CRANFIELD UNIVERSITY

SOPHIE CARR

INVESTIGATING THE APPLICABILITY OF BAYESIAN NETWORKS TO THE ANALYSIS OF MILITARY INTELLIGENCE

DEFENCE COLLEGE OF MANAGEMENT AND TECHNOLOGY

PhD THESIS

CRANFIELD UNIVERSITY

DEFENCE COLLEGE OF MANAGEMENT AND TECHNOLOGY

ENGINEERING SYSTEMS DEPARTMENT

PhD THESIS

Academic Year 2007-2008

Sophie Carr

Investigating the applicability of Bayesian networks to the analysis of military intelligence

Supervisor: Dr K McNaught

January 2008

ABSTRACT

Intelligence failures have been attributed to an inability to correlate many small pieces of data into a larger picture. This thesis has sought to investigate how the fusion and analysis of uncertain or incomplete data through the use of Bayesian Belief Networks (BBN) compares with people's intuitive judgements. These flexible, robust, graphical probabilistic networks are able to incorporate values from a wide range of sources including empirical values, experimental data and subjective values. Using the latter, elicited from a number of serving military officers, BBNs provide a logical framework to combine each individual's set of one-at-a-time judgements, allowing comparisons with the same individuals' many-at-a-time, direct intuitive judgements. This was achieved through a series of fictitious and historical case studies.

Building upon this work, another area of interest was the extent to which different elicitation techniques lead to equivalent or differing judgements. The techniques compared were: direct ranking of the variables' perceived importance for discriminating between given hypotheses, likelihood ratios and conditional probabilities. The experimental results showed that individuals were unable to correctly manipulate the dependencies between information as evidence accumulated. The results also showed varying beliefs about the importance of information depending upon the elicitation technique used. Little evidence was found of a high correlation between direct normative rankings of variables' importance and those obtained from the BBNs' combination of one-at-a-time judgements. Likelihood values should only be used as an elicitation technique by those who either regularly manipulate uncertain information or use ratios. Overall, conditional probability distributions provided the least troublesome elicitation technique of subjective preferences.

In conclusion, Bayesian Belief Networks developed through the use of subjective probability distributions offer a flexible, robust methodology for the development of a normative model for the basis of a decision support system for the quantitative analysis of intelligence data.

ACKNOWLEDGEMENTS

It is said that every journey starts with a single step. Although the research into this thesis has inevitably, had its ups and downs I can genuinely say that I am pleased I took the first and all subsequent steps. Throughout I have been very privileged to have received guidance and support from so many people, to whom I all owe a debt of thanks.

Firstly to my employers who showed great faith in me by supporting this work. I am fortunate enough to work in an environment which encourages research and provides the flexibility within a job to follow through from an initial spark of and idea. Of all the people I have worked with and for, specific thanks must be given to Professor Jim Moffat for his succinct and always constructive comments on my work.

A great deal of thanks must also go to my supervisor Dr Ken McNaught. Over the years he has continually provided the perfect balance of academic and moral guidance, support and encouragement. Often, I am sure he exercised great patience with me, for which I am extremely grateful. His infectious passion for research and genuinely helpful review sessions made this course of research a pleasure to undertake.

To my family and friends thanks must go for their seemingly never ending belief in me and for those days when they provided havens of normality from probabilities and thesis writing. Finally, to my husband and son thank-you for encouraging me to indulge in this research and spend seemingly endless hours reading and writing. Without the overwhelming and never ending supply of coffee, hugs and laughs you provided this thesis would not be what it is.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

GLOSSARY

H	Entropy
I	Mutual information
MaxMI	Maximum value of mutual information
MI	Mutual information
MinMI	Minimum values of mutual information
$P(x_i)$	Probability of outcome x
$u(x_i)$	Subjective utility of outcome x
W	Weights
?	Kendall's Tau
ACH	Analysis Competing Hypotheses
AD	Air Defence
AHP	Analytical Hierarchy Process
App	Apparent
ATK	Anti Tank Weapons
BBN	Bayesian Belief Network
BI	Business Intelligence
Cap	Capability
CIA	Central Intelligence Agency
CPM	Corporate Performance Management
DAG	Directed Acyclic Graph
DCI	Director of Central Intelligence
DCMT	Defence College of Management and Technology
DRSA	Dominance Based Rough set Approach
DSS	Decision Support System
EIS	Executive Information Systems
ER	Evidential Reasoning
FBI	Federal Bureau of Investigation
GB	Gigabyte
HUMINT	Human Intelligence
IMINT	Imagery Intelligence
ISR	Intelligence Surveillance and Reconnaissance

MAUT	Multi Attribute Utility Theory
MCDA	Multi Criteria Decision Analyis
MI5	UK security service which focuses solely upon domestic security.
MIS	Management Information Systems
NDM	Naturalistic Decision Making
ORS	Operation Research Section
PIP	Probabilistic Information Processing
Rep	Report
Recce	Reconnaissance
RPD	Recognition Primed Decision-Making
SC	Social Choice
SEU	Subjective Expected Utility
SIGINT	Signals Intelligence
SMART	Simple Multi Attribute Rating Technique
SNA	Social Network Analysis
SWOT	Strengths Weaknesses Opportunities and Threats
UAVs	Unmanned Air Vehicles
USA	United States of America
UK	United Kingdom
WMD	Weapons of Mass Destruction

CHAPTER 1 INTRODUCTION.

1.1 Background.

Since the attacks on September 11th 2001, it has become public knowledge that intelligence failures happen and when they do so the consequences can be catastrophic. Nevertheless, 'intelligence failure' is a phrase which must be used with some caution. Undoubtedly history is littered with examples of military intelligence failures. However, rarely is it the lack of information which prevented the unfolding events from being identified by intelligence analysts and government departments. All too often the signs of an impending event were picked up. As such, there was no 'failure' to gather the intelligence. What occurred was an inability to piece the information together; to correlate the many small data pieces into a larger picture from which direct action could have been taken. Indeed, consider the consequence of the Germans correctly interpreting all the information they had collated on the D-Day landings. In some part, the misinterpretation of information can, on this occasion, be put down to information passed on by double agents. With hindsight, the reason for misinterpreting information can always be found: an insufficient intelligence service (one of the main reasons for the failure to detect the impending attack on Pearl Harbour), underestimating enemy strength (the fall of Singapore in 1942) and simple failure to react to the known threat (Yom Kippur war). For a detailed account of these and other intelligence failures the reader is referred to Hughes-Wilson (2004).

Dame Eliza Manningham-Buller, Director General of the UK Security Service, beautifully illustrated why intelligence can rarely provide all the desired information: "Often difficult decisions need to be made on the basis of intelligence which is fragmentary and difficult to interpret. In sum, some is gold, some dross and all of it requires validation, analysis and assessment. When it is gold, it shines and illuminates, saves lives, protects nations and informs policy. When identified as dross, it needs to be rejected: that can take some confidence. At the end of the day, it requires people of integrity not only to collect it but also to prioritise, sift, judge and use it." (Dame Eliza Manningham-Buller in a speech to the Dutch Security Services at

Ridderzaal, Binnenhoff, The Hague, Netherlands, 1 September 2005, cited in Report into the Terrorist Attacks in London 7 July 2005, page 7).

1.2 Why do we collect intelligence?

In 'The Art of War', Sun Tzu makes a valuable observation about the importance of intelligence: "If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle." (Tzu, 1995, page 26).

The sentiment of the writing which is over 500 years old still holds relevance today. The collation, analysis and communication of intelligence is undertaken as an integral part of risk management. Commercial and military organisations utilise intelligence in an attempt to gain an advantage over their adversaries and support strategic decision-making. Based on the intelligence reports available, what enemy weakness can be exploited in order to maximise the desired gain or output? Possibly of more importance, however, is the use of intelligence reports to show where an organisation should be developing contingency plans against potential attack or loss. However, such insights come at a price. It is vitally important not to underestimate the financial, resourcing costs associated with the long term gathering of intelligence. However, the potential gains from correctly interpreting a situation and thus being the first to react can be equally substantial. It is difficult, and often painful, to place a cost on a missed opportunity.

Military intelligence is gathered from two main sources: open and covert. Within these the resources available can be further sub-divided into those provided directly by an individual (human intelligence) and those provided electronically (signals intelligence). Open source intelligence relates to information gathered from publicly available sources such as the world's media, academia and released government reports. Conversely, covert, or secret, intelligence is gathered without the knowledge of the person or organisation upon which the information is being collated. For both open and covert sources, the information gathered may be incomplete, hard to verify and have political or security implications attached to them. These latter two issues will dictate who the information can be disseminated to. Ultimately, who has access to the information dictates who will be able to make a decision based on the information. It is important to remember that intelligence reports are simply mere data until they are acted upon.

1.3 What happens when intelligence fails?

The most recent intelligence 'failures' which have been reported on the international level are: the attack on the World Trade Centre in 2001; the search for weapons of mass destruction (WMD) in Iraq; and the attacks in London on July 2005. The search for WMD in Iraq could be considered a traditional military intelligence operation. Conversely, the attacks on New York and London were asymmetric attacks carried out by people living amongst the civilian population. Predicting attacks by conventional forces is highly complex even when it is considered that the enemy mindset and general overall abilities are known. Therefore, predicting and preventing asymmetric attacks by groups of people whose overall access to technology and military training is uncertain adds additional levels of complexity. The ability to predict attacks in both cases will be based upon information which may be: fragmentary, incomplete, provided from diverse sources and presented over a sustained timeline. How can intelligence analysts be helped to improve their intelligence estimates across such a diverse area?

The intelligence cycle comprises five parts (Hughes-Wilson, 2004). Firstly the intelligence requirement is stated. Following this there is the collection, collation, interpretation and dissemination of the information. This thesis focuses on the analysis and subsequent interpretation of any collated information.

One of the first major reports on recent intelligence failures was provided by The National Commission on Terrorist Attacks Upon the United States (commonly referred to as the 9-11 Commission). The report presented the fact that intelligence suggesting an attack was becoming increasingly imminent had been collated. Between January and September 2001, the Presidential Daily Brief given to President

George W Bush contained 40 reports relating to Osama Bin Laden (9-11 Commission Report, 2004, page 254). The 9-11 report also comments that during the spring of 2001, the number of reports being received within the intelligence community relating to potential terrorist threats and attacks reached a peak not seen since the millennium. With hindsight, the Director of Central Intelligence (DCI) George Tenet is quoted in the 9-11 Commission Report (2004, page 259) as stating that 'the [intelligence] system was blinking red'.

The collation of the intelligence relating to the potential attacks of September 11th 2001 had been carried out by the fifteen intelligence agencies within the United States. The analysis of the available intelligence reports had been hampered by the fact that there was no way of 'pooling intelligence and using it to guide planning and assignment of responsibility for joint operations' (9-11 Commission Report, 2004, page 357). This weakness in the intelligence system meant information may not have been readily or easily accessible to personnel who would have had most need for it. Furthermore, within the intelligence community information is often distributed on a "need to know" basis. Indeed the 9-11 Commission noted that intelligence "agencies uphold a "need-to-know" culture of information protection rather than promoting a "need-to-share" culture of integration." (9-11 Commission Report, 2004, page 417). Analysts interviewed by the 9-11 commission reported frequent difficulties in accessing information on areas to which they were assigned to analyse. To some extent the answer to several of the above issues could have been developed during the 1990s. In 1998 the Federal Bureau of Investigation (FBI) devised a plan for a 'nationwide automated system to facilitate information collation, analysis and dissemination." (9-11 Commission Report, 2004, page 76) However, for numerous reasons including underinvestment and lack of human resources the plan did not succeed.

Overall, the 9-11 Commission report made a series of wide-ranging recommendations relating to changes to the intelligence community (many of the recommendations were passed in a USA Congress bill in December 2004) which lead to the largest restructuring of the community in nearly 60 years. However, the 9-11 Commission report did not make any recommendations relating to the intelligence process, nor the tools or techniques used within it. This is an interesting point to consider. The commission's final report notes that the intelligence community had failed to fundamentally adapt its gathering and analyses since the end of the cold war. Consequently, the intelligence process did not sufficiently consider the new threats posed by numerous entities which could attack the USA using advanced technology and potentially WMD (e.g. terrorist organisations, extremist factions, rogue nations). Simply put, the processes could not suitably analyse the new threats. Therefore, the USA intelligence community was not able to fully respond to the challenges of the twenty first century.

The second large intelligence failure within the last decade occurred both within the USA and British intelligence communities. The failure relates to the perceived presence and subsequent search for WMD within Iraq under the regime of Saddam Hussein. The failure to find WMD in Iraq led to major reviews on the available intelligence upon which government decisions and claims were made in both the USA (The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction, commonly known as the WMD report) and the UK (The Review of Intelligence of Weapons of Mass Destruction, also known and hereafter referred to as The Butler Review).

Within the UK, the main reasons stated by (at the time) Prime Minister Tony Blair for the invasion of Iraq was the belief that 'Saddam has continued to produce chemical and biological weapons, that he continues in his efforts to develop nuclear weapons, and that he has been able to extend the range of his ballistic missile programme….his [Saddam's] military planning allows for some of the WMD to be ready within 45 minutes of an order to use them." (Iraq's Weapons Of Mass Destruction – The assessment of the British Government, 2002, page 3) The United States government also "asserted that Saddam Hussein had reconstituted his nuclear weapons programme, had biological weapons and mobile biological weapon production facilities, and had stockpiled and was producing chemical weapons. All of this was based on the assessments of the U.S. Intelligence Community." (The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction 2005, page 3; hereafter referred to as the WMD report) However, following the invasion and subsequent search by allied forces the WMD report (2005, page 1) states that "not one bit of it [the assertion] could be confirmed when the war was over."

The WMD report asserts that the failure to find WMD is "in large part the result of analytical shortcomings; intelligence analysts were too wedded to their assumptions about Saddam's intentions. But it was also a failure on the part of those who collect intelligence… agencies collected precious little intelligence for the analysts to analyze, and much of what they did collect was either worthless or misleading. Finally, it was a failure to communicate effectively with policymakers; the Intelligence Community did not adequately explain just how little good intelligence it had--or how much its assessments were driven by assumptions and inferences rather than concrete evidence." (WMD report, 2005, page 3) Overall, the report notes that for the intelligence community to respond to the ever-changing threats it must become more integrated and flexible.

Of most interest to this thesis are the recommendations made by the WMD report (2005) in the area of intelligence analysis, specifically:

- "The [Intelligence] Community must develop and integrate into regular use new tools that can assist analysts in filtering and correlating the vast quantities of information that threaten to overwhelm the analytic process. Moreover, data from all sources of information should be processed and correlated community-wide before being conveyed to analyst." (WMD report, 2005, page 402).

- "The Director of National Intelligence [within the USA] should encourage diverse and independent analysis throughout the Intelligence Community by encouraging alternative hypothesis generation as part of the analytic process and by forming offices dedicated to independent analysis."(WMD report, 2005, page 405).

The above recommendations are interlinked: analysts did not consider a sufficient number of alternative hypotheses, nor were they able to clearly communicate the confidence associated to specific assessments. Based on these findings the WMD report notes that "As much as they hate it, analysts must be comfortable facing up to uncertainty and being explicit about it in their assessments" (WMD report, 2005, page 408).

These recommendations are supported by the findings of the Butler review which suggests a review of the intelligence assessment staff which "in particular considers whether they have available the volume and range of resources to ask the questions which need to be asked in fully assessing intelligence reports and in thinking radically." (The Butler Review, 2004, page 159).

Of course, the development and implementation of additional or novel tools and techniques must be supported by new training (currently out of the 16 week FBI training course, only three days are given to counter terrorism, 9-11 Commission Report, 2004).

Since 9-11 there have unfortunately been further high profile terrorist attacks across the globe including bombings in Bali, Istanbul, and Madrid. All of these attacks were seemingly planned and executed without active intervention from the local intelligence authorities. On July 7th 2005, London was also subjected to a terrorist attack culminating in the deaths of fifty-two people.

Following the attacks in London, the British Government Intelligence and Security Committee produced the Report into the London Terrorist Attacks on 7 July 2005. The report concedes that different decisions in the years prior to the attack could have increased the chances of identifying the planning and therefore preventing an attack. However, the report concludes that "in light of other priority investigations... the decision not to give greater investigative priority to these two individuals was understandable. In reaching this conclusion we have been struck by the sheer scale of the problem that our intelligence and security Agencies face and their comparatively small capacity to cover it." (Report into the Terrorist Attacks in London on July 7th 2005, 2006, page 16).

1.4 How does this thesis support the analysis of intelligence?

This thesis investigates the development of a support tool for analysts working with, and reasoning about information which is uncertain, incomplete or imprecise. Although a range of mathematical techniques may be used as the basis of such a tool, this thesis focuses on the use of Bayesian Belief Networks (BBN). These flexible, robust, graphical probabilistic networks are able to incorporate values from a range of sources including empirical values, experimental data and subjective values. Furthermore, the nature of the models requires all the associated uncertainties to be explicitly stated throughout the analysis procedure.

The use of a Bayesian methodology has another advantage in that it "provides a formal framework for assessing the odds of hypotheses in light of evidence." (Burns, 2006, page 1570). Consequently, a BBN will also show the change in likely outcome of the generated series of hypotheses based on the evidence available over a given timeframe. Of course, the BBN can only show the variation in those hypotheses identified by the model developer (the generation of hypotheses and development of BBNs are discussed in Chapter Three). That said, such a tool could indicate when a given hypothesis is increasing in likelihood even when an analyst (or team of analysts) perceives the situation to be moving in another direction. Such information could prompt the analysts to look again at the assumptions, data and information upon which any predictions have been based. This thesis considers various aspects of the development of such a support tool through the use of BBNs.

Overall it is important to remember that whilst BBNs have the ability to represent patterns of evidence in a wide range of scenarios, including asymmetric attacks, to be fully effective, senior management and decision-makers must understand the basic principles behind the technique and have confidence in the results obtained through the use of BBNs.

A support tool for analysts working with uncertain or incomplete information is an active field of research. The outputs of such research will begin to fulfil some of the recommendations highlighted from the WMD report (2005). This thesis asserts that an appropriate basis for the development of such a tool is BBNs. At its core, this research is a very practical application of mathematics in an area of huge relevance in today's ever changing world climate.

1.5 Why use Bayesian Belief Networks?

Bayesian Belief Networks are flexible, graphical models which incorporate formal probabilistic and statistical methodologies enabling uncertainty to be logically and consistently represented. The networks coherently update the beliefs in a set of given hypotheses as additional information is entered into the system. This unique combination of factors makes BBNs an ideal choice for use in systems designed for the fusion and analysis of uncertain or incomplete data. As such, they show use for the quantitative analysis of multiple intelligence reports. The use of Bayesian approaches within decision support systems is already well established within the medical profession (Luciani et al, 2007 Hamilton, 1994), and in diagnostic tools within the automotive and IT industries (Crossman, Hong, Murphey and Cardillo, 2003).

Whilst the specific mathematical basis for this research is the BBN which has its roots in probability theory, the wider aspects of this research rest in the area of decision analysis. This mathematical discipline encompasses many techniques for the structuring, representation and analysis of those aspects identified as the most important in a given decision situation. As such, decision analysis provides a means for individuals and organisations to assess important decisions within a structured, formal methodology.

Analysis of the decision situation is frequently undertaken utilising mathematical models. For example, the model could be a graphical representation of the problem space, such as a decision tree or influence diagram. Certain mathematical models, such as BBNs can be populated with probability estimates, either derived from historical data or elicited from personal beliefs. This allows for a quantitative analysis of outcomes related to each identified strategy or piece of information.

In addition to the mathematical techniques outlined above, this research has also considered the strands of psychology and sociology which describe the decision-making process.

1.6 Value of the research.

This thesis reports upon the development and use of BBNs for the fusion and assessment of intelligence reports within a series of fictitious and historical case studies. As such, all the data and evidence included within the scenarios is either fictitious or has been obtained from publicly available open source reference material. Not withstanding this fact, the findings of this research do show advantages in the use of BBN for the analysis of multiple intelligence reports. In particular, this research highlighted differences between direct intuitive judgements of complex events and normative combinations of judgements relating to simpler events.

For each case study considered, a supporting scenario and BBN were developed. Within the scenario, a series of intelligence reports were presented over a defined timeline. Initially, the research focused on how the participants' beliefs about the hypotheses changed in response to the information provided.

For each scenario, a series of conditional probability distributions and direct subjective estimates for various propositions involving diagnostic reasoning for each scenario were obtained. The probability distributions were used to populate the BBNs. Subsequently, the evidence presented within the scenario was sequentially entered into the networks. The BBN combined the various data and evidence available to calculate the Bayesian values relating to the direct probabilistic estimates.

The deviations between the Bayesian and direct values provided a measure of the participants' own logical inconsistencies.

The development of a BBN as the basis of a decision support tool for intelligence analysis is necessarily undertaken in conjunction with an appropriate elicitation technique. All elicitation techniques have their own strengths and weaknesses. Therefore, it is important to understand which technique should be used in a given situation. Consequently, the final part of this research centred upon an experiment comparing three elicitation techniques: direct rankings of the variables' perceived importance for discriminating between hypotheses, likelihood ratios and conditional probabilities. The aim was to ascertain the extent to which the different elicitation techniques lead to equivalent or different judgements. BBNs were used to provide a normative model for each participant which coherently combined the participant's initial judgements. This allowed for comparisons to be made with each participant's direct subjective estimates based on their intuitive reasoning.

1.7 Scenarios used within the research.

The scenarios used within this research are solely focused on military situations.

A fictitious scenario was developed for use in investigating the impact of information, as it became available on the participants' perception of a situation. The scenario developed considered a land-based section attack against an enemy outpost which may have been reinforced. It is important to note that this scenario only considered the UK army due to the influence junior army officers have upon the decision to engage in hostilities. All of the UK armed forces employ the concept of mission command. This philosophy, which states what to achieve and why, flows down from the Chief of Staff to the non-commissioned officers; to anyone who can influence mission success. British Air Power Doctrine AP300 defines mission command as being 'articulated through a statement of commander's guidance and intent together with articulation of subordinates' mission in the context of the overall plan.'(British Air Power Doctrine, 1999, page 1.3.4) Simply put the UK orders process describes a mission objective but is not prescriptive about how this should be achieved.

A military campaign can be broken down into a series of missions. Each of these mission will be subsequently divided until individual troop commanders receive their orders. It is at the level of troop commander (a troop commander normally manages a section of approximately 12 personnel) that this research initially focused. The decisions taken by troop commanders will impact upon their ability to achieve success and subsequently the higher mission, and finally upon campaign success.

This research required a scenario in which an individual could take a decision on when and how to enter hostilities. In addition to this the scenario had to contain a sufficiently small number of influencing factors affecting the decision as to be amenable to analysis. When considering the army it can be relatively simple to identify the factors influencing, for example, a section attack led by a 2nd Lieutenant who will decide when and how to attack. In comparison it is extremely complex to capture the multiple factors involved in the decisions taken by a Royal Navy Captain whilst executing their orders.

A historical scenario was used to further the investigations into the impact of information upon the perception of a situation and for the comparison of elicitation techniques. The case study chosen for use in this research was the historical case of the Arab Israeli conflict of 1973 (commonly known as the Yom Kippur War). Using open source material, a timeline based upon the main events observed leading up to the start of hostilities was developed. This scenario was chosen for the Israelis well documented incorrect interpretation of the gathered intelligence and subsequent conflict.

1.8 How Bayesian Belief Networks help generate alternative hypotheses.

The initial development of a BBN requires identification of all plausible hypotheses. The main node of interest within a BBN contains an exhaustive set of mutually exclusive hypotheses. If the analyst does not identify all possible hypotheses, the true answer may not be identified by the subsequent analysis procedure. As such the hypotheses generation procedure often requires analysts to ‘think outside of the box’. Indeed as already noted, support and encouragement for the generation of alternative hypotheses was included within the WMD report (2005). Some of the techniques used to generate hypotheses such as the analysis of competing hypothesis are presented in Chapters Two and Three.

1.9 How Bayesian Belief Networks help process intelligence.

The previous discussions have shown the limitations of the intelligence community to effectively manage the potentially overwhelming volume of available information. In essence the problem is one of preventing information overload. Information overload occurs when an individual is unable to cognitively, or emotionally, effectively manage all the data available to them. At this point an individual is unable to keep abreast of the situation and as such cannot take a fully informed decision.

To operate at peak performance a commercial organisation desires information which arrives at the right time and in the right format, matching the quality requirements of its potential users (Marcusohn, 1995, cited in Hall 1998). How can the intelligence community operate at peak performance when they have virtually no ability to influence any of the above factors within the information they receive? It is also important to remember that the processing of information within the intelligence community frequently refers to processing large volumes of disparate data.

Bayesian Belief Networks have several features which can assist with the processing of large volumes of information. The networks themselves are graphical and represent the current understanding of the situation based on the analysed information. The structure of a network is intuitive and even those with no experience of BBNs will be able to identify which factors have been included and which have been excluded from the analysis. In addition to this it is relatively easy to amend the structure of the network to include additional nodes as required.

BBNs can incorporate the uncertainty associated with information including the possibility of false positives and false negatives. Even information which does not support the current perception of the situation can be entered and its impact upon the belief of each competing hypothesis be easily evaluated. As information or intelligence becomes available it is easily integrated into a BBN. The networks logically and consistently fuse new information with the understanding of the situation immediately prior to the information being available through the use of probability theory. One main advantage of BBNs is that, depending on the information available the networks can be used to support different kinds of reasoning, e.g. for diagnosis or to predict future events.

1.10 How Bayesian Belief Networks can help analysts be explicit about the uncertainty within their assessments.

In his Review on Intelligence on Weapons of Mass Destruction Lord Butler states that "The most important limitation on intelligence is its incompleteness. Much ingenuity and effort is spent on making secret information difficult to acquire and hard to analyse. Although the intelligence process may overcome such barriers, intelligence seldom acquires the full story. In fact it is often, when first acquired sporadic and patchy and even after analysis may still be at best inferential." (The Butler Review, 2004, page 14) Two years after these words were published, the Report into the Terrorist Attacks in London on 7 July 2005 noted that "the issue of addressing the limitations of intelligence in intelligence assessments…had not yet been fully implemented." (Report into the Terrorist Attacks in London on 7 July 2005, 2006, page 32).

BBNs have defined techniques to incorporate for representing uncertainty associated with any information within a network. The processes required for formulating a network, deriving the required distributions and finally entering new information are all transparent processes open to audit and debate. Throughout the analysis procedure, analysts are constantly forced to consider the uncertainty associated with any inferences drawn from the results. The numerical results obtained from the network at the end of the process are defendable, open to scrutiny and amenable to sensitivity analysis.

1.11 Overview of the chapters.

Following on from this introduction, the thesis continues with a wide ranging literature survey focused on decision support systems within intelligence analyses. Consequently, this chapter begins with a discussion on decision-making and theories of decision-making models. For completeness a comparison of the normative decision-making model and Klein's descriptive Recognition Primed Decision model is made. This discussion also introduces common cognitive heuristics and how they lead to bias and error in human judgement. The second part of the literature review centres upon decision support systems: the need for them; potential designs and applications; evidence for the success and failures of decision support systems and the mathematical basis for the tools themselves are presented. Finally, the literature survey concludes by considering the tools and techniques commonly used for decision support systems within intelligence analyses.

Chapter Three introduces Bayesian Belief Networks and commences by presenting the development of the networks. In support of this thesis Chapter Three gives a mathematical description of how uncertainty is included within, and how new evidence propagates throughout the networks. The third chapter explains how this research interlinks with work already carried out in the field of tools for intelligence analysis.

A detailed presentation of the developed BBNs and case studies considered are presented within Chapters Four to Six. Within these Chapters the results of experiments are discussed alongside the associated sensitivity analyses. The performance of the BBNs and possible reasons for differences between the network and individual direct results obtained are discussed. Factors which were found to affect the model are presented along with strategies for improving the quality of overall results obtained.

Finally Chapter Seven considers potential future applications for Bayesian Belief Networks in tools for intelligence analyses and makes recommendations for the inclusion of these systems within the intelligence community.

CHAPTER 2 LITERATURE SURVEY.

2.1 Decision-making.

2.1.1 Introduction.

When a decisive turn of events leads to the winning of a strategic battle, social change or an increase in company profits, the decisions taken by those in authority are often deemed visionary or courageous. However, history is written by the victors, in the fullness of hindsight and out of the context in which the original decision was made. Often what is exalted or derided is the outcome of the decision, not the decision itself. There are all too many examples of 'good' decisions resulting in bad outcomes and 'bad' decisions leading to a desired outcome: many a man has made and lost a fortune from making first lauded and subsequently ruinous decisions. So what separates such decisions and their outcomes? Luck plays its part along with all the component parts of a decision which the individual cannot alter for example: terrain, weather, or unforeseen events. Importantly, and of interest to this thesis there is the impact of human judgemental error upon the decision-making process.

Many complex decisions may involve: numerous interacting factors, uncertainty, multiple stakeholders (with their differing perspectives and opinions), and subjective opinions from various experts. Often a complex problem will have no single clear outcome. Instead, the preferred outcome will require the decision maker to make trade-offs between the contributing factors. Thus selecting the 'right' course of action is not always easy as Heller (1989, page 25) writes about Montgomery "No decent person likes hurting other human beings, whether they come singularly or in battalions…..The really tough decisions taken by Montgomery included fighting the purely defensive action at Alam Halfa which halted Rommel in his tracks, and the refusal, against heavy pressure from Churchill, to launch El Alamein until the Eight Army was very good and very ready."

Montgomery had to take decisions in an incredibly complex environment, the output of which was perilous for the men under his command (and Montgomery was known for being careful with the lives of men under his command). The decisions taken by Montgomery beautifully illustrate the importance of the context in which a decision is taken. For as well as the geographical, logistical and myriad of factors which affected his fighting ability in North Africa, political pressure also came to bear.

Following D-Day (6th June 1944) some of the decisions taken by Montgomery's 21st Army Group were supported by the Number 2 Operational Research Section (No 2 ORS). For an engaging read of a large collection of the reports produced by No 2 ORS, the interested reader is referred to Copp (2000) and, for comparison to collated personal papers of Rommel by Hart (1953).

During the Second World War, and over previous millennia, wars have been fought against a known, identifiable enemy. Even when engaging multiple opposition forces, much was known about their location and aims. From this, a clearly defined end point – that at which victory would be achieved – could be defined. Today battles are fought in a more complex environment. Battles of attrition are rarely fought. Rather, highly trained yet irregular forces skilled in guerrilla warfare tactics are engaged across the globe. It is now far harder to define the end point of hostilities – many conflicts are fought not on territorial gain but upon ideological grounds for political advantage. For all of the dramatic changes in warfare, the military environment provides fascinating examples of decision-making in which (paraphrased from Noorderhaven 1995, page 2):

- The means of achieving an objective may radically alter (e.g. from an offensive to defensive stance) whilst the overall objective does not.
- Subjective assessments are based upon imperfect or incomplete information.
- The implementation of a strategy leads to unexpected results.
- The decision-making is not always based upon rational calculations but upon moral values, emotion and perhaps intuition.

2.1.2 Decisions – a cycle or process?

What is a decision? There are a multitude of answers. A logical explanation is that a decision is a rationally based choice between at least two alternative courses of action. If a decision-maker perceives only one choice of action is available, it can be argued that there is in fact no decision to be made. However, to have arrived at such a point is the outcome of many previous decisions.

This simple explanation of a decision conceals many of the intricacies and judgements within the intertwined decision-making and problem solving processes. Firstly a decision-maker must identify a need to make a decision. Often the decision-making process is instigated by a change in the status quo or driven by the desire to move towards an identified goal (e.g. improved company performance).

Simply identifying the need to make a decision does not guarantee that one will be made. Before the presumption of choice can be applied, the decision-maker must have considered the problem space and have identified a preferred end state (MacKenzie, 1975; March, 1994, and Hogarth 1980). Following this, the decision-maker must formulate and evaluate courses of action which could potentially achieve the desired end state. The cognitive workload required to assess numerous, possibly complex, courses of action should not be underestimated. Complex courses of action may not be easily compared and contrasted (Lindsay and Norman, 1977) particularly as the human short term memory can only store around seven pieces of information (Miller, 1956).

Once all the courses of action have been evaluated, it is critical that the decision-maker feels able to select and, crucially, implement the preferred course of action. Intent is rarely sufficient to gain a desired outcome. Once implemented, the progress of the chosen course of action should be monitored and controlled. There are a plethora of factors which may affect a decision which are hidden from the decision-maker until they unexpectedly alter the outcome and create uncertainty.

This procedurally based explanation of decision-making is contested by Hollnagel (2007, pages 4–5) who argues that decision-making should be considered as an activity based not upon the assumptions of a rational decision-maker but upon the assumptions that:

- "Decision making is not a discrete and identifiable event, but rather represents an attribution after the fact.
- Decision making is not primarily a choice among alternatives. It is very difficult in practice to separate decisions from what is otherwise needed to achieve a decision maker's objectives, that is, what is required to implement the chosen alternative.
- Decision making is not usually a distinct event that takes place at a specific point in time, or within a certain time window and which therefore can be dissociated or isolated – even if ever so briefly – from what goes on in the environment."

The consideration of decision-making as an activity, rather than a process moves the emphasis of a decision from being 'what should be done' to that of 'how and when' to undertake an activity.

2.1.3 Decision-making models.

Decision-making models are generally classified into three broad categories: descriptive, normative and prescriptive. It is argued that the latter of these descriptors is not in fact a decision-making model, but rather a framework which seeks to identify how decision-makers may be best supported (Bell, Raiffa and Tversky, 1988).

2.1.3.1 Descriptive decision-making models.

Developed within the field of psychology, descriptive decision-making models seek to portray how and why rational people make decisions. This immediately raises the question 'what is rational?'. Within the field of economics, rational people are taken to be those who seek to maximise their gain and as such seek to make optimal decisions. To genuinely make such a decision requires perfect knowledge of all possible alternatives and the probability of both the alternative and any associated outcome occurring (pure theory of rational choice). Rarely does this happen. Indeed in many situations, the cost and time requirement to gather the information required to provide all possible alternatives and their outcomes would, if not impossible, be prohibitively expensive. Hollnagel (2007, page 5) raises an important point on the topic of time and information in that "decision making, whatever it is, takes time and therefore logically requires that the information it uses remains valid whilst the decision is made."

In contrast to the theory of pure rationality, Simon (1957) made two important observations. Firstly, that individuals are only rational for part of the time and frequently make less than optimal decisions. Secondly, that individuals make use of a series of heuristics to simplify the decision-making process. These observations play an important part in Simon's widely accepted bounded rationality model (work for which he was awarded the Nobel Prize for Economics).

Bounded rationality accepts that both time and cost constraints can limit the data available upon which to base a decision. These constraints when combined with a decision-maker's intelligence, cognitive ability and perception of the situation prevent an optimal solution from being selected (Bazerman, 1986). The bounded rationality model simplifies the decision space by only requiring a decision-maker to consider the consequences of the current actions being considered. The requirement of perfect knowledge for alternatives and outcomes is removed. This allows uncertainty to enter into the decision-making process. Furthermore, by removing the necessity to select an optimal course of action, the assumption of bounded rationality allows for a satisficing decision to be taken (March and Simon 1958, and Simon, 1988). Such a decision does not seek the optimal course of action but rather that for which the predicted outcome exceeds a personally defined threshold for each attribute of the decision (note that the threshold level can vary between attributes).

The use of a threshold allows this theory to be applied in situations where alternative solutions can be assessed sequentially. Once a solution is found which surpasses the defined threshold for each attribute, the search stops and a satisificing solution has been found. It is feasible that an individual's thresholds will vary over time and, consequently, a decision may be defined as being acceptable, reasonable, rational or sensible when fully considered within the context in which it was taken. If the context or constraints in which it was taken are removed or misrepresented then the decision may no longer be thought of as sensible.

The choices taken by individuals within the constraints they find themselves in are defined by March (1994, page 2) as being conditional upon four basic questions:

- "The question of alternatives: What actions are possible?
- The question of expectations: What future consequences might follow from each alternative? How likely is each possible consequence, assuming that alternative is chosen?
- The question of preferences: How valuable (to the decision maker) are the consequences associated with each of the alternatives?
- The question of the decision rule: How is a choice to be made among the alternatives in terms of the values of their consequences?"

By providing a representation of what people actually do, descriptive models must consider a myriad of interlinking factors including (Bell, Raiffa and Tversky, 1988): perception of uncertainty, decision-making biases; how problems are decomposed and decisions made; the impact of tradition, intuition and culture. Of course not every model will consider all of these factors in detail. Indeed it may not be possible to incorporate some of the above ideas into computer models. Yet, all of these factors will impact upon real people's decisions and a descriptive model will be evaluated based upon how well its selected course of action matches those taken by individuals in the same circumstance.

2.1.3.2 Normative decision-making models.

In contrast to descriptive decision-making models, normative decision making models are based in philosophy, yet developed through the fields of economics and mathematics. These models do not seek to describe what real individuals do but rather provide a mathematical model of an idealised decision-maker taking an optimal decision. Essentially normative models show what a 'rational' decision-maker should do. The results of such a model provide a normative standard against which real life decisions can be evaluated and compared. Deviations from the normative standard are considered to have been caused by biases within individual decision-making.

The idealised representations of decision-making within normative models allows decisions to be assessed outside of their original, or any, context. Consequently, normative decision-making models do not consider an individual's core values, beliefs or perception. Such models do assume that the decision-maker is rational and in line with the economic definition of such a person, will seek to maximise their gain and minimise losses. One of the fundamental principles of normative models is that of transitivity. The property of transitivity should hold anytime objects / events are compared and its basic rules are:

- If A> B and B> C then A > C.
- If A = B and B = C then A = C.
- If A = B and C > 0 then A + C > B.

The above rules may be considered as 'just being common sense'. Indeed the use of these rules enables the development of a logical flow of choice and decisions within a mathematical framework. Yet, the principle of transitivity assumes an individual's preferences will not change which simply may not be true. Furthermore, individuals evaluate gains differently and not always numerically. So whilst these rules do form the basis of many decision models it can be seen that they do not always apply.

One commonly used mathematical technique for the analysis of normative decision-making models is utility theory (French, 1989). The psychological basis of utility theory is that a value, or utility, can be placed upon the perceived outcomes of each possible course of action. This classical economist view of utility, was expanded by von Neumann and Morgenstern (1944) to portray the decision-maker's attitude towards the concepts of risk and uncertainty within decisions. Leading on from this early work, Savage (1954) developed the Subjective Expected Utility (SEU) model in which the decision-maker assigns not only a utility to each identified outcome, $u(x_i)$, but also the probability that the outcome will occur, $P(x_i)$. The subjectivity of the measure lies in the fact that the probability of an outcome occurring is rarely known objectively, and therefore must be subjectively assessed by the decision-maker. Combining the utility and probability values of each subjective outcome i the subjective expected utility of each alternative is defined as:

$$\sum_i u(x_i)P(x_i)$$

Equation 2.1: Subjective Expected Utility.

Overall, the decision-maker would be expected to select the alternative with the highest SEU. One commonly used method to assign utilities is through the use of hypothetical lotteries. The use of such lotteries determines the shape of the utility function by varying the values included within the lottery. Based upon the decision-maker's stated preferences, the values associated with the lottery outcomes are altered until the decision-maker expresses no preference between the guaranteed result and that of the potential gamble. The shape of the finalised utility function is representative of the decision-maker's attitude to risk.

Of course, not all values associated with an outcome are monetary or have utilities which can be easily compared. The factors which must be considered when buying a house include not only cost but the size of the house, its location, local amenities and many more besides. A fair consideration of all these factors will require the use of Multi Criteria Decision Analysis (MCDA) (Belton, 1990; Guitouni and Martel, 1998, and Goodwin and Wright, 1999) techniques which include Multi-Attribute Utility Theory (MAUT) the Simple Multi Attribute Rating Technique (SMART), (Edwards 1971, and Edwards and Barron, 1994 advanced by Roberts and Goodwin, 2002) and the Analytical Hierarchy Process (AHP) (Saaty 1990 and Sinunay-Stern, Meherz and Hadad, 2000). Such techniques are compensatory (that is the overall strength of a potential solution is found through evaluation of its weak and strong attributes) and by evaluating all options simultaneously, not sequentially.

Although widely used, many decision-makers often express preferences, particularly in situations of risk, which are at odds with the fundamental principles of utility theory. In an extension of utility theory, Kahneman and Tversky (1979) developed prospect theory to take account of these sometime contradictory preferences.

Within utility theory a worth, as expressed through its utility, is associated with each outcome. Prospect theory replaces the notion of utility with a value expressed in terms of losses and gains relative to a reference point. The value of each alternative is assessed through a two stage process defined as editing and evaluation.

The focus of the editing phase of prospect theory is to 'frame' the decision which is, according to Tversky and Kahneman (1981, page 25) "the decision-maker's conception of the acts, outcomes, and contingencies associated with a particular choice. The frame that a decision maker adopts is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision maker." Importantly, decisions can frequently be framed in more than one way.

The framing of the decision also comprises two parts: simplification and coding. Firstly, the decision-maker creates a simplified mental model of the situation being considered. This is achieved through the use of heuristics such as cancellation (discarding common components of all potential courses of action and is related to the 'sure-thing' principle of utility theory as developed by Savage, 1954) and detection of dominance (which, if any, course of action is preferable in all aspects to every alternative (Noorderhaven, 1995).

Following the development of a simplified mental mode, the decision-maker codes each course of action. Wright (1985, page 5) defines coding as the "perception of a decision maker of each of the gamble outcomes as being either a gain or a loss, with a gain or loss being defined relative to a reference point." The reference point chosen by the decision-maker can be influenced by previous decisions, but is generally defined by a level of aspiration.

The overall value of a course of action is determined on the basis of a value function. A hypothetical function is shown below in Figure 2.1. It can be seen that the reference point is assigned a value of zero, above the reference point the function is concave, and below the reference point it is convex. Another important feature of the value function is that it is steeper for losses than for gains – this expresses the notion that individuals feel losses more keenly than gains.

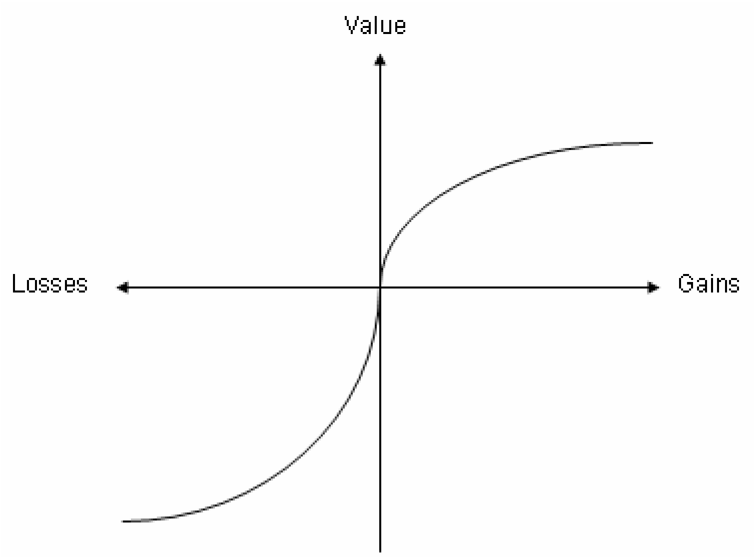


Figure 2.1: Hypothetical value function for prospect theory.

The positioning of the reference point also affects an individual's attitude to risk. If an outcome is considered to be a gain, then the decision-maker will tend to be risk averse. Conversely, for an outcome perceived to be a loss the decision-maker will be risk seeking. Prospect theory leads to a fourfold pattern of risk attitude, namely (Noorderhaven, 1995, page 90)

- "Risk aversion for gains of moderate to high probability.
- Risk seeking for gains of small probability.
- Risk seeking for losses of moderate to high probability.

- Risk aversion for losses of small probability."

An individual's attitude to risk will be influenced not only by whether the outcome is a loss or gain, but by the probabilities associated with the outcomes. In utility theory the probability is used to weight the utility. However, in prospect theory a value is multiplied by a decision's weight which is a function of but not the actual, probability. Overall decision weights tend to overweight small probabilities and underweight moderate and high probabilities (Tversky and Kahneman, 1981).

There are two more main differences between utility and prospect theory, both of which relate to the perception of risk: the certainty and pseudocertainty effect. The certainty effect is "a reduction in the probability of an outcome by a constant factor [and] has more impact when the outcome was initially certain than when it was merely probable." (Tversky and Kahneman, 1981, page 455, cited in Plous 1993 page 99) In essence, individuals prefer to remove rather than simply reduce risk. Linked to this effect is the pseudocertainty effect, which differs only from the certainty effect in that the certainty is apparent rather than real. Within the original work carried out by Tversky and Kahneman (1981), individuals were shown the potential outcomes of two treatments in response to an epidemic. In one scenario treatment A was presented only in terms of the number of survivors and treatment B in terms of the probabilities of those who would survive and those which would die. An alternative scenario presented treatment A only in terms of those which would die and treatment B again as probabilities of those which would die or survive. The actual values used in the experiment meant that each treatment had the same outcome in terms of those which would survive. However, the framing of the question revealed an interesting result about risk perception. In the first scenario, participants preferred treatment A, presented only in terms of survivors – the individuals were risk adverse. Conversely, in the second scenario, individuals were more risk-seeking, preferring the option B which gave the possibility of avoiding a perceived definite negative outcome.

2.1.3.3 Naturalistic Decision-Making.

In contrast to the study of decision-making in laboratory based situations, Naturalistic Decision-Making (NDM) is "the study of how people use their experience to make decisions in field settings" (Klein, 1997, page 11). Developed through extensive observations, NDM is mainly used in pressurised situations where experienced decision-makers are faced with ill-structured problems (Klein, Orasanu, Calderwood and Zsambok, 1993, Zsambok and Klein, 1997, and Shapira, 2002). One of the main differences between classical decision-making models and NDM is the assumed time available in which a decision can be made. The focus of NDM is upon time pressurised situations in which an expert simply 'knows what to do'. In contrast, normative and descriptive decision-making models centre on decisions in which the decision-maker has sufficient time to assess the available options.

Within NDM the initial focus of effort is upon assessing the situation to fully understand the problem with which the decision-maker is faced (for similar work see Cohen, Thompson, Adelman, Bresnick, Tolcott and Freeman, 1995). Subsequently, any potential courses of action are considered sequentially with options being rejected when a problem is identified, and eventually a satisficing, rather than optimal solution being accepted. In total, NDM comprises six main approaches (Shapira, 2002): Image theory, decision cycles, explanation based decision-making, cognitive continuum theory, the dominance search model and situation assessment and recognition.

This latter model is of particular interest to military decision-making. Much of Klein's work for the development of the Recognition Primed Decision-making (RPD) model was based upon observations of fire fighters, who often work in stressful, rapidly fluctuating situations. The fluid nature of the environment necessitates that the fire fighter continually evaluates the situation changing their priorities as required. Klein (1999) claims that experts (fire fighters or not) are able to quickly assess a situation based upon previous, similar situations encountered and hence can easily disregard irrelevant information and focus only on that which is important. However, a certain amount of deliberation may occur with decision-makers whose expertise or experience does not enable them to immediately recognise a situation and therefore be able to rapidly filter the available information. Potential courses of action which are not applicable to the situation being considered are swiftly ruled out and a course of action likely to succeed is selected. In essence, the expert is acting at an almost subconscious level as they simply know what to do (Klein, 1999). From the previous discussion, within RPD decision making is considered an activity, not a process (Hollnagel, 2007).

2.1.4 Cognitive heuristics and uncertainty.

2.1.4.1 Cognitive heuristics.

Following on from the work of Simon (1957), seminal work by Tversky and Kahneman (1974) reported on how three (robust) heuristics (representativeness, availability and adjustment and anchoring) were used to simplify the assessment and prediction of probabilities.

