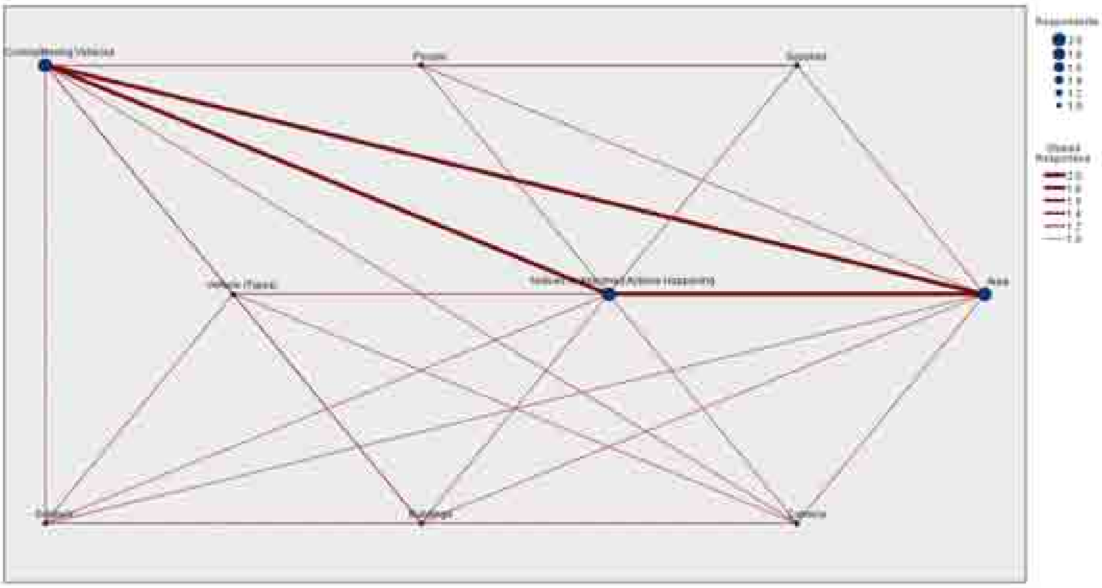

Fig. 5.3 Trial 4

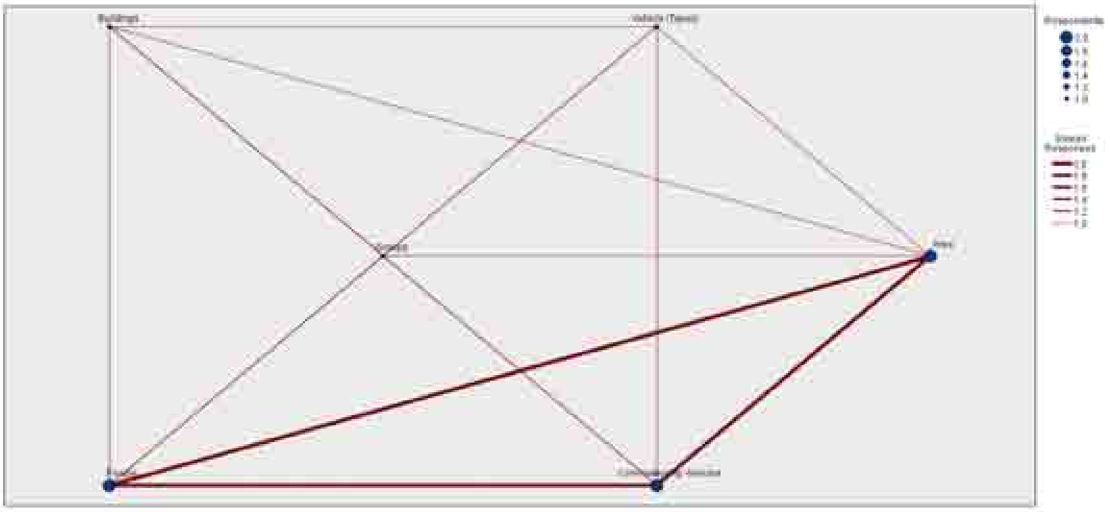

Fig. 5.4 Trial 5