For the most part, heuristics are seen as being helpful in decision-making (Hastie and Dawes, 2001). Not only are heuristics used to provide a simplified representation of a problem, they may also be considered as being both 'fast and frugal' (Gigerenzer, Todd and the ABC Research Group, 1999). This latter definition has been applied to heuristics for their ability to reduce the time taken to process the available (and sometimes incomplete) information associated with a decision. Yet, the prevalent use of heuristics throughout decision-making can also lead to the introduction of biases and errors within human judgements. Beginning with initially assessing the size of problem being faced, Nutt (1989, paraphrased from page 69) notes that:

- "Decision-makers recognize information selectively and give information that is readily available too much weight.
- Errors in recognizing and weighing information stem from:
  - Difficulty in accurately estimating the frequency of events.
  - Giving events with which one has a kinship too much weight.
  - Emphasizing information consistent with past experiences.

- Being drawn to vivid information and discounting the value of pallid information.
- Order effects, in which information that is initially recognised has more weight than information recognized later."

2.1.4.1.1 The availability heuristic.

The availability heuristic utilises the ease with which instances can be recalled to gauge the probability of an event occurring. This heuristic can be advantageous when considering large classes which are usually more easily recalled than instance from less frequent classes. That said, events which have an emotional attachment, or have vivid imagery associated with them will be more easily recalled than mundane every day events (Bazerman, 1986). Yet the ease and vividness with which an event is remembered does not mean these events have any relevance to the events for which an estimate is being derived (Lichtenstien, Slovic, Fishhoff, Layman, and Coombs, 1978). Therefore, it can be seen that the availability can introduce biases into judgements through (Tversky and Kahneman, 1974): retrievability of instances (affected by the decision-maker's familiarity with the class being considered and salience); the effectiveness of the search set (the search set identified is dependent upon the task and it may often be easier to identify contexts in which abstract concepts appear as opposed to placing concrete concepts in context); imaginability (when the frequency of a class is not assessed by the number of instances recalled but through the use of a given rule) and the illusory correlation (when the assessment of the frequency of two events occurring is based upon the perceived bond between them).

2.1.4.1.2 Representative heuristic.

Many probabilistic assessments require decision-makers to determine the relative frequency of an event or object from a given class. When arriving at an estimate, individuals commonly use the heuristic of representativeness and thereby base their value upon how similar, or representative, a person, event or object is to a specific class or 'type'. Consequently, this heuristic inevitably brings to the fore an individual's preconceptions about the characteristics of a given class.

The bias of representativeness can lead to people displaying gambler's fallacy, that is predicting that an event that has not occurred for a while is more likely to occur in the near future (Lindsey and Norman, 1977).

The gambler's fallacy may be observed when someone believes that an independent result is more likely to happen based upon a recent series of events, for example: the roulette wheel has stopped on three blacks in a row therefore it is more likely that on the next spin it will stop on a red. This, of course, is simply not true. The bias of base rate neglect arises when the underlying relative frequency of some condition is ignored or not taken fully into consideration. In addition to base rate neglect, the representativeness heuristic can also lead to insensitivity to sample size. Tversky and Kahneman (1974) note than when making an estimate, individuals effectively ignore data relating to sample size from which any information has been taken. Finally, the representative heuristic can lead to the regression fallacy. This fallacy occurs when an individual ascribes cause to an event for which none exists by failing to consider the natural variation in a series of events. Examples include the variation in the value of commodities in a stock market or grades of a student over differing semesters. The fallacy occurs when a prediction is made based upon an exceptional point in the series, for example a very high price for a given commodity. Most people naturally assume that future outcomes will be representative of past outcomes and assume that the exceptional point in the series will continue. When, over time the value returns to that which is considered more normal, individuals start to believe their actions created the peak value, when in fact it was not causal but simply a natural variation in the series.

2.1.4.1.3 Anchoring and adjustment.

The final heuristic considered by Tversky and Kahneman (1974) was that of anchoring and adjustment in which decision-makers arrive at an estimated value by adjusting their belief from an initial value, known as an anchor. As a result, different starting points will return different estimates – this is known as anchoring. The biases associated with these heuristics are: insufficient adjustment away from the initial anchoring value (which occurs both when the decision-maker is told the anchor value and when they are free to select their own) and in the evaluation of conjunctive and disjunctive events (overall, people tend to overestimate the probability of conjunctive events and underestimate the probability of disjunctive events).

Additional heuristics impacting upon individual decision-making are (Gross, 2001; Dawes, 1988, and Villejoubert and Mandel, 2002):

- Belief perseverance (the tendency to cling to a belief even in the face of contradictory evidence),
- Overconfidence,
- Expectations and hindsight,
- Entrapment (also known as the sunk cost fallacy, this is the feeling of no choice but to continue in an investment as the costs of withdrawal cannot be justified. For additional information refer to Arkes and Blumer, 1985),
- The inverse fallacy.

The overconfidence heuristic is often shown by individuals being overconfident that the true value they are estimating lies within the bounds they have provided. As such, events with a low probability of occurring are overestimated, and those with a high probability of occurring are underestimated (Lichtenstien et al, 1978) Furthermore, individuals tend to overestimate the probability of events which have a favourable outcome to them (also known as the misconception of chance, for additional information see Weinstein, 1980)

The last noted fallacy, the inverse fallacy, is of particular importance when seeking subjective probability judgements. When asked to provide the probability of an event given an observed piece of data, for example given that is has rained what is the probability that the grass is wet, individuals inadvertently respond with the probability that it has rained given that the grass is wet. Although this is a simple example, there are of course differences in the stated values. Ensuring that the required probabilities have been elicited for further use and analysis is important.

Although heuristics are a source of error in human judgement making, Goodwin and Wright (1999, page 263) note that the results of research into the psychological aspects of decision-making has "largely carried out on inexperienced decision makers carrying out artificial tasks in psychological laboratories, rather than real-world decision makers making real decisions." Perhaps then, the results of human judgement making are not that bad. Indeed, Stewarts, Roebber and Bosan (1997) found that there was negligible bias within human judgements on forecasts when high quality information was available and the forecast was based upon a predictable task.

2.1.4.2 Uncertainty.

The use of cognitive heuristics to simplify a problem and thus reduce cognitive workload comes to the fore in decisions involving uncertainty. Uncertainty can enter a decision in many ways including (Nutt, 1989): being unable to predict future conditions and the consequences of trends (e.g. an economic slow down) or the inability to evaluate alternative courses of action. In essence, uncertainty enters a decision when the decision-maker is unable to make predictions due to the desired information being unavailable (e.g. the information is incomplete, of low fidelity or may not yet have been collated).

A combination of ambiguity and uncertainty can lead to conflict (for example disagreement between stakeholders as to how to deal with uncertainty) and "tough" decisions (Nutt, 1989). Additional factors which complicate the decision-making process include time pressure, poorly defined goals and the presence of multiple stakeholders. The effects of uncertainty and time pressure impact a decision-maker's choice as when (Goodwin and Wright, 1999, page 65) "under pressure to reach a speedy decision, non-compensatory strategies become dominant and when information on an attribute is missing, people may *infer* [italics in original quote] an attribute value if they believe it is strongly related to another attribute. For example, price may be used to infer quality." Additional factors affecting time pressured decisions are the quality, timeliness and order in which information is presented (Horvitz and Barry 1995, Reneau and Blanthorne, 2001).

Of course, most real world decisions are taken with varying levels of uncertainty. When there is real difficulty in predicting the outcome of a decision, it is common to refer to the risk associated with the identified courses of action (Williams, 2007 and Radford, 1975). How the risks associated with a decision are viewed is dependent upon the framing of the problem being considered: is the problem couched in terms of losses or gains? (Purkitt and Dyson, 1987) Yet, the acceptance of risk (and thus definition of an acceptable level of risk) is dependent upon an individual decision-maker's personal perceptions of risk (Fischhoff, Lichtenstein, Slovic, Derby and Keeney, 1981). A review of normative models for risky decision-making is given in Gilhooly (1988).

2.1.4.2.1 Subjective probabilities.

One method of expressing risk and uncertainty is through the use of probabilities. There are two ways in which to assign probabilities: objectively through analysis of experimental or extant empirical data, or by subjective judgement. When uncertainty prevails within a problem, frequently no data exists from which to derive the required probabilities (such as the effect of locating high level nuclear waste as discussed in DeWispelare, Herren, Clemen, 1995). Consequently, decisions involving uncertainty often rely on the use of subjective probabilities. By relying upon their expertise, subjective probabilities elicited from experts are, obviously, quantitative expressions of belief personal to the individual in the context presented (Sanders and Ritzman 1992). Early work on subjective probability was developed by De Finetti in the 1920s in which he defined probability as being "simply a question of making mathematically precise the trivial and obvious idea that the degree of probability attributed by an individual to a given event is revealed by the conditions under which he would be disposed to be on that event" (De Finetti 1937, page 101). However, this definition rests upon the assumption that decision-makers are rational and will show coherence in any bets placed. It also assumes that uncertainty can be represented quantifiably through a probability (O'Hagan and Oakley, 2004).

Each of the many subjective probability elicitation techniques and methods differ in their approach to engaging an expert in order to elicit the required knowledge (van der

Gaag, Renooij, Witteman, Aleman and Taal, 2002; Hoffman, Shadbolt, Burton and Klein 1995; Wright and Ayton 1987; Hart, 1985; Shadbolt and Burton 1995, and Savage, 1971). There is no universally accepted technique, with the specific choice usually being left to the researcher. Many of the techniques are designed for the elicitation of rules to be used within expert systems and are therefore context specific (Moody, Will and Blanton 1996, and Oritz, Wheeler, Breeding, Hora, Meyer and Keeney, 1991). With this in mind, it is important to remember that there is evidence to suggest that probability assessments are not always invariant between elicitation techniques (Keren, 1991, and Pöyhönen and Hämäläinen, 2001). However, research by Breivik and Supphellen (2003) showed that method bias was not considered to be a serious problem when eliciting evaluative attributes. Therefore, whilst it is possible to elicit tacit knowledge, the techniques used should provide some validation of the experts (and novices) used in the process (Bradley, Paul and Seeman 2006).

There are two main types of elicitation techniques: natural and contrived (Shadbolt and Burton, 1995). Natural elicitation techniques include observing an expert in their work or the use of interviews (structured, semi-structured and unstructured) and protocol analysis (verbal and behavioural). To be of most use, the interview process must: define the aim and importance of the task; clearly define the required probability and identify and so remove cognitive biases in the responses before finally encoding and verifying the given probability distributions. (Spetzler and Stäel von Holstein, 1975 cited in Goodwin and Wright, 1999, and Lau and Leong, 1999). In essence, natural techniques use situations and expressions which are familiar to the expert. Contrived techniques use processes with which the expert may not be accustomed to, including conceptual mapping and goal decomposition.

Elicitation techniques can also be defined as those which require the expert to provide explicit statements of probability including the median and the range, or spread, of results about this value (commonly referred to as the inter quartile range, Garthwaite and O'Hagan, 2000, and Moody et al, 1996) and those which infer probabilities (often based upon the theoretical bets suggested by De Finetti) (Goodwin and Wright, 1999). Other commonly used techniques include direct numerical assessment, probability wheels and scaled probability bars. Wang, Dash and Druzdel (2002) developed a framework for the comparison of such commonly used elicitation techniques based upon machine learning of experts' beliefs. They concluded that the most accurate and least time consuming technique was a scaled probability bar. That said, the results from such comparison studies have not been conclusive. Renooij and Wittemen (1999, page 268, cited in Wiegmann, 2005 page 4) writes that "What is lacking are large multi-method studies where experts are asked to assess a large number of probabilities with every single method." Wiegmann (2005, page 4) goes on to state that "There is a need to determine which methods are best, or whether certain methods work better in different contexts."

Whichever technique is used, O'Hagan (1998) notes that eliciting expert beliefs should be done as simply as possible and in a way which is familiar to the expert. In addition to this any technique used should identify and separately elicit the main sources of uncertainty to reduce expert overconfidence (O'Hagan 1998). This is one of the main causes of poor quality subjective probability distributions (van Lenthe 1993). It is important that consistency checks are applied to the elicited distributions. Often, this can be achieved by calibrating the expert. In contrast to using an elicitation technique to minimise the biases within subjective values, O'Hagan and Oakley (2004) suggest explicitly stating the imprecision associated with subjective distributions. This could be achieved by specifically considering aleatory (caused by natural variation in the situation being assessed) or epistemic (caused by the expert having insufficient knowledge about a situation to provide the required value) uncertainty. Whilst the use of expert opinion can, in theory, reduce the instances of epistemic uncertainty it cannot reduce aleatory uncertainty.

2.1.4.2.2 Elicitation as part of risk analysis.

Risk analysis is a diverse field encompassing all areas of life from national defence, construction (Adams, 2006), power generation (Oritz, Wheeler, Breeding, Hora, Meyer, and Keeney, 1991), health, and transport to name but a few. Fundamentally, risk analysis is about quantitatively, or qualitatively, determining the likelihood of an event and its subsequent outcome occurring. For many situations, particularly when assessing the risk of rare or extreme events historical data may not be available, or if available be open to multiple interpretations (Parent and Bernier, 2003). Consequently, subjective opinion is commonly used within risk analysis.

When eliciting subjective values, care must be taken not only in the selection of experts but also on the interview procedure used. Cohene and Easterbrook (2005) developed a framework through which an appropriate interview could be designed, dependent upon the elicitation process to be used. Forester, Bley, Cooper, Lois, Siu, Kolaczkowski and Wreathall (2004) developed an elicitation technique specifically designed for use in estimating probabilities for unsafe human actions. The technique centres on the probabilistic risk assessment, the knowledge and experience of experts and the translation of information into probabilities which can then be used. Some verification of the experts used to provide estimates is gained through direct questioning of their expertise and knowledge relevant to the probability of interest. The approach can be applied to both groups and individuals. Galway (2007) suggests that any protocol aimed at eliciting probability distributions should elicit the minimum, upper, and most likely values to which a triangular distribution should be applied. In addition to this, Parent and Bernier (2003) write that subjective distributions can be usefully elicited by being based upon quantities with a practical application to an expert's own field.

Subjective values can be used to define prior distributions within Bayesian techniques to analyse risk. Wiegmann (2005) reports on the development of a methodology specifically focused on improving the accuracy of subjective expert values for use in risk analysis. The finalised tool integrated a variety of techniques to elicit values with (Wiegmann, 2005, page 7) "probability elicitation questions in frequency format, which has been shown to suppress overconfidence, base rate neglect, the conjunction fallacy, control bias and over-estimation of single-event probabilities." The tool elicited the required values based upon the expert's opinion of how a product influenced the level of accident precursors within the risk model. Comparing the output of a model against predictions made by experts is a common verification methodology. Laskey (1995) developed a framework designed to support the elicitation of probabilities based upon such a comparison. Through the use of sensitivity values the most influential parameters within a network were identified. Support was subsequently targeted at the elicitation of the values supporting these parameters.

The use of Bayesian techniques is particularly useful when data is scarce. Adams (2006) concluded that although experts appeared to have some difficulty in estimating intermediate and tail values of probability distributions, the values which were given could be used to develop prior distributions, and when coupled with sample information, this enabled a Bayesian analysis of risk. Coolen, Mertens and Newby (1992) used a Bayesian approach to risk analysis due to a lack of, and the poor quality of, available data.

For a comprehensive review of elicitation of subjective values in risk analysis and where appropriate their combination with Bayesian techniques, the reader is referred to Mosleh, Bier and Apostolakis (1988).

The following second section of this literature review moves on from the consideration of decision-making theories to discuss how decision-makers can be supported.

2.2 Decision support systems.

2.2.1 Is there a need for decision support systems?

Savešek, and Pavešic (2007, page 293) comment that "Complex systems and ill defined information are one of the most influential factors in the decision-making process. When making a decision, one has to face the uncertainty of future events and the uncertainties which accompany the transmission, transfer and reception of information." When faced with complexity and uncertainty, Jarupanthirun and Zahedi (2007, page 1533) note that "individual decision makers are often incapable of making the best decisions when the problem is complex".

Uncertainty, ambiguity and complexity are inherent parts of the military domain and intelligence analysis. The collation, fusing and interpretation of data to support informed decision-making has been an important part of the military campaign throughout history. Of course, there have been dramatic changes in the information available to commanders. Wellington relied mainly upon information he could see for himself. Today, intelligence can be gained from many sources. Such information can support a vast array of decisions through the provision of situational awareness, warning of potential or immediate threats through to analysis and identification of organised crime.

In a complex environment where priorities may rapidly change, decision-makers may feel compelled to respond quickly, frequently making strategic, tactical or operational decisions (Turbain, Aronson, Liang and Sharda, 2007). In such an environment, the speed of decision-making is limited by an individual's information processing capabilities (Silver, 1991).

The sheer volume of information, or the number of contributing factors create decisions which are simply too complex to be made without some form of support, whether that be procedural, informational or computational (Rhodes 1993). In such situations, there remains the possibility of automating the decision-making process. However, Zack (2007) remarks that this is only possible in situations where there is sufficient information about possible outcomes, and their likelihood, to be included within a model. This information is not always available, and in the absence of an automated decision-making rule set, the individual will use heuristics to simplify the decision problem. At this point a Decision Support System (DSS) may be of help.

By using a DSS a decision maker can "minimize the chances of (1) making a poor decision, (2) missing a good alternative, (3) making a decision at a bad time, or (4) focusing on the wrong issues or problems." (Carter, Murray, Walker and Walker, 1992). In his seminal work, Sprague (1980, taken from reprint in Sprague and Watson 1986, paraphrased from page 10) proposes 6 performance objectives for a DSS in that it should:

- Provide support for decision making, but with emphasis on semi-structured and unstructured decisions.
- Provide decision-making support for managers at all levels, assisting in integration between the levels whenever appropriate.
- Support decisions which are interdependent as well as those that are independent.
- Support all phases of the decision making process.
- Support a variety of decision making processes, but not be dependent upon any one.
- Be easy to use.

All of the benefits listed here are achievable by individuals given the appropriate amount of time and data. Yet a DSS can achieve these benefits quickly, accurately and cheaply. Even so, Reneau and Blanthorne (2001) note individuals are unlikely to base important decisions solely upon a statistical model. Therefore, the real benefit of using a DSS is to facilitate the combining of human and computational capabilities to enhance the decision-making process.

2.2.2 What is a decision support system?

The term DSS was initially used in the early 1970's. However, during the 30 year research into the field of DSSs no universally accepted definition has been agreed. Frada and George (2007, page 1647) write that "The term…DSS has since been treated as an "umbrella" term in that it represents a variety of techniques and technologies usually borrowed from a range of disciplines which aim at improving access to necessary information for more effective decision-making." Regardless of the tool, technique or process used, the general aim of a DSS is to provide access to "the right knowledge…to the right processors, in the right representations and at the right time." (Holsapple and Joshi, 2003, page 91 cited in Frada and George, 2007, page 1647)

Such broad definition of a DSS encompasses both ‘soft’ and ‘hard’ techniques within the realm of operational research. Carter et al (1992, page 9) separated all techniques into three broad areas:

- “Procedures (e.g., checklists) to organise his or her thinking or provide a framework within which specific analytical methods can be used.
- Methods and techniques (e.g., simulation) to help examine alternative solutions.
- Presentational forms (e.g., charts and graphs to help display and review data, inter relationships and outputs).”

Stenfors, Tanner Syrjänen, Seppälä, and Hasapalinna (2007) found that the most commonly cited DSSs were the simplest. Overall, SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis and those presented in spreadsheet applications were the most cited. This result is most likely due to the decision-makers who participated in the study citing the decision support systems they themselves used. More complex DSS (such as simulations or business intelligence systems) which may provide input for decisions taken at managerial or board level were not actually manipulated or run by the decision-makers questioned. Consequently, complex DSS were not cited as the most used. Undoubtedly, techniques such as SWOT analysis and brain storming are commonly used. However, the growth in computing power has inevitably lead to a focus on the development of interactive DSS software.

The first computer based DSS was developed during the Second World War to assist with the analysis of data in support of wartime needs (Silver 1991). However, it was not until the rapid development of computing capabilities in the 1970’s and onwards that computer based DSSs became viable. Edwards, Phillips, Hays and Goodman (1968) designed a probabilistic information processing system (PIP) to support decision-making. The system itself was Bayesian based and required participants to provide a likelihood ratio (probability of an event occurring given that the hypothesis is true versus the probability of the event occurring given that the hypothesis is not true) for each hypothesis being considered. Edwards et al (1968, page 248) write that “a computer aggregates these estimates by means of Bayes’ theorem…into a posterior distribution that reflects the impact of all available data on all hypotheses being considered. Such a system circumnavigates human conservatism in information processing, the inability of men to aggregate information in such a way as to modify their opinions as much as the available data justify.” Although the PIP was shown to be efficient and improve operator performance, it was restricted by the computational power available at that time.

Many early decision support systems built upon developments in Management Information Systems (MIS) which were capable of providing basic reports based upon the extant data collection protocols. DSSs differ from MISs in that whilst MISs respond to ‘what if’ queries with a rule based ‘if …then’ response “fundamental to the notion of a DSS is assistance provided in assessing the situation, identifying alternative courses of action, formulating the decision situation, structuring and analyzing the decision situation, and then interpreting the results of analysis of the alternatives in terms of the value system of the decision-maker.” (Sage 1991, pages 4-5).

From this work, if one relatively narrow definition of a computer based DSS persists, it is that it should contain a (Sage 1991, page 1):

- “Database management system.
- Model base management system.
- Dialogue generation and management system.”

This definition, although not incorrect, does not convey the wide range, nor the complexities of many computer based DSSs. Today, in addition to these three core components a DSS may have some form of knowledge or intelligence components (Turbain et al, 2007). Developments on DSS means they are now capable of fusing information from a range of sensors and data sources using differing media types (digital, graphical, maps, as well as numerical and written text). The DSS techniques include: decision models, visualisation technologies, data collection sensors, rule based engines (Bhargava, Power and Sun, 2007), as well as techniques from the arts such as story telling (Paradice, 2007) and artificial agents (Van Tol and AbouRizk, 2006, and Brynielsson, 2007).

The broad definition and relatively young age of the field of DSS has lead to a fragmentary approach to support decisions (Arnott and Pervan, 2008). Indeed, the field has many sub-divisions. In 1995, Eom carried out citation analysis and identified seven informal clusters of decision support systems. Building upon this work, and again trying to understand the discipline of DSS Arnott and Pervan (2008, pages 657-658) report the seven major sub-fields of DSS as being:

- "*Personal Decision Support Systems:* usually small scale systems developed for one manager, or a small number of independent managers to support a decision task;
- *Group Support Systems:* the use of a combination of communication and DSS technologies to facilitate the effective working of groups;
- *Negotiation Support Systems:* DSSs where the primary focus of the group work is negotiation between opposing parties;
- *Intelligent Decision Support Systems:* the application of artificial intelligence techniques to decision support;
- *Knowledge Management-Based DSS:* systems that support decision making by aiding knowledge storage, retrieval, transfer and application by supporting individual and organizational memory and intergroup knowledge access;
- *Data Warehousing:* systems that provide the large-scale data infrastructure for decision support;
- *Enterprise Reporting and Analysis Systems:* enterprise focused DSS including executive information systems (EIS), business intelligence (BI) and more recently, corporate performance management systems (CPM)."

2.2.3 Designing and basis of decision support systems.

It is clear that it is not possible to have a single definition of a DSS. The focus of much research is now on how a DSS can be designed to be of most use to a decision-maker. Arnott and Pervan (2008) identified a widening gap between research into

DSS and the actual practical use of DSS: only 10.1% of DSS developed were considered as having high or very high practical relevance. More worryingly 49.2% of DSS were considered to have low, or no practical relevance.

Even if the DSS is considered to have practical relevance, the system may still provide answers which are correct but are not palatable or acceptable to others. Outputs from a DSS which are at odds with the generally perceived view of a problem may be seen as a threat to an individual's credibility or simply be disregarded. This is an important point. The use of a DSS stops at the end of the analysis. A DSS does not implement a decision; the ultimate responsibility for taking a decision still rests with an individual. If they do not understand the results of the DSS, do not believe in its credibility or reliability then they may exercise their judgement and not use the results of the analysis. This is their prerogative and ultimately a decision will still be taken.

To try and mitigate such an outcome it is important that a DSS is not seen simply as a 'black box' into which data is fed and output magically produced. When determining how to design a DSS, Sage (1991, paraphrased from pages 161-162) suggests strong consideration is given to four issues:

- The specific task to be performed that influences the nature of the system's design.
- The environment into which the task is embedded, which includes both internal organizational issues and external issues surrounding an organization.
- The DSS user or problem solver's familiarity with the task at hand and the environment into which this task is embedded influences if the problems or tasks are structured or unstructured.
- The extent of designer knowledge and availability, and the extent to which there is a group within the organization that might potentially use a DSS.

These points are still relevant. Wherever possible, the decision-maker should be considered in all aspects of the design and development of a DSS. It is, of course, the decision-maker who drives the process, from data collection, through analysis and course of action selection. It is their knowledge and expertise which are critical to the use and guidance of the DSS and implementation of any outcome produced by the system. This will be most fruitfully achieved when the DSS complements an individual's decision-making style. In a global age the development of a DSS for an international company must take account of the varying national managerial and decision-making styles (Martinsons and Davison, 2007).

How then should DSS be developed? Paradice (2007) suggests that since complex decisions are intertwined with emotion and organisational culture, research should begin to focus on how emotion, passion and commitment impact upon problem formulation. Even your mood impacts upon how a DSS is viewed and utilised (Djamasbi, 2007). The results of further research in these areas will undoubtedly influence the design of future support systems. Hall and Davis (2007) observe that personal values are integral to behaviour and thus the search and selection of a course of action. Such values become of critical importance when the decision-maker is under time pressure or has uncertain information. Hall and Davis (2007) propose to support the decision-maker through guiding them to view the problem from a variety of perspectives (theoretical, social, political, religious, aesthetic and economic). Such an approach, it is argued, facilitates the development of a wide range of courses of action and ensures information is not disregarded.

Hosack (2007) considered if it were possible to influence value based decision-making in the context of a 'wicked' problem (that is a problem which has little clarity in either its definition, or end state which are defined in terms of good or bad). Through the use of feedback within a DSS, Hosack found it was possible to create some change in the decision-maker's behaviour although it was unclear if the most strongly held beliefs had been targeted for influence. The use of DSSs to tackle wicked problems was also considered by Mackenzie, Pidd, Rooksby, Sommerville, Warren and Westcombe (2006) who observed that the solution to a wicked problem is often achieved by a group and is essentially a social process. Through the use of a combination of brain storming, cognitive mapping and dialogue mapping Mackenzie et al (2006) captured the different views, opinions and arguments associated with the development of a solution or framing of a problem. Organisational solutions to wicked problems were also considered by Petkov, Petkova, Andrew and Nepal (2007). The representation of a problem from several perspectives was supported through the use of a range of MCDA and 'soft' thinking methods. Overall, they concluded that the number and type of techniques used should be dependent upon the complexity of the overall problem.

The use of a value-based approach could be very useful when developing organisational policy or other collaborative decision-making environments (many DSSs utilising MCDA techniques have been designed for strategic planning, Eom and Min 1999). When a group is required to arrive at a consensus, Lee, Lee, Kang and Baik (2007, page 3129) note that such situations are "highly dependent upon matters of perspective, values and opinions, all of which – being subjective in nature – are beyond the reach of existing formal decision technology." When a group individually rank a set of alternatives there may be a mix of subjective rankings (which can therefore be expressed as a value function) and judgemental rankings (which are assumed to have been assigned on a rational basis). Each of the many types of subjective and judgemental rankings have there own strengths and weaknesses. Yet, the different approaches can lead to misunderstanding in the rankings assigned by different members of the group. To facilitate reaching a consensus, and identifying logical reasoning or flaws in the rankings provided, Lee et al (2007) develop an 'a priori[1]' reference model which may be used to determine if the rankings being provided have been properly reasoned (through highlighting inconsistencies or logical flaws in an argument) or are based more on intuition. The model itself comprises of a logical ordering protocol and subsequent integration procedure. Unlike some models for group decision-making which apportion equal weighting (or power) to every individual within the group, Lee et al (2007, page 3143) used superimpositional ordering which "is based on the idea that the individuals or stakeholders may not have the same weight in an ordering operation because of their expertise, experience, representing power, ideological or academic agendas, etc." This approach should, it is

[1] An a priori model is that which has been developed purely from thought and reason.

noted remove the bias of groupthink. This bias occurs when a decision is taken without fully considering or evaluating the full range of alternative solutions or courses of action. A consensus is reached with minimum debate or argument. Such a situation arises when individual members of a group feel unable to express a view or idea which may be perceived to be 'at odds' with the commonly held view of the problem being considered (for a detailed overview of group think the reader is referred to Kramer, 1998).

An alternative approach to reducing groupthink is to assist a group in viewing a problem from an opposing view. This enables the group to assess and evaluate the assumptions upon which they have based their decision whilst allowing additional alternative courses of action to be identified. Techniques which facilitate this method are known as structured conflict approaches and include the use of a 'devil's advocate' and dialectic inquiry. Jarupathirun and Zahedi (2007) incorporated the dialectic process into a DSS to assist decision-makers faced with unstructured problems. The dynamic process not only countered group think but also supported the group in 'thinking out of the box' and assisted in the teaching of critical thinking.

Group decisions may also be facilitated by the use of Social Choice (SC) theory which supports the use of voting systems. Srdjevic (2007) reports on the use of social choice voting in conjunction with appropriate Multi Criteria Decision Analysis (MCDA) techniques (such techniques are commonly used for the structuring and evaluation of complex problems, often involving uncertainty, for an overview see Belton, 1990) to support decision-making. In conclusion, if the two techniques are implemented then "Reduced information may lead to different outcomes for different voting methods...However, the compensation for this drawback is there a good chance…that the goal will be achieved, and at least the best alternative will be recognised and posted to the first position on the list."

Complex decisions are often compounded by uncertainty or ambiguity. Zack (2007) differentiates between problems made under uncertainty (when information is lacking) and ambiguity (when knowledge is missing), concluding that computer based

DSSs are most appropriate to support decisions under uncertainty and human-centric approaches for ambiguous decisions.

How to successfully incorporate uncertainty into a DSS is dependent to some extent upon the intended use of the system. Poch, Comas, Rodriguez-Roda Sánchez-Marrè and Cortés (2004) comment that in addition to uncertainty, some problems have a multiplicity of scales e.g. an environmental problem which starts at a local scale may interact with both national and global environmental concerns. As such, what begins as a relatively small, defined problem may have to be considered within a much wider context. The overall problem is, in essence, a complex system of systems in which uncertainty and varying complexity must be incorporated into a DSS. A variety of techniques can be used to mathematically incorporate uncertainty into a DSS, many of which can be used in conjunction with MCDA techniques.

Fuzzy set theory is a mathematical framework which can be used when "knowledge and information about a system is incomplete and experience-based rather than systematic." (Savšek, Vezjak and Pavešic, 2007, page 293). Fuzzy logic, it is argued offers a good approximation of human reasoning in a DSS through its ability to represent situations which may not be expressed solely through probability: that is when the answer verbally would be 'maybe'.

The use of fuzzy logic in MCDA problems was also considered by Fan, Lii and Tzeng (2007). A common output of many MCDA problems are preference ordered sets which are generally amenable to data-mining techniques to provide useful information though the use of dominance based rough set approach (DRSA). Fan et al extend this work to consider the use of DRSA in cases of uncertainty where the results comparison tables may be incomplete or the values are able to take a fuzzy subset of each decision attribute. The result of the research was a series of logics representative of rules derived from preference ordered data tables.

Guo, Yang, Chin and Wang (2007) discussed the use of evidential reasoning (ER) to model and analyse both fuzzy (vague) information as well as uncertainty stemming from incomplete or absent information. Guo et al extended the ER algorithm (which uses the combination rule of the Dempster Shafer theory) to incorporate imprecise or incomplete weights assigned to attributes. The weight could have been assigned either directly, or through a pair wise comparison technique by either individuals or a group.

The focus of this thesis is, however, upon the use of Bayesian Belief Networks (BBN), which are part of classical probability theory. The combination of MCDA techniques with BBNs was researched by Fenton and Neil (2001) to combine the capability of reasoning under uncertainty whilst considering several attributes of reliability or safety. BBNs have also been used in time critical situations for the automation of decision-making (Bayse, Dean, Kirman and Lejter, 1992). The limitations of classical Bayesian decision theory to represent fuzzy variables are discussed in Wang, Qian Pagello and Pei (1996). These latter authors develop a system capable of incorporating uncertainty and presenting the possibility that two hypotheses may be plausible.

Burns (2006, page 1570) eloquently explains why Bayesian inference is potentially so powerful in DSSs, writing "Bayesian inference provides a formal framework for assessing the odds of hypotheses in light of evidence. This makes Bayesian inference applicable to a wide range of diagnostic challenges in the field of chance discovery, including…counter-terrorism." This literature survey will now turn to consider the use of DSSs for military applications, specifically military intelligence.

2.2.4 Applications of decision support systems to military intelligence.

Andriole (1989, pages 4-5) writes that "military decision support systems…are frequently operationally or procedurally embedded in weapons systems integral to combat effectiveness, perform critical tasks in time constrained situations, must be easy to use and must be thoroughly reliable." In addition to this use Carter et al (1992, page 11) report on the use of "real time DSSs that can support (for example) military field decision problems. Such systems can be used to correlate incoming intelligence and sensor data, perform a rapid evaluation of alternative options, and suggest the best route or weapons mix for a specific mission objective."

However, before any system can be developed, the tasks it is designed to support must be fully understood. A badly designed system can increase the time taken to complete a task, increase workload and heighten the chance of catastrophic errors (Pfautz and Roth, 2006). A variety of techniques can be used to gain the required information. Holt and Hazen (1988) concluded that structured observations provided an effective and informative method of collating information pertaining to an air defence officer's tasks. The approach taken helped clarify any ambiguity or uncertainty as to what tasks individuals undertook and where they could be best supported. Moynihan and Bowen (1987) sought to identify information requirements and identify potential areas for decision support systems for commanders working in command and control. Initially, Moynihan and Bowen (1987) developed an appropriate data collection methodology. The task sought to understand the decision-making processes used by the officers and their information and data needs. Pfautz and Roth (2006) used a cognitive engineering method (an iterative cycle of analysis, development and evaluation) to fully understand the characteristics of the intended DSS user and the context in which the system would be used. Subsequently, this information was used in the design of a DSS for stability and support military operations.

The ability to fuse and collate information from across the military environment has far reaching applications. However, this research has centred upon applications of DSSs for use in situations where an individual or group has an opportunity to review the available information. One obvious use for such an application is in the area of command and control (C2) from simulation (Moffat, 2000; Liao, 2000; Moffat and Witty, 2002, and Suzic, 2003) through to tactical decision aids (Vraneš, Lucin, Stanojevic, Stenvanovic, and Subašic, 1992). Toft (1997) writes that "a tactical decision is only optimal if it's conditional on a optimal strategy identified on the operational level" and as such, a DSS should be able to assess multi-level decision over varying timelines. Savšek et al (2007) developed a tactical decision aid based upon fuzzy logic, capable of showing the various states of units (for example the level of fuel, ammunition, number of soldiers) as a battle developed.

Information obtained from across the battlefield is part of the reality of ‘net centric warfare’ in which distributed systems are connected. The information from such systems should be able to provide those with access to the network a common understanding – or situational awareness - of the battle space. That said, the sheer complexity in creating a system capable of providing the information each different decision-maker needs at an appropriate time is immense. Arnbourg, Brynielsson Artman and Wallen (2000) note that such a system must have ‘information awareness’ and measure the information in the system on its precision (correctness of data), quality (fitness for purpose) and utility (expected benefit for use). How the impact of data quality and information load can be reduced is discussed in Cowie and Burstien (2007); Williams, Dennis, Stam and Aronson (2007), and Shankaranarayanan and Cui (2006).

One of the main functions of information collation and analysis is for the provision of situational awareness and the provision of timely warning of a potential attack. A lack of situational awareness limits mission performance, even in training situations (Worm, Jenvald and Morin, 1998).

BBNs have clear applicability for use in situational awareness due to their ability to update the belief in a hypothesis given some new evidence. This revised value could then be used in the projection of future events (Pew and Mavor, 2000). The detection of an intruder into a network was considered by Yuill, Wu, Settle, Gong, Forno, Huang and Asbery (2000). Their detection technique identified the areas of a network most likely to be under threat from information the attacker revealed about themselves. Scott (2004) used a Bayesian model-based approach for the development of a network intrusion detection system. Das (2000) utilised a BBN to represent the uncertainties in the development of situation assessment in order to update knowledge and decide upon a course of action. The identification of potential enemy courses of action is the responsibility of intelligence officers. Kem (2005) proposes a methodology for identifying and developing enemy courses of action which can be maintained once operations commence. The methodology builds up from an understanding of the enemy's capabilities and understanding of the purpose of their actions, to the identification of the centre of gravity and key decisive points likely to be attacked to achieve their objective. Falzon, Zhang and Davies (2000) used concepts such as the centre of gravity and decisive points for the development of a course of action underpinned from a policy analysis perspective.

Intelligence can also be used in the battlefield to provide situational awareness. The overall level of awareness is, of course, dependent upon the data and information collated by the various surveillance and reconnaissance platforms (such as unmanned air vehicles and satellites). Liao, Sun and Wang (2003) set out architecture to integrate the available knowledge and information in a decision support system capable of warning about potential or immediate threats. The use of DSS for event prediction was also considered by Jesse and Kalita (1997) through the development of a knowledge based suite of tools to rapidly and accurately carry out general intelligence analysis procedures.

More recently, Pfautz and Roth (2006) developed a visualisation tool for military intelligence analysts working on stability and support operations. Such operations have additional intelligence requirements to those of a more 'traditional' military campaign, including an understanding of the local, socio-economic and political landscape. There is a need in many such operations to "identify emergent patterns suggestive of likely future behaviour…to anticipate (and try to dissipate) the next 'flashpoint' or 'hot spot'" (Pfatuz and Roth, 2006 page 392) The finalised tool comprised a suite of techniques including BBN which supported both the capturing of vague information as well recording how any inferred values were derived.

Social Network Analysis (SNA) offers an alternative approach for a visualisation tool of potential use to intelligence analysts. Comprising of methodologies taken from psychology, anthropology and mathematics (including but not limited to graph theory, cluster analysis, algebra and statistics) a "social network consists of a finite set of factors and the relation or relations defined upon them" (Wasserman and Faust, 1994, page 20).

SNA works on the tenet that the way an individual behaves is dependent upon their wider social interactions. Therefore, by identifying the social network within which an individual moves it is possible to define flows of information not only between individuals but also groups, computers and organisations. It can be seen that such an approach has applicability to investigation into organised crime and terrorist groups. SNA has also been linked with Bayesian approaches to enhance the use of missing or uncertain data within the social network (Koskinen, Snijders, 2007 and Butts, 2003).

In a similar vein to SNA, shortest path algorithms can be used to identify the relationships between groups and thus highlight new lines of enquiry for intelligence analysts (Xu and Chen, 2004). The use of the internet to identify links between associates has also been investigated (Skillicorn and Vats, 2007) through the use of contextual search queries. The results of the search may be clustered into units, described by several key words. Overall, the clusters may be representative of the importance of various terms, or provide additional supporting evidence for trends already identified.

The building of networks showing causal relationships in complex situations has also been investigated through a combination of a Bayesian inference and influence diagrams (Rosen and Smith, 1996 and Rosen, Smith, Smith, and Maldony, 1998). The system, named CAST (Causal Strengths) is designed to support individuals who want to create causal models but feel unable to develop the conditional probability distributions necessitated by a full Bayesian Network. Alternatively, the developed influence diagram comprises of nodes and their associated cause-effect links. The ability of the cause to either promote or hinder the effect is portrayed by the value associated with each cause-effect link. Interestingly, although designed to represent cause-effect relationships, CAST elicits the baseline probability of an event occurring as though it were independent of the influences included within the diagram.

2.2.5 Evaluation of a decision support system.

Once a DSS has been developed or acquired it must be installed, tested and deployed. At this point it is possible to make an evaluation of the DSS – but how should success or failure be measured? Sprague and Carlson (1982, cited in Sojda, 2007 page 271) comments that "organizations building their first decision support system [should] recognize that it essentially is a research activity and that evaluation should center on a general "value analysis"." There is, of course, often the need to justify the expenditure on a new system. Frequently, this is achieved through a cost benefit analysis. Such a measure, however, does not indicate if the system is actually supporting decisions. There is always the possibility that the introduction of a system has decreased the quality of decisions being made. Indeed Wagner (1981, cited in Finlay, 1989 page 150) reports a study finding that around 80% of DSS users could not quantify the benefits of their DSS.

In response to this, Finlay (1989 page 150) lists the potential benefits of using a DSS as being: "Increase in the number of alternatives examined; better understanding of the business; fast response to unexpected situations; ability to carry out ad hoc analysis; new insights and learning; improved communication; control; cost savings; better decisions; more effective teamwork; time savings; making better use of data resources." In addition to these objectives a DSS impacts upon many aspects of the working environment from an individual's job to the structure of organisations. For example, the data made available from a new DSS may indeed facilitate some of the required analyses. This, however, could make an individual's job seem monotonous (Turbain et al 2007) Such an outcome is at odds with Czech and Dizek (1988) who perceive the successful application of a DSS to be that it stimulates the analysts involvement in a problem. Many of the benefits listed here will be hard to assess numerically and as such will most likely be assessed through anecdotal responses from the user.

Phillips-Wren, Mora, Forgionne and Gupta (2007, page 1) develop a framework for the assessment of intelligent DSSs by linking "the decision value of an intelligent DSS to both the outcome from, and process of, decision making and down to specific components of the intelligence DSS." Sojda (2007) comments that DSSs should undergo both a validation and verification processes. Preferably, the DSS should undergo empirical testing "which in some form is critical, and can range from experiments run against a pre-selected gold standard to more simple testing of system components. It is imperative to understand from an experimental and logical perspective, to what extent inferences can be made as a result of the validation. In the end, the question to answer is: Was the system successful at addressing its intended purpose? Often, searching for the right database for empirical evaluation can be as important as adequate decision support system development itself." (Sojda, 2007, page 275).

The next and final section of this literature review will focus on the intelligence cycle and highlight where DSS tools may be of most use.

2.3 Military intelligence.

2.3.1 Why collect intelligence?

The Department of Defense in the USA defines intelligence as "information and knowledge obtained through observation, investigation, analysis or understanding" (Chizek, Elsea, Best, Bolkcon, 2003, page 2). The definition makes an important point: intelligence is not collected: information and data are received from a variety of sources. Only after analysis or interpretation does such disparate information become intelligence. The costs associated with all aspects of intelligence, from the collation of raw data to the dissemination of the final intelligence report are phenomenal, so what impact does intelligence have to warrant this expenditure?

Gudgin (1989, page 71) writes that "…the prime function of a Military Intelligence Organization, at whatever level, is to give advice concerning the armies of the enemy or potential enemy; this was laid down in 1904, for the British Army at least, in the report of the Esher committee into the working of the Mobilization and Intelligence Department during the Boer War, and has remained as a clear-sighted definition of the military intelligence ever since." Keegan (2003, page 370) remarks that the ideal of military intelligence as being when one side is “privileged to know the other’s intentions, capabilities and plan of action in place and time – how, where, what and when – while its opponent neither knew as much in return nor, that his own plans were uncovered.”

However, Keegan (2003) observes that the outcomes of very few conflicts since the end of the Second World War can be shown to have been influenced by the availability of intelligence. In contrast there are many historical examples, from the Roman conquests, through to the Napoleonic Wars and within the World Wars of intelligence providing the key to winning a battle. Furthermore, Keegan notes that even battles considered to have been turned upon intelligence made available to commanders (such as the Battle of Midway) often have additional, underlying supporting factors (prior to the Battle of Midway, a USA submarine accidentally strayed into the path of Japanese carriers forcing the redeployment of a dive bombing squadron). That said, historically during and prior to the World Wars, intelligence was mainly used to provide a short term tactical gain.

Of course, this is still true. Yet, in today’s relatively peaceful world the majority of intelligence is focused on continuous processes aimed at providing security and prosperity (Keegan, 2003, and Gill and Phythian, 2006). Eisenhower (MacCloskey, 1967 page 7, cited in Treverton, page 73) defined his intelligence requirements thus: “In war nothing is more important to a commander than the facts concerning strength, dispositions, and intention of his opponent, and the proper interpretation of those facts. In peacetime, however, the necessary facts are of a different nature. They deal with the conditions, resources, requirements and attitudes prevailing in the world. They are essential to the development of policy to further our long-term national security and best interests.”

Another fundamental change from the historical battles such as those of Napoleon is the lack of a clearly defined threat. Until the end of the Soviet Union, adversaries tended to be in the form of another nation whose doctrine, ideology and capabilities were relatively well known. In contrast to this Treverton (2003) observes that threats now stem from changing demographics, economic concerns, the availability of asymmetric warfare (from both rogue states and non-state actors such as fanatical religious groups and terrorist groups), and organised crime (including drug trafficking and nuclear material). These threats no longer focus solely upon military targets but at National Infrastructures and all members of society. Intelligence is no longer used by Governments just to prevent a military surprise, indeed the UK has MI5 which is an intelligence organisation focusing solely upon domestic security.

How then does intelligence support security? Gudgin (1989, page 75) writes that "…the best insurance against war occurring is to be prepared for it; and one of the most essential preparations is to know as much as possible about one's potential enemy. Intelligence can provide this information, can define the risks ahead and can estimate the cost of providing warnings against each of them…" Intelligence should, according to Treverton (2003) be able to keep abreast of military capabilities, politics and economies of major powers and in addition to this provide understanding of a given situation. In summary the preparedness of a military for a conflict is dependent upon its intelligence estimates, without it a Nation is at the mercy of better informed adversaries (Odom, 2004).

Yet the future is uncertain. Gazit (1989, page 61) comments that "…commanders of today…find it hard to accept a situation in which nobody can foretell the future for them. Many of them hope, or delude themselves, that the intelligence system serving them can fulfil this purpose….intelligence must limit itself to two areas:

- Specifically stating what may be expected, based on hard information about the other side's resolutions;
- Presenting the possibilities based on knowledge of the other side's general intentions and the optimal technical feasibilities at its disposal."

Handel (1989, page 196) elegantly writes that "the sole reason…the intelligence community exists [is] for the purpose of reducing uncertainty on political and military issues. Only very rarely can ambiguity and uncertainty be eradicated….In the world of intelligence, even technical data concerning performance or number of weapons, let alone less quantifiable issues such as intentions, military doctrine, and morale, cannot be objectively assessed – which means that clear agreement on their 'meaning' cannot be reached."

Overall the final output of intelligence "is a better understanding in the heads of people who must act or decide" (Treverton, 2003, page 107), and thus supporting optimal actions.

2.3.2 Intelligence sources.

There are a variety of sources upon which an intelligence estimate can be based. The source which provides the majority of information is known as 'open source' or information which is available in the public domain. A wealth of information can be gleaned from reviews of technical journal such as those produced by Jane's Information Group, media coverage of politicians and simply surfing the internet. Although open sources information is readily available and its collection easy, deciding what information should be believed and what should be rejected is far more complicated.

The remaining main sources of information used in intelligence estimates are all covert and encompass: human observation (HUMINT), electronic and communication signals (SIGINT), and photography and imagery (IMINIT).

Of these, the oldest form of intelligence is HUMINT which has been used throughout history. There are many forms of human intelligence including defectors from adversary nations, recruited agents, those coerced into providing information through blackmail and information obtained through the interrogation of prisoners of war. All such sources are capable of knowingly, or unwittingly, providing false or biased information.

Keegan (2003, page 28) writes that "the history of 'how, what, where, when' in military intelligence is…largely one of signal intelligence." Signal intelligence, once the domain of intercepting written orders, today focuses on electronic signals and the technology required to intercept often encrypted, messages. Techniques used in this domain range from the tapping of telephones, to the 'bugging' of rooms through to complex military listening stations. The sheer volume of information sent electronically (over e-mail, telephone, radio and secure military communication systems) is vast and can create an information glut. Information collected in this manner is of course, also open to misinterpretation, for example when intercepting a signal the listener may not know who is talking – is the conversation between two people who have access to the truth or are they simply speculating?

The use of imagery for intelligence purposes started in earnest with the advent of flight. Advances in technology have led to the use of satellites and intelligence, surveillance and reconnaissance (ISR) assets such as unmanned air vehicles (UAVs). ISR assets are increasingly used in military domains due to their flexibility. Satellites may only cover an area of interest for a matter of minutes every few hours, and due to their known orbits it is possible to hide assets from their sight. In contrast, the mobile nature of ISR assets enables them to provide mission specific information about an enemy to the decision-maker relatively quickly (Chizek et al, 2003). Of course, the goal of real-time intelligence is the ability to communicate quickly and securely. Gill and Phythian (2006) note, however, that one of the main limitations of IMINT is its limited ability to reveal the intent of an adversary.

Often, to provide a rounded intelligence estimate will require the fusion of several forms of intelligence – for example an image of movement supported by either SIGINT or HUMINT.

2.3.3 The intelligence cycle.

The production of an intelligence estimate is not, according to Gill and Phythian (2006) an exact science but an imprecise art, dealing only in probabilities, many of which are subjectively assessed. Handel (1989, page 188) writes that intelligence has "some quantifiable dimensions and requires decisions and forecasts to be made under conditions of pressure and uncertainty where failure is immediately reflected in negative results…failure can lead to severe criticism, even punishment, while success is taken for granted."

The process used by the Central Intelligence Agency (CIA) to produce intelligence estimates is a 5 stage cycle, shown in Figure 2.2 (copied from Gill and Phythian, 2006 page 3). However, this definition, according to Gill and Phythian (2006, page 3) "cannot fully capture the dynamic impact of intelligence's impact on the external environment. A better way of viewing the intelligence process in order to capture this dynamic fully is to adopt the concept of a system that includes feedback." The system they propose is shown in Figure 2.3 (copied from Gill and Phythian, 2006 page 4).

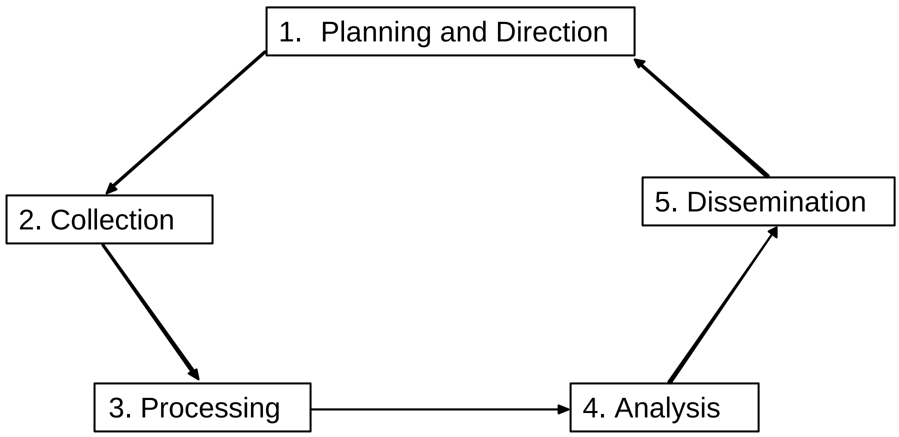


Figure 2.2: The CIA intelligence cycle.

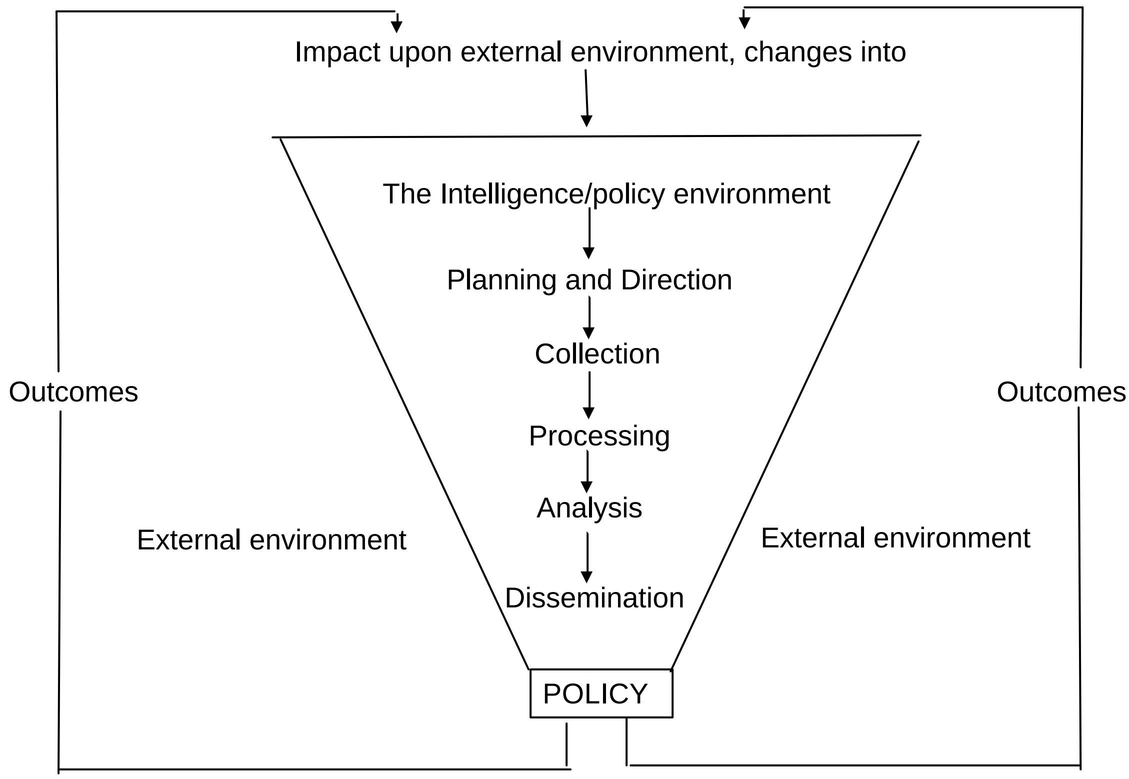


Figure 2.3: The intelligence process.

The planning and direction of intelligence is inextricably linked to National Policies. However, Treverton (2003, page 106) observes "policy officials seldom have the time or patience to articulate their information requirements precisely. Nor do most of then know enough to task intelligence operators effectively should they find the time to try. "More on Iran" or "better stuff on Saddam Hussein's intentions"". This is the level at which most policy officials express their intelligence needs. In the military context information requirements vary depending upon the operational type and are submitted, in the USA military, through priority intelligence reports which drive data collection. If too many are placed, the assets to collect the data are simply not available and it can be hard to show a direct link between the intelligence provided and subsequent decisions made (Nelson, 2004).

The subsequent phase of the production of an intelligence estimate is the collection of information and data with all its inherent biases, uncertainties and vagaries. MI5

(MI5 website 2008) notes that "If intelligence is worth recording we ensure that this is done accurately, clearly marking its origin and authenticity, and ensuring that it can be retrieved swiftly. If we assess that a particular threat needs to be investigated, then resources are deployed to obtain further intelligence." All of the requested or required information is rarely available all at the same time. Inevitably, information is fragmentary and provided over a period of time in small, discrete pieces some of which contradict each other. The aim is to fuse together all this available information to provide a consolidated view of the situation in a timely and clear format. This is achieved through the analysis and interpretation of the collated information.

Within the UK, the Defence Intelligence Staff (DIS) define the assessment phase of the intelligence process as "judging the authenticity and reliability of new information and its relevance to existing intelligence. Assessments focus on probable and possible outcomes, to provide the best available advice for developing a response or resolution. They are continually adjusted in light of new intelligence or events." (Defence Intelligence, Ministry of Defence website 2008).

There are two main types of analysis: tactical which focuses on short term gains and strategic which considers longer term aims. Ideally an intelligence agency will provide a balanced assessment of each area however, this is rarely achieved. Often, the imperative need of the near term requirement is answered, sometimes at the detriment to the longer term strategic view. Gill and Phythian (2006) describe three main types of strategic intelligence reports:

- Basic-descriptive reports which provide an overall assessment of other governments, militaries, or markets. The report is usually based upon open source material with some covert information to add value.
- Current reportorial which includes the latest available information (a commonly cited example of such a report is the USA daily Presidential briefing).
- Speculative – evaluative which attempts to estimate or assess possible futures.

Of course, the collation of data and its interpretation into intelligence does not in anyway guarantee that good decisions will follow. It is the prerogative of the recipient to consciously disregard the intelligence presented to them. However, Gaitz (1989) notes that for good decisions to occur, the intelligence estimate must be produced on time and subsequently be read and understood by its intended recipient. To facilitate the acceptance of an intelligence estimate, due care and consideration must be given to the language and presentational format of the estimate – the customer must want to read the report and trust those who produced it (Gill and Phythian, 2006). In this vein, Schum (2005) discusses the use of narrative accounts for the presentation of the results of the intelligence process, including common errors in the building of narratives.

2.3.4 Why does intelligence fail?

Intelligence failure is an emotive phrase. Given the uncertainties inherent in making intelligence estimates, it is almost inevitable that at some point, the process will not correctly predict an event and a surprise will occur.

The military theorist Clausewitz held intelligence in disdain. His suspicions of intelligence stemmed from the fact that "many intelligence reports in war are contradictory; even more are false, and most are uncertain…in short, most intelligence is false, and the effect of fear is to multiply lies and inaccuracies" (Clausewitz, 1976, page 136). Khan (1986) in reviewing Clausewitz's opinion of intelligence comments that Clausewitz's considered poor intelligence to contribute to the friction of war, and intelligence failures to reduce both physical fighting strength (by capabilities not being in the appropriate place) and the psychological capabilities of commanders (through for example over-estimation of enemy strength or the belief that information reinforces preconceived ideas).

Within the UK, intelligence failures are rarely openly discussed and infrequently attributable to a single point in the intelligence cycle. Intelligence failures can occur simply by not collecting the appropriate data such that indicators and warnings are not seen. In part though, the blame for such a failure rests with the policy makers and commanders directing the intelligence gathering collection. There are also those failures which stem from inadequate analysis and processing of information and data.

The intelligence process is open to many cognitive biases including the assumption that the adversary is rational, or has the same values as the analyst. What is perceived as being an acceptable course of action is affected by the culture of the country under consideration. In addition to this, if an adversary feels that there is no option, even though a territorial war cannot be won, the possibility of a political or moral victory will be sufficient to start hostilities.

Another cognitive bias which affects intelligence failure is a strongly held preconceived idea, particularly if this mental model has been shown to be correct. Consider the lead up to the Yom Kippur war of 1973, even though indicators and warnings suggested an attack was imminent, the general Israeli belief was that Egypt was simply not capable of launching such an attack. Each indicator was explained away, and the overall picture not identified. Some of the lack of acceptance of Egypt's ability could be explained by the 'ethnocentric' bias (Handel, 1989) – simply Israel believed in her own superiority to her surrounding Arab Nations. Of course, when such a mindset is in place, it can be hard to evaluate the situation from a different perspective, and even when this occurs the results may be refuted and ignored.

In addition to these biases, it is easy to believe that an event will not occur because it is a very risky option, or that it is simply impossible to achieve. Handel (1989, page 244) comments that "the greater the risk, the less likely it seems to be, and the less risky it actually becomes. Thus the greater the risk, the smaller it becomes."

Are these failures solely the responsibility of the intelligence agency? Of course the intelligence agency must be responsible for the information it produces. However, the processing of information does not stop at the production of the final report. It is then passed to a decision-maker who must read and process the information before deciding how to act.

Therefore, the failures of intelligence are not solely attributable to intelligence agencies, Betts (1978, cited in Gill and Phythaian, 2006, page 104) reports "In the best-known cases of intelligence failure, the most crucial mistakes have seldom been made by collectors of raw information, occasionally by professionals who produce finished analyses, but most often by the decision-makers who consume the products of intelligence services. Policy premises constrict perceptions, and administrative workloads constrain reflection. Intelligence failure is political and psychological more often than organizational." This view is supported by Schmitt (1997) in the military operating environment. By reviewing known 'intelligence failures' of U.S intelligence during the Korean war, Schmitt (1997, page 65) writes that "to blame intelligence primarily for our military misfortunes is to fail to understand the fundamental capabilities and limitations of intelligence and to fail to understand the hierarchical relationship between intelligence and the command and operations activities it supports……If intelligence "fails", as sometimes it must, it rarely does so alone. In the end, intelligence failures are failures of command."

2.3.5 How can intelligence assessments be supported?

Handel (1989) defines the three elements necessary for successful intelligence as being: data, analysis and a lack of political interference in the system. The focus of this thesis is not upon how to support or improve the collection of data, nor to discuss the political aspects of intelligence estimates. However, this work centres solely on the analysis and interpretation of collected data.

Handel (1989) suggests that the difficulties in providing intelligence estimates could be reduced through three methods: removal of all human biases and perceptions; by taking each reported threat seriously and taking appropriate action; finally by the introduction of reforms to improve the objectivity of the intelligence decision-making process. Of most interest to this work is supporting the analysis process through the reduction of human biases and perceptions within the analytical process.

One technique used to do just this is the use of indicators and warnings. "Essentially, the purpose of the method is to help the…analyst pick and choose the significant from the massive amounts of ambiguous and possibly conflicting data that would be abundantly available in crisis situations." (Handel 1989, page 247) This analysis "warns of impending attack…[and] depends upon the ability to predict enemy activity, based on enemy plans, doctrine and observed exercises and training." (Chizek, 2003, page 28). Potential indicators may be departure of foreign citizens, unusual concentration of forces, or changes in radio communications. Of course, such indicators may not be applicable on their own forces which are concentrated for many reasons other than the start of hostilities. Furthermore, if the indicators are known by an adversary they can be manipulated. Patterns of force mobilisation and concentration can become normal, such that an intelligence agency no longer considers them an indication of attack.

Even when the indicators and warnings do show attack is imminent, if this view goes against a long held belief, the intelligence can be rejected. When using indicators and warnings, the context in which they are seen is absolutely critical to their interpretation.

The subjective nature of the intelligence process does lend itself to the application of structured processes. However, decision aids are not widely used within the intelligence community. This is quite surprising given the requirement for (sometimes inexperienced) analysts to fuse large volumes of qualitative and quantitative data under time pressure into an intelligence estimate. That said, one widely used reference book for the structuring of problems and limiting the effects of cognitive biases in intelligence estimates is that of Heuer (1999).

In his 2005 paper, Heuer outlines many of the established procedures used to "help analysts question assumptions and adjust their mindsets. These include:

- *Red Cell Analysis* –trying to predict the behaviour of others by putting yourself in their shoes.
- *Devil's Advocacy* – assigning someone to challenge a single, strongly held consensus by building the best possible case for an alternative explanation.

- *Brainstorming* – an unconstrained group process for generating new ideas and concepts.
- *What If? Analysis* – taking as a given that an unexpected event has occurred and then trying to explain how it could have come about, known as thinking backwards.
- *Alternative Futures Analysis* – applies the collective knowledge and imagination of a group of experts to identify driving forces that are likely to shape an issue and how these different forces, when given different weights, might play out in plausible future scenarios.
- *Analysis of Competing Hypothesis*."

The Analysis of Competing Hypotheses (ACH) is a structured process which aids analysts in questioning basic assumptions and testing a series of generated hypotheses. The process begins with a brainstorming session to generate a complete set of hypotheses, all of which receive equal treatment. Subsequently a matrix is created and populated with evidence both supporting and refuting each identified hypothesis. Finally, the most probable hypothesis is that with the least contradictory evidence against it, not that with the most supporting evidence.

In addition to these techniques Tatarka (2002) includes the use of decision support tools in the forms of graphical displays or sequential presentation of information and the use of training for the production of more accurate intelligence estimates.
Of course, no process will guarantee that the resulting intelligence assessment is correct. What a structured process or decision support system does provide is an audit trail of how the estimate was developed. What assumptions were used, where the uncertainty in the estimate lies and, crucially communicate the magnitude of uncertainty in the intelligence estimate to the decision-maker.

2.3.6 Chapter summary.

This literature survey started by discussing the main types of decision-making models used in modern psychology and the common decision-making biases. It is evident that in situations in which time pressure, complexity and uncertainty are inherent, individual's may often benefit from some form of decision support.

As presented in the second part of this survey, decision support can take many forms, from simple brain storming sessions through to complex, bespoke pieces of computer software. The specific type of support needed is dependent upon the not only the decision itself (its structure, its attributes etc) but also the number of people involved in making the decision.

In a military environment there are potentially vast amounts of uncertain or incomplete data from a variety of sources. Once collated, the data must be analysed. In certain circumstances some of the decisions based upon the collated data may be automated (for example in sensor allocation or weapon firing). However, many cannot. In such a situation, analysts may benefit from a support tool. The next chapter of this thesis will introduce one of the main techniques used for collating data which contains uncertainty and will argue for the use of Bayesian Belief Networks in DSSs.

CHAPTER 3 BAYESIAN BELIEF NETWORKS.

3.1 Bayes theorem.

Initially defined by the Reverend Thomas Bayes and presented to the Royal Academy in 1796, two years after his death, Bayes' theorem is simple, elegant and powerful. Defined as:

$$P(B \mid A,c) = \frac{P(A \mid B,c)P(B \mid c)}{P(A \mid c)}$$

Equation 3.1: Bayes' Theorem.

The theorem relates the conditional and prior probabilities of two events. In the above equation, $P(A \mid c)$ and $P(B \mid c)$ are defined as prior probabilities. They represent the probability of an event occurring when only background or contextual knowledge is known, e.g. the probability of your garden having wet grass in the morning. No observation or information on the actual event is known, with the above example you will not have seen a weather report for the previous evening, or opened the curtains yet. Within the theorem $P(A \mid c)$ acts as a normalising constant.

The probabilities $P(A \mid B,c)$ and $P(B \mid A,c)$ are conditional probabilities. These values represent the probability of an event (say having wet grass in the morning) given that another event has occurred (knowing that it rained during the night). Consequently to assess these conditional probabilities requires some understanding of the causal relationships between events (in this simple example that rain makes grass wet).

The above explanation shows that Bayes' theorem separates out the effect of information on a probability into two sources: the prior probabilities and the probabilities in the light of any available evidence (Pew and Mavor, 2000). Indeed, Bayes' theorem is a logical method for the calculation of probability and is an example of probabilistic inference. Such techniques are used in a wide variety of applications including, and of most interest to this work, the prediction of events. However, probabilistic inference techniques are only applicable in situations where the events being observed, or the evidence made available can be expressed in a probabilistic manner.

Assuming the required probabilities have been obtained, Bayes' theorem expresses the probability that a given hypothesis is true based upon the available information. New observations or information will lead to a revision in the probability associated with the hypothesis. However, the variation in probability should be "something that changes opinion rather than as a measure of determining the ultimate truth." (The Economist, 2000). This theorem has obvious strengths in many fields, most notably medicine and forensic scientific investigations (for additional information on the application of BBN in these and other areas the reader is referred to Golub, 1997).

3.2 Brief history of probabilistic models.

The description of Bayes' theorem indicated that in order to ascertain the conditional probabilities, some knowledge of the causal relationships between the events being considered is required. The first causal model to be developed is attributed to Sewall Wright in the 1920s (Pearl, 1997). Working in the field of genetics, Wright graphically depicted the linkages between the markings on guinea pigs through their parentage.

Since this first causal model, many others have been developed, including probabilistic networks. Probabilistic networks show the dependence (though not necessarily causality) between a set of defined variables and their associated probability distributions. Early attempts to incorporate uncertain reasoning, within expert systems which were generally rule based (for example: if event X occurs then the response is Y) involved the use of so-called certainty factors and the expert system MYCIN (Buchanan and Shortliffe, 1984) provides the best known example. However, results from using this approach only coincide with results from probability theory under a very limited set of conditions. This, coupled with the emergence of probabilistic networks as a coherent approach to reasoning under uncertainty, eventually led to their decline.

There are three broad categories of mathematical techniques used to represent uncertainty: (Greenberg 2007, paraphrased from pages 14-16):

- Basic measurements including: classical measures of probability; possibility based on fuzzy sets and the Dempster-Shafer theory of evidence.
- Risk Metrics including the concepts of utility, credibility and loss.
- Modelling and analysis paradigms incorporating amongst others the theories of expected value, minimax, regret and mean risk models.

The initial successor to rule-based expert systems combined classical probability theory and decision theory. Known as normative expert systems, these models were able to incorporate uncertainty. This, undoubtedly, enhanced the overall capability of the expert systems. It also, however, led to the creation of very computationally heavy models. Eventually, normative expert systems became unworkable.

It was not until the seminal work of Pearl (1986) that the use of classical probability in expert systems became viable. The development of computationally efficient algorithms such as Pearl's (1982) and Lauritzen and Spiegelhalter (1988) supported the development of large-scale, complex probabilistic networks. Researchers such as Eric Horvitz, David Heckerman and Jack Breese particularly advanced the use of probabilistic models known as Bayesian Belief Networks (BBN). Such networks use the laws of probability for predictive (e.g. if it has rained then the probability the grass is wet is…) and Bayes' theorem for diagnostic reasoning (e.g. the grass is wet, therefore the probability it has rained is…)

3.3 Definition of a Bayesian Belief Network.

A Bayesian Belief Network (BBN) is a probabilistic framework represented through a directed acylic graph (DAG) and a set of probability distributions. An acyclic graph is one which contains no loops and therefore, it is impossible to trace a path through the graph and return to your starting place. A BBN is comprised of three distinct parts:

- Variables (which may be discrete or continuous) which in a BBN are represented within nodes.
- The dependence (though not necessarily causality) between nodes is shown by the directed edges or arcs within a BBN.
- Each node has an associated set of probability distributions which acts as the mechanism for the incorporation of uncertainty within the BBN.

A simple pictorial representation of the development of a BBN from these parts is shown below in Figure 3.1:

The representation of problems through the use of graphs has clear strengths. Firstly, it enables a clear, concise, intuitive presentation of the variables and causal relationships within a problem domain. Essentially, the topology of a BBN shows the qualitative knowledge about the situation being considered. Experts on the situation represented within the BBN are able to easily review the BBN and comment upon the inclusion, and exclusion, of relevant variables (nodes) and dependencies (arcs). For example, in the simple network shown in Figure 3.1 a review could suggest alternate reasons for the grass being wet and therefore, the inclusion of nodes to represent, for example: the presence of dew, burst water pipes, etc. This can be achieved without the need for reviewers to understand the, sometimes complex, underpinning mathematical techniques. Any alterations to the BBN, such as the inclusion or removal of additional nodes and arcs is easily achieved within available BBN software packages such as HUGIN or NETICA.

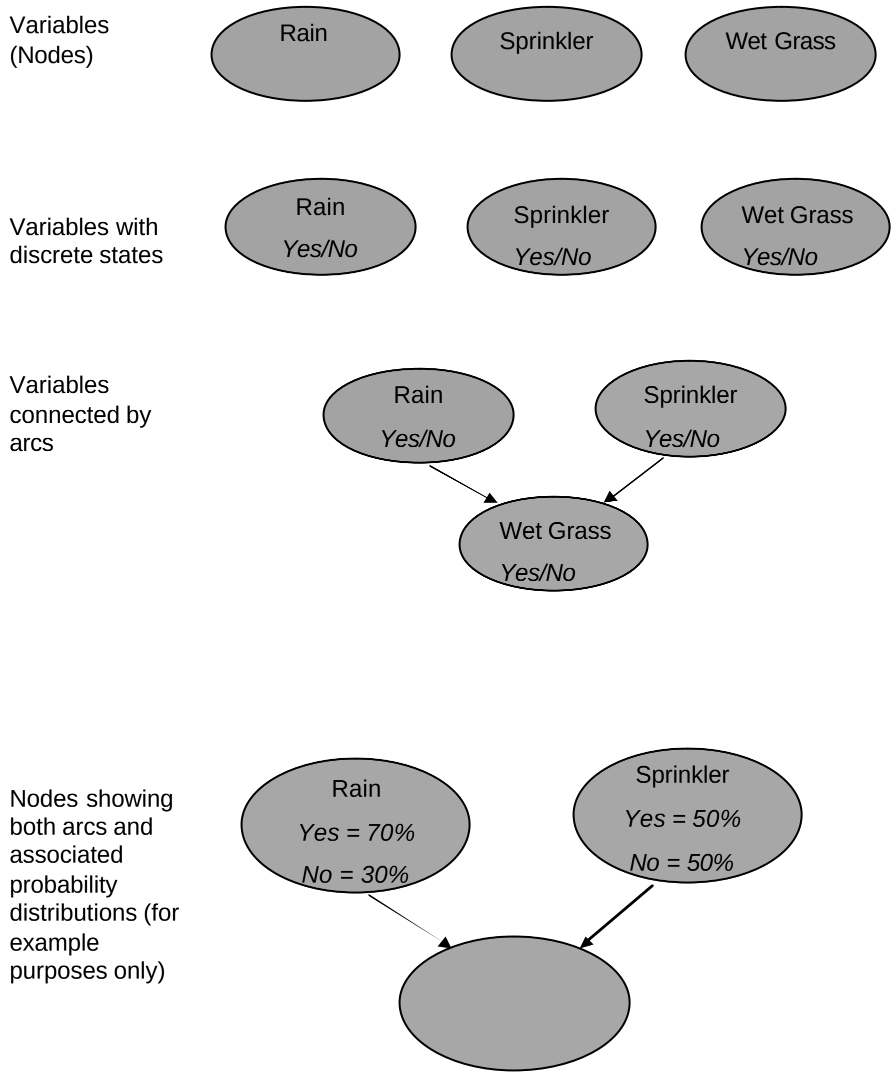


Figure 3.1 Example construction of a BBN from variables, arcs and probability distribution.

It is important to note that the derivation of the conditional probabilities associated with each node (each of which may be defined independently) may not be trivial. Conditional probabilities may be obtained from a variety of sources including experimentation, analysis and subjective opinion. This allows for the best available distributions to be entered at the time the network is developed. Consequently, variables with ill defined probability tables can be included within the network from its conception. For such variables subjective opinion may provide a probability distribution when no other exists. Yet, the use of subjective values can lead to the inclusion of cognitive biases such as: anchoring and adjustment, representativeness and overconfidence. Therefore, it is important to maintain detailed documentation of why subjective opinion was used and how the values were elicited. If at all possible subjective values should be put through a validation process. Of course, if more refined probability distributions do become available, they can be easily entered into the BBN. Overall, this approach ensures that uncertainties, as expressed within the probabilities, are stated explicitly throughout the analysis process.

As a BBN is acyclic, the network is a static representation in time of a given probability distribution. Whilst evidence may update the beliefs within the network, the observation of events will not alter the actual structure of the network: there is no opportunity for feedback. The ability of evidence to affect variables over time is however incorporated into dynamic BBNs.

3.4 Developing a Bayesian Belief Network.

Research papers rarely seem to discuss the development of the BBN they are presenting. It is hard to find information on how researchers identified the variables and dependencies to be included within a given network. This may be symptomatic of the lack of defined procedure for the development of a BBN. Hence, each researcher must identify the variables and dependencies of interest to themselves before structuring a BBN.

To structure a BBN which is an accurate representation of a complex problem may require many revisions. Once the variables to be included has been agreed, the dependencies between the variables can be depicted in three ways (as shown in Figure 3.2, based on Neil, Fenton and Nielson, 2000). These representations, known as d-connections, form the basis of information flow within the network. Any variables within a network which are not d-connected are said to be d-separated.

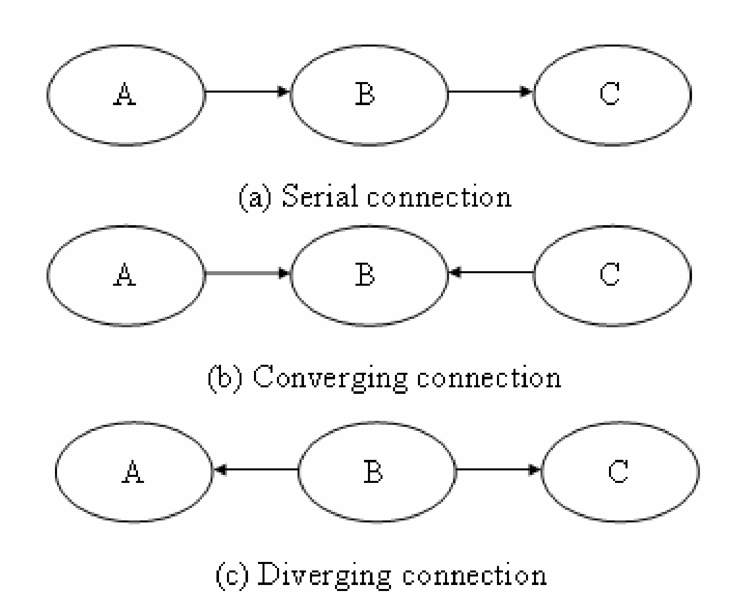


Figure 3.2: Serial, converging and diverging d-connections..

Expanding upon each of these connections, Figure 3.3 presents an example serial connection.

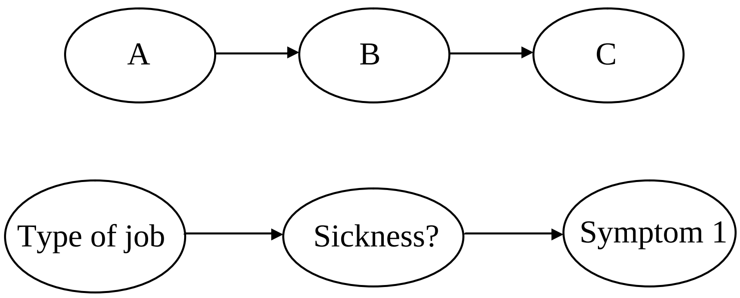


Figure 3.3 Example of a serial connection

Mathematically, in a serial connection, node B is conditionally dependent upon A, and node C upon B. In this example, a symptom of a sickness is dependent upon being employed in job 'J' and the diagnosis of a sickness is dependent upon the observation of 'symptom 1'. Evidence available at either end of a serial connection (here nodes A and C or type of job and symptom 1) will update the belief at the node(s) in-between (the belief in sickness?). Nodes at either end of the connection are normally dependent, that is evidence entered at these nodes will revise the belief in the other (for example knowing the type of job revises the belief in the individual having a symptom 1). However, these nodes (A and C, or in the example type of job and symptom 1) become conditionally independent if the connecting node (B, or sickness) is known for certain and becomes instantiated. (knowing a sickness has been observed subsequently means that knowing someone is employed in a given job will not revise the belief in an individual having symptom 1)

Figure 3.4 presents an example of a converging connection.

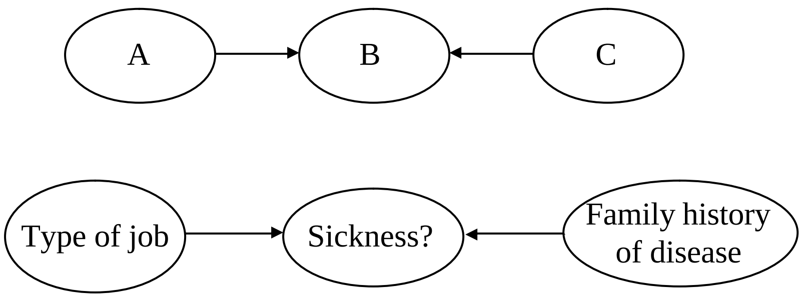


Figure 3.4 Example of converging connection

In a converging connection node B is dependent upon nodes A and C: that is, being diagnosed with a sickness is dependent upon observing type of job and any family history of disease. Consequently, evidence entered at node B will update both nodes A and C (diagnosing a sickness updates the belief in observing both of the type of job and belief in family history of disease). Evidence entered at node A will update B but will not feed up to node C. Lines of communication between A and C will only open when there is direct evidence on node B (or any of node B's descendants). When such information becomes available, nodes A and C become conditionally dependent. For example, knowing type of job, will update the belief in a sickness but not in the family history of disease. Only the certain diagnosis of a sickness will allow information on the type of job to update the belief in any potential family history of disease.

A diverging connection maybe (Figure 3.5):

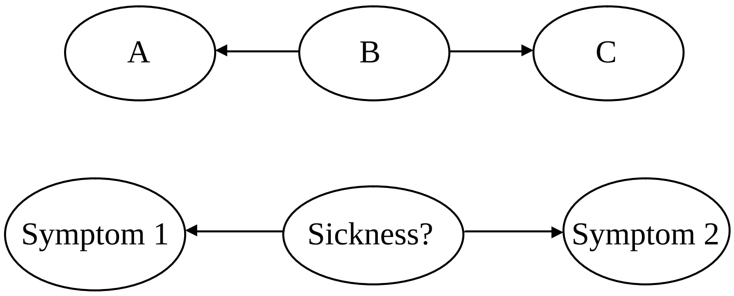


Figure 3.5 Example of a diverging connection.

In the example shown in Figure 3.5 example, both symptoms are dependent upon the likelihood of a sickness. Subsequently, information which updates the belief in a sickness, updates the belief in the presence (or absence) of symptoms as appropriate. Correspondingly, evidence entered at the end nodes (A and C) will inform the network about the state of node B and subsequently revise the belief in the remaining child node. However, when node B (sickness) is known for certain and becomes instantiated, nodes A and C (observing symptoms 1 and 2) become conditionally independent. At this point, evidence entered at either end of the connection will not update the belief in the remaining end node vice versa.

Figure 3.6 shows the previous connections combined into a single network. The three types of connections discussed can be used to support the development of a network. For example, with a given connection, if evidence becomes available, does it update the beliefs within the nodes the developer anticipated, or did the evidence lead to revision in unexpected nodes? Such assessments can lead to the restructuring of networks. Of course, it is prudent to submit a BBN to a review procedure.

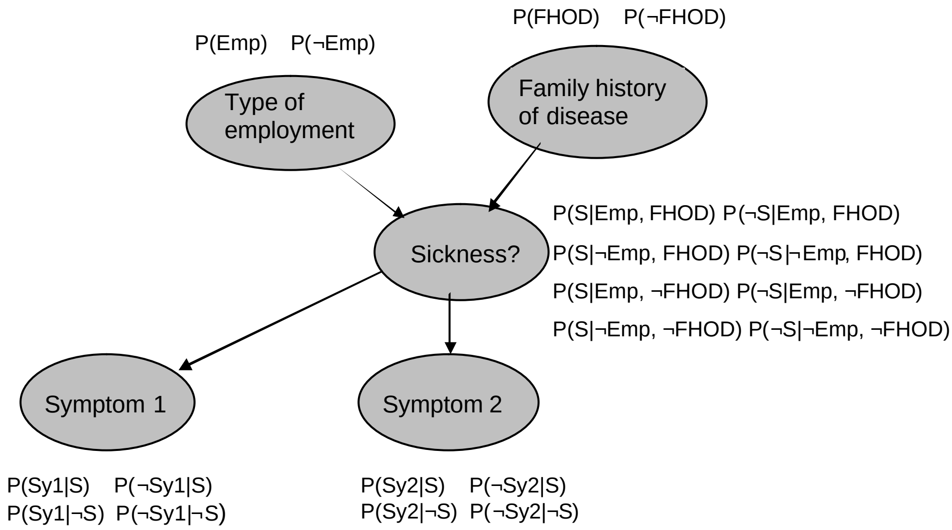


Figure 3.6 An example of a Bayesian Belief Network.

Upon agreement of the network structure, each variable within the BBN must be populated with its associated conditional probability distribution. Figure 3.6 also shows the values required to fully populate the BBN shown. Within this figure, the nodes 'type of employment' and 'family history of disease' are root nodes, that is no arcs feed directly into them, all the other nodes are child nodes. For root nodes, the probability distribution is simply the prior probability distribution, which for this example are the probabilities of:

- Working, or not working in given type of employment,
- Having, or not having a family history of a disease.

The conditional probability table for a child node includes the probability of the node attaining each of the defined states conditioned upon all possible combinations of its parent nodes. Thus, for the node 'symptom 1' the conditional probability distribution comprises the probabilities of:

- Displaying symptom 1 given a sickness (P(Sy1|S))
- Not displaying symptom 1 given a sickness (P(¬Sy1|S)
- Displaying symptom 1 given no sickness (P(Sy1|¬S))
- Not displaying symptom 1 given no sickness (P(¬Sy1|¬S))

Even within this relatively simple example, considering all the relevant factors to assess each of the required probabilities is not trivial. It is easy to see how the addition of only a small number of additional nodes can rapidly increase the number of probabilities required to fully populate a network and increase the burden of calculations across the network.

3.5 Propagation.

BBNs can quickly and efficiently calculate the required probability distributions and subsequently update the network values accordingly in response to many different types of queries. This process is referred to as propagation, and is the main strength of a BBN.

BBNs exploit the conditional independence within their structure to reduce both the number of:

- Probabilities required to fully specify a BBN, thus reducing the burden of elicitation.
- Computations required to answer many potential queries.

The exploitation of conditional independence means that the joint probability distribution is not always needed to answer a query within a BBN. This enables BBNs to be computationally efficient. However, if required, the joint probability distribution may be calculated through the use of the chain rule over the conditional probability tables. For a small BBN, it may be possible to calculate the joint probability distribution manually over the whole network and then marginalise to answer a specific query.

The problem of propagation in general is one which is non-deterministic polynomial time hard (NP hard). This is a class of problem which is intrinsically harder than those which can be solved by a non-deterministic Turing machine in polynomial time. The optimal solution to such problems is very computationally heavy (no algorithms for the rapid solutions of an NP hard problem exists) as all possible solutions must be fully assessed before selecting that which is indeed optimal.

Early propagation algorithms used a local distributed message passing architecture (Pearl, 1982 and Kim and Pearl, 1983 both cited in Pearl, 2000). These initial ideas were expanded and developed by Pearl (1986) himself and by Lauritzen and Spiegelhalter (1988). The latter of these authors developed the now commonly used method of junction-tree propagation. This algorithm, which is applicable to many probabilistic networks, decomposes a DAG into a series of sub-sets (based upon the d-connections within the network), known as cliques within which information flow is relatively contained. Subsequently, the cyclic links between cliques are eliminated to create a junction tree.

Within a junction tree information is passed between the cliques and not throughout the whole network. This enables rapid, efficient propagation of evidence and belief updates through the network. For a comprehensive overview of the calculations used within a BBN, the reader is referred to Pearl (1988) and Jensen (2001). Efficient propagation algorithms are readily available in BBN software such as HUGIN and NETICA.

An example of information flow within a BBN is shown below in Figures 3.7 and 3.8. The updating of beliefs within the network is triggered by evidence entering the network (Figure 3.7). Subsequently, in accordance with the laws of probability revision messages are sent out across the network (Figure 3.8).

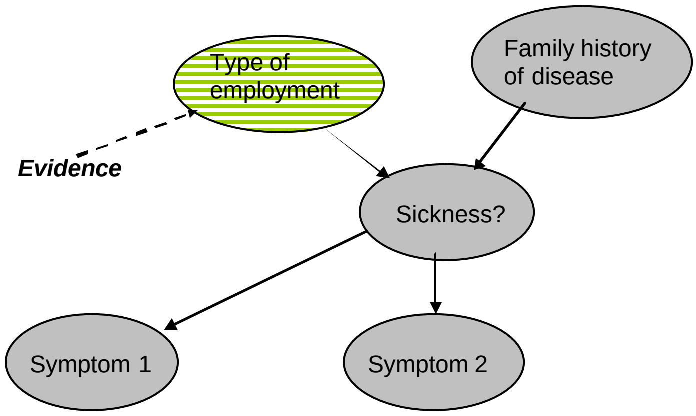


Figure 3.7: Evidence enters the BBN.

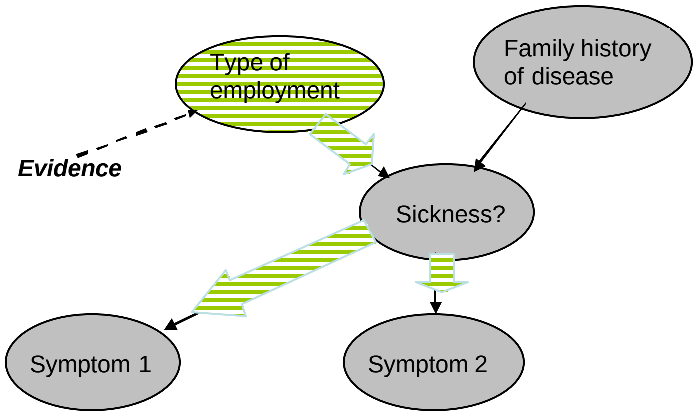


Figure 3.8: Node at which evidence enters updates and subsequently sends out belief revision messages.

As the belief messages reach the various nodes, they too update their belief distributions. Propagation terminates when no nodes remain to return a belief message. For a detailed description of propagation within a BBN the reader is referred to Pearl (1988).

3.6 Large-scale Bayesian Belief Networks

BBNs are applicable to a wide range of problems due to their capacity for diagnostic (from an observed event, identify the reason for it having occurred) and predictive (from an observation predict what will happen) reasoning. The representation of some problem domains can lead to BBNs (Neil, Fenton and Nielson 2000):

- Which are small enough for the causal directions on the edges to be obvious.
- Where the actual inferences made can run counter to the edge directions.

When considering many real-life complex problems, the causal directions between nodes may not always be obvious. It should be remembered that some dependencies may be domain dependent. However, although "“cause to effect” and “effect to cause” are mathematically equivalent…applying uniform interpretations are critical if we are to build large scale networks with meaningful semantics.” (Neil et al, 2000).

Complex problems, with their many interlinked variables may require a large-scale BBN. Potential improvements to the structuring of such networks may be through the use of automated techniques. Research by Klopotek (2005) focused on the development of a BBN comprising of up to 100,000 nodes through analysis of available text. Nevertheless, automated techniques still require human intervention in order to identify those BBN structures which are not realistic representations of the problem domain. Whilst it is feasible to develop large-scale BBNs it is important to note that the propagation algorithms developed by researchers such as Pearl (1986) become inefficient on such networks.

3.6.1 Network fragments.

An alternative approach to automated techniques is the structured development of large-scale BBNs through the use of ‘network fragments’. Pioneered by Laskey and Mahoney (1997), a network fragment is defined as “a set of related variables together with the knowledge about the probabilistic relationships among the variables.” (Laskey and Mahoney, 1997, page 335) The development of a large-scale BBN is achieved through the combination of smaller network fragments each developed by experts within a specific field on a problem domain. Such an approach complements scenarios in which there is a great deal of uncertainty about which variables may be important. As work progresses, variables identified as being important may simply be linked into the network, thus overcoming any initial lack of clarity. The capability to rapidly incorporate new variables is a great strength when undertaking real-time analysis.

Recent advances in the use of network fragments by Neil and Fenton (2005) considers the building of large-scale networks through the use of idioms. Differing from earlier work, the network fragments "…represent very generic types of uncertain reasoning…interested only in the graphical structure and not in any underlying probabilities and so for this reason an idiom is not a BN as such, but simply the graphical part of one." (Neil and Fenton, 2005, page 2). It is postulated that the use if idioms could be used to rapidly develop BBNs based upon a specific type of reasoning.

The structuring of a large-scale BBN must be carried out with consideration given to the elicitation of the required conditional probability tables. A variety of probability derivation techniques may be used, including the Noisy-OR distribution (which reduces the probabilities to be elicited), and the Noisy-MAX distribution (Jurgelenaite and Lucas, 2005, and Neil and Fenton, 2005).

3.6.2 Object-oriented methodologies.

Object-oriented approaches have been applied to the development and use of BBNs. The basic component of an object oriented approach is an object which at the simplest level, with respect to a BBN, would represent a variable (node). At a higher level of abstraction, an object could represent a network fragment which Weidl, Madsen and Israelson (2005) refer to as an instance object. Object oriented approaches provide for the inclusion of objects (sometimes hidden) within other objects. This approach enables a hierarchical structure of a problem to be developed (for example: nodes to network fragments to the overall BBN). Dawid, Mortera and Vicard (2005) make the point that a modular, hierarchical structure may be easily refined through the inclusion of objects, or the refinement of existing ones. As such, a model can be developed at an appropriate level of abstraction for the situation being analysed. Koller and Pfeffer (1997) developed an object oriented approach starting from the use of a stochastic functional language. Their work identified three main advantages to the approach, namely the:

- Development of model fragments which could be reused in a semantically meaningful way.
- Simple construction and modification of appropriate BBN fragments.
- Overall reduction in model complexity.

The structuring of a problem within an object oriented approach is facilitated by the existence of pattern catalogues which "have proved very useful to capture flexible and reusable object-orientated structures, promoting also a shared vocabulary for developers to talk about object-orientated design issues." (Berdún, Pace, Amandi and Campo, 2008). Such libraries do provide a researcher with some defined procedure for the development of a BBN. Their use, however, is dependent upon the researcher identifying the correct pattern for use in the problem of interest.

Initial work to extend the use of BBNs focused on the use of logic-programming like rules. However, the use of network fragments and object-oriented methods have found particular applicability within the fields of knowledge management, decision support systems and reasoning under uncertainty (Valtorta, Byrnes and Huhns, 2005). The flexibility of object-oriented technology was used by Wang and Chien (2003) to develop internet based group decision support systems allowing for the collation and sharing of ideas and opinions. Object oriented techniques have also been used to support the search for information at a suitable level of detail in time critical decisions (Tseng and Gmytrasiewicz, 2006).

Applications of object-oriented techniques within the military environment have been as diverse as modelling military organisation and behaviour (Suzic, 2003) to supporting the specifications required for future information fusion (Laskey and Costa, 2007). Current research on the use of object oriented Bayesian Networks to fuse information has also supported their use in providing situational awareness and intelligence analyses (Liao et al, 2003 and Pfatzu and Roth, 2006).

3.7 Bayesian Belief Networks in military intelligence analyses.

3.7.1 Background.

Intelligence analysis is not a precise science. Zlotnick (1972, page 43) succinctly states that "The very best that intelligence can do is to make the most of the evidence without making more of the evidence than it deserves." The use of probabilistic reasoning within intelligence analysis ensures:

- The explicit definition of uncertainties within the analysis from the beginning.
- The logical and consistent manipulation of uncertainty.
- Each piece of evidence contributes to the final outcome on its own merits.
- Multiple hypotheses are considered in a structured, fair manner.

Zlotnick (1967) was one of the first researchers to report upon the use of Bayes' theorem in predictive intelligence analysis simulations. Overall, Zlotnick (1967) considered the main strengths of using a Bayesian approach was its ability to support internally consistent analysis and concluded that "Mathematical processing will not become an alternative to present methods of intelligence analysis. It will become a reliability check on present methods. It will help show the plausibility of conclusions which the intelligence analyst would not otherwise recognize as compatible with the evidence and his own inner logic. It will tell the analyst: if you interpret the evidence in this way, then here is the conclusion you should probably reach." (Zlotnick, 1967, page 12).

The use of Bayes' theorem as a predictive intelligence technique was also considered by Fisk (1972). Within a realistic setting of a potential attack by the Union of the Soviet Socialist Republics (USSR) against China's nuclear capabilities, Fisk (1972) compared intelligence forecasts made through both the traditional and a Bayesian approach. Concluding, Fisk (1972) reported that whilst the Bayesian approach was not shown to be significantly more accurate, such an approach generated a clear audit trail of the uncertainties associated with information and, importantly, how and why such values had been allocated.