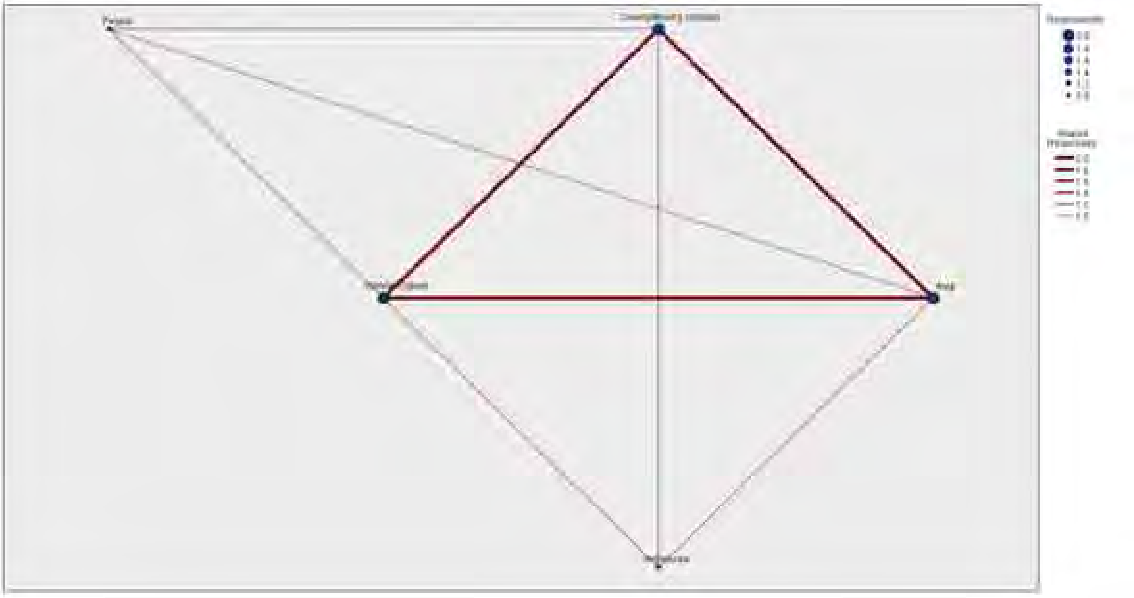

Fig. 5.5 Trial 6

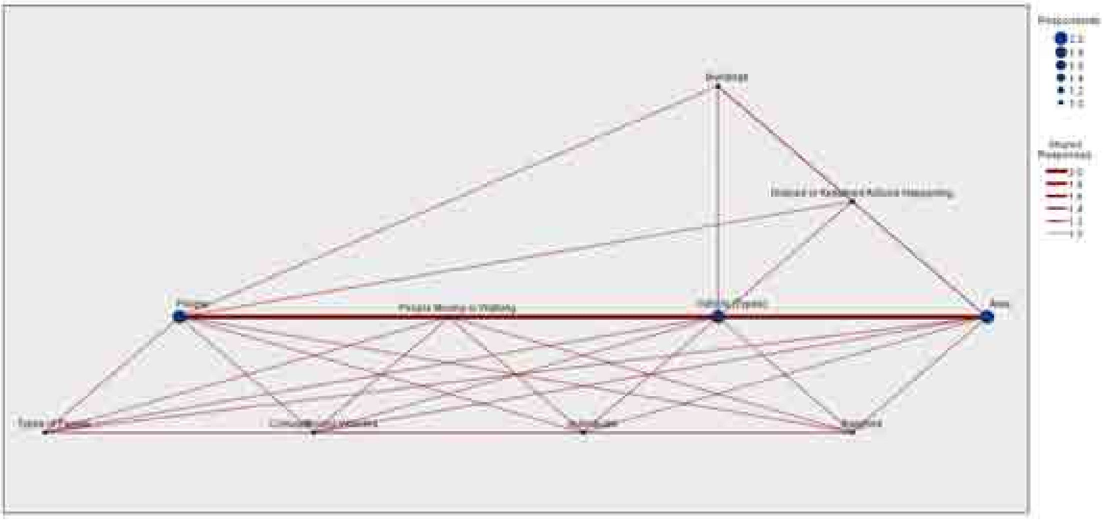

Fig. 5.6 Trial 7

6.0 Text Analysis