Schweitzer (1976) further extended the use of Bayes' theorem as a predictive methodology through its application to political analysis. During a relatively long term study (spanning three years from 1974 to 1976), Schweitzer applied Bayesian techniques in three relevant scenarios in which hostilities may commence (North Vietnam, Sino-Soviet relations and the Arab Israeli conflicts). The experiment required participants to give initial prior probabilities for a set of defined hypotheses. Subsequently, participants were provided with intelligence feeds based on openly available data sources. Based on the information provided, participants revised their belief in the probability distributions for the hypotheses set. Bayes' theorem was used to manipulate the probabilities given by the participants. During the course of the experiment no hostilities commenced and as such, the research could not assess the predictive utility of the methodology. Nevertheless, Schweitzer (1976) was able to report on the general advantages and limitations of the technique. In addition to those already outlined, Schweitzer (1976, paraphrased from page 40) identified the following advantages:

- The process provides a repeatable audit trail as to how a given assessment was made.
- Analysts must consider how intelligence affects the set of hypotheses and not just the assumed most likely hypothesis.
- Allows judgements to be made and reported on a numerical basis.
- The process was shown to be less conservative than informal opinions and moved the assigned probabilities away from 50/50 faster and farther then overall subjective judgements.
- Periodic assessments provide a degree of assurance that the problem is being monitored.

The following disadvantages to the use of a Bayesian technique were also identified as paraphrased from Schweitzer (1976, page 41):

- The question being considered must be formulated in mutually exclusive categories.
- The question itself must be amenable to definition as a series of hypothetical outcomes.
- A flow of data relevant to the question must be available for probability revisions to be made. Scarcity of data makes the overall technique less reliable.
- The question must not relate to an event which is largely chance or random.

For a further detail on these points, the reader is referred to Schweitzer (1976).

It is of interest to note that the work by Zlotnick, Fisk and Schweitzer was not computer based. Today, there is an increasing use of information technology, which is frequently networked, to turn large volumes of disparate data into pertinent information.

3.7.2 Predictive intelligence analysis.

The use of Bayesian methodologies within intelligence analysis continues to be an area of active research. In 2002 Paté-Cornell presented a BBN designed to support the analysis and fusion of disparate intelligence information. Within the research, two distinct issues relating to the fusion of information was noted. Firstly, that relevant, accurate information must be communicated in a timely manner to those who require it. Secondly and, as Paté-Cornell (2002, page 445) writes "perhaps more difficult, merging the content of the signal, some "sharp" and some "fuzzy" some dependent and some independent into useful information."

The use of Bayesian updating to fuse information to provide a quantitative intelligence assessment was further considered by McLaughlin and Paté-Cornell (2005). The results of the work (which was set within the search for WMD within Iraq) demonstrated that Bayesian analysis could be an effective tool for the support of intelligence analysis. Overall, the advantages identified by McLaughlin and Paté-Cornell (2005) support those outlined by Schweitzer (1976).

3.7.3 Identifying deception.

An important part of the potential knowledge to be gained about a situation is if the opposition is undertaking a programme of deception. Deceptions play on people's preconceptions and inability to systematically consider all the explanations for the evidence they observe. Within the intelligence community, deceptions which go undetected can lead to catastrophes. In response to this, Stech and Elsasser (2007) developed a semi-automated tool to support tactical decision-making through detecting deception and thus aiding intelligence analysis based on an extended version of Heuer's Analysis of Competing Hypotheses (ACH) (Heuer, 1999). Their work varies from many of its contemporaries in its consideration of specifically searching for signs of deception and counter deception. The result of the work is a tool which "yields the states that must be hidden, observations that have no probative value to the observer, and the states that one might simulate to deceive an adversary." (Stech and Elsasser, 2007).

3.7.4 Analysis of Competing Hypotheses and BBN.

ACH is a commonly used technique to ensure all hypotheses are considered equally. Part of the ACH process is the development of a matrix detailing the identified hypotheses and the identified supporting evidence. This matrix can be represented within a Bayesian Belief Network (for details of such an approach the reader is referred to Valtorta, Dang, Goradia, Huang and Huhns, 2005) enabling the logical reasoning under uncertainty to be assessed and dependencies among hypotheses to be revealed. However, before the BBN can be used, the prior probabilities must be evaluated and entered to allow for the calculation of accurate posterior probabilities. It is important to remember that individuals are prone to confirmation bias, and hypotheses are often generated based on the available evidence. Therefore, additional techniques may be required, alongside the use of the ACH to ensure that as many plausible hypotheses as possible are developed.

With specific reference to intelligence analysis, an additional interesting extension to Heuer's ACH has been in combination with the use of subjective logic (Pope, Jøsang, and McAnally, 2006). The combination of the two techniques "allows analysts to include weak diagnostic information that could be used to expose possible deceptions and provide more balanced conclusions as to expected outcomes." (Pope at al, 2006). A possible extension to such work would be the combination of ACH and BBN to indicate to analysts that their mental model of a situation being considered maybe incorrect and additional information should be sought.

3.8 Chapter summary.

This third chapter of the thesis has presented the methodological basis, development and application in military intelligence of Bayesian Belief Networks. Their robust mathematical basis, ease of building and ability to accept data of varying certainty makes them an ideal tool for use in military intelligence applications.

The following chapter presents an initial investigation into the development of a DSS based upon BBNs.

CHAPTER 4 : INVESTIGATIONS INTO THE APPLICABILITY OF BBNs AS DECISION SUPPORT SYSTEMS.

4.1 Introduction.

Chapter Three introduced Bayesian Belief Networks (BBNs) and discussed their application within military intelligence analysis. There is a growing interest in the development of decision support systems designed to help individuals who have to collate, analyse and draw conclusions from large volumes of disparate data. Such systems are not intended to replace the human analyst, merely to enhance their capabilities by providing a clear audit trail and a second opinion. If both the analyst and decision support system arrive at the same conclusion there is extra credence in the assessment made. However, any inconsistencies between the two results should lead to the analyst re-evaluating their work. This will ensure the analyst understands why there is a difference within the results which subsequently reduces the chance of an oversight or error within the analysis on their part.

The research presented within this Chapter adds to the body of knowledge on the potential use of BBNs for the quantitative analysis of military intelligence. It is important to note that the BBN network developed here was not intended to provide detailed analysis of the raw data streams provided from surveillance equipment. Alternatively, the developed BBN is designed to support a commanding officer or intelligence analyst through their decision-making cycle. It is anticipated that the use of the BBN will ensure appropriate consideration of all available information and assist the individual in drawing their own informed conclusions about the situation being faced.

For the purposes of this research, the use of BBNs for intelligence analyses is presented within a fictitious military scenario. The developed scenario considers a land-based section attack against an enemy outpost which may be reinforced. The focus of the research reported within this Chapter was the impact information had, as it became available, on the participants' perceptions of the situation. Specifically, it was of interest to determine if individuals could logically and consistently combine various probabilities in order to determine their own direct subjective probability of an hypothesis being true. Numerous psychological studies have shown that this is an area where humans often perform weakly due to numerous subconscious biases (for further descriptions of common decision-making biases and heuristics see Chapter Two).

Therefore, the research presented here sought to investigate:

- The impact of additional information upon individual perceptions of a situation.
- The ability of individuals to logically and consistently combine the available evidence.
- The ability of individuals to assess the discriminative value of information sources.

4.2 Scenario and Bayesian Belief Networks.

The scenario developed as part of this research is representative of a land based attack. In the scenario, the United Kingdom (UK) had been deployed and tasked with protecting a country with which it has good relations and is of strategic importance to the UK. The mission was to advance and defeat all enemy outposts which were within territorial borders and over section[2] strength, to secure the area of interest and the lines of communication.

Within the scenario, (fully detailed in Annex A) participants were told that a reconnaissance group reported two enemy outposts, believed to be of section strength in the area of interest. There was, however, some ambiguity. Following further reconnaissance reports, it became highly probable that one of the two outposts had been reinforced. Participants were instructed to imagine that they were heading towards one of the located outposts. Each of the two outposts had an equally likely

[2] A 'section' consists of two fire teams totaling 8 men (in the USA referred to as Squads)

chance of being reinforced. This gave a fifty percent probability that the participant was advancing towards the reinforced outpost.

Ambiguity within the scenario enabled participants to reach their own conclusions relating to: the potential threat posed by the outpost as well as the level of technical support and the tactics available to themselves. Since the focus of the experiment was the subjective assessments, any inconsistencies between assumptions made by individual participants did not impact upon the results obtained in any subsequent analysis. Comparisons were only made between each participant's direct subjective values and the corresponding Bayesian values. Had the participants been logical and consistent in their assessment of the probabilities there would be little difference between the values. Any difference between the values provided a measure of their own inconsistencies.

A BBN was constructed for use with the land based attack scenario. Detailed in Figure 4.1 the main node of interest was the hypothesis node ' Reinforced'. The states of this node relate as to whether or not the outpost (to which the participant was advancing) was in fact reinforced.

Within the scenario, three types of military assets could be reported as having been sighted at the outpost, namely:

- Anti-tank weapons (ATK).[3]
- Armour.[4]
- Mortars.[5]

The reporting of any of these assets is dependent upon whether or not the outpost had been reinforced. This can be seen in Figure 4.1 where the hypothesis node is the

[3] Within the UK Army the main anti-tank weapon is Javelin.
[4] Armour covers a wide range of capabilities including tanks, armoured vehicles such as the Warrior and reconnaissance armoured vehicles such as the Scimitar.
[5] Mortars are used to provide fire support in exact locations as requested by the commanding officer.

parent not to the nodes representing the presence of each asset. The numbers shown in Figure 4.1 are for illustrative purposes only.

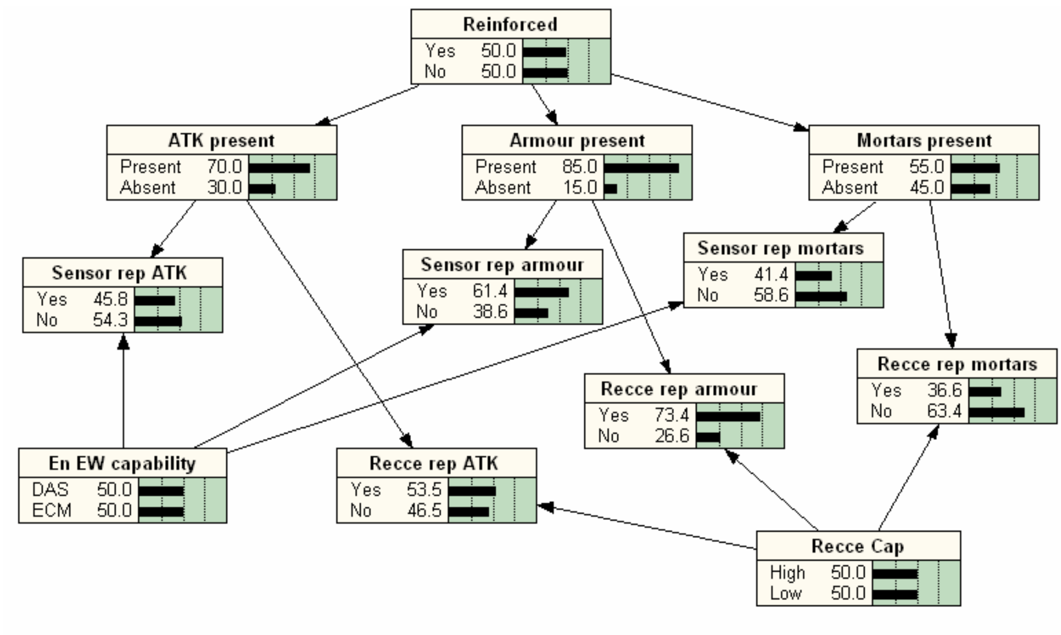


Figure 4.1: A Bayesian Belief Network of the relationships between the enemy outpost being reinforced and the intelligence reports available [6]**.**

Within a military situation, a commander or analyst rarely has access to the 'ground truth'. Therefore, it was imperative that the BBN include a representation of:

- Reconnaissance capability (Recce Cap).
- Enemy electronic warfare capability (En EW capability) as this will impact upon the utility of sensors.

It is this 'middle layer' of nodes relating to the reporting of assets present at a reinforced outpost that were of most interest to this part of the research.

As part of the experiment, participants were told the level of reconnaissance capability (high or low), to be expected from their own side which reflected good or poor

[6] ATK is an abbreviation of Anti Tank Weapons. Rep is an abbreviation of Report. Recce is an abbreviation of reconnaissance. Cap is an abbreviation of Capability.

visibility conditions. They were also told of the enemy's electronic warfare capability. Various pieces of intelligence were then provided to the participants in the form of reconnaissance group or electronic sensor reports.

To elicit the conditional probabilities required to populate the developed BBN, an experiment was developed which comprised two questionnaires (detailed in Annex A) both of which elicited a series of numerical values. The first questionnaire focused on the conditional probability distributions required to populate the BBN. The second questionnaire posed a series of questions, the answers to which could be logically derived from the values provided in the conditional probability distributions. However, the participants were not made aware of this fact. The questions posed required the participant to undertake mainly diagnostic reasoning (e.g. considering the scenario, what in your judgement is the probability that the sensor reports mortars being present at the enemy outpost?). As part of the experiment, participants were also asked for their experience of working with intelligence reports which may have developed their skills in combining multiple data sources.

To provide a proof of principle execution of the scenario, BBN and developed questionnaires, several pilot studies were undertaken. All participants were serving military officers undertaking study at Cranfield University, Defence College of Management and Technology. The pilot studies provided an opportunity to refine the presentation and explanation of the experiment to the participants and provided several interesting results.

It was concluded that the experiment should only be completed by serving U.K. Army officers. Furthermore, to ensure that all the officers could understand the scenario and had an equal level of training from which to develop a response participants had to hold the rank of Major or above. Firstly, to have reached the rank of Major the participants will have worked within the army for a substantial length of time. This provided a level of experience upon which the participants may call to provide answers posed within the questionnaires. However, it was also important that the participants had a common level of training. Whilst this is initially provided by officer training at the Royal Military College at Sandhurst, following this officers undertake a period of initial officer training in their own area (e.g. logistics) and then subsequently specialise (e.g. in catering) before taking captaincy exams. Following training at Sandhurst, the first common training for all officers occurs at the level of Major. Therefore, to ensure a consistency in the level of expertise, gained both through experience in serving in the military and formal training, all participants were required to hold the rank of Major.

The pilot studies also revealed the importance of running the experiment, whenever possible 'face to face'. This approach allowed for any points of ambiguity to be clarified as the experiment progressed

Building upon the results of the pilot study, an experiment was conducted with 13 participants, all holding the rank of Major or above. All participants were serving Army Officers undertaking post-graduate study at Cranfield University, at the Defence College of Management and Technology. Whenever possible, experiments were conducted 'face to face' with participants. However, this approach was not always possible. In such cases, the experiments were sent out via e-mail and participants were encouraged to ask for clarification. It was noted that the difference in data collection techniques may impact upon the overall results obtained.

The conditional probability distributions were used to populate the developed BBN for each participant. The combined the available probability distributions to calculate the Bayesian (normative) probability of the outpost being reinforced. The BBN was interrogated by entering evidence (such as low reconnaissance capability, sensors report mortars at outpost) into the network. The values of interest calculated by the BBN was compared against the equivalent direct responses provided by the participants (given in the second questionnaire detailed within Annex A). The deviation between the Bayesian and corresponding direct probability was calculated and provided a measure of the participants own individual inconsistencies. These deviations were analysed to determine if any trends or biases were apparent.

4.3 Results and discussion.

4.3.1 Investigations into the probability of an outpost being reinforced.

The main focus of this investigation was the effect of additional information on each participant's perception of the outpost being reinforced. Intelligence reports acted as the mechanism to provide additional information on the reported sightings of assets at an outpost. Within the scenario, the assets potentially sighted were: armour, anti-tank weapons (ATK) and mortars.

Representative results for an individual participant are shown in Figure 4.2. This figure clearly shows a continual increase in the probability of the outpost being reinforced for both the Bayesian and direct subjective probability. The result shows fairly close direct subjective and Bayesian probabilities when only one piece of information is available. However, as additional information becomes available, the participant begins to over-adjust the probability of the outpost being reinforced. This is shown by the increasing difference between the direct subjective and Bayesian probabilities. The over-adjustment is suggestive of the participant being unable to correctly take account of the interdependencies between the various assets sighted at the outpost.

For the results shown, the Bayesian values, derived from the supporting BBN, support the belief of the participant: the outpost is probably reinforced. However, a decision support system based upon this BBN could note caution as to the magnitude of the likelihood of this compared to that perceived by the participant. This could assist the participant's decision on how to commence any attack against the outpost as well as affecting attitude to the status of the second reported outpost.

.

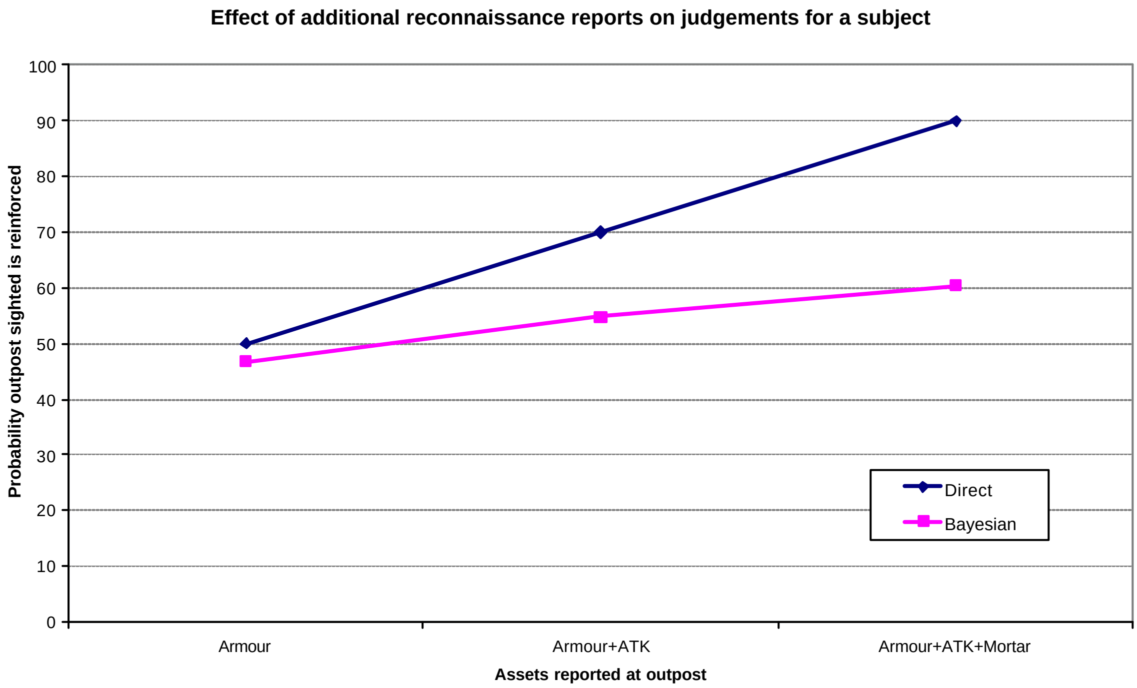


Figure 4.2: Effect of additional reconnaissance reports on the Bayesian and direct probabilities of the outpost being reinforced.

Following the result shown in Figure 4.2 it was of interest to determine if these results were replicated across the group of results obtained. This was achieved by calculating each participant's deviation (difference between the direct subjective and Bayesian probability) as each intelligence report became available. Subsequently, the mean deviation for the group of participants was calculated as each intelligence report was obtained. This result, along with the maximum and minimum deviation within the group as each intelligence report became available, is shown in Figure 4.3. It can be seen that there is a continual increase in the deviation between the direct and Bayesian probability as additional assets are sighted at the outpost supports the findings shown in Figure 4.2. Interestingly, there was little difference in the highest calculated deviation at each intelligence report (range from 34 to 39), however, the minimum calculated deviation ranged from -57 to -2.

The results shown in both Figures 4.2 and 4.3 could be indicative of the representative heuristic. This bias is most likely to have been caused by the participant's mental model of the situation being different from that shown in the BBN, for example other hitherto unvoiced factors affect their perception of the required probabilities.

It is considered possible that the participant's perception of the size of the enemy force located at the outpost was affecting their estimates of the likelihood of the outpost being reinforced. An outpost that is manned at section strength would be unlikely to be supported by all three assets of armour, ATK and mortars. Hence, if all three assets are located at the outpost it is possible that there would be more than a section of men (and possibly other assets) located there which have not yet been seen by the reconnaissance group.

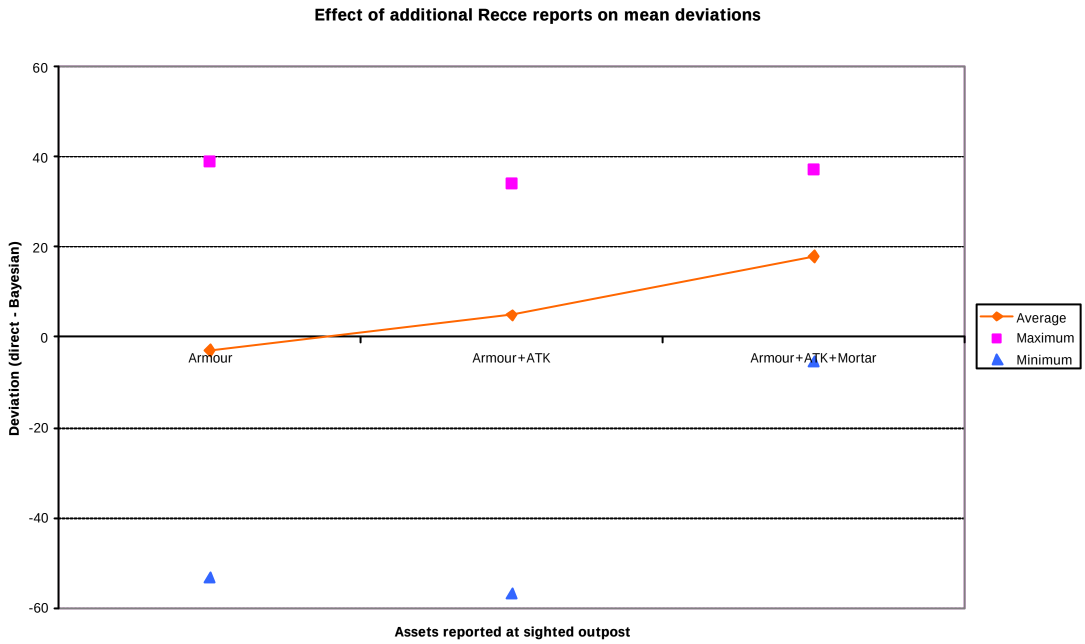


Figure 4.3: Effect of additional reconnaissance reports on the mean deviation of the probability of the enemy outpost being reinforced.

Due to the small number of experimental results it was not possible to assume that the Central Limit Theorem would apply. It was also not possible to assume normally distributed data, hence a non-parametric statistical test was used. The Wilcoxon signed rank test (Panik, 2005) was selected for use and applied to the calculated deviations. The Bonferroni correction (Berk and Carey, 1999) was applied to attain an overall significance level of 5%.

Based on the initial sample size of thirteen participants, three questions produced deviations between direct and Bayesian probabilities which were found to be of statistical significance. The deviations of interest related to the probabilities when the reconnaissance capability was low and that:

- ATK was present given that the reconnaissance group report ATK present, Reconnaissance capability low.
- ATK was present given that the reconnaissance group reports no ATK, Sensors report no ATK.
- The outpost was reinforced given that the reconnaissance group reports ATK mortars and armour.

The collated results were also analysed for evidence of the inverse fallacy. This heuristic occurs when individuals confuse the probability of event A occurring given event B already has (written as $P(A \mid B)$ and for example may be the probability that the grass is wet given that it has rained) and its inverse, the probability of event B occurring given event A already has (written as $P(B \mid A)$ and continuing the example may be the probability that it has rained and therefore the grass is wet). Consequently, individuals inadvertently give the wrong probability.

As previously explained, the conditional probability distributions were used to populate the developed BBN for each participant. The BBN was interrogated by entering evidence through a series of available intelligence reports into the network. The normative values calculated by the BBN were compared against the equivalent direct responses provided by the participant based upon the information available. The conditional probability tables also allowed for the calculation of the inverse fallacy probability for each of the interrogations entered in the BBN and for which a direct response had also been provided.

A comparison of the probabilities obtained from the interrogation of the BBN against the equivalent direct responses elicited from the participants found one direct match between the results. This match, in which the participant is said to have provided a normative, or Bayesian response, is unlikely to have been caused by the participant being perfectly rational. The more probable explanation is that the probability calculated by the BBN was caused either by rounding errors, or have been inevitable due to the conditional probability tables elicited from the participant. Indeed, had the participant been perfectly rational, more direct matches would have been anticipated.

4.3.2 Further investigations into the probability an outpost is reinforced.

It was considered important to investigate further the conjecture that when participants were asked "what is the probability that the outpost has been reinforced?" the response being given was actually a response to the question "*by how much* has the outpost been reinforced?” Based on the scenario, it was necessary to determine what force strength the participants perceived to be located at the outpost if all three assets were reported as being present. Furthermore, based on this perceived force size what, if any, other assets would the participants expect to be present at the outpost which may have not yet been identified. It is entirely feasible that the participants had a preconceived idea of the strength of the outpost and the additional information simply served to reinforce their initial opinion.

To fully investigate this conjecture a slight alteration was made to the subjective conditional probability questionnaire. An additional question was included which was designed to elicit the participant’s perception of the force present at the enemy outpost. No changes were made to the BBN developed in support of the experiment. The experiment was subsequently re-run with an additional ten participants.

The results obtained disclosed that nine of the ten participants considered that an outpost of section strength (as included in the experimental scenario) would have some form of anti-tank weapon. Therefore, the sighting of anti-tank weapons should not impact upon the probability of the outpost being reinforced. Further to this, all the participants believed that there was at least a 50% probability of the outpost having being reinforced by the enemy, prior to it have being located by a reconnaissance group. Hence, a reinforced outpost will have signs of additional support. In total seven participants believed that the outpost would be manned by more than a section of men. Overall, the perceived additional support anticipated as being present at the outpost ranged from men and assets taken from a mortar platoon to a full company of soldiers.

In line with the earlier analysis, the BBN developed in support of the experiment was populated and interrogated for each individual participant. Again, for each participant the deviation (the difference between the direct subjective and Bayesian probability) as each intelligence report became available was ascertained. Subsequently, the mean deviation for the group of participants was calculated as each intelligence report became available. This result is shown in Figure 4.4. As seen in the earlier study, the sighting of a single asset was associated with only a small deviation between the direct subjective and Bayesian probability. However, as additional assets were sighted at the outpost, the deviations between the two probabilities increased.

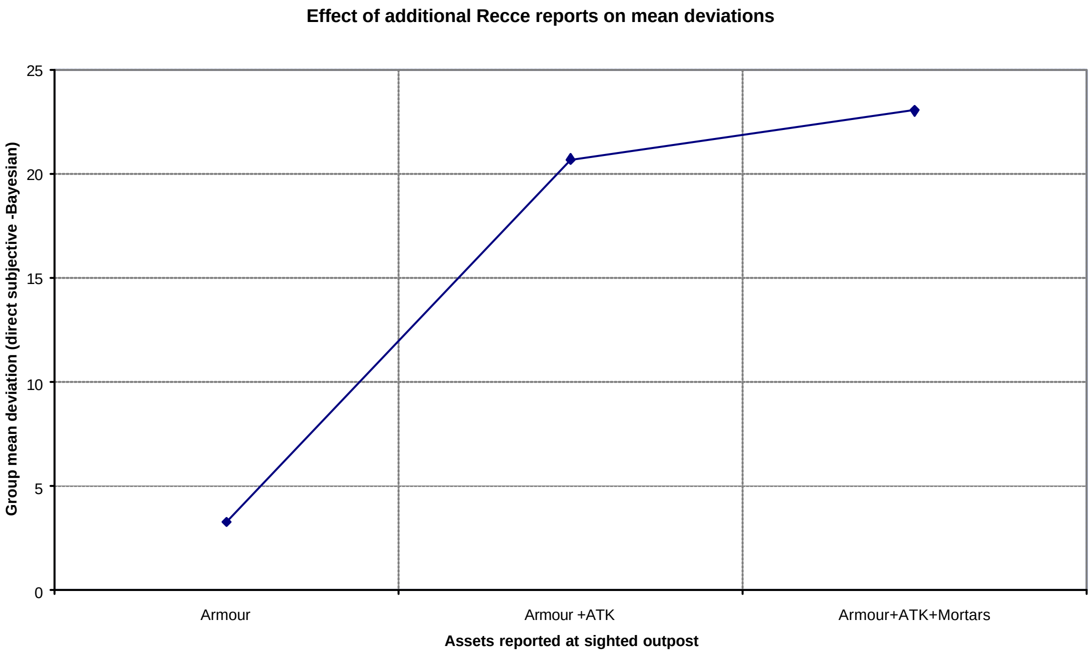


Figure 4.4: Effect of additional reconnaissance reports on the mean deviation of the probability of the enemy outpost being reinforced.

As previously mentioned, nine of the participants expected ATK to be present at the outpost. Therefore, the sighting of both armour and ATK contained within the second intelligence report should not have led participants to substantially revise the probability of the outpost being reinforced from the value assigned when only armour had been sighted. However, as the intelligence reports included the sighting of additional assets at the outpost, participants seemed unable to correctly manipulate the dependencies between the assets. As such, as additional information became available, participants began to over-estimate the probability of the sighted outpost being reinforced.

Overall, the analysis of the results obtained from the revised experiment showed that the participants' perception of the enemy force present at the outpost had no discernible effect on the measured difference between the Bayesian and direct probabilities obtained. An alternative reason for the observed deviations may be due to the structure of the network. It is considered that the most likely explanation of the deviation is the inability of the participants to correctly manipulate the interdependencies between the available intelligence sources. This resulted in the participants over-estimating the probability that the outpost was reinforced. This could subsequently impact upon the decisions taken on how to mount an attack against the outpost.

4.3.2.1 Preferred intelligence report.

Following on from the above results, investigations were carried out into the participants' preferences for additional information during the revised experiment. Participants were questioned as to whether they would prefer a reconnaissance or sensor report to decide if:

- Armour was present at the outpost.
- The outpost was reinforced.

Eight of the ten participants questioned, preferred the reconnaissance group report to assist the decision of whether armour was present at the outpost.
Figure 4.5 shows the probability of armour being present at the enemy outpost based on a positive sensor report. Analysis of the most valuable report was undertaken

using the ‘odds ratio’ which is the ratio of probability of A divided by the probability of not A. In this case, the odds ratio is the ratio of the probabilities of armour present/armour not present. This ratio based on the probabilities of the sensor report of armour being present at the outpost is 83.1/16.9=4.29. The corresponding ratio of the probabilities of armour pesent/not present based on a reconnaissance report is 9.83 as shown in Figure 4.6. Consequently, the reconnaissance report is more valuable.

All the participants networks were interrogated to determine which report was the most valuable. The results of the analysis are shown in Table 4.1 with the most influential report being highlighted in bold. Of the eight participants who stated a preference for the reconnaissance report to determine the probability of armour being present at the outpost, seven were found to have this as most influential report in their BBN.

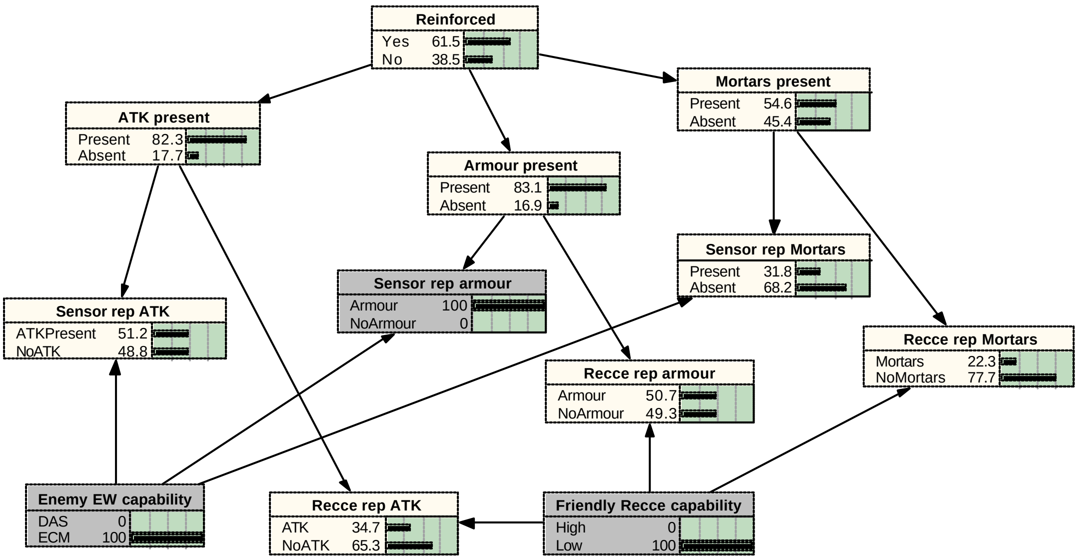


Figure 4.5: Probability armour is present at the enemy outpost based on low reconnaissance capability, presence of enemy electronic counter measures and a sensor report of armour.

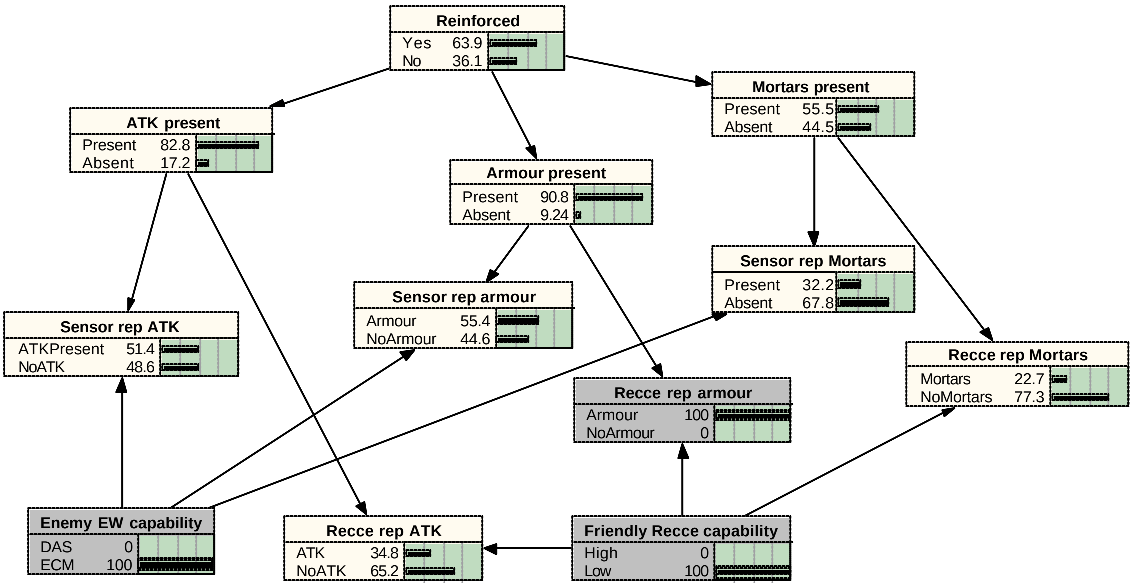


Figure 4.6: Probability armour is present at the enemy outpost based on low reconnaissance capability, presence of enemy electronic counter measures and a reconnaissance report of armour.

	Ratio of armour present/not present based upon reports from**:**	
Participant	Sensors	Reconnaissance Group
1	**36.04**	5.99
2	**14.85**	5.58
3	1.86	**4.65**
4	0.43	**3.29**
5	1.84	**5.17**
6	1.99	**3.72**
7	3.31	3.31
8	0.67	**1.56**
9	4.92	**9.83**
10	1.78	**17.09**

Table 4.1: Ratio of probabilities of armour present/armour not present.

In Table 4.1, a ratio of 1 implies that the report from the sensor did not affect the probability of armour being present at the outpost. A ratio of less than 1 indicates that the probability of armour being present at the outpost is highest. Conversely, a ratio of less than 1 reveals the higher probability to be that armour is not present at the outpost. Participants 1 and 2 show very high ratios for the sensor report ratio of the probabilities of armour present /armour not present. These ratios were caused by the conditional probability distributions elicited from the participants and used to populate their own BBN. The distributions provided by participant 1, when combined in the BBN calculated that for: low reconnaissance capability, with enemy EW having ECM capabilities and the sensor report indicated armour, the probability of armour being present at the outpost was a 97.3% and a 2.7% probability of armour not being present. Interestingly, participant 7 shows no difference in the ratio for the probabilities of armour present / armour not present for the sensor or reconnaissance report. This means that each report had an equal impact upon the probability of armour being present based upon their elicited conditional probability distributions (a positive report on armour from either the sensors or reconnaissance group gave a probability that armour being present as 76.8% and of not being present as 23.2%)

Participants were also questioned as to whether they would prefer a reconnaissance or sensor report to decide if the outpost was reinforced. Seven of the ten participants preferred a reconnaissance group report. Table 4.2 shows the ratio for the probabilities of an asset being present/not present for both the sensor and reconnaissance group reports. For each participant, the most valuable report for each asset is shown in bold. For example, considering participant 1, the most influential report of the sighting of ATK and armour comes from sensor reports with a reconnaissance group report having most influence on the belief as to whether or not the outpost was reinforced with mortars.

In contrast, the results for participant 2 show the sensor report being the most valuable with respect to deciding if ATK were present and a reconnaissance group report being most influential for the presence of armour. Interestingly, participant 2 shows no difference in the value between a reconnaissance or sensor reports on the presence of mortars. Participant 7 show no difference in the value between the intelligence reports for all three assets. As discussed in the results presented in Table 4.1, this is due to the conditional probability tables provided by the participant giving equal importance to the two sources. As such, a positive report from each source makes an equal impact upon the probability of armour being present or not present at the outpost. Interestingly, for the reports on armour the reconnaissance report was the most valuable in 9 of the BBNs for determining if the outpost was reinforced.

	Report on ATK		Report on Armour		Report on mortars	
Participant	Sensor	Recce	Sensor	Recce	Sensor	Recce
1	**1.06**	1.05	**1.22**	1.84	2.14	**2.57**
2	**1.36**	1.22	**1.47**	1.30	1.10	1.10
3	1.00	1.00	1.00	**1.26**	1.00	1.00
4	1.00	1.00	1.00	**2.65**	1.00	**1.09**
5	1.02	**1.05**	1.25	**1.79**	1.00	1.00
6	0.95	**1.22**	1.02	**1.20**	1.04	**1.12**
7	1.14	1.14	1.56	1.56	1.25	1.25
8	1.0	**1.16**	1.00	**1.70**	1.00	1.00
9	**1.22**	1.19	1.60	**1.77**	1.72	**2.23**
10	**1.83**	1.68	0.78	**1.69**	0.69	**0.90**

Table 4.2: Ratio of probabilities of outpost reinforced/outpost not reinforced.

A ratio of probabilities equal to 1 implies that the report did not effect the probability of the outpost being reinforced (which thus remained at 50%) A ratio of probabilities which is less than 1 indicates that the reporting of an asset in such a report (for example participant 6 with the sensor report on ATK) reduces the probability of the outpost being reinforced to below that of the outpost not being reinforced.

As for the initial experiment, due to the small number of experimental results it was not possible to assume that the Central Limit Theorem would apply. It was also not possible to assume normally distributed data, hence a non-parametric statistical test was used. The Wilcoxon signed rank (Panik, 2005) test on the observed deviations was calculated. The Bonferroni correction (Berk and Carey, 1999) was applied to attain an overall significance level of 5%. Two questions led to deviations which were found to be statistically significant:

- If the Recce group report ATK weapons and armour but does not report mortars, what is the probability that the outpost has been reinforced?
- If the Recce group report ATK weapons, mortars and armour, what is the probability that the outpost has been reinforced?

4.3.3 Selection of the most appropriate intelligence report.

A further experiment was conducted to investigate the selection of a preferred intelligence report. Two additional changes were made to the modified experiment used to investigate the preferred intelligence report, namely:

- At no detriment to the experimental results and in order to shorten the time required to complete the experiment, a reduced set of conditional probability tables were elicited.
- To facilitate analysis an additional two questions were included:
  - If you could have either the sensor report or the reconnaissance report, which would you choose to help you decide whether armour was present at the outpost?
  - If you could have any single sensor or reconnaissance report, which would you choose to help you decide whether the outpost had been reinforced?

The six possible answers to the second question listed above were the reconnaissance or sensor report on: armour, ATK weapons or mortars. Of the seven participants, who completed the experiment, six indicated that they would prefer the reconnaissance group report to help them decide if armour was present at the outpost. However, the results actually showed that for six of the participants it was in fact the sensor report which was the most valuable.

It is unclear as to why there is a difference between the preferred report started by the participant through direct questioning and that indicated by the BBN based on their elicited conditional probabilities. It is possible that the participants did not have experience of working with reports collated from electronic sensors. Hence, the participants may prefer to rely on highly skilled reconnaissance teams, with whom they may have served for a substantial length of time.

4.4 Conclusions and chapter summary.

This research has sought to investigate the following objectives:

- The impact of additional information upon individual's perceptions of a situation.
- An individual's ability to logically and consistently combine the available evidence.
- The ability of individuals to assess the discriminative value of information sources.

The developed decision support system was not intended to tell an individual which decision to take: the output of the system is not the 'right or wrong' decision. The decision support system introduces objectivity and an audit trail of how a decision was taken. It is anticipated that this will strengthen the decision-making process by ensuring a fair assessment of all the available information and thereby reducing the inherent risk within the decision-making process.

To support the research into the above objectives a fictitious scenario representative of a land based attack and supporting BBN were developed. The basis of the scenario was an attack against an enemy outpost which, although perceived to be manned at section strength, had a fifty percent probability of having been reinforced. The focus of the research was the impact of additional information, as it became available during the situation, had on an individual's belief that the outpost had been reinforced. A Bayesian Belief Network was used to provide a normative model of the situation and provide the Bayesian probability of the outpost being reinforced.

Analysis of the results highlighted a sharp increase in the mean deviation between the direct and Bayesian probabilities when two or more assets were sighted at the outpost. This is likely to have been caused by participants incorrectly manipulating the dependencies between the assets sighted at the outpost. The deviations may also have been caused by differences between the structure of the BBN and the participants mental model of the situation. Therefore, additional research must carefully consider the development and structure of any BBNs. In conclusion, the availability of additional information did alter the participants' view of the situation. However, they were unable to logically and consistently combine the available information to make accurate assessments of the situation being faced. What remains unclear is the impact the decision support system would have had on subsequent decisions relating to courses of action on the attack of the outpost. The results show that the participants had strong mental assumptions relating to the outpost and the importance of information received from human intelligence sources. It is possible that the results presented from the network would not have changed the opinion of the participants.

The participants in the experiment tended to prefer the reconnaissance report to the sensor report when offered a choice. However, the BBNs developed from the participants' conditional probability tables frequently showed the sensor report to be the most valuable. It is unclear why this trend was observed but may have been caused by participants' careers leading them to have had more contact with reconnaissance reports than sensor reports.

The results obtained indicated that participants were not able to give sufficient credence to information received by electronic methods. It was unclear why this trend was observed. However, there is merit in the development of a system which could logically and consistently combine both human and signal intelligence to aid analysts' understanding of a situation and ensure accurate consideration of all available intelligence. As such, it can be concluded that there is merit in the design of a BBN decision support system which could support intelligence analysis by:

- Ensuring reports from all sources are fairly considered and evaluated.

- Indicate, based on the available information the most plausible hypotheses.

This latter point would, in theory enable a decision-maker to identify when their view of the situation may require re-consideration or when additional information would be required. These points will be investigated further in Chapter Five.

CHAPTER 5 : CASE STUDY: APPLICATION OF BBNs AS DECISION SUPPORT TOOLS TO THE ARAB ISRAELI CONFLICT OF 1973.

5.1 Introduction.

Chapter Four of this thesis investigated areas where a BBN used as a DSS could be of most use. The results of the investigation concluded that:

- Experimental participants were unable to logically and consistently combine multiple data sources.
- There was evidence to suggest participants were not able to give proper attention to data from electronic sources.
- There was merit in the development of a decision support system able to indicate the most plausible hypothesis and how this may change over time with the available information and derived subjective conditional probability distributions.

To further investigate and address these strands of research, this chapter presents a BBN as a decision support system within the context of the main events leading up to the start of the Arab Israeli conflict of 1973. Through the use of hypothetical probability estimates the case study shows where and how such a tool could have been of use to the Israeli intelligence forces by addressing the above points of interest.

The timeline used for the analysis is comprised from entirely open source data and, of course, has been constructed with the luxury of hindsight. Therefore, the work presented here is in no way meant to provide a comparison of the decisions taken during the actual events discussed here. Those intelligence analysts and military personnel present at the time of the events would have had access to additional, classified information. Furthermore, any decisions taken at that time would have been made under the pressures and political constraints of the time. However, substantial effort has been given to ensuring that the timeline, BBN and conditional probabilities used for this research are as realistic as possible. Overall, the purpose of the research is to illustrate the potential benefits and strengths, as well as any weaknesses in the use of BBNs for intelligence analyses.

5.2 Development of the timeline and Bayesian Belief Networks of the main events leading up to the start of the 1973 Arab Israeli conflict.

Prior to the start of the 1973 conflict, commonly known as the Yom Kippur war, Israel was renowned for its exceptional intelligence service and capabilities. How then, was the nation taken by a surprise attack? Hughes-Wilson (2004, page 257) notes that following the conflict of 1967 that Israel "underestimated her enemy".

Indeed hindsight had shown that there was no failure to gather intelligence which suggested an impending attack. Troop movements, the location of bridging equipment and the call up of reserves amongst other signals were all noted. The information was there. It was "Israel's overconfidence that had prevented accurate intelligence assessments." (Bickerton and Klausner, 1991. page 178). The combination of overconfidence and underestimating the enemy "led directly to the [second] Israeli mistake: a curious inability to draw the right conclusions from a given set of facts. Thus the simultaneous build-ups in Syria and Egypt never appear to have been linked. The likelihood of an attack was just plain ignored. There seemed to be an absolute assumption that the Arabs could not attack until Israel's own political criteria for any Arab attack had been met." (Hughes-Wilson, 2004, page 258).

The purpose of this chapter is not to give a detailed account of the history leading up to, nor the facts of, the Yom Kippur war. An overwhelming body of literature exists on these topics. However, the interested reader is referenced in particular to Hughes-Wilson (2004), Bickerton and Klausner (1991) for excellent comprehensive accounts of the war from the both the military and political perspectives. In addition to these Wagner (1974) provides a fascinating analysis of the decisions taken by the Israeli Government in 1973.

Based on an open source literature review, the research presented here considers, two collectively exhaustive and mutually exclusive hypotheses. Together, these hypotheses represent the intent of the Egyptian and Syrian coalition at various points from 1967 to 1973. The hypotheses considered were:

Hypothesis A Hostile: The Egyptian and Syrian coalition will undertake a hostile attack against Israel to achieve their aims.

Hypothesis B Peaceful: The Egyptian and Syrian coalition will not undertake a hostile attack against Israel to achieve their aims.

Table 5.1 presents the main military and political events leading to the start of hostilities in the Arab Israeli conflict of 1973. Whilst many sources are available, the information included below is a summary of the information given in Buckwalter (2002).

Date	Action
Background	During the Six Day War Egypt suffered territorial losses. Following this there was a period of 'no war – no peace' which greatly affected Egypt's economy through the loss of revenue from the closure of the Suez Canal. President Sadat of Egypt felt he had to take action. President Sadat of Egypt and President Assad of Syria agreed a simultaneous attack on two fronts against Israel with the aim of removing Israelis from the captured lands (Operation Badr). Egypt secured additional arms in exchange for an extension of the Soviet Egyptian Naval Agreement.
Dec-71 to May-73	Egypt creates numerous minor and three major military scares in Israel by: Mobilising sections of their reserves; Moving bridging material to the Suez canal; Preparing crossing places and Concentrating tanks, artillery and troops.
Apr-73	Syria and Egypt receive 1st installment of air defence missiles.
May-73	Egyptian mobilisation creates major scare in Israel.
Mid-73	Doubts are spread about the efficacy of the air defence missile force. Media rumors of Arab disunity and friction with the Russians are not discouraged. Threats on the use of an oil embargo continue. Damming report on feasibility of canal crossing leaked to Israeli intelligence and building of an oil pipeline near intended war zone publicised.
Jul-73	Syria and Egypt receive 2nd installment of air defence missiles.
13-Sept-73	Syrian fighters attack Israeli reconnaissance flight.
Sept-73	Syria increases strength opposite the Golan heights.
24-Sept-73	CIA passes intelligence to Israel noting discrepancies between Egypt's build up to exercise and build ups to previous exercises.
25-Sept-73	King Hussein of Jordan flies to Israel saying that the Syrian deployments were a precursor to war.
27-Sept-73	Egypt mobilises large number of reserves.

Date	Action
30-Sept-73	Egypt mobilises another large number of reserves and announces the demobilisation of the 27th Sept reserves. Israeli intelligence receives reliable human intelligence that the Egyptian exercise will end in an actual crossing of the Suez Canal.
1-Oct-73	Egyptian planned exercise (Tahrir 41) commences.
2-Oct-73	Israel receives intelligence on Syrian movements of bridging equipment, fighter aircraft and surface to air missile batteries. Egypt mobilises bridging equipment and crossing spots.
3-Oct-73	Egypt tell Soviet Union of the planned attack against Israel.
4-Oct-73	Israeli Air reconnaissance over Sinai reveals an unprecedented build up of Egyptian forces. Also noticed is a Soviet airlift heading for the region.
5-Oct-73	Human intelligence source sends Israeli intelligence code word for imminent war. Israeli intelligence receives reports that Soviet naval vessels are departing.
5-Oct-73	Israeli cabinet meet to discuss what evacuation of the Russian civilians means.
6-Oct-73	Israeli human and signal intelligence sources leave no doubt that Arab hostilities were imminent. 1400: Start of hostilities. Israel is surprised.

Table 5.1: Timeline of main representative events providing a potential indication of attack against Israel from 1967-1973.

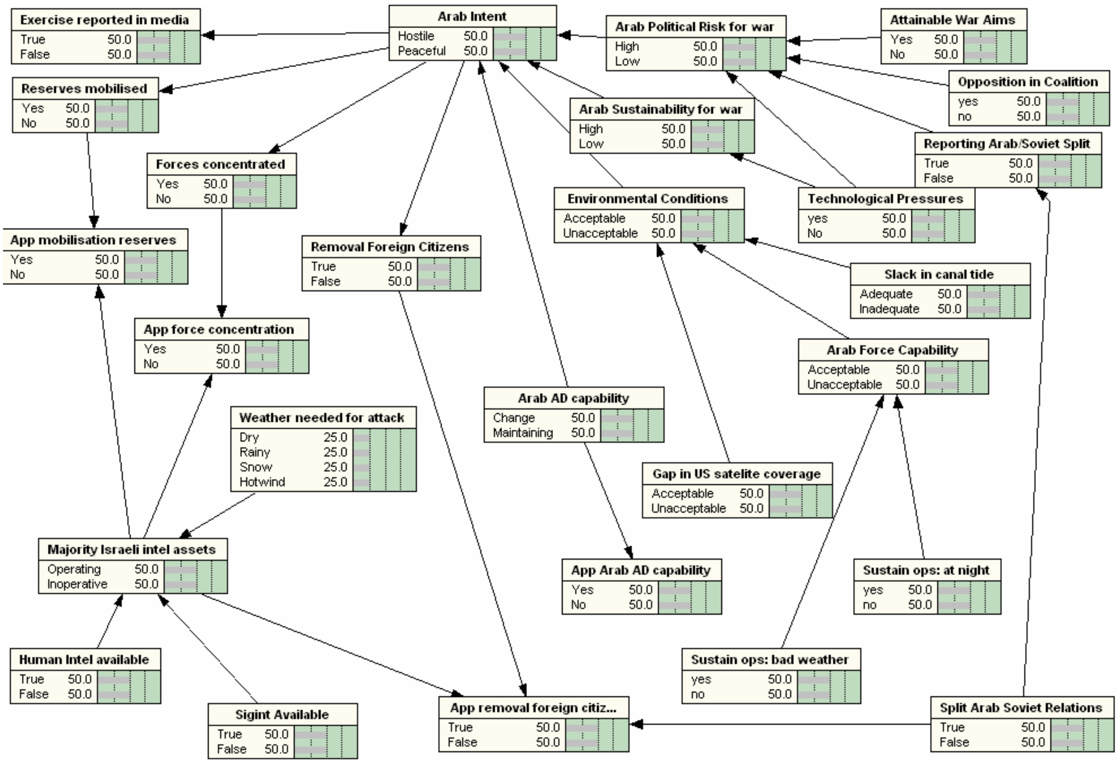


Figure 5.1: Initial BBN developed to support quantitative analysis of events potentially indicating an attack against Israel in 1973[7].

Based upon information in the developed a timeline, a BBN was developed to support the research, as shown in Figure 5.1. Since the focus of this work was the use of BBNs for intelligence analysis it was of most interest to develop the network from the point of view of an Israeli intelligence officer, as it was the Israelis who were surprised at the start of the 1973 hostilities.

Had this network actually been developed between 1967 and 1973 it would have been done so against the territorial gains Israel had made during the Six Day War of 1967. Following this success, the Israeli intelligence and government developed very firm views on what conditions had to be met before Egypt could re-commence hostilities. Specifically these were: The purchase of aircraft capable of providing deep attack as well as air defence capabilities and the belief that Syria would not commence hostilities without being in a coalition with Egypt.

[7] App is a shortened version of apparent. AD is a shortened version of Air Defence. Ops is a shortened version of operations. SIGINT refers to signal intelligence.

The BBN in Figure 5.1 incorporates these beliefs and other potential military cues as to the intent of the Egyptian government. For example, the overall capability of a fighting force is affected by their ability to operate at night and in bad weather. A force which is only fully effective on a clear day will have a limited set of courses of action available to them.

There are however, other general factors which impact upon a nation's ability to commence hostilities, not least of which is the fact that war is expensive. Often going to war requires a government to increase military spending as a percentage of the gross domestic product. Whilst going to war may support the defence industry, it can create hardships and poverty within the warring nations through impacting upon areas such as the stock market, trade routes and tourism. It is important to remember that between 1967 and 1973 the Egyptian economy was weakened through a loss of revenue following the closure of the Suez Canal. In addition to this, conflict within the Middle East carries the possibility of affecting the global supply of oil through cuts in production or an embargo the effects of which would be seen worldwide.

Going to war also carries political risk. At the lowest level this could be the inability to achieve national cross party support for the government commencing hostilities. However, this risk can escalate and could result in outcomes including: possible sanctions against a nation, the alienation of countries previously considered friendly and potential retaliation attacks. All of these issues could lead to destabilisation of the local region.

5.3 Refinement of the Bayesian Belief Network developed in support of the main events leading up to the start of the 1973 Arab Israeli conflict.

The BBN shown in Figure 5.1 is relatively complex due to the number of events and relationships it represents. However, before it could be used for any analyses, all the required conditional probability distributions had to be obtained. For the purposes of this research, the probabilities were elicited from a military historian working at the Defence College of Management and Technology over a series of four meetings.

The first meeting provided an opportunity to present, discuss and refine the BBN and supporting timeline. The discussions centred on the structure of the network and as such no conditional probabilities were elicited. During the meeting it was agreed to simplify the BBN shown in Figure 5.1 through, in essence the removal of all nodes explicitly representing the political, economic and environmental factors.

The nodes removed were done so for several reasons. Firstly, there were certain nodes which could not be adequately expressed in probabilistic terms namely: attainable war aims, an Arab coalition (noted as opposition in coalition in the network) split in Arab/Soviet relations, and the reporting of a split in Arab/Soviet relations. Secondly, the review process highlighted several nodes which it was considered unlikely that an Israeli intelligence officer would have known about. These nodes mainly related to the environmental conditions that the Egyptian and Syrian coalition desired to commence hostilities including: sufficient slack in canal tide, gaps in USA satellite coverage, Arab force capability, ability to sustain operations at night, and the ability to sustain operations in bad weather. Thirdly, those nodes relating to the political risk and ability to sustain war were removed. It was not considered possible to represent the myriad interlinking factors which affect theses nodes within the BBN. Furthermore, even if the BBN were to incorporate such a representation, it is considered unlikely that intelligence officers could obtain the required probability distributions. Finally, the node relating to the reporting of exercises within the media was removed. An exercise requires the mobilisation of forces, equipment and the concentration of forces and, as such, an exercise can actually represent the start of hostilities.

In addition to the removal of the above nodes, those nodes relating to the availability of intelligence were amended. The nodes relating to human and signals intelligence (HUMINT and SIGINT) were changed from whether the intelligence was available to whether it was credible. As such, the summation of the available intelligence now gave the overall credibility of the intelligence.

None of the above changes to the developing BBN were considered detrimental to the aim of there search. The resulting BBN is shown in Figure 5.2 in which all values are purely illustrative.

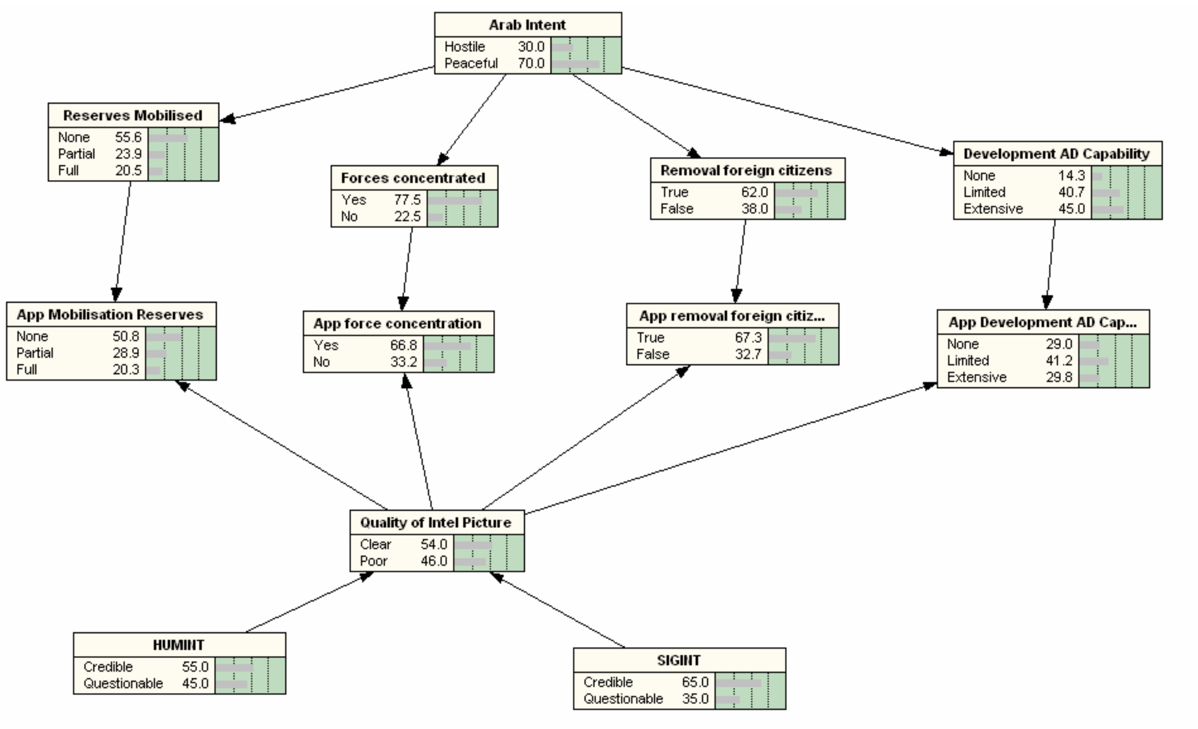


Figure 5.2: Revised BBN developed to support quantitative analysis of events potentially indicating an attack against Israel in 1973.[8]

At the second meeting, the BBN was populated with fictitious numbers in order for the military historian to observe how the introduction of evidence altered the network values. This enabled a review of the relationships presented within the network. The military historian was asked to verify that the values within the network moved in an 'intuitive' way.

The following, third, meeting focused on the most appropriate method for the military historian to express numerical values, e.g. as percentages, frequencies, fractions or using a pointer on a sliding scale. This third meeting also provided an additional opportunity for the military historian to review the model. The outcome of the third

[8] App is a shortened version of apparent. AD is a shortened version of Air Defence. HUMINT refers to Human Intelligence. SIGINT refers to Signals Intelligence.

meeting was an agreement to elicit the required conditional probability distributions as percentages. Subsequently, all the required probabilities were elicited at the fourth and final meeting. This was the longest of all the meeting and lasted several hours. In addition to the conditional probability distributions, the military historian was asked to directly state a range of probabilities based upon a variety of available information sources, namely the probability of hostile intent given that there was clear intelligence and an apparent:

- Extensive build up of an air defence capability.
- Partial mobilisation of reserves.
- Concentration of forces.
- Departure of foreign citizens.
- Extensive air defence capability and partial mobilisation of reserves.
- Extensive air defence capability, partial mobilisation of reserves and concentration of forces.
- Extensive air defence capability, partial mobilisation of reserves, concentration of forces and the departure of foreign citizens.

Using the elicited conditional probability distributions, the BBN was populated. Subsequently, the network was interrogated by the sequential instantiation of the appropriate nodes relating to the simplified timeline of events shown in Table 5.2. The main focus of the work was to observe the impact of additional, cumulative information on the likelihood that Arab intent was hostile. The BBN calculated the Bayesian probability of hostile intent through combining the various data and evidence available based on the military historian's subjective probabilities.

Date	Action
Apr-73	Syria and Egypt receive 1st installment of air defence missiles.
May-73	Egyptian mobilisation creates major military scare in Israel.
Jul-73	Syria and Egypt receive 2nd installment of air defence missiles.
2-Oct-73	Israel receives intelligence on Syrian movements of bridging equipment, fighter aircraft and air defence batteries. Egypt mobilises bridging equipment and crossing spots.
4-Oct-73	Israeli Air reconnaissance over Sinai reveals an unprecedented build up of Egyptian forces. Also noticed on this date is a Soviet airlift heading for the region

Table 5.2: Simplified timeline of main representative events providing a potential indication of attack against Israel from 1967-1973.

A series of questions were composed and subsequently put to the historian, the answers to which could be logically deduced through interrogation of the BBN. The historian, however, was unaware of this. Therefore, when the questions were directly posed to the historian the results given were intuitive. The deviations between the Bayesian values and directly elicited subjective probabilities for the various questions posed were calculated. The differences were a measure of the historian's own individual inconsistency. As such, the deviations were analysed to determine if any trends or biases were present within the results.

5.4 Results and discussion.

5.4.1 Proof of principle of derived BBN.

Firstly, the key events in Table 5.2 were sequentially and cumulatively entered into the BBN. No assumptions were made relating to the state of the nodes referring to human and signal intelligence, or subsequently the quality of the intelligence picture. Table 5.3 summarises the results.

Events observed	Probability Arab intent was hostile (as %)
Initial belief that hostilities would occur	30
Apparent limited build up of an air defence capability	32
Apparent limited build up of an air defence capability and mobilisation of reserves	44.6
Apparent mobilisation of reserves and an extensive build up of an air defence capability	47
Apparent mobilisation of reserves, an extensive build up of an air defence capability and a concentration of forces	51.5
Apparent mobilisation of reserves, an extensive build up of an air defence capability concentration of forces and the departure of foreign citizens	61.0

Table 5.3: Probability of Arab intent being hostile based on evidence available in timeline.

The results in Table 5.3 show that as the number of apparent events observed accumulated, the probability that Arab intent was hostile increased. It is important to remember that the Bayesian values shown above are dependent upon the elicited prior probability distributions. Consequently, the values shown are subjective and made in full consideration of all the available information with the luxury of hindsight. Therefore, these values should not be used in any comparison against decisions taken during the actual events being considered. Nevertheless, the results do show that BBNs have applicability in the use of assessing the likelihood of various hypotheses based on the continual updating of information.

5.4.2 Investigations into trends and biases.

As part of the experiment, the military historian provided direct estimates for a series of propositions which are shown in Table 5.4. All the propositions assumed that clear intelligence was available and started by assuming only a single piece of intelligence was available. Subsequently, two, three and finally four pieces of intelligence became simultaneously available. Each proposition was entered into the BBN and the differences between the direct subjective estimates and the Bayesian values were assessed.

The results generally show a noticeable difference between the Bayesian and direct estimates for a given proposition. In the main, the results in Table 5.4 show an increase in the calculated deviation as additional simultaneous pieces of information became available. These findings are shown graphically in Figure 5.3.

The result shown in Figure 5.3 presents a continual increase in the probability of the coalition intent being hostile for both the Bayesian and directly elicited probabilities. When only two pieces of information were available, there is fairly close agreement between the two sets of values. As additional information becomes available, the historian begins to over adjust the probability of hostile intent. In Figure 5.3 this is shown by the increasing difference between the direct subjective and Bayesian probabilities. The over adjustment in suggestive of the historian being unable to correctly take account of the interdependencies between the intelligence reports provided.

	Probability value		
Probability that Arab intent was hostile given:	**Direct**	**Bayesian**	**Direct-Bayesian**
Clear intelligence and an extensive build up of an air defence capability	50	36.6	13.4
Clear intelligence and a partial mobilisation of reserves	45	44.4	0.6
Clear intelligence and an apparent concentration of forces	70	34.1	35.9
Clear intelligence and an apparent departure of foreign citizens	35	39.6	-4.6
Clear intelligence, an extensive air defence capability and an apparent partial mobilisation of reserves	60	51.9	8.1
Clear intelligence, an extensive air defence capability, an apparent partial mobilisation of reserves and an apparent concentration of forces	80	56.6	23.4
Clear intelligence, an extensive air defence capability, an apparent partial mobilisation of reserves an apparent concentration of forces and the apparent departure of foreign citizens	90	66.6	23.4

Table 5.4: Comparison of Bayesian and Direct values for the main propositions of interest.

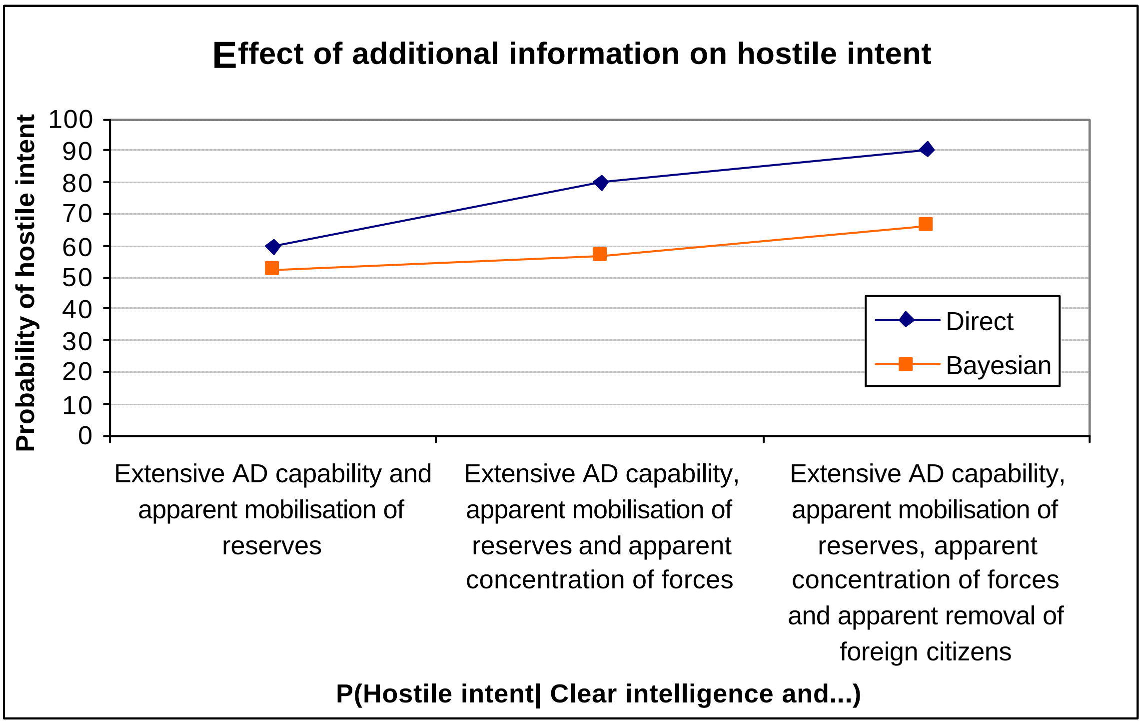


Figure 5.3: Effect of additional information on the Bayesian and directly elicited probabilities of coalition intent being hostile.

5.4.3 Relative importance of potential information sources.

The final area of interest was that of assessing the relative importance of each potential information source. The results presented in Chapter Four showed that the majority of participants were unable to correctly assess the information report which has most impact upon the likelihood of a set of hypotheses. The experiment reported on in Chapter 4 focused specifically on the selection of reconnaissance or sensor reports. In the experiment presented here the emphasis was not upon *how* the information was collected but upon *what* the information related to. That is, which of the potentially available pieces of information (Mobilisation of troops; Concentration of forces; Development of an air defence capability; Departure of Soviet citizens) provided the most information relating to the possible intent of the Egyptian Syrian coalition?

As part of the experiment the military historian was requested to rank, in order of influence, the potential information sources available for assessing the intent of the

Arab coalition. Table 5.5 presents the direct rankings provided by the military historian are shown alongside those obtained from the interrogation of the BBN.

	Results obtained from:	
Influence ranking	**Direct questioning**	**Network interrogation**
1	Development of an air defence capability	Mobilisation of reserves
2	Mobilisation of reserves	Departure of foreign citizens
3	Apparent concentration of forces	Development of an air defence capability
4	Departure of foreign citizens	Apparent concentration of forces

Table 5.5: Ranking of importance for available information.

It can be seen that there is some inconsistency between the rankings provided verbally by the military historian and those derived from the network populated with the distributions provided by the same person. It is plausible that when providing the direct probabilities for various diagnostic propositions, the military historian was subconsciously giving preference to intelligence from the various sources in a way not consciously expressed through the elicitation of the conditional probability distributions.

5.5 Conclusions.

This analysis has presented a BBN as a DSS to support intelligence analysis in the historical context of the main events leading up to the start of the Yom Kippur War. Whilst proof of concept has been shown, the results depend upon the prior probabilities elicited and used to populate the network. Therefore, these values must be carefully obtained and assessed. The work has not sought to challenge those decision made during the real time occurrence of the events considered which would have been made in a different environment with additional resources.

The results further indicate that as multiple information sources become available individuals are unable to correctly manipulate the interdependencies between the sources. In addition to this, inconsistencies were revealed between the relative importance of each potential information source as defined by an individual and the BBN.

Overall, the results suggest that intelligence analysts may benefit from a system capable of combining different streams of information to calculate Bayesian values for a given hypotheses set. The derivation of the network would provide a clear audit trail for the hypotheses and intelligence nodes included and rejected. The techniques used to obtain the conditional probability distributions used to populate the BBN are transparent and create a clear audit of the distributions used.

In addition to this, should the network provide unexpected results, the analyst would be forced to reconsider the values and assumptions behind it. This could help prevent subconscious preferences being given to particular intelligence sources. Of more concern would be how to ensure a thorough review of the BBN should it provide the expected answer.

CHAPTER 6 : SUBJECT LIKELIHOOD ELICITATION: A COMPARISON OF DIRECT RANKING, LIKELIHOOD RATIOS AND CONDITIONAL PROBABILITIES FOR USE IN DECISION SUPPORT SYSTEMS.

6.1 Introduction.

The experiments presented within Chapters Four and Five investigated the applicability of Bayesian Belief Networks (BBN) for the quantitative analysis of intelligence reports. So far it has been concluded that to be of greatest benefit any system developed to support such a process must:

- Logically and consistently combine multiple intelligence reports obtained from both human and signals intelligence sources.
- Assist in determining the relative importance and influence of each information source in discriminating between a set of known hypotheses.
- Have a structure which is open to scrutiny and provides a clear audit trail of the assumptions, hypotheses and data used.
- Logically and coherently update the belief for each hypothesis considered as new evidence or data enters a network.

The development of a BBN which achieves all of these goals and makes use of subjective probability distributions requires an appropriate elicitation technique. Inevitably, all such techniques have strengths and weaknesses. Therefore, it is important to fully understand which elicitation technique should be selected in any given situation. Advancing the body of knowledge in this area, this chapter describes the development of a pilot study and subsequent experiment to compare: direct ranking of the variables' perceived importance for discriminating between given hypotheses, likelihood ratios and conditional probabilities. The focus of this experiment was the extent to which the different elicitation techniques lead to equivalent or different judgements.

A BBN was created in support of the experiment to act as a normative model. An individual network was created for each participant based on their own conditional probabilities. Comparisons were made between the results of the normative model and the participant's direct subjective estimates based upon their intuitive reasoning.

6.2 Development of experiment.

6.2.1 Design of the BBN.

Building upon previous work, the experiment to compare elicitation techniques utilised the scenario based on the main events observed in the lead up to the Arab Israeli conflict of 1973 (as detailed in Chapter Five). The timeline developed in support of the scenario was used without alteration, and for ease of reference is repeated in Table 6.1. The accompanying BBN, also developed in Chapter 5, is also repeated for clarity in Figure 6.1.

Event number	Date	Action
1	Apr-73	Syria and Egypt receive 1st instalment of air defence missiles.
2	May-73	Egyptian mobilisation creates major military scare in Israel.
3	Jul-73	Syria and Egypt receive 2nd instalment of air defence missiles.
4	2-Oct-73	Israel receives intelligence on Syrian movements of bridging equipment, fighter aircraft and air defence batteries. Egypt mobilises bridging equipment and crossing spots.
5	4-Oct-73	Israeli Air reconnaissance over Sinai reveals an unprecedented build up of Egyptian forces. Also noticed on this date is a Soviet airlift heading for the region.

Table 6.1: Simplified timeline of main representative events providing a potential indication of attack against Israel from 1967-1973.

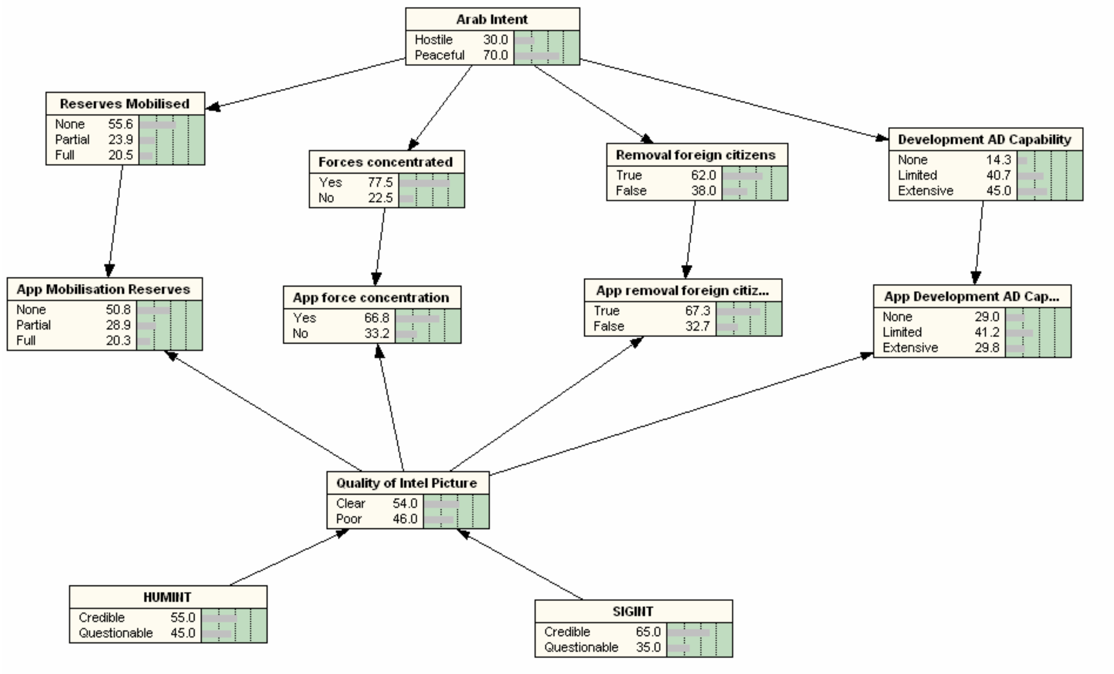


Figure 6.1: BBN developed to support quantitative analysis of events potentially indicating an attack against Israel in 1973.[9]

As noted in Chapter 5, the time required to elicit the required probability distributions to fully populate the network through consultation with a military historian was substantial and occurred over a series of meetings. Importantly, the military historian only provided probability distributions and was not required to directly rank the variables within the network, nor provide estimates of likelihood ratios. The participants for this experiment were students at the Defence College of Management and Technology. As such, the experiment was generally completed either at the end of a lecture or during a break in teaching. Therefore, the experiment was designed to take no more than a maximum of 30 minutes to complete. Considering this, the network shown in Figure 6.1 was considered to contain too many nodes and conditional probability distributions to be of practical use in a 30 minute experiment which contained three different elicitation techniques. The experiment would simply take too long to complete.

[9] App is short for apparent. AD is short Air Defence. HUMINT refers to Human Intelligence. SIGINT refers to Signals Intelligence.

As the focus of the experiment was not if the network could perfectly predict the probability of hostile intent, but on how the differing elicitation techniques considered may lead to equivalent or differing judgements, it was considered prudent to simplify the BBN. This was achieved through the removal of various nodes and thus reduced the overall elicitation burden placed on participants.

Firstly, the nodes ‘HUMINT’, ‘SIGINT’ and ‘Quality of Intelligence Picture’ were removed. The removal of these nodes actually established a degree of flexibility within the network. Either, these values could be explicitly stated as part of the experiment, or participants could be allowed to make their own inferences as to the overall quality of the intelligence picture (this option utilised the premise that each participant would consistently use their assumed level on the quality of intelligence picture throughout the answers they provided). In either case, the overall quality of the intelligence picture was assumed to remain constant throughout the experiment. As such, the removal of these nodes did not impact upon the analysis undertaken to assess the impact of the different elicitation techniques on the judgements made by the participants.

Secondly, the four nodes relating to the ‘apparent’ observation of events (e.g. apparent mobilisation of reserves) were also removed. The removal of these nodes was undertaken through consideration of the aim of the experiment. At its core, the scenario focused upon an interpretation on a series of events. The network in Figure 6.1 requires the participants to separate out the probabilities of the events actually occurring, and subsequently the probability that the event would be observed and reported. Although this level of abstraction provides greater detail to the network, it did not directly support the aim of this final experiment which was to compare the judgements made by participants through differing elicitation techniques. To provide sufficient data for analysis, the participants simply had to provide their assessment of the situation based upon the available information. Therefore, removing the distinction between ground truth and observed events simplified the BBN without compromising the results obtained.

The removal of all nodes relating to the quality of intelligence and the apparent observation of events ensured that the BBN was sufficiently small for all three elicitation techniques to be used within a 30 minutes experiment.

Whilst reviewing the BBN, some consideration was also given to changing the hypotheses used within the BBN to:

Hypothesis A: The Egyptian and Syrian coalition is capable of hostile attacks against Israel but is seeking to hide this fact.

Hypothesis B: The Egyptian and Syrian coalition is continuing a programme to become capable of undertaking a hostile attack against Israel.

Hypothesis C: The Egyptian and Syrian coalition is not interested in undertaking a hostile attack against Israel.

These hypotheses were used in a brief pilot study, completed by six post graduate students at the Defence College of Management and Technology (four of whom were serving military officers). The results of the study identified that individuals were not able to clearly discriminate between a coalition which has hostile intent but is seeking to hide this fact (and hence there would be at best limited indications of the real intent) and a coalition which has hostile intent but has not yet the capability to achieve their aim. Therefore, a more plausible set of hypotheses was considered to be:

- The Egyptian and Syrian coalition intent was hostile.
- The Egyptian and Syrian coalition intent was peaceful.

The finalised BBN incorporating all of the above amendments is shown in Figure 6.2 (note values are for illustration only).

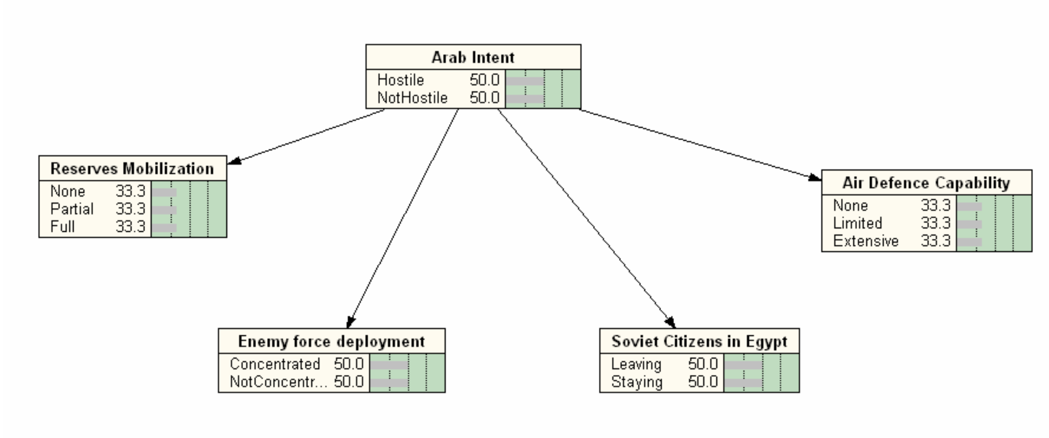


Figure 6.2: Finalised BBN developed for the elicitation experiment.

6.2.2 Design of the experimental questionnaire.

The experiment consisted of a single questionnaire comprising three distinct parts, each one relating to a specific elicitation technique (as detailed in Annex B). Section one focused on the direct ranking of variables. The results given in this section revealed participants' perceived importance for each variable (potential information source within the BBN) with respect to the hypotheses considered. The experiment itself had four potential information sources which relate to events seen within the scenario timeline (Table 6.1):

- Mobilisation of reserves.
- Enemy force concentration.
- Soviet citizens leaving Egypt.
- Air defence capability.

To rank the potential information sources, firstly participants decided which information sources were the most and least important in uncovering the opposition's intent. Next, participants ranked the remaining two information sources between the most and least important. Weights were assigned to the four rankings using a linear scale. The least and most important sources were allocated weights of zero and one hundred respectively. Subsequently, participants placed marks on the linear scale representing the relative position of the remaining two ranked sources. The linear scale consequently contained four marks: zero, one hundred and the two marks made by the participant. It was clearly explained to the participant that the gaps between the marks on the scale represented the relative difference in importance of each potential information source. Participants were encouraged to vary the position of the middle ranks until the gaps between all four variables were perceived to be correct. Only then were the middle rankings assigned numerical weights.

The second section of the experiment concentrated on eliciting likelihood ratios. Likelihood ratios are an individual's assessment of the ratio: probability of an event occurring given that the hypothesis is true versus the probability of the event occurring given that the hypothesis is not true. A ratio must be assigned for every event and hypothesis being considered.

Likelihood ratios represent the relative likelihood of each event occurring and the importance, or impact, that information about the event's occurrence has on the set of hypotheses. The events being considered, which for the purposes of this research are in a defined timeline (Table 6.1), may or may not support a hypothesis. Furthermore, the events being considered are not independent. Therefore, the likelihood ratios must incorporate the fact that the observed event may be dependent upon events previously seen and accounted for. Such an approach should prevent the double counting of information. However, this approach also leads to one of the main weaknesses of the technique. For a timeline incorporating many events, an assessment of a likelihood ratio for a given event and hypothesis must be made based on all the information available from every previous event. For a long chain of events, it can be seen that this would become an increasingly complex process, possibly beyond the capabilities of a single individual. This view is not shared by McLaughlin and Paté-Cornell (2005). When considering the use of likelihood ratios to elicit probabilities against the commonly used technique of conditional probabilities, McLaughlin and Paté-Cornell (2005, page 4) state that "…for an expert to estimate the conditional probability of an event given a hypothesis *and* (italics in original quote) all previously available information is generally not more complex that thinking about p(E|A) (defined in the paper as the likelihood function) alone. In fact, intuitively, it may be the most natural way to process the information."

Within the context of this experiment participants were asked for the likelihood ratio for each event within the timeline given all previous events. For example, the likelihood ratio for event two: Egypt orders a partial mobilisation of forces, is given in the knowledge of a reported instalment of air defence missiles. A likelihood ratio of 1 implies that an event is equally likely to occur whether the Arab coalition's intention is imminently hostile or not, given all previous events observed. A ratio greater than 1 suggests that an event is more likely to occur when hostilities are imminent. Conversely, a ratio less than 1 represents a belief that the event being considered is more likely to occur when hostilities are not imminent.

The third and final section of the experiment centred on straightforward conditional probability distributions. Often probability distributions for Bayesian Belief Networks (BBN) are constructed from collated data. However, within the applications of BBNs considered here, it is inevitable that the data required to construct all of the required probability distributions will not be available. This necessitates the use of subjective probability distributions. Whilst subjective probabilities incorporate the contextual knowledge of the expert from whom they are elicited, the values given are prone to include well-known cognitive biases (as presented in Chapter Two).

In total, two sets of subjective probabilities were elicited: the conditional probability distributions required to populate the supporting BBN; and a series of direct estimates requiring diagnostic reasoning based on the events shown in Table 6.1. The conditional probability distributions were used to populate a BBN for each participant. The BBN combined the available evidence to calculate the Bayesian (normative) probability of Arab coalition (Egyptian and Syrian) intent. Any deviation between the Bayesian and direct subjective probability provided a measure of the participant's own individual inconsistencies. The deviations were analysed to determine if any trends or biases were present within the results.

6.3 Experiment into subjective value elicitation: A comparison of direct rankings, likelihood values, and conditional probabilities for use in decision support systems.

6.3.1 Experimental methodology.

An experiment was conducted at the Defence College of Management and Technology. 38 serving UK army officers undertaking a period of post graduate study complete the experimental questionnaire (as detailed in Annex B). Questionnaires are commonly used elicitation technique when investigating cause and effect relationships, or eliciting procedures. Once fully designed, the standard format of the questionnaire facilitates the analyses of (potentially) large volumes of data. The questionnaires were not administered whilst the participants were undertaking their normal duties and as such the use of questionnaires cannot be considered as being a natural technique in the formal sense. However, the questionnaires contained three different methods for the expression of their expertise (direct rankings, likelihood ratios and conditional probability distributions). As such, the questionnaire should have provided an opportunity for the participant to express their opinion in a manner with which they were familiar, and was therefore natural to them.

None of the officers were familiar with intelligence work. Building upon the findings of the earlier experiments all 38 participants were met face to face. All participants undertook the experiment at the same time. The experiment was presented and each elicited technique was briefly explained. Subsequently, the participants completed the questionnaire by themselves without discussion with those around them. However, participants were encouraged to seek clarity on any points of ambiguity with the experiment facilitators. Many questions were raised and answered on an individual basis. All responses were completed within 30 minutes. Although the participants were encouraged to seek clarification, in total 31 of the 38 sets of results were amenable to analysis. The remaining 7 sets of results were not amenable to analysis due to either being returned incomplete, or incorrectly answered (for example conditional probability distributions which in total exceeded 100%).

6.3.2 Direct ranking and weighting of potentially available information sources.

6.3.2.1 Analysis of rankings.

The calculated (normative) Bayesian rankings (based upon the mutual information between each source of evidence and the hypothesis node) were compared with the given direct rankings. The results are shown in Table 6.2. Identical Bayesian and direct results are highlighted through the use of shading.

Of the 31 direct responses analysed, the majority of participants considered information relating to the:

- Mobilisation of reserves to be the most influential piece of evidence (16 of 31 responses).
- Concentration of forces to be the second most influential piece of evidence (14 of 31 responses).

Initial comparisons of the direct rankings against the BBN rankings showed similar aggregates. In total 12 individuals directly ranked the mobilisation of forces as the most important piece of information and the concentration of forces as the second. Correspondingly, an interrogation of all the BBNs found eight networks which also ranked the mobilisation of forces as the first and concentration of forces as the second most influential pieces of evidence. However, the direct and BBN results generally corresponded to different people. Furthermore, strikingly only two participants (participants 14 and 17 in Table 6.2) gave the same four rankings when questioned directly and through interrogation of their BBN.

Participant	Results from	Development of Air defence capability	Mobilisation of reserves	Concentration of forces	Departure of foreign citizens
1	Direct	4	2	1	3
	Bayesian	4	1	2	3
2	Direct	3	2	1	4
	Bayesian	2	1	4	3
3	Direct	4	1	2	3
	Bayesian	4	3	1	2
4	Direct	1	2	3	4
	Bayesian	1	3	2	4
5	Direct	3	1	2	4
	Bayesian	4	3	2	1
6	Direct	4	1	2	3
	Bayesian	4	3	1	2
7	Direct	4	1	2	3
	Bayesian	2	1	4	3
8	Direct	1	2	3	4
	Bayesian	2	1	3	4
9	Direct	3	4	2	1
	Bayesian	1	2	4	3
10	Direct	3	2	1	4
	Bayesian	3	1	2	4
11	Direct	4	1	2	3
	Bayesian	3	1	2	4
12	Direct	3	1	2	4
	Bayesian	2	1	3	4
13	Direct	4	1	3	2
	Bayesian	1	3	4	2
14	Direct	4	1	3	2
	Bayesian	4	1	3	2
15	Direct	3	1	2	4
	Bayesian	3	2	4	1
16	Direct	4	1	3	2
	Bayesian	4	1	2	3

Participant	Results from	Development of Air defence capability	Mobilisation of reserves	Concentration of forces	Departure of foreign citizens
17	Direct	4	2	1	3
	Bayesian	4	2	1	3
18	Direct	4	3	1	2
	Bayesian	4	3	2	1
19	Direct	1	2	3	4
	Bayesian	2	1	3	4
20	Direct	4	2	1	3
	Bayesian	4	3	1	2
21	Direct	3	1	2	4
	Bayesian	2	3	1	4
22	Direct	3	2	1	4
	Bayesian	4	2	1	3
23	Direct	2	4	3	1
	Bayesian	1	3	2	4
24	Direct	4	1	2	3
	Bayesian	1	4	2	3
25	Direct	1	3	2	4
	Bayesian	3	1	2	4
26	Direct	4	1	2	3
	Bayesian	2	1	3	4
27	Direct	3	1	2	4
	Bayesian	1	2	4	3
28	Direct	3	2	1	4
	Bayesian	4	1	2	3
29	Direct	4	2	1	3
	Bayesian	4	1	2	3
30	Direct	2	1	3	4
	Bayesian	4	1	2	3
31	Direct	4	1	2	3
	Bayesian	2	1	4	3

Table 6.2: Example Results of direct ranking and BBN rankings of potentially available information sources.

Table 6.3 shows the joint probability distribution of the direct and Bayesian rankings, with the conditional probability distributions presented in Table 6.4.

	Direct rankings			
Bayesian rankings:	**1**	**2**	**3**	**4**
1	9.68	9.68	2.42	3.23
2	7.26	5.65	8.87	3.23
3	5.65	4.03	9.68	5.65
4	2.42	5.65	4.03	12.90

Table 6.3 Joint probability distribution (as percentages) of the Bayesian and direct rankings.

	Direct rankings			
Bayesian rankings:	**1**	**2**	**3**	**4**
1	38.71	38.71	9.68	12.90
2	29.03	22.58	35.48	12.90
3	22.58	16.13	38.71	22.58
4	9.68	22.58	16.13	51.61

Table 6.4: Conditional probability distribution (as percentages) between direct and Bayesian rankings.

It is interesting to note that the closest agreement between the Bayesian and direct rankings occurred for placing the least influential piece of information.

To assess the overall correlation (or agreement) between the direct and Bayesian rankings, the non-parametric measure of Kendall's tau statistic (Equation 6.1) was used.

$$t = \frac{4P}{n(n-1)} - 1$$

Equation 6.1: Kendall's Tau.

Within equation 6.1, n is the number of items being compared and P is the sum, over all ranked items, of items ranked after a given item by both rankings. An example calculation of Kendall's tau statistic using example participant results is shown in Table 6.5.

	Concentration of forces	Mobilisation of reserves	Departure of foreign citizens	Development of Air defence capability
Direct	1	2	3	4
Bayesian	2	1	3	4
Cumulative value of P	2	4	5	5

Table 6.5: Example calculation of Kendall's tau statistic.

Table 6.5 presents the potential information sources in direct rank order (columns moving left to right) alongside the corresponding Bayesian ranks. Utilising the order given by the direct rankings, P is calculated cumulatively based upon the Bayesian rankings. Thus, the first direct ranked potential information source is 'concentration of forces' which has a corresponding Bayesian rank of 2. Consequently, of the three columns to the right of 'concentration of forces' two will have higher ranks (i.e. 'departure of foreign citizens' ranked 3rd and 'development of an air defence capability' ranked 4th). Therefore, the value of P is initially set at two. Subsequently, the second direct ranked potential information source 'mobilisation of reserves' is assessed. With a corresponding Bayesian ranking of 1, the remaining two columns to the right both have higher Bayesian rankings. The value of P is therefore increased to 4. Next, the column with the 3rd direct ranking (departure of foreign citizens) is reviewed. This potential information source also has a Bayesian ranking of 3. At this point there is only one remaining column to the right of the table which has a Bayesian rank of 4. Hence, the value of P increases to 5. The remaining and final column of the table does not change the value of P as there are no columns to the right with which to base an assessment. Therefore, the final calculated value of P is 5.

The value of n in the above example is 4 (the number of potential information sources to be ranked). Entering the values of n and P into the equation for Kendall's tau gives:

$$t = \frac{4 \times 5}{4(4-1)} - 1 = 0.667$$

Kendall's tau statistic was calculated for every participant, the results of which are shown in Figure 6.3. The overall median value of the statistic was 0.33. It can be seen that in total 13 of the 31 results had a value greater than +0.6 which is indicative of a good agreement between the direct and Bayesian rankings. Of these, however, only two had a tau statistic greater than 0.9 which represents strong agreement between the rankings. That some participants exhibit good agreement between the rankings is not surprising given what is being compared. What is surprising is that there are not more of them and that the agreement is not much stronger.

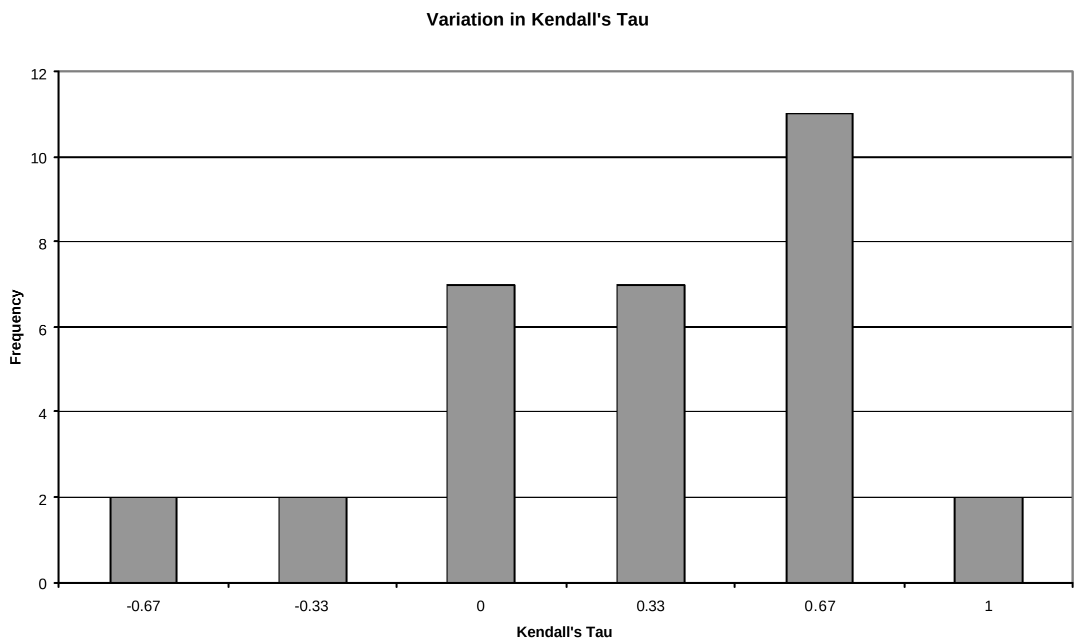


Figure 6.3: Variation in Kendall's tau statistic.