Expert Observers (Helicopter Pad)

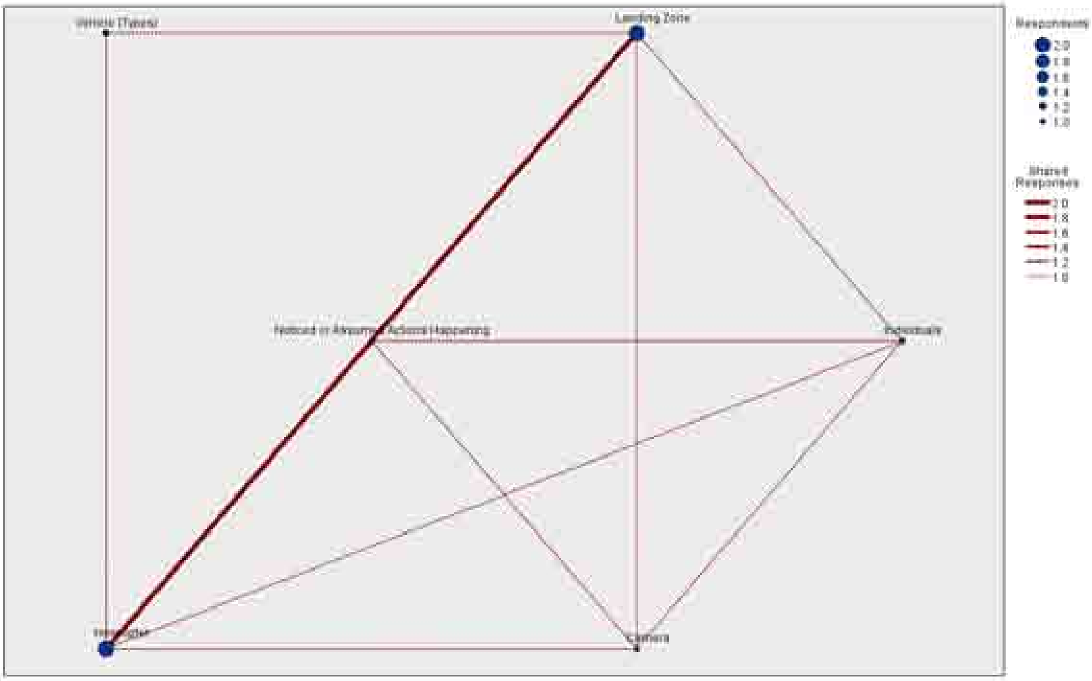

Fig. 6.1 Trial 2

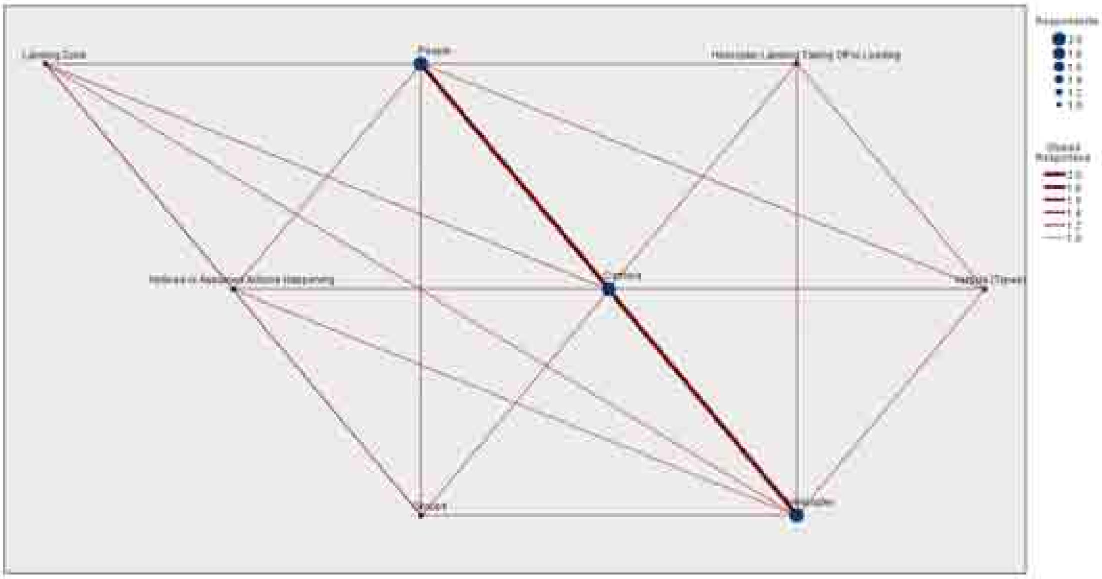