6.3.2.2 Analysis of weightings.

This analysis was undertaken through the calculation of the sum of squared rank deviations. The calculation was based upon the participant's stated direct weightings and the BBN's normative weighting of the potential information sources (based on the mutual information between each source and the hypothesis node).

From the field of information theory, mutual information and entropy measure different aspects of uncertainty. Entropy, as defined by Shannon (1949), is a measure of the uncertainty associated with a given variable. The entropy H of a probability distribution of X is expressed as:

$$H(X) = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i)$$

Equation 6.2: Entropy.

Where X is a discrete random variable taking values $\{x_1, x_{2,} ... x_n\}$.

When interpreting entropy, the higher the entropy value, the more uncertainty there is. If a value is known for certain then the entropy is zero. If all possible outcomes are equally likely then the probability distribution becomes uniform and the entropy has its maximum value which may simply be calculated as $\log(n)$. When looking at a BBN the calculation of expected entropy change in the hypothesis node is known as the value of mutual information. In essence, mutual information details how much information one event or piece of evidence provides about another. Within the context of this experiment, mutual information calculates the expected reduction in the uncertainty of a hypothesis resulting from evidence from a potential information source and can therefore be used to used to rank the 'worth' of each intelligence report in removing the uncertainty associated with Arab intent.

The mutual information between a node X and an intelligence source Y, is represented as $I(X;Y)$ and is the average reduction in entropy in node X based upon a given instantiation in the intelligence source, that is: $I(X;Y) = H(X) - H(X \mid Y)$. Where

$$H(X \mid Y) = -\sum_{j=1}^{m}\left[\sum_{i=1}^{n} P(x_i \mid y_j)\log_2 P(x_i \mid y_j)\right]P(y_j)$$

Equation 6.3: Mutual information.

Based upon a participant's response, an example calculation of the value of mutual information follows. For this example we will consider the value of *H(Arab Intent | enemy force deployment).* The node of enemy force deployment has two states: concentrated and not concentrated. Figure 6.4 shows the experimental BBN for a participant before any information has been entered into the network.

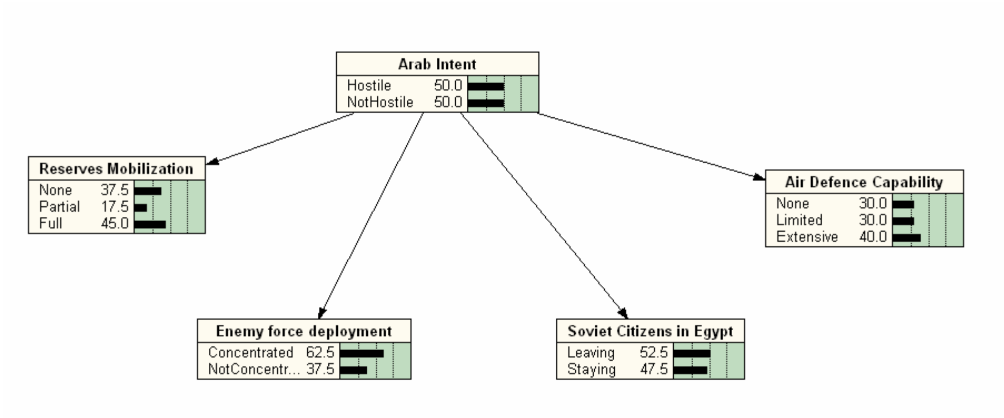


Figure 6.4: Example participant BBN.

Subsequently, Figure 6.5 presents the same BBN with the the node enemy for concentration instantiated as concentrated.

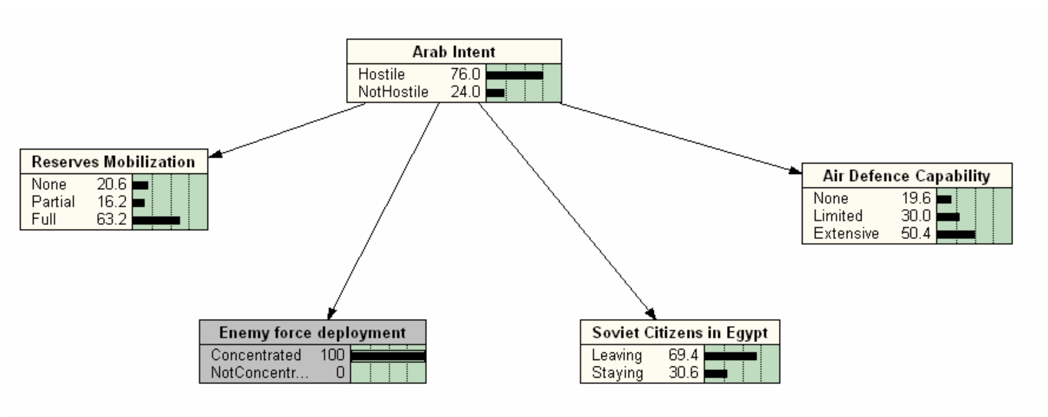


Figure 6.5: Example participant BBN with enemy force deployment node instantiated as concentrated.

From this figure the value of *H(Arab Intent | enemy force deployment = concentrated):*

$$= -\left[0.76\log_2(0.76) + 0.24\log_2(0.24)\right]$$
$$= -\left[-0.30 - 0.49\right] = 0.79$$

Figure 6.6 again shows the same network, but this time with the node enemy force concentration instantiated as not concentrated. Thus the value of *H(Arab Intent | enemy force deployment = not concentrated)* is:

$$= -\left[0.0667\log_2(0.0667) + 0.933\log_2(0.933)\right]$$
$$= -\left[-0.26 - 0.09\right] = 0.35$$

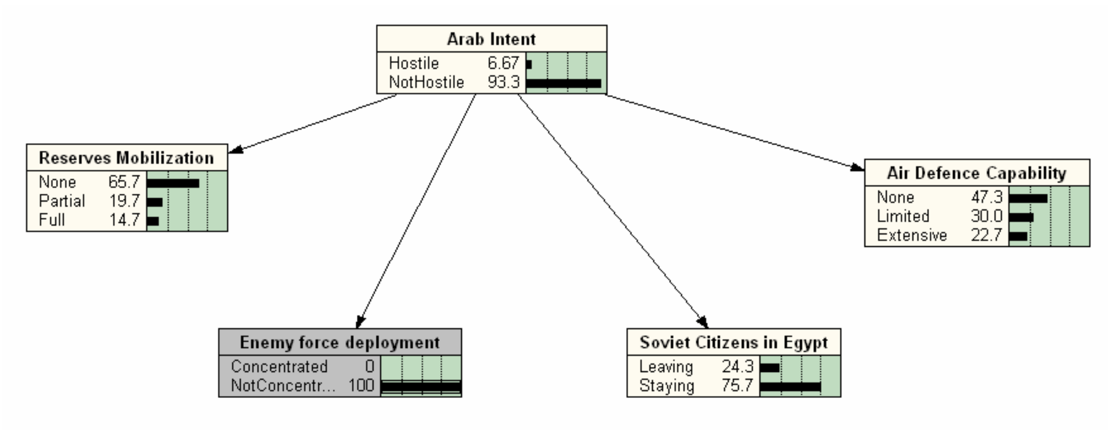


Figure 6.6: Example participant BBN with enemy force deployment node instantiated as not concentrated.

The probability of forces being concentrated is 62.5% and the probability of force not being concentrated is 37.5%, thus:

H(Arab Intent | enemy force deployment) $= (0.625 \times 0.79) + (0.375 \times 0.35) = 0.625$

The entropy value associated with the hypothesis node (based on the equal priors for each state of 50%) is 1. Therefore, overall the value of mutual information is:

I(Arab Intent; Enemy force concentration) $= 1 - 0.625 = 0.375$

Using this value of mutual information, the reduction in entropy at the Arab Intent node is 37%. That is, knowing the true state of the node enemy force concentration reduces the uncertainty in Arab intent by an average of 37%. By repeating similar calculations for the other three intelligence sources the following values of mutual information were obtained:

- Mobilisation of reserves: 46.86%
- Enemy force concentration: 37.5%
- Soviet citizens leaving Egypt: 33.23%
- Air defence capability: 18.05%

The sources with the highest and lowest values of mutual information were assigned weights of one hundred and zero respectively. The weights, W_i, of the other sources were determined such that:

$$W_i = \frac{MI_i - MinMI}{MaxMI - MinMI}$$

Equation 6.4: Weighting of mutual information.

For the participant used in the explanation of mutual information, the scaled weights and subsequent calculation of the sum of squared standard deviation is shown in Table 6.6. This value was used to assess the correlation between the direct and Bayesian weightings. As each participant only ranked four variables many standard rank correlation techniques, such as Spearman's coefficient were not truly applicable.

An example calculation of the weights and sum of squared rank deviation, for the participant used in the previous examples, is shown in Table 6.6. The cumulative results for all participants is shown in Figure 6.7.

Node	Value of mutual information	Scaled mutual information	Direct probability	Mutual info. ranking	Direct ranking	Rank deviation	Sum of squared rank deviation
Development of an air defence capability	46.86	100.00	0.00	1	4	3	18
Mobilisation of reserves	37.06	65.98	50.00	2	3	1	
Concentration of forces	33.23	52.69	100.00	3	1	-2	
Removal of foreign citizens	18.05	0.00	65.00	4	2	-2	

Table 6.6: Example sum of squared rank deviations (based upon direct and Bayesian weightings) for potential information sources.

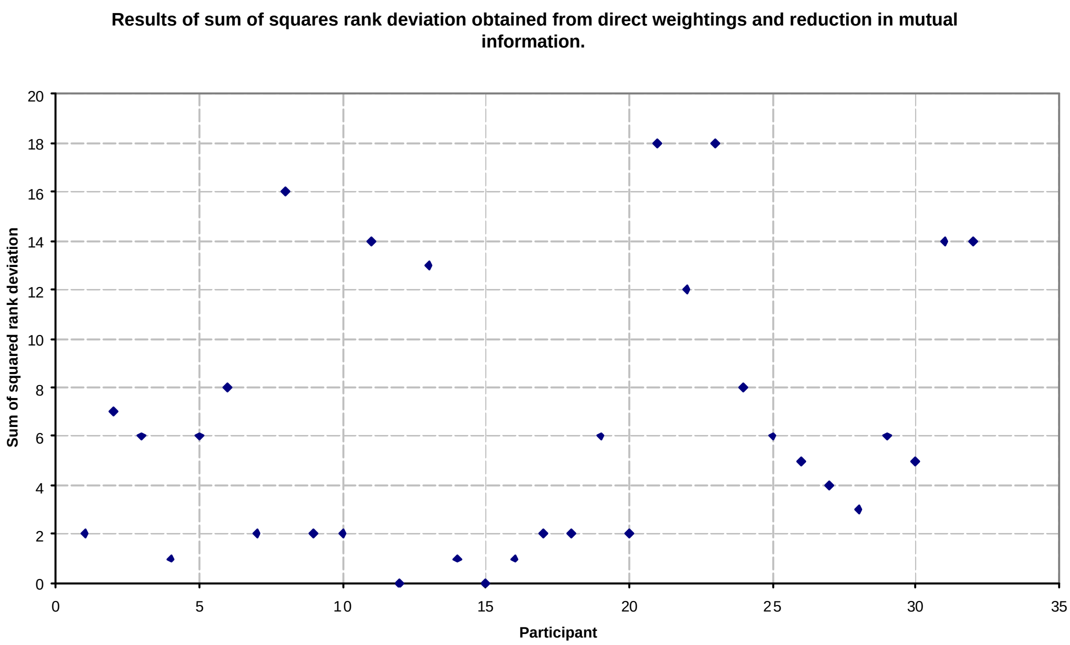


Figure 6.7: Sum of Squared results (mutual information and direct weightings) of potentially available information sources providing an indication of Hypothesis A.

Overall, Figure 6.7 presents a wide variation in the calculated results. It can be seen that calculated deviations range from zero to 18. This higher deviation is indicative of substantial disagreement between the direct and Bayesian weightings. Indeed, only two participants had zero deviation. The use of direct rankings and weightings is frequently proposed as a quick method for the development of a BBN (such as Fenton, Neil and Gallan, 2007 and Neil and Fenton, 2005). However, this research indicates such an approach may lead to substantial, yet unintentional discrepancies within the network.

6.3.2.3 Likelihood ratios of potentially available information sources.

The second part of the analysis focused on the given likelihood ratios provided by participants for each event within the timeline.

Figure 6.8 presents the variation in probability (based on likelihood ratios) for hypothesis A (Egyptian and Syrian coalition intent is hostile). The predominant trend in Figure 6.8 is for an increase in the probability of hostile intent as each subsequent event is observed. Nonetheless, there is a small group of results for which the accumulation of observed events actually leads to a decrease in the probability of hostile intent. The presence of this group within the data raises questions about the reliability of likelihood ratios as an elicitation technique.

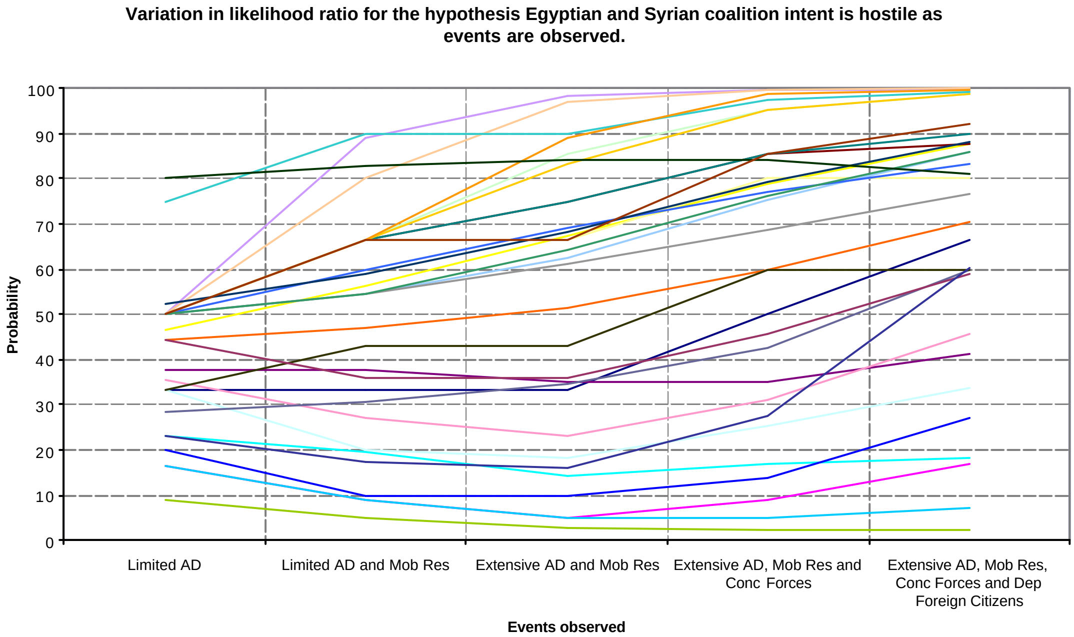


Figure 6.8: Variation in likelihood ratios for Hypothesis A.

It is, of course, entirely possible that some participants genuinely believed that the accumulation of events observed within the timeline reduced the probability of hostile intent. After all, nations develop air defence capabilities during times of peace and reserves may be mobilised for planned military exercises. That said, it is perhaps more plausible that these participants were unclear about the required ratio relating the two distinct hypotheses. It was important to try and assess if either of these situations had occurred. As such, for each participant who consistently provided likelihood ratios of less than one (and therefore a decreasing probability of hostile intent), at each event within the timeline a comparison was made between the stated likelihood value and the participant's equivalent direct subjective probability and normative value provided by their own BBN.

For every participant whose likelihood ratios indicated a decreasing probability of hostile intent, their equivalent direct subjective estimates and values obtained from their BBN showed an increase in the probability of hostile intent. As such, it was decided that these participant showed a level of rationality as in each of the elicitation techniques used, the probabilities elicited consistently moved in a given direction (albeit in two of the cases the probability of hostile intent increased and in the third it decreased). Furthermore, in two of the elicitation techniques, the probability of hostile intent increased as the events in the timeline were observed. These findings supported the view that participants were unclear about how to derive the required likelihood ratio. As such, for each participant who had consistently given likelihood ratios of less than one, the reciprocal value was calculated (thus giving an increasing probability of hostile intent as events were observed, inline with the results from the additional elicitation techniques). The result of the analysis, with this correction applied is shown in Figure 6.9.

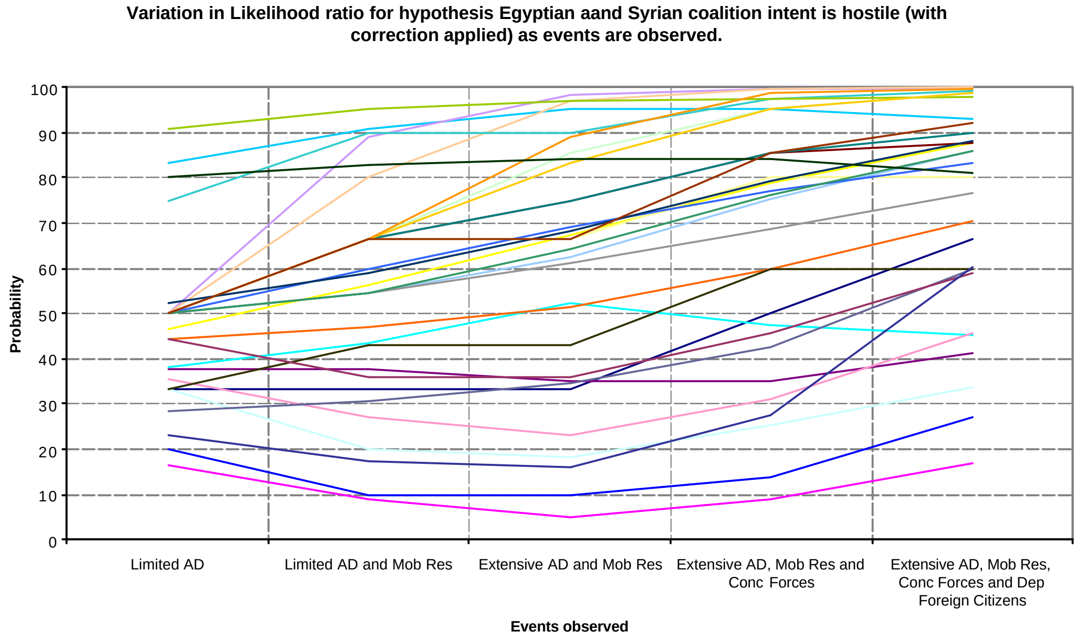


Figure 6.9: Variation in likelihood ratios for Hypothesis A with correction applied to continual likelihood ratios of less than one.

The prevailing trend depicted within Figure 6.9 appears more intuitively correct. As events are observed, there is a corresponding increase in the belief that the Egyptian and Syrian coalition intent is hostile. It is important to remember that this work has been carried out with the luxury of hindsight which permits an intuitive assessment of the expected results. Should this technique be used in a previously unseen circumstance, the developer of a support system would have to be sure that the correct ratios, and not the inverse values, had been given. This could be achieved, for example, through the use of leading questions (note alternative methods could be used) to ascertain how the participants anticipate the probabilities will move. A comparison of this and the given values would reveal any immediate discrepancies.

As previously noted, some research has considered likelihood ratios as being the most natural way in which to process information (McLaughlin and Paté-Cornell, 2005). The experiment presented here centred on a relatively short timeline comprising only five main events. Accordingly, participants may have been able to assess all the available information when determining the likelihood ratios. Yet, even in this relatively simple situation, 29% (9 out of 31) of participants appear to have provided the inverse ratios to those stated in the experimental questionnaire. It is of interest to note that during the conduct of the experiment, it was the elicitation of likelihood ratios which took the most time and raised the most questions. A great deal of effort was spent explaining the meaning of the ratios and the principle behind their derivation. Several participants felt unable to provide values and as such were unable to complete the questionnaire. Many of the participants expressed concern over their ability to provide the requested ratios at all, let alone with any level of confidence. The general impression during this part of the experiment was not one of ease, nor that thinking in such a way was indeed natural. Therefore, the use of this technique in the development of any decision support system would require a detailed elicitation questionnaire to ensure that the required likelihood ratios, and not the inverse values, had been provided.

6.3.2.4 Conditional probability distributions.

The final section of the analysis centred on the elicitation of conditional probability distributions. Each participant's stated distributions were entered directly into their own BBN (as shown in Figure 6.2). Once populated, the events comprising the timeline (detailed in Table 6.1) were sequentially entered into the BBN as evidence. Correspondingly, the beliefs in the hypotheses were updated in the light of this evidence. The results calculated by the BBN are shown in Figure 6.10 and 6.11.

The main trend observed in Figure 6.10 is for increasing probability of hostile intent as events in the timeline were sequentially observed. The corresponding decrease in the probability of peaceful intent is shown in Figure 6.11. These trends support those seen in the results of the likelihood ratios. The average Bayesian probability as each event was observed is shown in Figure 6.12. Interestingly, the reporting of reserve troops being mobilised (in addition to the development of an extensive air defence capability) did not alter the average probability in the coalition's intent. Only when forces are concentrated does the average belief in hostile intent substantially alter.

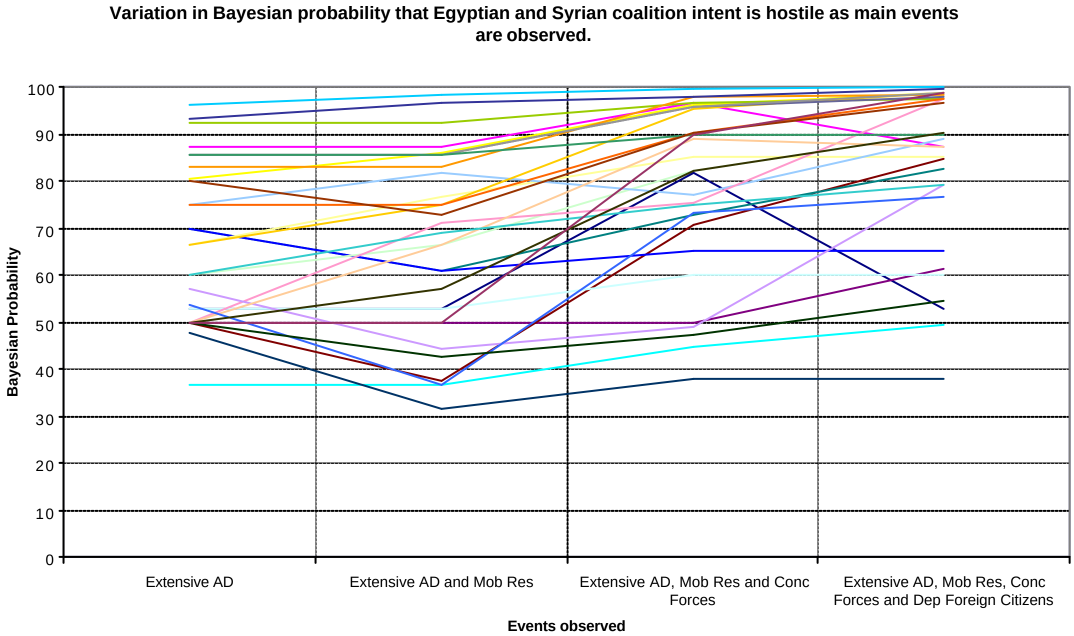


Figure 6.10: Variation in Bayesian probability for Hypothesis A.

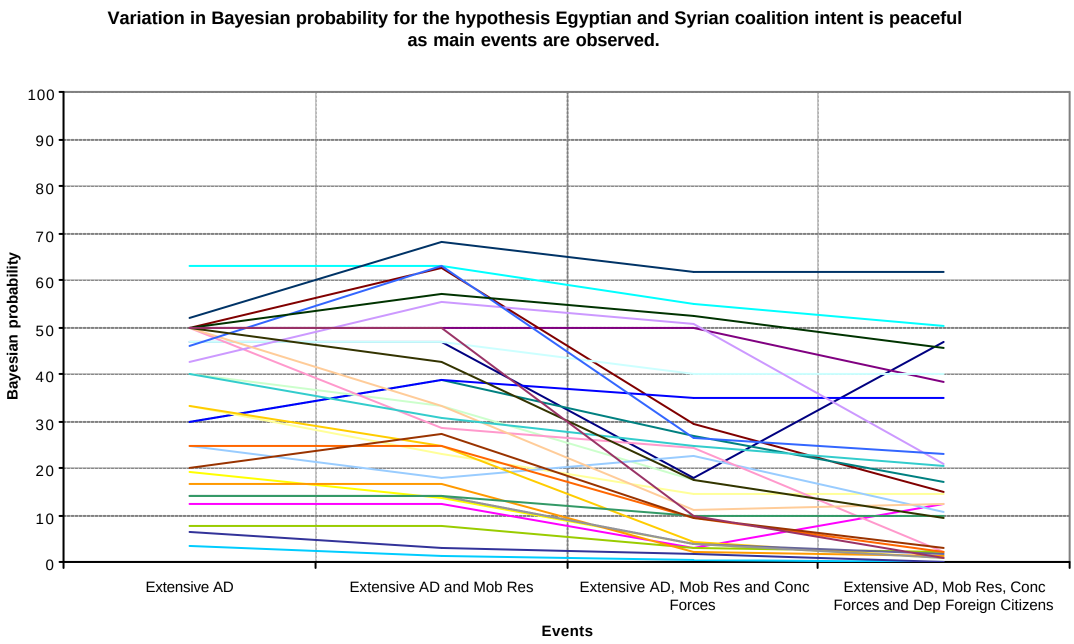


Figure 6.11: Variation in Bayesian probability for Hypothesis B.

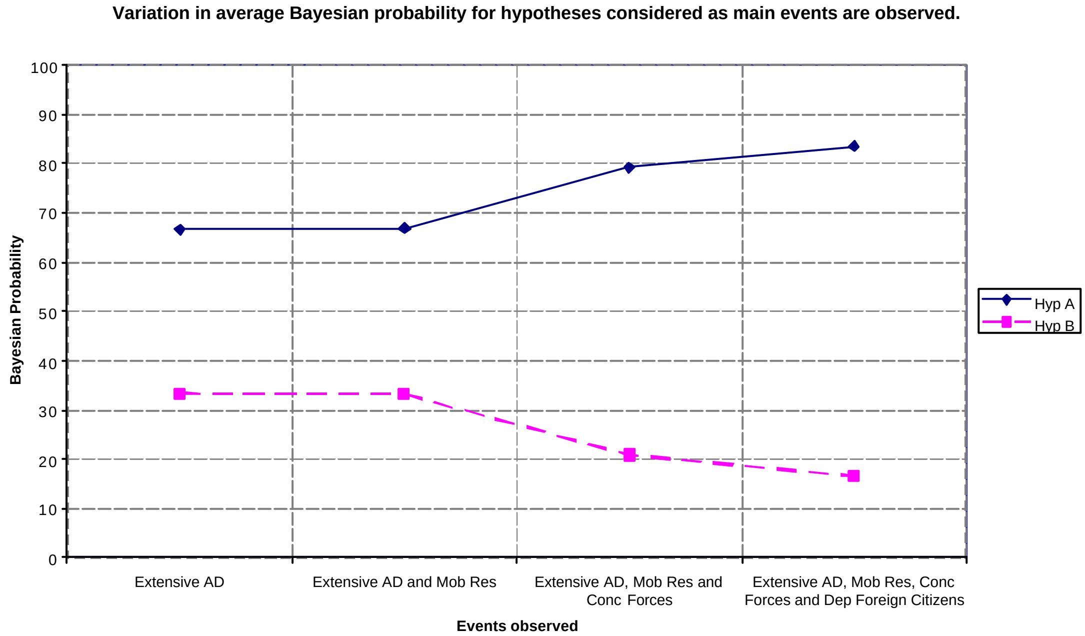


Figure 6.12: Variation in average Bayesian probability for hypotheses A and B.

As part of the experiment, participants were asked series of questions requiring direct probabilistic estimates using intuitive diagnostic reasoning. All the direct estimates were based around the events included within the scenario outline (as detailed in Table 6.1). For example, the participants were asked to consider 'what is the probability that the Arab coalition intent is hostile given that Egypt has':

- Extensively built up its air defence capability.
- Extensively built up its air defence capability and ordered a partial mobilisation of reserves.
- Extensively built up its air defence capability, ordered a partial mobilisation of reserves and concentrated its forces.

The direct estimates were compared against the corresponding Bayesian values calculated by the participant's corresponding BBN. Figure 6.13 shows the variation between the direct and Bayesian values for the hypothesis of hostile intent. The deviations ranged from -83.3% to +53.2%. Some instances of an exact match between the direct and Bayesian value were seen. However, it is unlikely that these were due to the participants being perfectly logical. It is likely to have been caused either by the structure of the network, or have been inevitable due to the conditional probability distributions used to populate the BBN. It can be seen that as there is an accumulation of events observed, the trend is for the results to move from a negative to positive deviation.

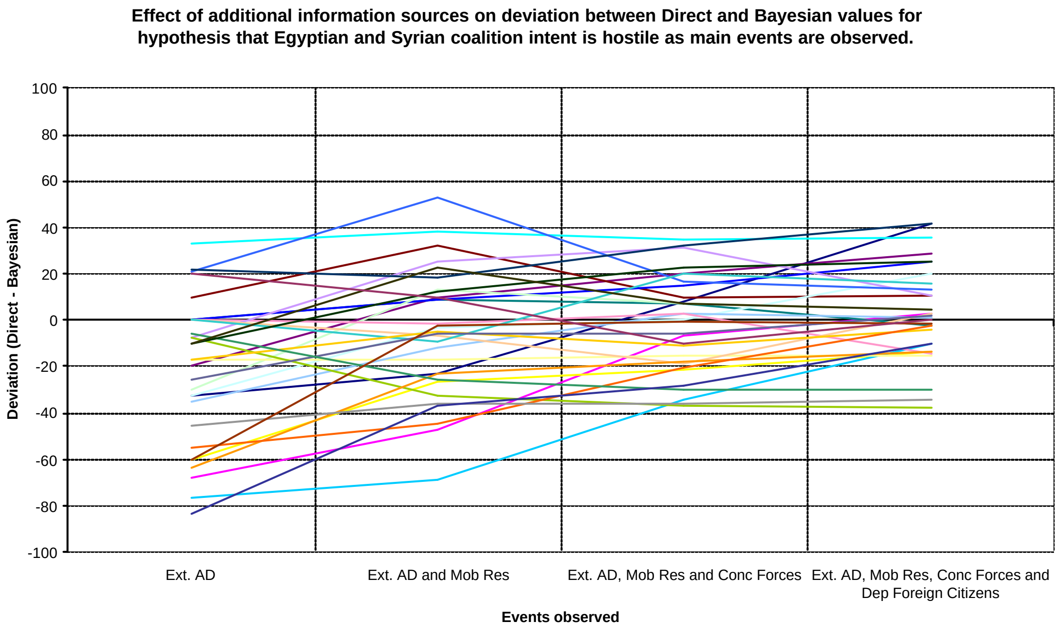


Figure 6.13: Variation in deviations (Direct –Bayesian) for Hypothesis A.

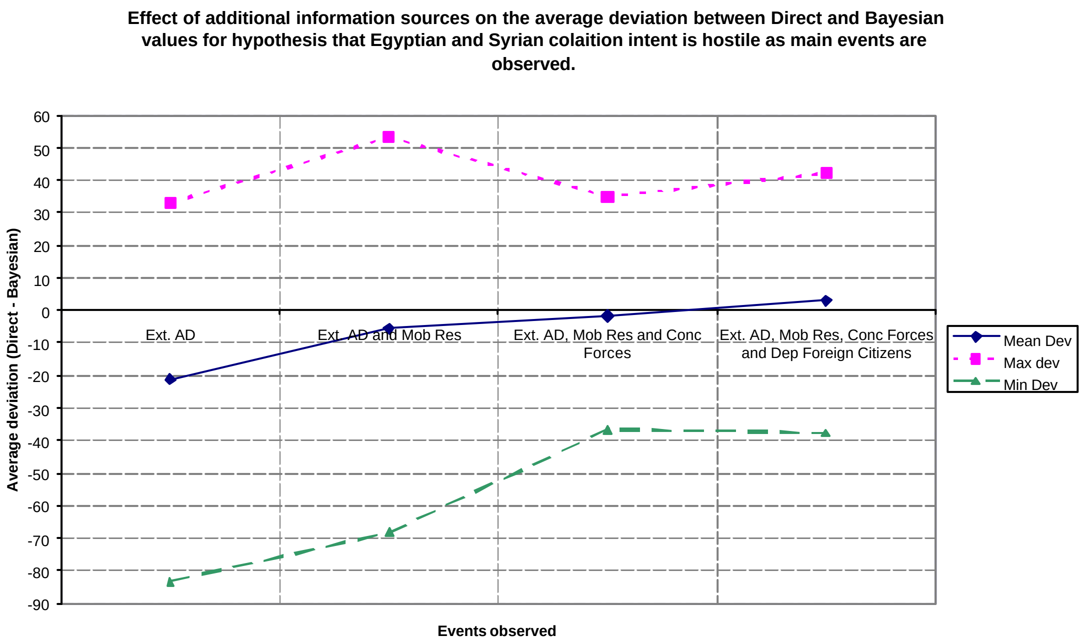


Figure 6.14: Variation in average deviation (Direct –Bayesian) for Hypothesis A.

The average deviation at each observed event is shown on Figure 6.14, alongside the overall spread of deviations (maximum and minimum). It can be seen that as an accumulation of events are observed, there is a trend for the overall spread of results to decrease. The mean deviation can been seen to follow the trend for moving from a negative to positive deviation.

Based upon the values seen in Figure 6.13 and 6.14, the Wilcoxon signed rank test was applied to each participant's deviation between their direct and corresponding Bayesian value. Using a null hypothesis that the deviations have a median value of zero, two deviations were found to be statistically significant namely when an:

- Extensive build up of air defence was been reported.
- Extensive build up of air defence, a mobilisation of reserves, a concentration of forces and the departure of Soviet Citizens were reported.

Negative deviations result from a larger Bayesian value than direct. That being so when only one or two events were observed most participants tended to underestimate the probability of hostile intent. However, as additional events were observed and reported, the deviation kept moving in the same direction. Over the course of the timeline to some extent, it is almost inevitable that the deviation between the Bayesian and direct values will decrease due to constraints placed upon the probabilities.
Figure 6.15 shows the changes in deviations between the first observed event (extensive air defence) and the last event (extensive air defence, mobilisation of reserves, concentration of forces, and departure of foreign citizens).

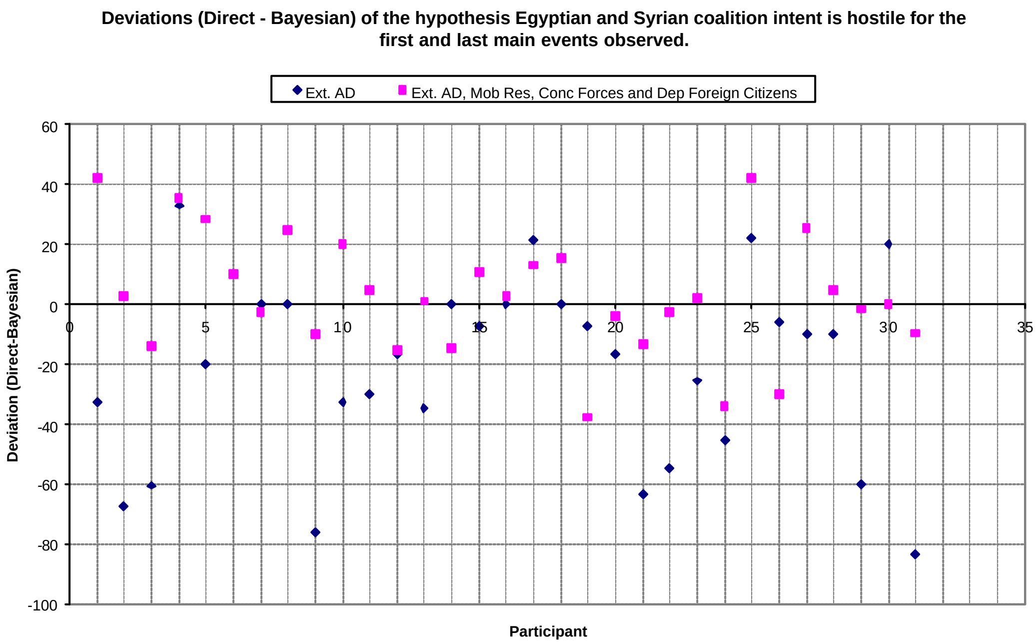


Figure 6.15: Variation in deviations (Direct –Bayesian) for Hypothesis A for the first and last main events observed.

What is particularly interesting about this plot is that for 23 of the 31 participants, the first deviation is less (i.e. more negative) than the final deviation. As seen in the average deviations (Figure 6.14), for most of the participants the first deviation is less (i.e. more negative) than the final deviation. This warranted additional investigation. It was not possible to assume that the Central Limit Theorem would apply to the data set. Neither could it be assumed that the data were normally distributed. Using a null hypothesis that the difference between the two deviations has a median value of zero, the Wilcoxon signed rank test was applied to 31 results with a correction of continuity equal to -0.5. The resulting value was $z = -2.68$ which showed statistical significance at just over the 0.01 level for a two tailed non-directional test.

Finally, it was of interest to compare participants' Bayesian range of probabilities (defined as the difference between the maximum and minimum probabilities) with their direct range of probabilities. Using a null hypothesis of a median difference in range of zero, the Wilcoxon signed rank test was applied to the 31 sets of results. There was insufficient evidence to reject the null hypothesis as the returned value of p was 0.085 which is higher than the 5% significance level. That said, the computed confidence interval (-1.0, 14.6) has a confidence level of 95.1%. As such, the null hypothesis may still be considered questionable. A larger test group may be needed to demonstrate significance.

Interestingly during the conduct of the experiment, the elicitation of conditional probabilities raised the least concerns from participants even though it constituted the longest section of the questionnaire. The most likely reason for this is that participants are used to expressing values in terms of percentages (or frequencies) as opposed to ratios or rankings. So although this section was the longest it was possibly seen as the most straightforward to complete and therefore raised the least questions. Whilst the questionnaire must be refined and developed for the situation of interest, once completed the elicitation procedure provides a clear audit trail. The development of the BBN is relatively simple and intuitive to review. In addition to this, the structure of a BBN may be easily manipulated, allowing expansion or contraction of the modelled domain.

However, any changes may require additional conditional probabilities to be elicited. Furthermore, should the network provide unexpected results, the analyst would be forced to reconsider the values and assumptions behind it. This could help prevent subconscious preferences being given to particular intelligence sources.

Consideration of the results obtained from the conditional probabilities show that participants were not always able to coherently assess the information available to them. When only one or two events were observed (and thus reported) participants generally underestimated the probability of hostile intent with their direct estimates. However, as additional events were reported participants were unable to correctly manipulate the interdependencies between the events. Consequently, participants began to overestimate the probability of hostile intent. In a military situation such under and overestimation of the likelihood of hostilities commencing could have undesirable consequences.

6.3.2.5 Comparison of likelihood ratios and direct estimates requiring diagnostic reasoning.

In theory, if participants were perfectly coherent then the same results would have been elicited throughout the experiment. However, as deviations in the results have been seen it was important to determine if these deviations were random or systematic. Therefore, building upon the comparison of likelihood and direct probabilistic estimates based upon intuitive diagnostic reasoning against a normative model, it was of interest to compare the likelihood and direct values against one another. The calculated deviations are shown in Figure 6.16. The average deviation and overall spread of deviations calculated at each event in the timeline are shown in Figure 6.17.

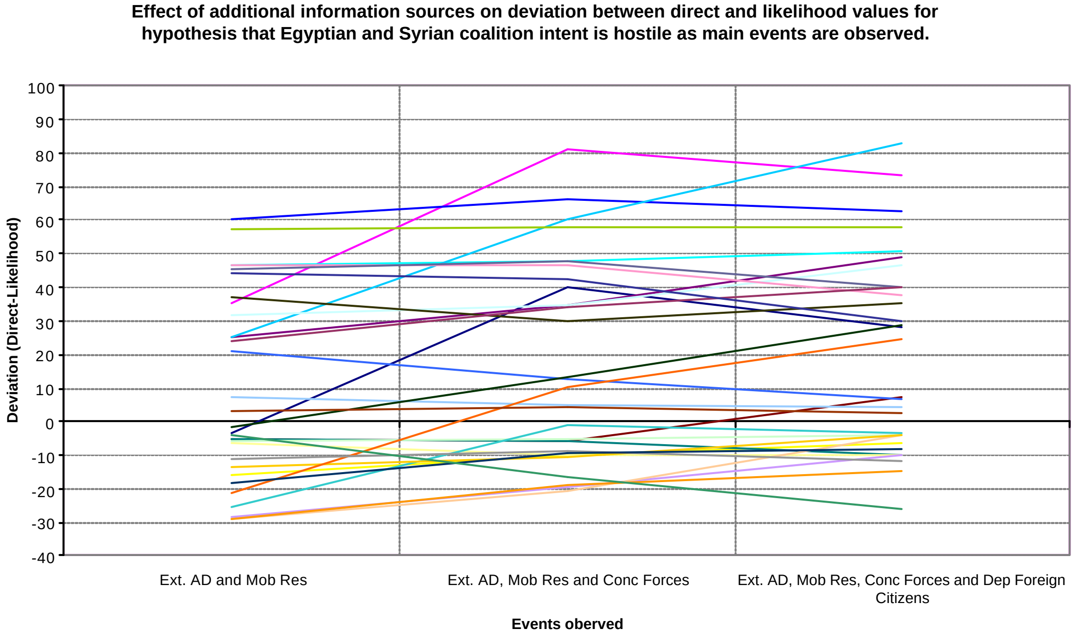


Figure 6.16: Variation in deviations (Direct –Likelihood) for Hypothesis A.

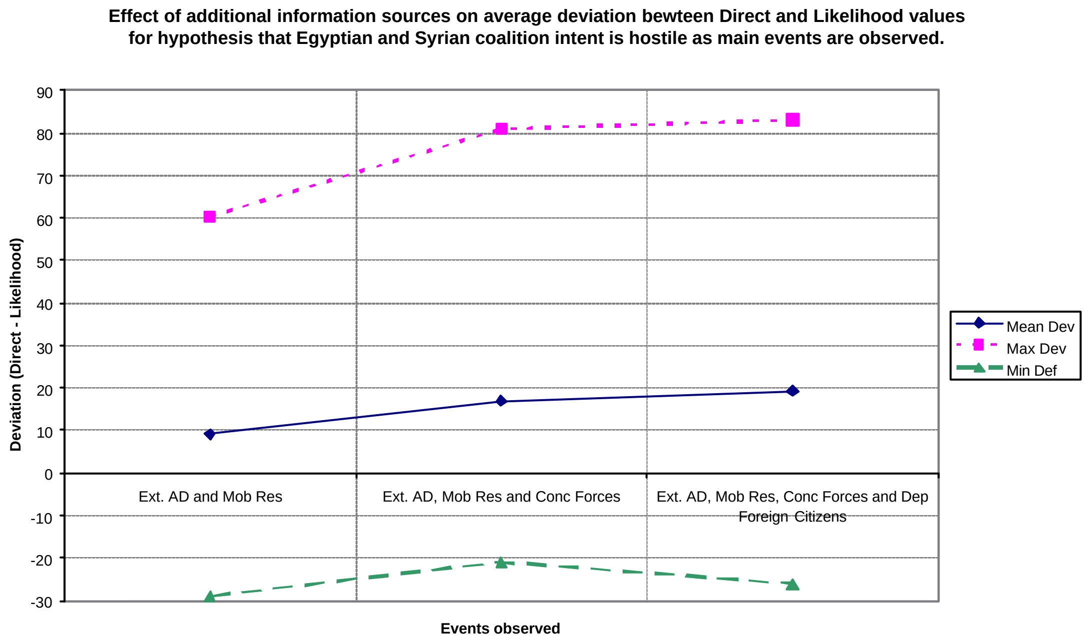


Figure 6.17: Average variation in deviations (Direct –Likelihood) for Hypothesis A.

Both Figures 6.16 and 6.17 show that as additional information becomes available at each event, the deviation between the direct and likelihood ratio value increases. This can easily be seen by the variation in the mean deviation in Figure 6.16. It is interesting to note that the average deviation lies closest to the lower end of the spread of results across the events observed within the timeline. However, the overall trend seen in Figures 6.16 and 6.17 supports the results seen in the comparison of the direct and normative values.

6.4 Experiment into subjective value elicitation with intelligence personnel: A comparison of direct rankings, likelihood values, and conditional probabilities for use in decision support systems.

6.4.1 Introduction.

The participants in the main experiment were not experienced in intelligence analysis. Therefore, it was considered important to repeat the experiment with a group of participants who have current knowledge in this area. In total five serving military officers completed the experiment via e-mail. Participants were encouraged to seek clarification on any ambiguous points. No questions were raised. An additional sixth participant completed the experiment who had extensive knowledge of intelligence analysis but was no longer a serving military officer. This final participant completed the experiment through a face to face interview. Once completed, this participant was shown the populated BBN and was given an opportunity to revise their conditional probability distributions. It is accepted that this could create a difference in the results obtained.

6.4.2 Ranking and weightings of potentially available information sources.

As in the main experiment, the two most influential intelligence reports were the mobilisation of reserves and enemy force concentration. However, in contrast to the main study, the majority (four) responses placed enemy force concentration as the most influential, not the mobilisation of reserves.

This final experiment also showed surprisingly little agreement between the direct and BBN normative rankings based upon mutual information. Indeed, interrogation of all the BBNs found only one network which ranked the concentration of forces as the most influential intelligence report. The results of the interrogation revealed that four (the majority) of BBNs ranked the mobilisation of results as the most influential potential intelligence source. This is supported by Kendall's tau statistic which showed no indication of a strong agreement between the rankings (the highest value calculated was 0.33).

Further to this, the direct and normative weightings associated with the potential information sources were analysed using the sum of squares rank deviations. The results varied from 2 to 16. Overall, the use of directly ranking nodes as a technique for the rapid development of a BBN could introduce substantial discrepancies within the results derived from the network.

6.4.3 Likelihood ratios of potentially available information sources.

These results showed substantial differences from the trends seen in the main experiment. The overall trend for each participant is an increase in the probability of hostile intent as events are observed. There are no examples of a continual, or overall, decrease in the probability of hostile intent as events were observed and reported. This result is illustrated in Figure 6.18 which shows the average probability based upon the likelihood ratio as each event in the timeline unfolds.