Fig. 6.2 Trial 3

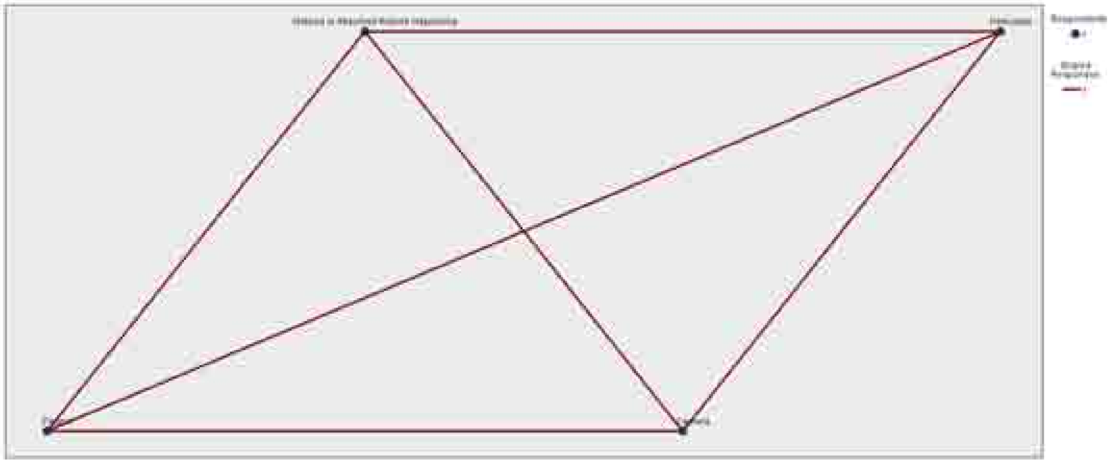

Fig. 6.3 Trial 4

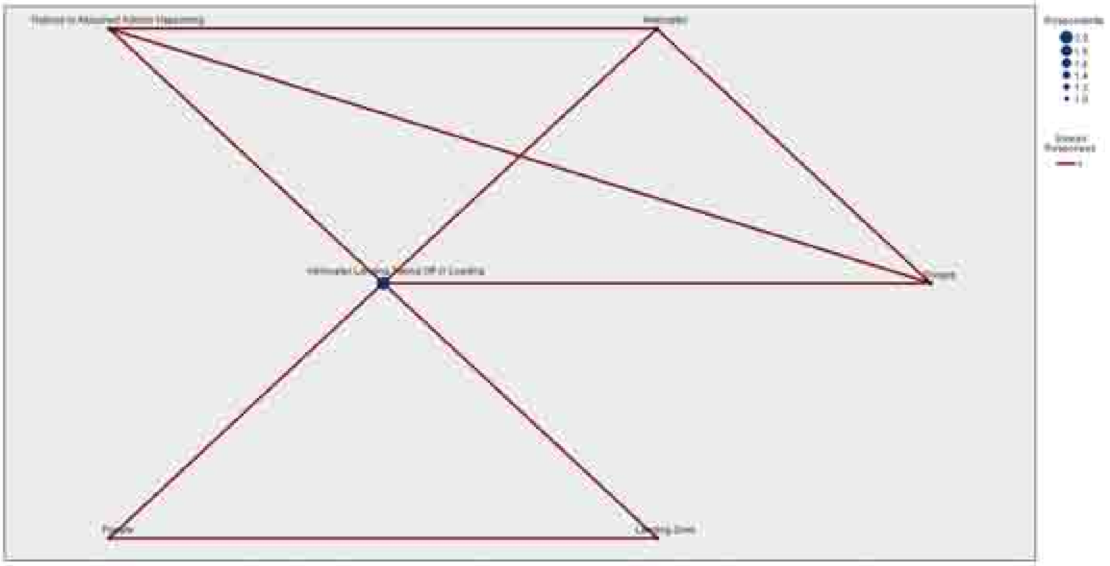