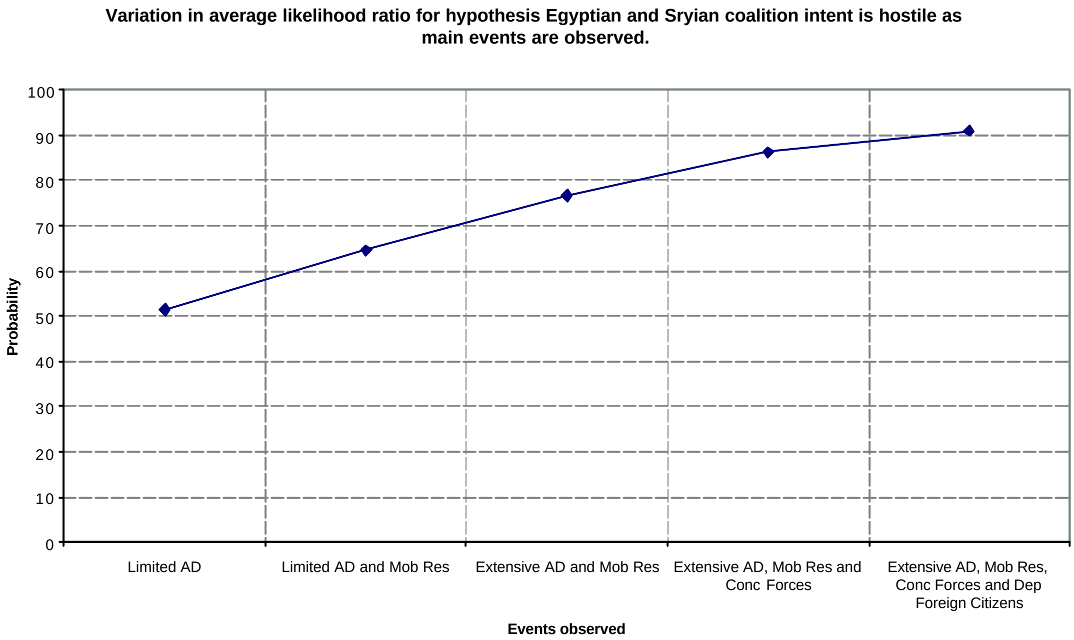


Figure 6.18: Average variation in likelihood values for the Hypothesis A.

It was previously postulated that likelihood ratios consistently below one were caused by the participants being unclear about the required ratio relating the two distinct hypotheses. The fact that all participants in this experiment seemed to have understood the ratios required may be due to their relevant experience in handling intelligence information. In comparison to the previous experiment, participants did not express the same levels of concern or raise as many questions whilst completing this section of the experiment. As such, it is possible that likelihood ratios may be an acceptable elicitation technique for use with individuals who are used to manipulating intelligence data or work with ratios and mathematics as part of their daily role.

6.4.4 Conditional probability distributions.

The variation in Bayesian probability of hostile intent is shown in Figure 6.19, which has two engaging features. Firstly, the observation of the second event in the timeline splits the results into two groups. Three of the participants perceived the observation of partially mobilising reserves following the development of an air defence capability decreases the probability of hostile intent. Yet, for the remaining participants this observation increases the probability of hostile intent. For five of the participants all the remaining events observed increase the probability of hostile intent. The remaining participant decreased the probability of hostile intent upon the observation of forces concentrating, but increased the probability following the report of foreign citizens departing.

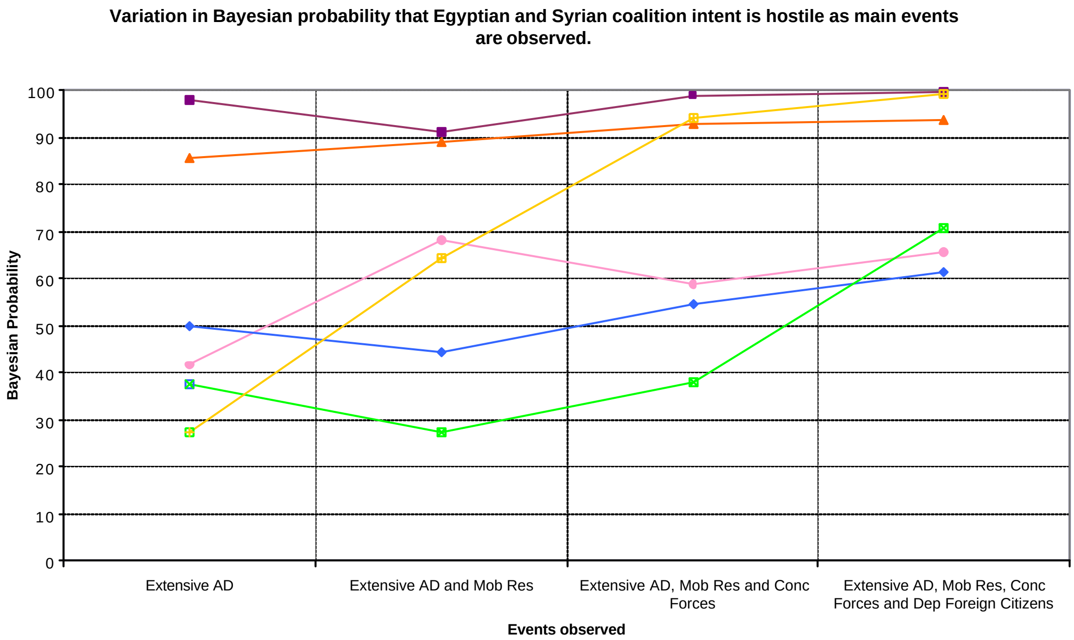


Figure 6.19: Variation in Bayesian probability for Hypothesis A.

The second observation of note in Figure 6.19 is that each participant has a higher probability of hostile intent when all four events have been seen than at the start of the timeline. The average variation for both hypotheses is given in Figure 6.20. As seen in the previous larger scale experiment, the observation of reserves being partially mobilised did not substantially alter the average probability of intent. Only when, in addition to this, were forces seen to concentrate was the probability of intent seen to alter.

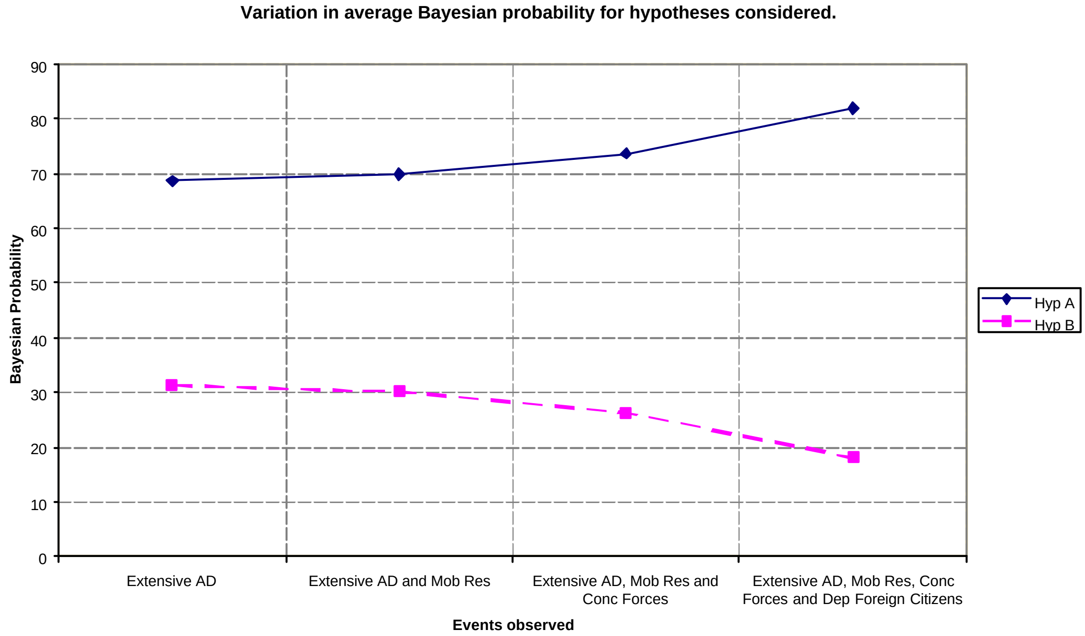


Figure 6.20: Variation in average Bayesian probability for Hypotheses A and B.

Subsequently, the participants' direct probabilistic estimates based on diagnostic reasoning of events within the timeline were compared against normative values calculated from their own BBN. Figure 6.21 presents the deviations between these two values for each event within the timeline. The trend supports that seen in the previous experiment (Figure 6.13) that is for a move in deviations from negative to positive as events unfolded. This was potentially explained by participants underestimating the probability of hostile intent when only one or two events has been observed, yet overestimating the probability when three or more events had been observed. This trend was not repeated in the results of this experiment. What can be seen is the movement of each participant's deviation towards zero as the events are observed. As previously explained, this result is almost inevitable due to the constraints placed upon the numbers available.

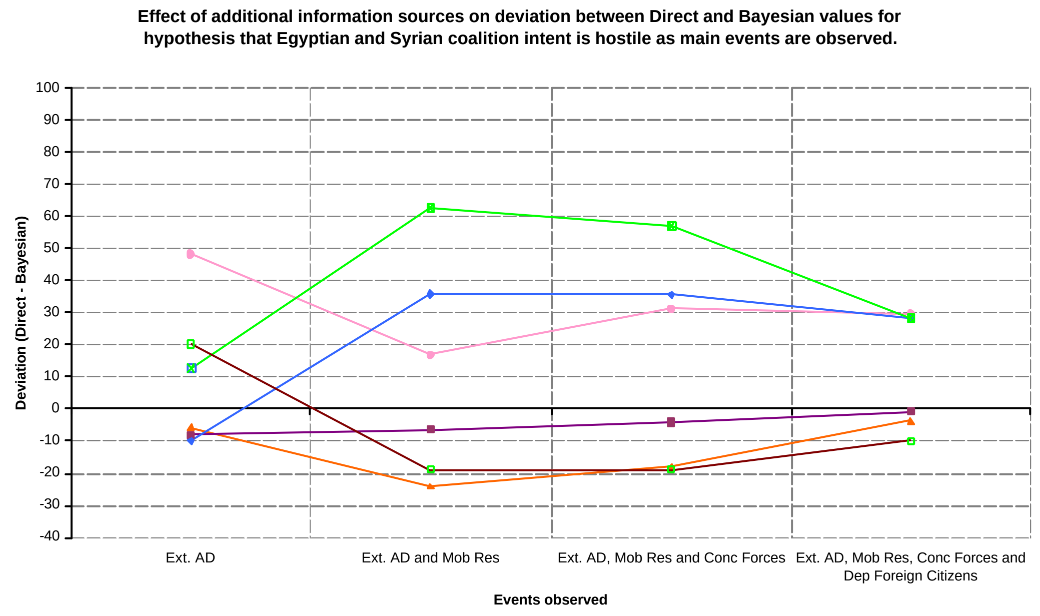


Figure 6.21: Variation in deviations (Direct – Bayesian) for Hypothesis A.

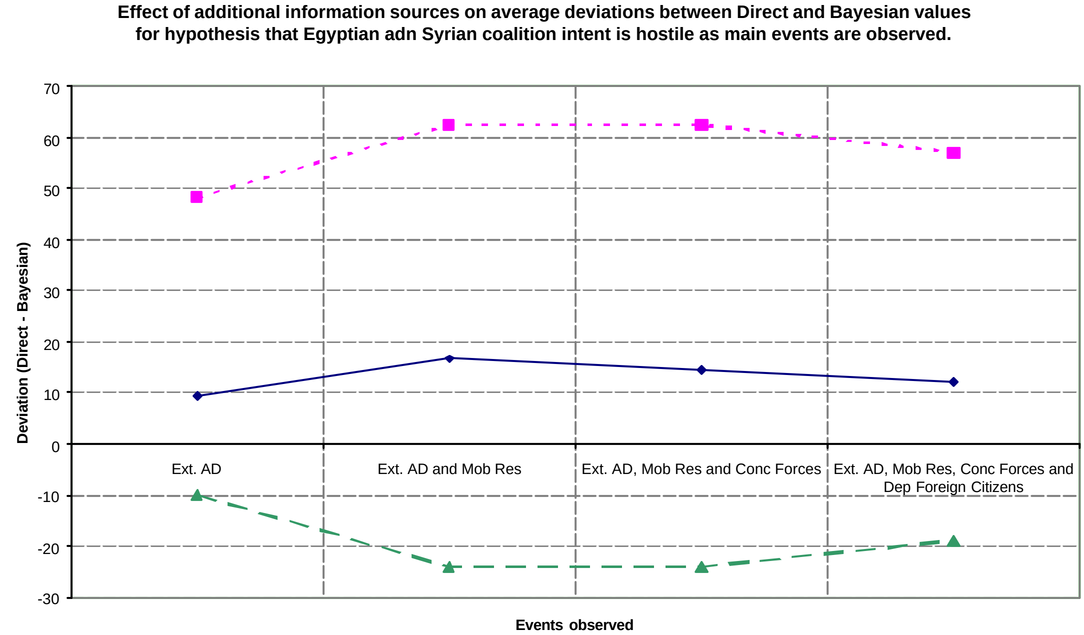


Figure 6.22: Mean variation in deviations (Direct – Bayesian) for Hypothesis A.

Figure 6.22 presents the average deviation between the directly estimates based on diagnostic reasoning and the corresponding normative Bayesian values. For comparison, the overall spread of results is also shown in Figure 6.22. The average deviations shown in Figure 6.22 are smaller than the corresponding average deviations seen in the main experiment (Figure 6.14) and overall are closer to the lower part of the spread of results. This raises the possibility that participants who regularly manipulate intelligence data may be more consistent than those who do not.

6.4.5 Comparison of likelihood ratios and intuitive diagnostic probabilities.

Finally, a comparison was made between the likelihood and direct estimates probabilities based upon diagnostic reasoning. The deviations between these values for the hypothesis of hostile intent are shown in Figure 6.23, with the average deviation and spread of results shown in Figure 6.24. It is clear that the shape of Figure 6.23 is similar to that of 6.21, with the deviations moving towards zero as events are observed. The average deviations between the direct conditional probabilities and likelihood ratios gave the smallest deviations seen during the research (all below 5%).

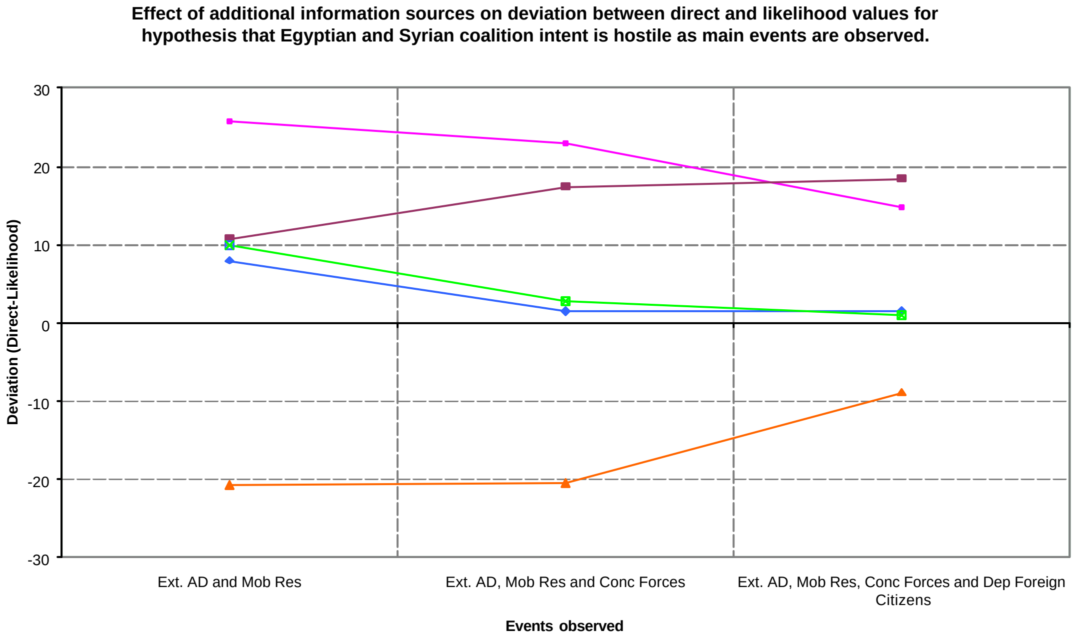


Figure 6.23: Variation in deviations (Direct – Likelihood) for Hypothesis A.

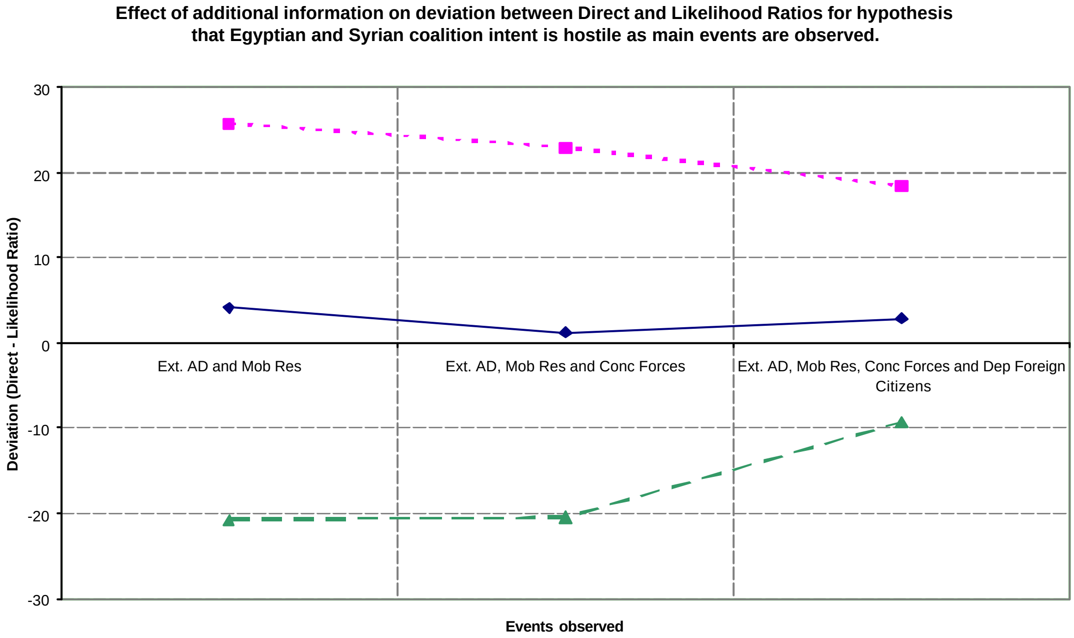


Figure 6.24: Average variation in deviation (Direct - Likelihood) for Hypothesis A.

It was of interest to further analyse the deviations between the likelihood values and direct probabilistic estimates from diagnostic reasoning. Based upon the absolute average deviation between these values the results from both 6 participants in the final study and the 31 participants in the larger study were used in a two tailed Mann-Whitney non-parametric test. The returned p-value was 0.1435 which is greater than the critical test value of 0.05 indicating no significant differences at the 5% level. This result does not prove the null hypothesis of no difference between the deviations seen in the results of the participants whom had experience of working with military intelligence and those who did not. However, the null hypothesis is plausible. To demonstrate a statistically significant difference between the two groups would, it is suggested, require a larger sample of results from officers with intelligence experience.

6.5 Conclusions and chapter summary

6.5.1 General conclusions

This thesis argues that BBNs provide a suitable basis for the development of a system capable of quantitative analysis of intelligence reports. One of the main issues associated with the development of a BBN is the construction of the probability distributions used to populate the network. Often, the required distributions may only be derived subjectively. Therefore, to further the body of knowledge available in this area, an experiment has been conducted to compare three elicitation techniques: direct ranking of the variables' perceived importance for discriminating between given hypotheses, likelihood values and conditional probabilities. The focus of this experiment was the extent to which the different elicitation techniques lead to equivalent or different judgements.

Based upon the main events observed leading to the start of the Arab Israeli conflict of 1973, a timeline and supporting BBN were developed. The BBN provided a normative model against which the subjective values were compared and analysed for indications of biases and numerical trends.

Two experiments were conducted. The first generated 31 sets of results from serving military officers studying at the Defence College of Management and Technology. Subsequently, a second experiment was completed by 6 participants, all of whom had extensive knowledge of military intelligence. Five of these participants were currently serving officers.

6.5.2 Rankings

Comparisons were made between the direct rankings and the BBN rankings (based upon mutual information between each source and the hypothesis node). Both experiments placed the mobilisation of reserves and the concentration of forces as the two most influential nodes. The larger experiment considered the mobilisation of forces as the most influential. Conversely, the experiment comprising of those with intelligence experience placed the concentration of forces as being the most influential information in this situation. In both experiments there was little agreement between the participants direct and associated BBN rankings. The lack of correlation between the rankings was further shown through the calculated values of Kendall's tau statistic. Overall, only two participants had a tau statistic value greater than 0.9 which is representative of strong agreement.

6.5.3 Weightings.

The analysis was undertaken through the calculation of the sum of squared rank deviations, based upon direct and BBN weightings. The BBN weightings were based upon mutual information between each source and hypothesis node. A wide variation in the calculated results was seen from zero to 18. High deviations are representative of substantial disagreement between the direct and Bayesian weightings. The use of direct ranking and weightings is proposed as a quick method for the development of a BBN. However, this research indicates such an approach may lead to substantial, yet unintentional discrepancies within the network.

6.5.4 Likelihood ratios.

The overall trend was for an increase in the probability of hostile intent as events within the timeline were observed. However, within the larger experiment there were a small group of participants for which the accumulation of observed events actually lead to a decrease in the probability of hostile intent. For these participants, a comparison of the likelihood and direct conditional probabilities highlighted a major discrepancy. For each result in which the likelihood values indicated a decrease in hostile intent as events were observed, the corresponding direct conditional probabilities showed and increase in the probability of hostile intent. Therefore, it is concluded that a group of participants provided the inverse likelihood ratio to that requested within the experiment. Interestingly, no examples of the inverse ratio were found in the experiment conducted by those who have experience of military intelligence.

It is concluded that the use of likelihood ratios are really only applicable to those who have experience of working with intelligence data, or with ratios as part of their work.

For those not familiar with the derivation of such ratios, the elicitation procedure caused concern over their ability to provide the required value with any level of certainty. If likelihood ratios were used to provide the subjective values within a BBN, the elicitation procedure would require additional information to ensure the ratios had been given in the required format. For example, the procedure could include leading questions to ascertain which how the participants anticipate the probabilities will move. A comparison of this and the given values would reveal any immediate discrepancies.

6.5.5 Conditional probability distributions.

Each participant's conditional probability distributions were entered directly into their own BBN. Once populated, the events comprising the timeline were sequentially entered into the BBN as evidence. This served to update the probability of each hypothesis. The main trend observed was for an increasing probability of hostile intent as the events were observed. It is of interest to note that the observation of the second event (the partial mobilisation of reserves following a limited air defence capability) did not substantially alter the probability of hostile intent. Only when, in addition to these events, was a concentration of forces observed did the probability of hostile intent noticeably increase.

Participants also provided a series of direct probabilistic estimates using diagnostic reasoning. All direct estimates were based upon events within the scenario timeline. The results shows a move from negative deviation (caused by the Bayesian value being the largest) to positive (caused by the direct probability being the highest). That being so when only one or two events were observed most participants tended to underestimate the probability of hostile intent. Participants appear unable to correctly manipulate the dependencies between the observed events within the intelligence reports. For example, the concentration of forces and mobilisation of reserves may be indicative of an attack or equally be part of a planned military training exercise. The departure of foreign citizens could be due to a national holiday, expulsion or be based on insider information. Yet the combination of these factor creates a context in which there are dependencies between this information. The context is key to understanding the dependencies between the intelligence reports received.

Over the course of the timeline the deviations tended towards zero. To some extent, it is almost inevitable that the deviation between the Bayesian and direct values will decrease due to constraints placed upon the probabilities. The elicitation of conditional probability distributions raised the least questions of any section of the experiment. This is most likely due to participants being used to expressing probabilities as percentages or frequencies.

In summary, the results of the analysis show that:

- As evidence accumulates participants are unable to correctly manipulate the dependencies between information in intelligence reports within the context being considered.
- The results suggest participants indirectly expressed varying preferences as to the importance of information depending upon the elicitation technique used.
- There is some evidence to suggest correlation between direct and normative rankings, however correlation is not as high as might be expected.
- Participants who were not regularly exposed to intelligence information, nor regularly used ratios were not able to easily, nor effectively express probabilities through the use of likelihood ratios.
- Participants who are regularly exposed to intelligence data provided more consistent values across the elicitation techniques used than those who did not have the exposure to intelligence data.
- Conditional probability tables provide a consistent methodology for the elicitation of subjective preferences.
- Bayesian Belief Networks offer a flexible, robust methodology for the development of a normative model upon which a decision support system for the quantitative analysis of intelligence data may be built.

CHAPTER 7 CONCLUSIONS.

Intelligence failures are rarely caused by a lack of information. Often, the failure is caused by an inability to piece information together; to correlate the many small pieces of data into a larger picture. Essentially, the collation, analysis and communication of intelligence is an integral part of risk management: what enemy weakness can be exploited, what contingency plans should be developed? There is an almost ever-increasing complexity to military intelligence analysis. In today's rapidly changing environment there is a need to predict asymmetric attacks from disparate groups of people whose overall access to technology and military training is uncertain. Such predictions will be based upon information which may be fragmentary, incomplete, imprecise and provided from diverse sources over a sustained timeline.

Reports into the major 'intelligence failures' of the past decade have all suggested improvements within the field of intelligence analysis. The 9-11 Commission report (2004) made a series of wide-ranging recommendations relating to changes in the structure of the intelligence community. However, the report did not make any recommendations relating to the intelligence process, nor the tools or techniques used within it. Conversely the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction made specific recommendations relating to the intelligence analysis process. The WMD report (2005) states that the failure to find WMD was due to a combination of factors: insufficient information, insufficient number of alternative hypotheses considered, lack of clear communication about the uncertainty and confidence in any predictions made. Within the UK, the Butler report suggested a review of intelligence staff with the specific aim of determining if they have all the resources required to complete their work.

This thesis has investigated the development of a support tool for analysts working with, and reasoning about information which is uncertain, incomplete or imprecise. It has been argued that the basis of such a tool should be a Bayesian Belief Network (BBN)

A BBN is a probabilistic framework represented through a directed acylic graph and a set of probability distributions. These flexible, robust networks are able to incorporate values from a range of sources including empirical values, experimental data and subjective values. Key advantages of BBNs are their ability to show the variation in the probability distribution over a series of hypotheses and support both diagnostic and predictive reasoning. BBNs have defined techniques for representing uncertainty associated with information within a network. Furthermore, the processes required for formulating a network, deriving the required distributions and finally entering new information are all transparent processes open to audit and debate. The procedures necessitate explicit consideration of uncertainty on both the information used in and inferences drawn from the analysis. All of the numerical results obtained from the BBN at the end of the analysis are defendable, open to scrutiny and amenable to sensitivity analysis.

Inevitably, all techniques have strengths and weaknesses. With respect to a BBN, the main weaknesses are ensuring all plausible hypotheses have been identified and captured within the network; and the use of subjective values may require too much time from domain experts and incorporate cognitive biases into the network.

This thesis has considered various aspects of the development of a decision support tool based upon a BBN through both fictitious and historical case studies. For each case study considered, a supporting scenario and BBN were developed. Within the scenario, a series of intelligence reports were presented over a defined timeline.

Military situations often involve complex decisions. Decisions may be complex for many reasons such as the requirement to consider many interlinking factors. In addition to having to fully evaluate many factors, which in itself is complex, additional complexity may arise from information on the factors containing uncertainty or being incomplete. Further complexity may arise from the individuals involved in making a decision; from the number of stakeholders required to reach a consensus to the emotions and personal values driving each individual stakeholder. It is important to remember that the outcome of military decisions may be perilous. In such environments, a decision may only be seen as acceptable, reasonable or rational when fully considered within the context in which it was taken. If there is real difficulty in predicting the outcome of a decision, it is common to refer to the risk associated with the identified courses of action. Whilst how the risks are viewed is dependent upon the framing of the decision, the acceptance of risk is personal to each decision-maker. One of the most common methods for expressing risk is through probabilities.

This thesis elicited the required subjective values through the use of specifically designed questionnaires. Experimental participants were, whenever possible, met on a face to face basis in order to discuss the scenario and any points of ambiguity relating to the required subjective values. The participants were required to explicitly state their subjective values, as opposed to having the required values derived from theoretical bets. To facilitate the experiment both the BBN and questionnaire used were kept as simple as possible. It is an accepted weakness of the experimental methodology that no measure of the participants' uncertainty in the subjective values they provided was elicited, nor were any of the participants calibrated. This weakness increased the chance of the participants being overconfident in the values they gave. As part of the experiment, participants were encouraged to provide probabilities in the format with which they felt most comfortable either as a stated number or a frequency. The presentation of values as frequencies has also been shown to reduce errors within the stated answers.

Although no calibration of the participants answers were sought, participants were questioned to ascertain if they had any previous experience of working with intelligence reports. Any relevant experience may have developed their skills at combining multiple data sources and uncertainty.

Decision support systems have applicability in situations where the sheer volume of information or number of contributing factors creates a situation in which an individual cannot make an informed decision alone. This research has centred upon military intelligence analysis in which an individual has an opportunity to review the information as it becomes available. In today's relatively peaceful world the majority of intelligence is focused on continuous processes aimed at providing security and prosperity.

Situational awareness is an area in which BBNs have clear applicability due to their ability to update the belief in hypotheses given some new evidence. Any evidence may be reported in a variety of sources. Often, the provision of a rounded estimate will require the fusion of several forms of intelligence – for example, an image of troop movements supported by additional human or signals intelligence. The revised estimate of an outcome can be used in the prediction of future events.

Of course, the collation of data and its interpretation into intelligence does not in anyway guarantee that good decisions will follow. Intelligence analysis is prone to many cognitive biases including the belief that the opposition is rational, over riding strongly held preconceived ideas, and the belief that risky options will simply not occur. Information to inform the decision-making process may be presented in the form of an intelligence report. Whilst the timely provision of an intelligent report presented in a format which is easily understood by the recipient facilitates the reading of the report, it does not ensure that the information is understood or accepted. Outputs from a DSS which are at odds with the generally perceived view of a problem may be seen as a threat to an individual's credibility or simply be disregarded. This is an important point. The use of a DSS stops at the end of the analysis. A DSS does not implement a decision; the ultimate responsibility for taking a decision still rests with an individual. If they do not understand the results of the DSS, do not believe in its credibility or reliability then they may exercise their judgement and not use the results of the analysis. This is their prerogative and ultimately a decision will still be taken.

Within the field of DSSs for military intelligence, this work has mainly investigated human biases within the analysis process. The developed DSS could be used with existing knowledge of indicators and warnings to identify when observed events may be suggestive of behaviour which is at odds with the generally held perception of a situation.

The DSS investigated as part of this research was not intended to replace the human analyst, merely to enhance their capabilities by providing a clear audit trail and second opinion. If both the analyst and DSS arrive at the same conclusion there is extra credence in the assessment made. However, any inconsistencies between the two results should lead to the analyst re-evaluating their work. This will ensure the analyst understands why there is a difference within the results which subsequently reduces the chance of an oversight or error within the analysis on their part.

The first part of this research sought to investigate:

- The impact of additional information upon an individual's perceptions of a situation.
- The ability of individuals to logically and consistently combine the available evidence.
- The ability of individuals to assess the discriminative value of information.

The scenario used for the experiments in supports of these objectives was a fictitious land-based section attack against an enemy outpost which may be reinforced. The experiment comprised of two questionnaires. One elicited subjective probability tables in the causal direction of influence. The other elicited direct subjective probabilities for various propositions requiring diagnostic reasoning. A BBN was used to provide a normative model of the situation and provide the Bayesian probability of the propositions of interest.

The results found that as additional information became available the participants' view of the situation did alter. However, it was clear that participants were unable to logically and consistently combine the available information within the intelligence reports to make accurate assessments of the situation being faced. The results also showed participants stated a preference for HUMINT over SIGINT. However, for many participants, the normative result indicated that the information received from SIGINT sources was in fact the most valuable in removing uncertainty about the situation.

It was unclear why participants were not able to give sufficient credence to information received from electronic sources. This has implications relating to the design of a DSS. Overall, the results of the first set of experiments concluded that there was merit in the use of BBNs as a DSS to support intelligence analysis by:

- Having a structure which is open to scrutiny and provides a clear audit trail of the assumptions, hypotheses and data used.
- Logically and consistently combining multiple intelligence reports obtained from a variety of sources thus ensuring all reports are fairly considered and evaluated.
- Logically and consistently update the belief for each hypothesis considered as new evidence or data enters a network.

This latter point would, in theory enable a decision-maker to identify when, in their opinion, additional information should be sought to clarify a situation.

To further investigate the results of the initial investigations a BBN and scenario were developed within the context of the main events leading up to the start of the Arab Israeli conflict of 1973. The purpose of the experiment was to illustrate the potential benefits and strengths, as well as any weaknesses in the use of BBNs for intelligence analysis. For the purposes of the experiment, the BBN was refined and populated over a series of meetings with a military historian at the Defence College of Management and Technology.

As the focus of this part of the research was the use of BBNs for intelligence analysis, it was of most interest to develop the network from the point of view of an Israeli intelligence officer, as it was the Israelis who were surprised at the start of the hostilities. The results obtained supported the result seen in previous experiments that as multiple information sources became available the historian was unable to correctly manipulate the interdependencies between the information sources. In addition to this, the analysis revealed inconsistencies between the relative importance of each possible information source as directly ranked by the historian and the normative rankings provided by the BBN.

In both experiments stated preferences for reports and the considered relative importance of information sources varied between the direct and normative results. Therefore in addition to the previously stated requirement of a DSS the system must also:

- Assist in determining the relative importance and influence of each potentially available information source in discriminating between a set of known hypotheses.

This thesis argues that BBNs provide a suitable basis for the development of a system capable of quantitative analysis of intelligence reports. One of the main issues associated with the development of a BBN are the probability distributions used to populate the network. Often, the required distributions may only be derived subjectively. The development of a BBN making use of subjective probability distributions requires an appropriate elicitation technique. Advancing the body of knowledge in this area, two experiments were conducted to compare: direct ranking of the variables' perceived importance for discriminating between given hypotheses, likelihood ratios and conditional probabilities. The focus of the experiments was the extent to which the different elicitation techniques lead to equivalent or different judgements.

Building upon early work the experiments used the BBN and scenario based on the main events leading to the start of the Arab Israeli conflict of 1973. The first experiment generated 31 sets of results from military officers studying at the Defence College of Management and Technology. Subsequently, a second experiment was completed by 6 participants, all of whom had extensive knowledge of military intelligence. Five of these participants were currently serving officers.

Comparisons were firstly made between the direct rankings and the normative BBN rankings (based upon mutual information between each source and the hypothesis node). Surprisingly little agreement was found between the direct and normative rankings. In total only two participants had a direct match between all four direct and normative rankings. Investigations into the correlation between the two sets of results found a weaker association than might have been expected. The use of direct ranking and weightings is proposed as a quick method for the development of a BBN. However, this research indicates such an approach may lead to substantial, yet unintentional discrepancies within the network.

The overall trend in the results of likelihood values was for an increase in the probability of hostile intent as events were observed. However, there were a small group of participants (all of whom did not have experience of military intelligence) for whom the accumulation of evidence actually led to a decrease in the probability of hostile intent. For these participants, a comparison of the likelihood and direct probabilities highlighted a major discrepancy. For each result in which the likelihood values indicated a decrease in hostile intent as events were observed, the corresponding direct probabilities showed an increase in the probability of hostile intent. Therefore, it is concluded that this group of participants provided the inverse likelihood ratio to that requested within the experiment.

Consequently, it is concluded that the use of likelihood ratios for the derivation of probability distributions for use in BBNs should only be used by those who have experience of working with intelligence data, or with ratios as part of their work. For those not familiar with the derivation of such ratios, the elicitation of likelihood ratios caused uncertainty and created the possibility of the inverse ratio to that requested being stated. If individuals who are not confident with the use of ratios are required to state likelihood ratios then this should be conducted with the elicitation of additional information to ensure the required ratios have been stated. For example, the procedure could include leading questions to ascertain how the participants anticipate the probabilities will move. A comparison of this and the given values would reveal any immediate discrepancies.

The final section of the analysis centred on the elicited conditional probability distributions. Each participant's stated distributions were entered directly into their own BBN. Once populated, the events within the timeline were sequentially entered into the BBN as evidence. The general trend was for an increasing probability of intent as more events were observed. Interestingly, the observation of the second event (the partial mobilisation of reserves following a limited air defence capability) did not substantially alter the probability of hostile intent. Only when, in addition to these events, a concentration of forces was observed did the probability of hostile intent noticeably increase.

Analysis of the results showed most participants underestimated the probability of hostile intent with their direct judgements when only one or two events had been observed. However, as additional information was received, participants seemingly became unable to correctly manipulate the interdependencies between the reported events. This resulted in the participants over estimating the probability of hostile intent with their direct judgements relative to the normative values. Although just failing to show significance at the 5% significance level, there is some suggestion from the data that the range of direct judgements might be greater than the range of Bayesian normative values and this deserves further investigation. It is the context in which the events are observed which holds the key to their interpretation.

In summary the results show that participants indirectly expressed varying beliefs about the importance of information depending upon the elicitation technique used. Little evidence was found of a high correlation between direct rankings of variables' importance and those obtained from the BBNs normative results. Participants who were not regularly exposed to intelligence information, nor regularly used ratios were not able to easily, nor effectively express probabilities through the use of likelihood ratios. Conditional probability distributions provided the least troublesome technique and showed the smallest deviations between the direct and normative values assessed. It is concluded that BBNs do provide a flexible, robust methodology for the development of a normative model upon which a decision support system for the quantitative analysis of intelligence reports may be built.

The development of a BBN provides a clear audit trail of the assumptions, variables and data used within the creation of an intelligence assessment. The networks are easy to construct, review and refine in readily available software such as NETICA. The finalised network is capable of showing the variation in the probability distribution over a series of hypotheses and support both diagnostic and predictive reasoning. However, the network can only show those hypotheses which have been identified. To maximise the utility of the BBN, careful consideration must be given to the use of an appropriate technique to support hypothesis generation. Within the field of intelligence analysis one of the most commonly used techniques is that of the analysis of competing hypotheses.

BBNs have defined processes for the incorporation of uncertainty. Furthermore, the processes for formulating a network, deriving the required distributions and finally entering new information requires explicit consideration and communication of uncertainty on both the information used in and inferences drawn from the analysis.

Within the field of intelligence analysis it is possible that a BBN will require input from numerous experts. As a result the final network may be composed of smaller networks (possibly network fragments), each developed by a specialist. This is supported by the ability of a BBN to incorporate a range of probability distributions. Each section of the network could be populated through a different elicitation technique. Those with knowledge of intelligence or ratios may prefer to use likelihood ratios. However, the use of conditional probability distributions was found to produce the least differences between normative and direct values.

One of the main limitations of this research is the number of participants who have experience of handling intelligence data. Undoubtedly an area for further research would be to conduct a large scale experiment into the comparison of elicitation techniques. Such an experiment could consider additional elicitation techniques to those investigated within this research and determine if the results support the findings reported here.

The results of the analysis have shown statistically significant deviations between the military officers directly elicited probabilistic judgements and the corresponding normative combinations of their complex judgements. This conclusion could be used to further guide the development of a support tool for intelligence analyses. Potentially, the results shown here could assist in the creation of an explanation facility within a support tool capable of assisting analysts surprised by the outcome of some analysis. The explanation facility may help clarify why the resulting trend has been seen given the accumulation of evidence and suggest which, if any, probability distributions the analyst may wish to review. To this end, had more time been available for this research it would have been interesting to further investigate the results from the 6 participants with experience of military intelligence. This would have been achieved by gaining feedback from the participants on the BBN developed with their conditional probability distributions and gain their views on if the network does indeed behave as they envisaged. The results could further show what support an analyst may need to review probability distributions used within a BBN.

An additional area for consideration would be the training requirements for analysts who would build and use the networks. The construction of a BBN within software such as NETICA is, to a great extent, intuitive. It should be made apparent to both analysts and the customers of intelligence reports that the figures produced by the networks are not absolute. The strength of a BBN is not in a single numerical value, but in its ability to show variations in trends and identify when observed events begin to suggest an outcome at odds with the perceived wisdom. How to build trust in such a system and ensure its correct use is a topic for further research.

REFERENCES

Adams, F. (2006) Expert Elicitation of Bayesian Analysis of Construction Contracts Risks: An Investigation. *Construction Management and Economics*, Vol. 24, Issue 1 January, pages 81 – 96.

Andriole, S.J. (1989) *Handbook of decision support systems.* Blue Ridge Summit, PA: Tab Books Inc.

Arkes, H. & Blumer, C. (1985) "The Psychology of Sunk Cost" *Organizational Behavior and Human Decision Process* 35, Pages 124-140

Arnbourg, S., Brynielsson, J., Artman, H., & Wallenius, K. (2000) Information awareness in command and control: Precision, Quality and Utility. *In Proceedings of the Third International Conference on Information Fusion, 2000, (FUSION 2000),* Vol. 2, Pages THB1/25 – THB1/32, Paris, 10th – 13th July.

Arnott, D. & Pervan, G. (2008) Eight key issues for the decision support systems discipline. *Decision Support Systems,* 44, Pages 627 – 657.

Basye, K., Dean, T., Kirman, J., & Lejter, M. (1992) A decision-theoretic approach to planning, perception, and control. *IEEE Expert*, *7*(4):58-65, August.

Bazerman, M.H. (1986) *Judgement in Managerial Decision Making*. New York, John Wiley and Sons Inc, 1986.

Bell, D.E., Raiffa, H., & Tversky, A. (Eds) (1988) *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, Cambridge University Press, Cambridge, UK, 1988.

Belton, V. (1990) *Multi Criteria Decision Analysis (MCDA) – Practically the Only Way to Choose.* Operational Research Society, Birmingham. L C. Hendry and R. W. Eglese.

Berdún, L., Pace, J A.D., Amandi, A., & Campo, M. (2008) Assisting novice software designers by expert designer agent. *Expert Systems with Applications Vol. 36, No. 2, 2008.*

Berk, K.N., & Casey, P. (1999) *Data Analysis with Microsoft Excel.* Boston: PWS Publishing Co.

Betts, R.K. (1978) Analysis, war and decision: why intelligence failures are inevitable, *World Politics,* 31, 1, Pages 61-89.

Bhargava H.J., Power D.J, & Sun D. (2007) Progress in web-based decision support systems. *Decision Support Systems, Vol. 43, Issue 4, August 2007, Pages 1083-1095.*

Bickerton, I.J. & Klausner, C.L. (1991). *A concise history of the Arab Israeli conflict.* Englewood Cliffs, NJ: Prentice Hall.

Bradley, J.H., Paul, R., & Seeman, E. (2006) Analyzing the structure of expert knowledge. *Information and Management* 43, Issue 1, January Pages 77-91.

Breivik, E., & Supphellen, M. (2003) Elicitation of product attributes in an evaluation context: A comparison of three elicitation techniques. *Journal of Economic Psychology* 24 Pages 77-98.

British Air Power Doctrine AP3000. Third Edition Ministry of Defence. Available at http://www.raf.mod.uk/downloads/ap3000.cfm. Last accessed 6th January 2008.

Brynielsson, J. (2007) Using AI and games for decision support in command and control. *Decision Support Systems* 43, Pages 1454 – 1463.

Buchanan, BG and Shortliffe, EH (1984) Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley, Reading, MA

Buckwalter, D.T. (2002) *Case Studies in Policy Making and Implementation*. 3Williams (Ed) 6th Edition Naval War College 2002. Available at http://www.au.af.mil/au/awc/awcgate/navy/pmi/1973.pdf. Last accessed 6th January 2008.

Burns, K. (2006) Bayesian inference in disputed authorship: A case study f cognitive errors and a new systems for decision support. *Information Sciences*, 176, Pages 1570 – 1589.

Butts, C.A. (2003) Network inference error and information (in)accuracy: a Bayesian approach. *Social Networks*, Vol 25, Issue 2, May Pages 103 – 140.

Carter, G.M., Murray, M.P., Walker, R.G., & Walker, W.E. (1992) *Building Organizational Decision Support Systems.* San Diego: Academic Press Inc.

Chizek, J.G; Elsea, J., Best, Jr, R.A., & Bolkcom, C. (2003) Military Transformation. *Current Issues in Intelligence , Surveillance and Reconnaissance*. Novinka Books, New York, 2003.

Clausewitz (1976) *On War*. As translated and edited by Paret, P. and Howard, M. (Eds) (1976). Princeton: Princeton University Press.

Cohen, M.A., Thompson, B.B., Adelman, L., Bresnick, T.A., Tolcott, M.A., & Freeman, J.T. (1995) *Rapid Capturing of Battlefield Mental Models*, (Technical Report 95-3 United States Army Research Institute) Arlington, Virginia: Cognitive Technologies Inc.

Cohene, T., & Easterbrook, S. (2005) Contextual Risk Analysis for Interview Design, in *Proceedings on the 13th IEEE International Conference of Requirements Engineering*, 29th August to 2nd September, Pages 95 to 104.

Coolen, F.P.A., Mertens, P.R., & Newby, M.J. (1992) A Bayes-competing risk model for the use of expert judgment in reliability estimation. *Reliability Engineering & System Safety*, Vol. 35, Issue 1, Pages 23-30.

Copp, T. (Ed) (2000) *Montgomery's Scientists: Operational Research in northwest Europe – the work of No. 2 Operational Research Section with 21 Army Group, June 1944 to July 1945*. Wilfred Laurier University, Waterloo.

Cowie, J., & Burstein, F. (2007) Quality of data model for supporting mobile decision making. *Decision Support Systems*, 43, Pages 1675 – 1683.

Crossman, J.A.; Hong, G., Murphey, Y.L.; Cardillo, J. (2003) Analyzing the structure of expert knowledge feature extraction and quasi-optimal feature selection. *IEEE Transactions on Vehicular Technology*. Vol. 52, Issue 4, July Pages: 1063 – 1075.

Czech, D.R., & Dizek, S.G. (1988) Artful Analysis – the key to successful warranties. In *Proceedings of the IEEE Engineering Management Conference 1998, 'Engineering Leadership in the 90's'*, Dayton, Ohio, USA, 24th – 26th October, Pages 51 – 58.

Das, B. (2000) Representing Uncertainties in Bayesian Belief Networks. *DSTO-TR-0918*. Available at: http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/4269/1/DSTO-TR-0918%20PR.pdf Last accessed 3rd May 2008.

Dawes, R.M. (1988) *Rational Choice in an Uncertain World*. Orlando, USA, Harcourt Brace College Publishers.

Dawid, A.P., Mortera, J., & Vicard, P. (2005) Object-orientated Bayesian networks for complex forensic DNA profiling problems. *In Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, January 3--6, Key West, Florida, (edited by Christopher M. Bishop and Brendan J. Frey.)

De Finetti, B. (1937) Foresight: its logical laws, its subjective sources, *Annales de l'Institut Henri Poincaré*, as translated in 1964, in Kyburg and Smokler, editors, *Studies in Subjective Probability*. New York: Wiley.

Defence Intelligence, Ministry of Defence. Available at:
http://www.mod.uk/nr/exeres/e12ee169-9451-44d0-a847-adc664a562c8.htm
Last accessed 9th January 2008.

Djamasbi, S. (2007) Does positive affect influence the effective usage of a decision support system? *Decision Support Systems*, 43, Pages 1707 – 1717.

DeWispelare, A.R., Herren, L.T., & Clemen, R.T. (1995) The use of probability elicitation in the high-level nuclear waste regulation program. *International Journal of Forecasting*, Vol. 11, Issue 1, Pages 5-24.

Edwards, W. (1971) Social Utilities. *Engineering Economist*, Summer Symposium Series, 6 Pages 119-129.

Edwards, W., & Barron, F.H. (1994) SMARTs and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement, Organizational Behaviour and *Human Decision Processes, 60,* Pages 306-325.

Edwards, W., Phillips, L.D., Hays, W.L., & Goodman, B.C. (1968) Probabilistic Information Processing Systems: Design and Evaluation. *IEEE Transactions on Systems, Science and Cybernetics*, Vol. SSC-4, No. 3, September, Pages 248 – 265.