Fig. 6.4 Trial 5

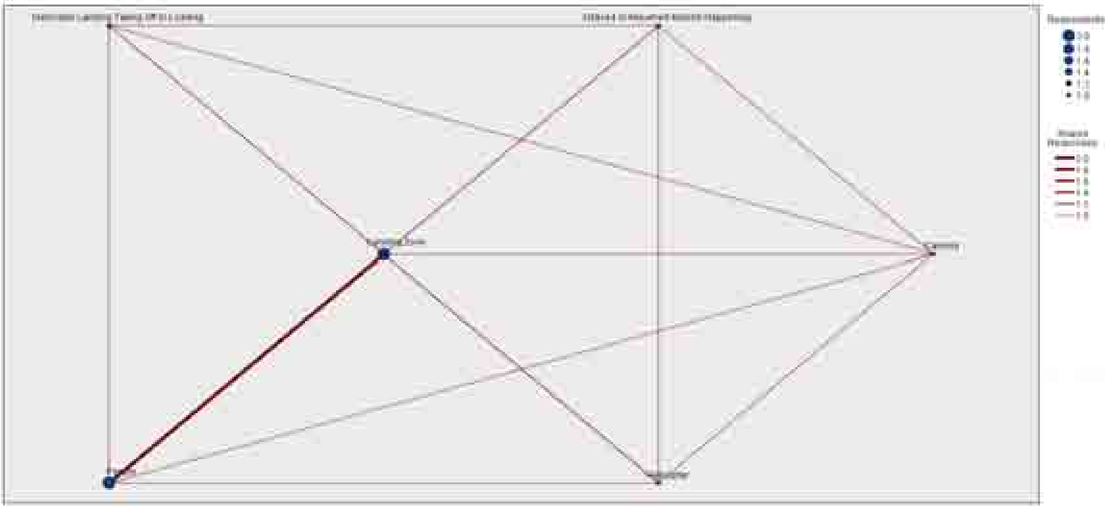

Fig. 6.5 Trial 6

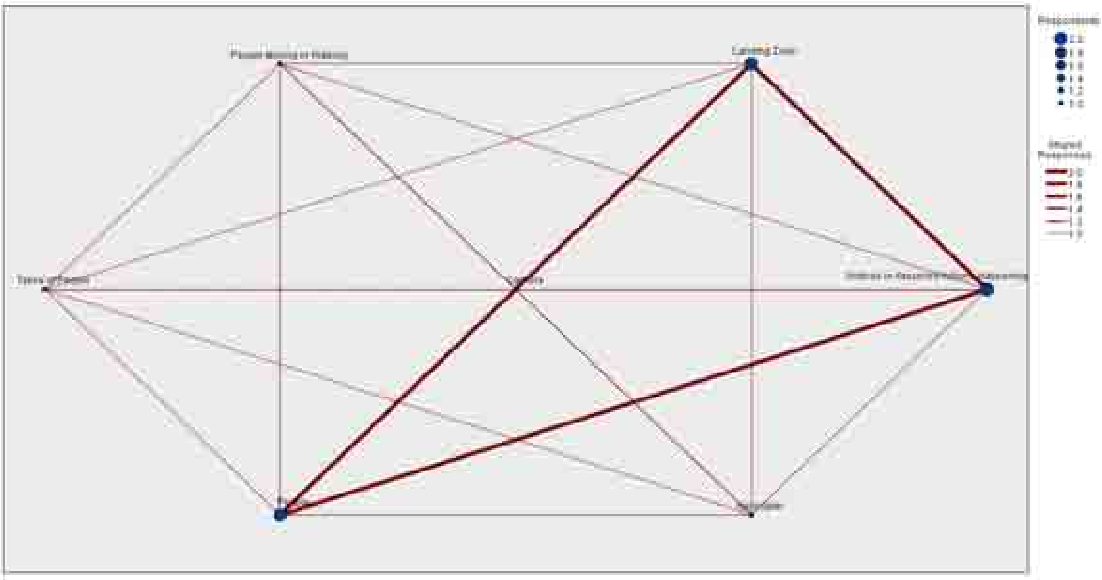

Fig. 6.6 Trial 7