Eom, S.B. (1995) Decision Support Systems Research: Reference Disciplines and a Cumulative Tradition. *Omega International Journal of Management Science,* Vol 23, No. 5, Pages 571 – 523.

Eom, S.B., Min, H. (1999) The contributions of Multi-Criteria Decision Making to the Development of Decision Support Subsystems Specialities: An Empirical Investigation. *Journal of Multi-Criteria Decision Analysis*, Volume 8, Pages 239-255.

Falzon, L. (2006) Using Bayesian network analysis to support centre of gravity analysis in military planning. *European Journal of Operational Research,* 170 Pages 629 – 643.

Falzon, L., Zhang, L., & Davies, M. (2000) A policy analysis approach to operational level courses of action analysis. *15th International Command and Control Technology Symposium*, Canberra Australia. Available at: http://www.dodccrp.org/events/5th_ICCRTS/papers/Track2/018.pdf Last accessed 6th January 2008.

Fan, T-F., Liu, D-R., & Tzeng, G-H. (2007) Rough set-based logics for multi criteria decision analysis. *European Journal of Operational Research*, 182, Pages 340-355.

Fenton, N.E., & Neil, M. (2001) Making Decisions: Using Bayesian Nets and MCDA. *Knowledge-Based Systems* Vol. 14, Pages 307-325.

Fenton, N.E., Neil, M., & Gallan, J. (2007) Using Ranked nodes to model qualitative judgements in Bayesian Networks in *IEEE Transactions on Knowledge and Data Engineering* Vol. 19, No. 10, Pages 1420-1432.

Finlay, P. (1989) *Introducing Decision Support Systems*. Oxford: NCC Blackwell.

Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L. & Keeney, R. L. (1981). *Acceptable risk*. New York: Cambridge University Press.

Fisk, C.E. (1972) The Sino-Soviet Border Dispute: A Comparison of the Conventional and Bayesian Methods for Intelligence Warning, *Stud. Intel*. V16:2-53-62 (Spring)

Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., Kolaczkowski, A., & Wreathall, J. (2004) Expert elicitation approach for performing ATHEANA quantification. *Reliability Engineering and Systems Safety* 83 Pages 207-220.

French, S. (1989). *Readings in Decision Analysis*. London, Chapman and Hall.

Frada, B., & George, W. (2007) Decision Support in an uncertain and complex world (Editorial). *Decision Support Systems,* 43 Pages 1647 – 1649.

Galway, L.A. (2007) *Subjective probability distribution elicitation in cost risk analysis*. Rand Technical Report TR 410. Accessed at: http://www.rand.org/pubs/technical_reports/2007/RAND_TR410.pdf last accessed 6th January 2008.

Garthwaite, P.H., & O'Hagan, A. (2000) Quantiyfying expert opinion in the UK water industry: an experimental study. *The Statistician* 49, Part 4, Pages 455-477.

Gazit, S. (1989) *Intelligence Estimates and the decision-maker*. In Handel M I (Ed) (1989) *Leaders and Intelligence.* Frank Cass and Company Limited, London, Pages 261-287.

Gigerenzer, G., Todd, P.M., & the ABC [Center for Adaptive Behavior and Cognition] Research Group (1999). *Simple Heuristics that make us smart*. New York: Oxford University Press.

Gilhooly, K.L. (1988) *Thinking: Directed, Undirected and Creative*, 2nd Edition, London: Academic Press.

Gill, P., & Phythian, M. (2006) *Intelligence in and insecure world*. Ploity Press, Cambridge.

Golub, A L. (1997) *Decision Analysis An Integrated Approach*. John Wiley and Sons Inc.

Gross, R. (2001) *Psychology*. *The Science of Mind and Behaviour*. London, Hodder and Stoughton Educational.

Goodwin, P., & Wright, G. (1999) *Decision Analysis for Management Judgement*: Chichester: John Wiley and Sons.

Greenberg, H.J. (2007) Representing Uncertainty in Decision Support. *OR/MS Today*, June, Pages 14-16. Available at: http://www-math.cudenver.edu/~hgreenbe/papers/Greenberg07ORMSuncertainty.pdf last accessed 6th January 2008.

Gudgin, P. (1989) *Military Intelligence The British Story*. Arms and Armour Press, London.

Guitouni, A., & Martel, J.M. (1998) Tentative Guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research* Vol 109 Pages 501-521.

Guo, M., Yang, J-B., Chin, K-S., & Wang, H. (2007) Evidential reasoning based preferences programming for multiple attribute decision analysis under uncertainty. *European Journal of Operational Research*, 182, Pages 1294-1312.

Handel M I., (1989) *War, Strategy and Intelligence*. Frank Cass and Company Limited, London.

Hall, H. (1998). Networked information: dealing with overload. *In Proceedings of Information Scotland 1997, Strathclyde Business School,* November 4th Pages. 37-44. Paisley: LACIG Scotland. Available at http://www.dcs.napier.ac.uk/~hazelh/esis/hh_paper10.html. Last accessed 6th January 2008.

Hall, D.J. & Davis, R.A. (2007) Engaging in multiple perspectives: A value-based decision-making model. *Decision Support Systems*, 43, Pages 1588-1604.

Hamilton, P.W., Anderson, N., Bartels, P.H., & Thompson, D. (1994) Expert system using Bayesian Belief Networks in the diagnosis of fine needle aspiration biopsy specimens of the breast. *J Clin Pathol* April; 47 (4): Pages 329-326

Handel M I (Ed) (1989) *Leaders and Intelligence*. Frank Cass and Company Limited, London,.

Hart, A. (1985) Knowledge elicitation: issues and methods. (1985) *Computer Aided Design,* Vol. 17, No. 9, November Pages 455-462.

Hart, B.H.L. (1953) *Rommel Papers*. Collin's: St James's Palace London.

Hastie, R. & Dawes, R.M. (2001). *Rational Choice in an Uncertain World*. Thousand Oaks, CA: Sage.

Heller, R. (1989) *The Decision Makers. The People and Methods of Great Achievement*. Heller Arts Ltd.

Heuer, Jr, R.J. (1999) *Psychology of Intelligence Analysts*, Centre for the Study of Intelligence, available from www.odci.gov/csi/pubs/html. Last accessed 6th January 2008.

Heuer, Jr, R.J. (2005) Limits of Intelligence Analysis. *Orbis*, *Winter*, Pages 75-93.

Holt J (1988) Assessing the need for decision support systems. *European Journal of Operational Research*, 37 Pages 73-82.

Hoffman, R.R., Shadbolt, N.R., Burton. A.M., & Klein, G. (1995) Eliciting Knowledge from Experts: A Methodological Analysis. *Organizational Behaviour and Human Decision Processes*, Vol. 62, No. 2, May, Pages. 129-158.

Hogarth, R.M. (1980). *Judgement and Choice. The Psychology of Decision*. New York, John Wiley and Sons.

Hollnagel, E. (2007) *Decision about "What" and Decisions about "How" in Decision Making in Complex Environments* Cook M, Noyes J and Masakowski Y (Eds) Aldershot: Ashgate Publishing Limited.

Holsapple, C., & Joshi, K.D. (2003) A knowledge management ontology, in: Holsapple (Ed.), Handbook on Knowledge Management, *Knowledge Matters*, Vol 1, Springer, Pages 90 – 124, Chapter 6.

Holt J.R., & Hazen C.M. (1988). A procedure to identify applications of expert systems in air forcebase facilities operations and maintenance. In the *Proceedings of Engineering Management Conference 1998. Engineering Leadership in the 1990's. IEE Conference*. Conference held at Dayton, Ohio, USA, 24-26th October, 1988.

Horvitz, E., & Barry, M. (1995) Display of Information for time-critical decision-making in *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, Montreal, August, Pages 296-305. Morgan Kaufmann: San Francisco

Hosack, B. (2007) The effect of system feedback and decision context on value based decision-making behavior. *Decision Support Systems*, 43, Pages 1605 – 1614.

Hughes-Wilson, J., Colonel. (2004) *Military Intelligence Blunders and Cover-Ups*. Carroll & Graff.
Iraq's Weapon's of Mass Destruction – The assessment of the British Government. 24th September 2002. Available at http://www.number10.gov.uk/output/page271.asp

Jarupathirun, S., & Zahedi, F. (2007). Dialetic decision support systems: Systems design and empirical evaluation. *Decision Support Systems* 43, Pages 1553 – 1570.

Jensen, F.V. (2001) *Bayesian Networks and Decision Graphs.* New York: Springer

Jesse, L.A., & Kalita, J.K. (1997) Situation Assessment and Prediction in Intellligence Domains. *Knowledge-Based Systems*, 10, Pages 87-102.

Jurgelenaite, R., & Lucas, P.J.F. (2005) Exploiting causal independence in large Bayesian networks. *Knowledge Based Systems* Vol 18 Pages 153-162.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*, 313-327.

Keegan, J. (2003) *Intelligence in War. Knowledge of the enemy from Napoleon to Al-Qaeda.* Hutchinson, London.

Kem, J. (2005) An Adaptive Methodology for Developing Enemy Courses of Action. *Military Intelligence Professionals Bulletin;* Jan – Mar; 31, 1; Pages 13 – 17. ProQuest Military.

Keren, G. (1991) Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica* 77, Pages 217-273.

Khan, D. (1986) *Clausewitz and Intelligence*. In Handel M I (Ed). Clausewitz and Modern Strategy, Frank Cass and Company Limited, Oxon, Pages 117-126.

Kim, J.H., & Pearl, J. (1983) A computational model for combined causal and diagnostic reasoning in inference systems. *In Proceedings IJCAI-83*, Pages 190-193, Karlsruhe, Germany.

Klein, G.A. (1989). *Recognition-primed decisions.* In W.B. Rouse (Ed.*) Advances in man-machine system research* Vol. 5, Pages.47-92. Greenwich, CT: JAI Press.

Klein, G.A. (1997) *The current status of the naturalistic decision making framework*. In Flin R, Salas E, Strub M, Martin L (Eds) *Decision Making Under Stress Emerging Themes and Applications.* Burlington, Ashgate Publishing Company.

Klein, G.A. (1999) *Sources of Power: How people make decisions*. MIT Press..

Klein, G.A., Orsanu, J., Calderwood, R., & Zsambok, C. E. (Eds) (1993) *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Klein, M.R., & Methlie, L.B. (1995) *Knowledge-based Decision Support Systems with Applications in Business, Second Edition*. Chichester: John Wiley and Sons Ltd.

Klopotek, M.A. (2005) Very large Bayesian multinets for text classification. *Future Generation Computer Systems* 21 Pages 1068-1082.

Koller, D., & Pfeffer, A. (1997) Object Oriented Bayesian Networks. *Proceedings of the Thirteenth Annual Conference on Uncertainty on Artificial Intelligence.* August 1st-3rd, Brown University, Providence, Rhode Island, USA, Pages 302 - 313. Morgn Kauffman Publishers Inc, San Francisco.

Koskinen, J.H., & Snijders, T.A.B. (2007) Bayesian inference for dynamic social network data. *Journal of Statistical Planning and Inference,* 137 Pages 3930 – 3938.

Kramer, R.M. (1998) Revisiting the Bay of Pigs and Vietnam Decisions 35 Years Later: How Well Has the Groupthink Hypothesis Stood the Test of Time? *Organizational Behaviour and Human Decision Processes,* Vol 73, Nos 2-3, February-March Pages 236-271.

Laskey, K. (1995) Sensitivity Analysis for probability assessments in Bayesian Networks. *IEEE Transactions on System, Man and Cybernetics,* Volume 25, Number 6, June, Pages 901-909.

Laskey, K.B. & Mahoney, S.M. (1997) Network Fragments: Representing Knowledge for Constructing Probabilistic Models. *Proceedings of the Thirteenth Annual Conference on Uncertainty on Artificial Intelligence.* August 1st-3rd Brown University, Providence, Rhode Island, USA, Pages 344 - 341. Morgan Kauffman Publishers Inc.

Laskey, K. B., & da Costa P.C.G. (2007) Probabilistic Ontology for Net-Centric Fusion. *In the Proceedings of the 10th International Conference of Information Fusion,* Quebec, Canada, 9-12th July Pages 1 – 8.

Lau, A., & Leong, T. (1999) PROBES: A framework for probability elicitation from experts. In *AMIA Annual Fall Symposium,* Pages 301-305, 1999. Available at: http://www.amia.org/pubs/symposia/D005714.PDF Last accessed 6th January 2008.

Lauritzen, S.L., & Spiegelhalter, D.J. (1988) Local computations with probabilities on graphical structures and their application to expert systems (with discussion). *Journal of the Royal Statistical Society,* Series B, 50(2), Pages 157 – 224.

Lee, D.H., Lee, K.C., Kang, S., & Baik, J. (2007) A priori ordering protocols to support consensus-building in multiple stakeholder contexts. *Information Sciences,* 177, Pages 3129 – 3150.

Liao, S-H. (2000) Case-based decision support system: Architecture for simulating military command and control. *European Journal of Operational Research,* 123 Pages 558-567.

Liao, S-H., Sun, B-L., & Wang, R-Y. (2003) A knowledge-based architecture for planning military intelligence, surveillance and reconnaissance. *Space Policy,* 19 Pages 191-202.

Lichstenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Coombs, B. (1978) Judged Frequency of Lethal Events, *Journal of Experimental Psychology: Human Learning and Memory,* 4, Pages551-578.

Lindley, D.V. (1985) *Making Decisions Second Edition*. Published John Wiley and Sons Limited, Great Britain.

Lindsey, P.H., & Norman, D.A. (1977) *An Introduction to Psychology, Second Edition.* Academic Press, Inc (London) Ltd.

Luciani, D., Cavuto, S., Antiga, L., Massimo, M., Monti, S., Pistolesi, M., & Bertolini, G. (2007) Bayes pulmonary embolism assisted diagnosis: a new expert system for clinical use. *Emerg Med J,* 24: Pages 157-164.

MacColskey, M. (1967) *The American Intelligence Community,* New York: Richards Rosen Press.

MacKenzie, W.J.M. (1975) *Choice and decision From Power, Violence, Decision,* Pergdrine Books 1975 Pages 176-9. Reprinted in McGrew A, Wilson M.J. (Eds) *Decision Making Approaches and Analysis,* Manchester University Press (1985) Pages 16-18.

MacKenzie, A., Pidd, M., Rooksby, J., Somerville, I., Warren, I., & Westcombe, M. (2006) *Wisdom*, decision support and paradigms of decision making. *European Journal of Operational Research*, 170, Pages 156 – 171.

March, J.G. (1994) *Bounded Rationality, Ambiguity, and the Engineering of Choice*. In *Decision making: Descriptive, normative, and prescriptive interactions* Edited by Bell D. E., Raiffa H, Tversky A, Cambridge University Press Pages 33-57.

March, J.G., & Simon, H.A. (1958), *Organizations*, New York: John Wiley & Sons

Marcusohn, L.M. (1995). The information explosion in organisations: a review of individuals' information overload. *Swedish Libraries Research*, 3/4, Pages 25-41.

Martinsons, M.G., & Davison, R.M. (2007) Strategic decision-making: Comparing American, Japanese and Chinese Management. *Decision Support Systems*, 43, Pages 284 – 300.

McLaughlin, J. & Paté-Cornell, M.E. (2005) A Bayesian Approach to Iraq's Nuclear Program Intelligence Analysis: A Hypothetical Illustration. *In Proceedings of International Conference on Intelligence Analysis.* 2nd-4th May, Mclean, VA. Available at: https://analysis.mitre.org/proceedings/Final_Papers_Files/85_Camera_Ready_Paper.pdf. Last accessed 6th January 2008.

Miller, G.A. (1956): The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, *Psychological Review*, Vol. 63 Pages, 81-97.

MI5 (2008) http://www.mi5.gov.uk/output/Page55.html

Moffat, J. (2000) Representing the command and control process in simulation models of conflict. *Journal of the Operational Research Society* April Vol. 51, No. 4 Pages 431 – 439.

Moffat, J., & Witty, S. (2002) Bayesian Decision Making and Military Command and Control. *Journal of the Operational Research Society*, July, Vol. 53, No. 7, Pages 708-718.

Mosleh, A., Bire, V.M., & Apostolakis, G. (1988) A critique of current practice for the use of expert opinions in probabilistic risk assessment. *Reliability Engineering and Systems Safety*, 20, Pages 63-85.

Moody, J.W., Will, R.P., & Blanton, J.E. (1996) Enhancing knowledge elicitation using the cognitive interview. *Expert Systems with Applications*, Vol. 10, No. 1, Pages. 127-133.

Moynihan, P., & Bowne, K. (1987) A research study of human decision-making in higher level military command and control. *European Journal of Operational Research*, 31 Pages 284 – 291.

Neil M., Fenton N. E. (2005) Improved Methods for building large-scale Bayesian Network *Presented at The Third Bayesian Modeling Applications Workshop, Uncertainty in Artificial Intelligence (UAI) 2005* Statement of interest available at: http://www.dcs.qmw.ac.uk/~norman/papers/neil_fenton_UAI_2005.pdf last accessed 6th January 2008.

Neil, M., Fenton, N., & Nielsen, L. (2000) Building Large-Scale Bayesian Networks. *The Knowledge Engineering Review*, Vol 15. Issue 3. September, Pages 257-284.

Nelson, J.A. (2004) Do Steady State PIRSs Work in Stability Operations? Answering the Commander's Intelligence and Decisionmaking Needs. *Military Intelligence Professional Bulletin*, Jul-Sep; 30, 3, Pages 37 – 42.

Noorderhaven, N. (1995) *Strategic Decision Making*. Addison Wesley, Printed Guildford, Great Britain.

Nutt, P.C. (1989) *Making Tough Decisions*. San Franscisco, Jossey-Bass Inc Publishers.

Odom, W.E. (2004) *Fixing Intelligence for a more secure America. 2nd Edition*. Yale University Press, New Haven.

O'Hagan, A. (1998) Eliciting expert beliefs in substantial practical applications. *The Statistician* 47, Part 1, Pages 21-35.

O'Hagan, A., & Oakley, J.E. (2004), Probability is perfect, but we can't elicit it perfectly, *Reliability Engineering and System Safety*, 85, Pages.239-248.

Oritz, N.R., Wheeler, T.A., Breeding, R.J., Hora, S., Meyer, M.A., & Keeney, R.L. (1991) Use of expert judgement in NUREG-1150. *Nuclear Engineering and Design*, 126, Pages. 313-331.

Panik, M.J. (2005) *Advanced Statistics from an Elementary Point of View*. Boston: Elesevier/National Academic Press Inc.

Paradice, D. (2007) Expanding the boundaries of decision support systems. *Decision Support Systems*, 43, Pages 1549 – 1552.

Parent, E., & Bernier, J. (2003) Encoding prior expert judgements to improve risk analysis of extreme hydrological events via POT modelling. *Journal of Hydrology*, Vol 283., Issues 1-4, December Pages 1 – 18.

Paté-Cornell, M.E. (2002) Fusion of Intelligence Information: A Bayesian Approach. *Risk Analysis* Vol. 22, No. 3. Pages 445-454, June.

Pearl, J (2000). *Causality: Models, Reasoning and Inference*. New York: Cambridge University Press.

Pearl, J. (1997) Bayesian Networks. UCLA Cognitive Systems Laboratory, Technical Report (R-246), Revision I, July 1997. In *MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA, 1999. Available at http://ftp.cs.ucla.edu/pub/stat_ser/R246.pdf Last accessed 6th January 2008.

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems.* San Mateo: Morgan Kaufmann.

Pearl, J. (1986) .Fusion, propagation and structuring in belief networks. *Artificial Intelligence,* Vol. 29, Pages 241 – 288.

Pearl, J. (1982) Reverend Bayes on inference engines: A distributed hierarchical approach. *Proceedings American Association of Artificial Intelligence National Conference on AI,* Pittsburgh, PA, 133--136.

Petkov, D., Petkova, O., Andrew, T., & Nepal, T. (2007) Mixing Multiple Criteria Decision Making with soft systems thinking techniques for decision support in complex situations. *Decision Support Systems,* 43 Pages 1615 – 1629.

Pew, R.W., & Mavor, A.S. (2000) *Modelling Human and Organisational Behaviour: Application to Military Simulations*. Washington: National Academy Press.

Pfautz, J., & Roth, E. (2006) Using cognitive engineering for system design and evaluation: A visualization aid for stability and support operations. *International Journal of Industrial Ergonomics*, 36, Pages 389 – 407.

Phillips-Wren, G., Mora, M., Forgionne, G.A., & Gupra, J.N.D. (2007) An integrative evaluation framework for intelligent decision support systems. *Article in press, European Journal of Operational Research,* doi:10.1016/j.ejor.2007.11.001.

Plous, S. (1993). *The psychology of Judgement and Decision Making*. New York: McGraw-Hill, Inc.

Poch, M., Comas, J., Rodriguez-Roda, I., Sánchez-Marrè, M., & Cortés, U. (2004) Designing and building real environment decision support systems. *Environmental and Modelling Software*, 19, Pages 857-873.

Pope, S., Jøsang, A., & McAnally, D. (2006) Formal Methods of Countering Deception and Misperception in Intelligence Analysis. *In Proceedings of 11th ICCRTS Coalition Command and Control in the Networked Era Conference* 26 to 28 September, Cambridge, United Kingdom Available at http://sky.fit.qut.edu.au/~josang/papers/PJM2006-ICCRTS.pdf Last accessed 6th January 2008.

Pöyhönen, M., & Hämäläinen, R.P. (2001) On the convergence of multiattribute weighting models. *European Journal of Operational Research*, 129, Pages.569-585.

Purkitt, H.E., & Dyson, J.W. (1987). *An Experimental Study of Cognitive Processes and Information in Political Problem Solving*. In Rohrmann B., Beach L. R., Vlek C., Watson S. R. (Eds) Advances in Decision Research. *Selected/Edited Proceedings of the 11th Conference on Subjective Probability, Utility and Decision Making.* Cambridge, UK. 23 – 27 August 1987 pages 201 – 243.

Radford, K.J. (1975) *Managerial Decision Making*, U.S.A: Prentice Hall.

Reneau, J.H., & Blanthorne, C. (2001) Effects of information sequence and irrelevant distracter information when using a computer based decision aid. *Decision Sciences*, Winter, Vol 32, Issue 1, Pages 145-163.

Renooij, S. & Witteman, C.L.M. (1999) Talking probabilities: Communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning, 22, Pages 69-194.*

Report into the Terrorist Attacks in London on 7 July 2005. The Stationary Office (May 2006) Also available at http://www.official-documents.gov.uk/document/cm67/6785/6785.pdf

Rhodes, P.C. (1993) *Decision Support Systems.* Henley on Thames: Alfred Waller Limited Publishers.

Roberts, R., & Goodwin, P. (2002) Weight Approximation in Multi-attribute Decision Models. *Journal of Multi-Criteria Decision Analysis.* Vol. 11, Pages 291-303.

Rosen, J., & Smith, W. (1996) Influence Net Modelling with Causal Strengths: an Evolutionary Approach. *In Proceedings of Command and Control Research and Technology Symposium*, Pages 699--708. Naval Post Graduate School, Monterey, CA, USA.

Rosen, J.A., Smith, W.L., Smith, E.S., & Maldony, Jr, M.A. (1998) Planning with Influence Net Modelling: An Architecture for Distributed, Real-Time Collaboration. *66th Military Operations Research Society Symposium, 23rd – 25th June*, Naval Postgraduate School in Monterey, California.

Saaty, T.L. (1990) *The Analytical Hierarchy Process*, Pittsburgh: RWS Publications.

Sage, A.P. (1991) *Decision Support Systems Engineering.* New York: John Wiley and Sons Inc.

Sanders, N.R., & Ritzman, L.P. (1992) The need for contextual and technical knowledge in judgemental forecasting. *Journal of Behavioural decision making,* Vol 5, Pages 39-52.

Savage, L.J. (1954). *The Foundations of Statistics*. New York, NY: Wiley

Savage, L.J. (1971) Elicitation of Personal Probabilities and Expectations. *Journal of American Statistical Association.* December, Vol. 66, No. 336, Pages 783-801.

Savšek, T., Vezjak, M., & Pavesic, N. (2007) Fuzzy tress in decision support systems. *European Journal of Operational Research*, 174, Pages 293 – 310.

Schmitt, J.F. (1997) What is an intelligence failure? A case study of Korea 1950. *Marine Corps Gazette*, Oct; 81, 10 Pages 60 – 65.

Schum, D.A., (2005) *Naratives in Intelligence Analysis: Necessary but often dangerous.* Available at: http://www.evidencescience.org/pubs/pubs_detail.asp?pubID=83 Last accessed 3rd May 2008.

Schweitzer, N. (1976) Bayesian Analysis for Intelligence: Some Focus on the Middle East. *Stud. Intel.* V20: 2 Pages 31-44 (Summer 1976) available at: https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/pdf/v20i2a03p.pdf Last accessed on 11th October 2007.

Scott, S.L. (2004) A Bayesian paradigm for designing intrusion systems. *Computational Statistics and Data Analysis.* Vol 45, Issue 1, 28th February, Pages 69 – 83.

Shadbolt, N. & Burton, M. (1995) *Knowledge Elicitation: A Systematic Approach.* In *Evaluation of Human Work: A practical ergonomics methodology, 2nd and Revised Edition* Wilson J. R., Corlett E. (Eds) Pages 406 – 440. London: Taylor and Francis.

Shapira, Z (Ed) (2002). *Organizational Decision-Making.* Cambridge University Press.

Shankaranarayanan, G., & Cai, Y. (2006) Supporting data quality management in decision-making. *Decision Support Systems*, 42 Pages 302 – 317.

Shannon, C.E., Weaver, W. (1949). *The Mathematical Theory of Communications*, The University of Illinois Press, Urbana, Illinois. First published in 1949, republished in paperback in 1969.

Silver M. S. (1991) *Systems that support decision-makers.* Chichester: John Wiley and Sons Inc.

Simon, H.A. (1957) A. A Behavioral Model of Rational Choice, in *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley.

Simon H.A. (1988) *Rationality as process and as product of thought. In Decision making: Descriptive, normative, and prescriptive interactions* Edited by Bell, D.E., Raiffa, H., & Tversky, A. Cambridge University Press, Pages 58-77.

Sinuany-Stern, Z., Mehrez, A., & Hadad, Y. (2000) An AHP/DEA methodology for ranking decision making units. *Intel Trans in Op. Res* Vol 7 Pages 109-124.

Skillicorn, D.B. & Vats, N. (2007) Novel information discovery for intelligence and counterterrorism. *Decision Support Systems,* 43, Pages 1375 – 1382.

Sojda, R.S. (2007) Empirical evaluation of decision support systems: Needs, definitions, potential methods, and an example pertaining to waterfowl management. *Environmental Modelling and Software,* 22 Pages 269 – 277.

Spetzler, C.S., & Stäel von Holstein, C.A. (1975) Probability Encoding, *Decision Analysis, Management Science,* 22, Pages 340-352.

Sprague, R.H. (Jr)., & Watson, H.J. (1986) *Decision Support Systems: Putting Theory into Practice.* Englewood Cliffs, New Jersey: Prentice-Hal Inc.

Sprague, R.H. (Jr)., & Carlson, E.D. (1982) *Building Effective Decision Support Systems:* Prentice-Hall. Englewood Cliffs, New Jersey.

Sprague, R.H. (Jr)., *A framework for research on decision support systems*. In Fick G., and Sprague, R.H., (Ed) (1980) Decision Support Systems: Issues and Challenges, London: Pergamon Press.

Srdjevic, B. (2007) Linking analytical hierarchy process and social choice methods to support group decision-making in water management. *Decision Support Systems* 42 Pages 2261-2273.

Stewart, T.R., Roebber, P.J., & Bosan, L.F. (1997). The importance of the task in analyzing expert judgement. *Organizational Behaviour and Human Decision Process*, Vol 69, No. 3, March Pages 205-219.

Stech, F.J. & Elsaesser, C. (2007) Midway Revisited: Detecting Deception by Analysis of Competing Hypotheses. *Military Operations Research*, V12, N1, Pages 35-50.

Stenfors, S., Tanner, L., Syrjänen, M., Seppälä, T., & Hasapalinna, I. (2007) Executive views concerning decision support tools. *European Journal of Operational Research*, 181, (2007), pp 929 – 938.

Suzic, R. (2003) Generic Representation of Military Organisation and Military Behaviour: UML and Bayesian Networks. *In Proceedings of the NATO RTO Symposium on C3I and M&S Interoperability*, Analya, Turkey, October. Available at http://www.nada.kth.se/~rsu/SuzicGenericRep.pdf Last accessed 6th January 2008.

Tatarka, C.J. (2002) Overcoming Biases in Military Problem Analysis and Decision Making. *Military Intelligence Professional Bulletin, Vol.28, Issue 1, January – March.* Available at: http://findarticles.com/p/articles/mi_m0IBS/is_1_28/ai_82351480 Last accessed 3rd May 2008.

The Butler Review: Review of Intelligence of Weapons of Mass Destruction.(2004) The Stationary Office (14 July 2004). Also available at http://www.butlerreview.org.uk/ last accessed 6th January 2008.

The Final Report of the National Commission on Terrorist Attacks Upon the United States. W W Norton &Co Ltd, 31st July 2004. Also available at: http://www.9-11commission.gov/ Last accessed 6th January 2008.

The Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction. (2005) US Independent Agencies and Commissions (March 31, 2005). Also available at http://www.wmd.gov/report/ last accessed 6th January 2008.

Thierauf, R.J. (1982) *Decision Support Systems for Effective Planning and Control: A Case Study Approach*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

The Economist (2000) In Praise of Bayes (philosophy of mathematician and statistician Thomas Bayes). *The Economist 356,* No 81900 (30th September): Pages 83-84.

Toft, N. (1997) Model based decision support in sow herds with emphasis on linking decisions with different time horizons. *Proceedings of the Dutch/Danish Symposium on Animal Health and Management Economics,* Copenhagen, January 23 – 24, Dina Notat No 56, Pages 55-60.

Treverton, G.F. (2003) *Reshaping National Intelligence for an age of Information.* Cambridge University Press, New York, 2003.

Tseng, C-C., & Gmytrasiewicz, P.J. (2006) Real-time decision support and information gathering system for financial domain. *Physica A.* 363 Pages 417-436.

Turbain, E., Aronson, J.E., Liang, T-P., & Sharda, R. (2007) *Decision Support and Business Intelligence Systems.* New Jersey: Pearson Prentice-Hall.

Tversky, A., & Kahneman, D. (1974) Judgement under uncertainty: heuristics and biases. *Science*, Vol. 185, Pages 1124-31.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, Pages 435-458. Repeated in Wright, G (Ed) (1985) *Behavioural Decision Making*, New York: Plenum Press, Pages 25-42.

Tzu Sun. (1985) *The Art of War*. Foreward by James Clavell. Published Hodder and Stoughton, Great Britain.

Valtorta, M., Byrnes, J., & Huhns, M. (2007) Logical and Probabilistic Reasoning to Support Information Analysis in Uncertain Domains. *In the Proceedings of the Third Workshop on Combining Probability and Logic (Progic-07)*, Canterbury, England, September 5-7th.

Valtorta, M., Dang, J., Goradia, H., Huang, J., & Huhns, M. (2005) Extending Heuer's Analysis of Competing Hypotheses Method to Support Complex Decision Analysis. *In Proceedings of the 2005 International Conference on Intelligence Analysis (IA-05)*, Washington, D.C., May 2nd-4th, Available at https://analysis.mitre.org/proceedings/Final_Papers_Files/192_Camera_Ready_Paper.pdf Last accessed 6th January 2008.

van Lenthe, J. (1993) ELI: Interactive Elicitation Technique for subjective probability distributions. *Organizational Behaviour and Human Decision Processes,* 55, Pages 379-416.

Van der Gaag, L.C., Renooij. S., Witteman, C.L.M., Aleman, B.M.P., & Taal, B.G. (2002) Probabilities for a probabilistic network: a case study in oesophageal cancer. *Artifical Intelligence in Medicine,* 25, Pages 123-148.

Van Tol, A.A., & AbouRizk, S.M. (2006) Simulation modeling decision support through belief networks. *Simulation Modelling Practice and Theory,* 14, Pages 614 – 640.

Villejoubert, G., & Mandel, D.R. (2002) The inverse fallacy: An account of deviations from Bayes' theorem and the additivity principle. *Memory and Cognition,* 30 (2) Pages 171 – 178.

Von Neumann, J., & Morgenstern, O. (1944) *Theory of Games and Economic Behaviour*. Princeton: Princeton University Press.

Vraneš, S., Lucin, M., Stanojevic, M., Stenvanovic, V., & Subašic P. (1992) Blackboard metaphor in tactical decision making. *European Journal of Operational Research* 61 Pages 86-97.

Wagner, A.R. (1974) *Crisis Decision-Making Israel's Experience in 1967 and 1973.* Praeger Publishers, New York, USA.

Wagner, G.R. (1981) Decision Support Systems: The Real Substance,"
INTERFACES (11:2), April 1981

Wang, K-J., & Chien, C-F. (2003) Designing and Internet-based group decision support system. *Robotics and Computer Integrated Manufacturing* 19 Pages 65-77.

Wang, H., Dash, D., & Druzdzel, M.J. (2002) A Method for Evaluating Elicitation Schemes for Probabilities. *IEEE Transactions on Systems, Man and Cybernetics,* Vol 32, Issue 1, February, Pages 38-43.

Wang, X-G., Qian, W-H., Pagello, E., & Pei, R-Q. (1996) On the uncertainty and ignorance of statistical decision and evidence combination. *In proceedings of the IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems,* Washington, DC, USA, 8th – 11th December, Pages 166-173.

Wasserman, S., & Faust, K. (1994) *Social Network Analysis*. Cambridge University Press, England.

Weidl, G., Madsen, A.L., & Israelson, S. (2005) Applications of object-orientated Bayesian networks for condition monitoring, root cause analysis and decision support on operation on complex continuous processes. *Computers and Chemical Engineering* 29 Pages 1996-2009.

Weinstein, N.D. (1980) Addiitonal notes on unrealistic optimism about future life events. *Journal of personality and social psychology,* Vol 39, No 5, Pages 806-820.

Wiegmann, D.A. (2005) *Developing a Methodology for Eliciting Subjective Probability Estimates During Expert Evaluations of Safety Interventions: Applications using Bayesian Belief Networks.* Aviation Human Factors Division: Institute of Aviation, University of Illinois. Final Technical Report AHFD-05-13/NASA-05-4. Available at:
http://www.humanfactors.uiuc.edu/Reports&PapersPDFs/TechReport/05-13.pdf
Last accessed 6th January 2008.

Williams, D.J. (2007) *Risk and Decision Making in Decision Making in Complex Environments* Cook M, Noyes J and Masakowski Y (Eds) Aldershot: Ashgate Publishing Limited.

Williams, M L., Dennis, A.R., Stam, A., & Aronson, J.E. (2007) The impact of decision support systems use and information load on errors and decision quality. *European Journal of Operational Research*, 176, Pages 468 – 481.

Worm, A., Jenvald, J., & Morin, M. (1998) Mission efficiency analysis: evaluting and improving tactical mission performance in high-risk, time-critical operations. *Safety Science*, 30, Pages 79 – 98.

Wright, G. (1985) *Behavioral Decision Making*. New York: Plenum Press

Wright, G., & Ayton, P. (1987) Eliciting and Modelling Expert Knoweldge. *Decision Support Systems*, 3, Pages 13-26.

Xu, J.J., Chen, H. (2004) Fighting organised crimes: using shortest path algorithms to identify associations in criminal networks. *Decision Support Systems* 38, Pages 473 – 487.

Yuill, J., Wu, F., Settle, J., Gong, F., Forno, R., Huang, M., & Asbery, J. (2000) Intrusion-detection for incident-response, using a military battlefield-intelligence process. *Computer Networks* 34 Ppages 671 697.

Zack, M.H. (2007) The role of decision support systems in an indeterminate world. *Decision Support Systems* 43 Pages 1664-1674.

Zlotnick, J. (1967) A theorem for prediction. *Stud. Intel*. V11:4-1-12 (Winter) Available at https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/pdf/v11i4a01p.pdf. Last accessed 6th January 2008.

Zlotnick, J. (1972) Bayes' Theorem for intelligence analysis. *Studies in Intelligence*, Vol. 16, No. 2 (Spring 1972).. Available at: https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/pdf/v16i2a03p.pdf. Last accessed 10th October 2007.

Zsambok, C.E., & Klein, G. (Eds) (1997) *Naturalistic decision making*. Mahwah NJ: Lawrence Erlbaum Associates.

BIBLIOGRAPHY

Alter, S. (1996) *Information Systems: A Management Perspective*, 2nd Edition. Menlo Park, Ca: The Benjamin/Cummings Publishing Company Inc.

Antunes, C.H., & Dias, L.C. (2007) Managing uncertainty in decision support models. Forward to the special issue. *Decision Support Systems*, 43, Pages 1451 – 1543.

Applegate, L.M., Bailey, G., Banks, C.G., Song, J., Takriti, S., Tong, B.Q.M. (2004) *Navigating Information overload in the extended enterprise.* IBM Consulting Services Architecture of an on demand business session, May 17-18th Available at http://www-935.ibm.com/services/us/imc/pdf/g510-3946-navigating-info-overload.pdf. Last accessed 6th January 2008.

Burstein, F., & Widmeyer, G.R. (2007) Decision Support n an uncertain and complex world. *Decison Support Systems*, 43, Pages 1647 – 1649.

De Finetti. B. (1974) *Theory of Probability, Vol 1*, New York: John Wiley and Sons.

Johnston, R. (2003) Developing a Taxonomy of Intelligence Analysis Variables. *CIA Studies in Intelligence VOL. 47, NO. 3, UNCLASSIFIED EDITION.* Available at: https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol47no3/article05.html Last accessed 6th January 2008.

Korver, M., Lucas, P.J.F. (1993) Converting a Rule Based Expert System into a Belief Network. *Medical Informatics*, Vol. 18, No. 3, Pages. 219 - 241.

Leal, A. (1982) *Evaluating the Effectiveness of Military Decision Support Systems: Theoretical Foundations, Expert System Design, and Experimental Plan.* U. S. Army Research Institute for the Behavioural and Social Sciences, Research Note 83-18, September 1982. Available at: http://www.stormingmedia.us/08/0803/A080331.html Last accessed 8th January 2008.

Leonhardt, K.A. Maj. (2007) *Why Israel was Surprised in October 1973*. Marine Corps University Command and Staff College Thesis 1990. Available at when accessed on 15th July 2007 at:
http://www.globalsecurity.org/intell/library/reports/1990/LKA.htm

Neave, H.R., & Worthington, P.L. (1988) *Distribution Free Tests*. Unwin Hyman Ltd, London.

Nunnink, J., & Pavlin, G. (2006) Towards Robust State Estimation with Bayesian Networks: A New Perspective on Belief Propagation. *In the Proceedings of 9th InternationalMIS Conference on Information Fusion*, Tokyo Pages 1 - 8.

Oliver, R.M. & Smith, J.Q. (Eds). (1990) *Influence Diagrams, Belief Nets and Decision Analysis*. Published John Wiley and Sons New York.

Pearson, E.S., & Hartley, H.O. (1976) *Biometrika Tables for Statisticians Volume II.* University Printing House, Cambridge.

Takarka, C.J. (2002) Overcoming Biases in Military Problem Analysis and Decision-Making. *Military Intelligence Professional Bulletin*, Jan-Mar 28, 1 Pages 8 – 10.

Waddington P. (1997) *Dying for information? A report on the effect of Information Overload in the UK and the world*. Presented at the CNI Conference: Beyond the Beginning: The Global Digital Library. QE II Conference Centre, London June 16-17th. Conference report available at:
http://www.cni.org/regconfs/1997/ukoln-content/repor~13.html

ANNEX A: Land based scenario.
An enemy force invaded a country that the UK is friendly with, and is of strategic importance to, the UK. During its advance, the enemy force gained as much land as feasibly possible for use in subsequent negotiations for territorial gain. However, the strategic reason for the hostile act is still not fully understood.

In response to the invasion, the UK deployed a battle group to repel the enemy force; the operation was not carried out as part of a coalition. A battle group comprising of 12 Mechanised Brigade[10], 19 Mechanised Brigade[11] and 16 Air Assault Brigade was deployed[12].

The intent of the battle group was to defeat all the remaining enemy outposts within the territorial borders of the friendly country. To achieve this 12 Mechanised Brigade and 19 Mechanised Brigade conducted an operation to destroy any remaining enemy outposts, to ensure the security of the area and the lines of communication. Earlier operations were of high intensity to ensure the successful defeat of the enemy force. It was anticipated that the destruction of remaining enemy outposts would involve close quarters fighting.

During the operation, the reconnaissance group reported two enemy outposts, believed to be of section strength, in the area of interest. However, there was some ambiguity as following further reports it became highly probably that one of the outposts sighted has been reinforced.

The participant's mission was to advance and defeat all enemy outposts, which were within territorial borders and over section strength, to secure the area of interest and the lines of communication. The mission had to be completed within 24 hours of its start. The mission was given high priority, as the Battle Group could not advance to the meeting point until the area was secured.
To complete the mission the following assets are available to you:

[10] Comprises armoured medium and light roled troops.
[11] On 1st January 2005 19 Mechanised Brigade became the 19 Light Brigade.
[12] An air manoeuvre brigade.

- Human intelligence from the Battle Group reconnaissance group
- 2 companies and 2 squadrons forming part of 12 Mechanised Brigade
- Support from 1 AS90 battery and a mortar platoon
- Engineering plan is to provide support to the advance. Once the advance is complete, priority will switch to mobile support to the secured route
- Electronic warfare capabilities are available.

Participant response sheet (conditional probabilities).

The weather on the day of the advance is rain in the local area giving poor visibility. Friendly forces have **LOW RECCE CAPABILITY.** It is also known that the enemy has an **EW** capability.

<u>For questions 3-8 the probabilities you give to each row should sum to 100</u>

E.g.

		What is the probability of	
When enemy ATK is:	**And friendly recce capability is**	**Friendly recce sighting of ATK**	**NO friendly recce sighting of ATK**
Present	High	50	50

For question 2, please enter the probability of each event occurring. The rows will not sum to 100.

E.g.

What is the probability at the en location that:		
ATK is present	**Armour is present**	**Mortars are present**
50	60	70

DEFINITIONS:

DAS CAPABILITY: Defensive aid suite technology available for use.

ECM CAPABILITY. ESM/ECM/EPM available for use.

REINFORCED EN LOCATION. This relates to an outpost which has received reinforcements from the enemy force and has more assets or personnel present than those anticipated.

SENSOR information received by any method other than the human eye

HIGH RECCE CAPABILITY all Recce capability originally available are still in service and appropriate for current climatic conditions.

LOW RECCE CAPABILITY Some assets originally available may be out of service or not appropriate for current climatic conditions.

1. **For this question please give your answer as a percentage.** In your judgement what is the probability of enemy outposts being reinforced before they are definitely located by the BG recce patrol?

2. **FOR THIS QUESTION EACH ROW DOES NOT SUM TO 100**. For each enemy outpost please enter the probability of each asset being present**.**

	What is the probability that:		
For an enemy outpost that is	**ATK is present**	**Armour is present**	**Mortars are present**
NOT Reinforced			
REINFORCED			

3. Each row should sum to 100.

		What is the probability of:	
When enemy ATK is:	**And friendly recce capability is**	**Friendly recce sighting of ATK**	**NO friendly recce sighting of ATK**
Present	High		
Absent	High		
Present	Low		
Absent	Low		

4. **Each row should sum to100**.

		What is the probability of:	
When enemy armour is:	**And friendly recce capability is**	**Friendly recce sighting of armour**	**NO friendly recce sighting of armour**
Present	High		
Absent	High		
Present	Low		
Absent	Low		

5. **Each row should sum to100**.

		What is the probability of:	
When enemy Mortars are:	**And friendly recce capability is**	**Friendly recce sighting of Mortars**	**NO friendly recce sighting of Mortars**
Present	High		
Absent	High		
Present	Low		
Absent	Low		

6. **Each row should sum to 100**.

		What is the probability of:	
When enemy EW capability is:	**And enemy ATK is**	**Friendly sensors show enemy ATK present**	**Friendly sensors DO NOT show enemy ATK present**
DAS	Present		
DAS	Absent		
ECM	Present		
ECM	Absent		

7. **Each row should sum to 100.**

		What is the probability of:	
When enemy EW capability is:	**And enemy armour is**	**Friendly sensors show enemy armour present**	**Friendly sensors DO NOT show enemy armour present**
DAS	Present		
DAS	Absent		
ECM	Present		
ECM	Absent		

8. **Each row should sum to 100.**

		What is the probability of:	
When enemy EW capability is:	**And enemy Mortars are**	**Friendly sensors show enemy Mortars present**	**Friendly sensors DO NOT show enemy Mortars present**
DAS	Present		
DAS	Absent		
ECM	Present		
ECM	Absent		

Participant response sheet (consistency and causal reasoning)

The weather on the day of the advance is rain in the local area giving poor visibility. Friendly forces have **LOW RECCE CAPABILITY.** It is also known that the enemy has an **EW** capability.

For questions 1,2,4,5,6,8,9,10-18, please give in YOUR judgement the single probability of the event occurring.

For questions 3 and 7 the sum of the probabilities you give to parts a) and b) should sum to 100.

DEFINITIONS

DAS CAPABILITY: Defensive aid suite technology available for use.
ECM CAPABILITY. ESM/ECM/EPM available for use.
REINFORCED ENEMY LOCATION. . This relates to an outpost which has received reinforcements from the enemy force and has more assets or personnel present than those anticipated.
SENSOR information received by any method other than the human eye
HIGH RECCE CAPABILITY all Recce capability originally available are still in service and appropriate for current climatic conditions.
LOW RECCE CAPABILITY Some assets originally available may be out of service or not appropriate for current climatic conditions.

1. Considering the scenario, in your judgement what is the probability that ATK is present at the enemy outposts?

2. Considering the scenario, in your judgement what is the probability that the enemy outposts have been reinforced?

3. **In this question, the response to parts a) and b) SUM to 100.**

If the **enemy** has ECM capabilities, and considering the scenario, what is the probability that the sensor report for the enemy outposts:

a. Indicates ATK is present

b. Indicates ATK is NOT present

4. Considering the scenario, in your judgement what is the probability that armour is present at the enemy outposts?

5. Considering the scenario, in your judgement what is the probability that Recce Gp ***reports*** armour present the enemy outposts?

6. Considering the scenario, in your judgement what is the probability that the sensors report armour present at the enemy outposts?

7. **In this question, the response to parts a) and b) SUM to 100.**

If the Recce Gp have low visibility conditions and the Recce Gp report does not mention armour, what is the probability that at the enemy locations:

a. Armour is present

b. Armour is not present

8. Considering the scenario report in your judgement what is the probability that Mortars are present at the enemy outposts?

9. Considering the scenario, in your judgement what is the probability that the Recce Gp reports Mortars being present at the enemy outposts?

10. Considering the scenario, in your judgement what is the probability that the sensors reports Mortars being present at the enemy outposts?

11. If the Recce Gp reported no artillery, and no Mortars, but did report the presence on armour, in your judgement what is the probability that the enemy outpost under observation has been reinforced?

12. If the Recce Gp reported no Mortars but did report the presence of both ATK AND armour, in your judgement what is the probability that the enemy outpost under observation has been reinforced?

13. If the Recce Gp reported the presence of artillery, armour and Mortars, in your judgement what is the probability that enemy outpost under observation has been reinforced?

14. If the Recce Gp reported ATK was not present at the enemy outpost but the sensor report DID indicate the presence of artillery, in your judgement, what is the probability of ATK being present at the enemy outpost?

15. If the Recce Gp reports ATK is present at an enemy outpost, but the sensor report DID NOT indicate the presence of artillery, in your judgement, what is the probability of ATK being present at the enemy outpost?

16. Both the Recce Gp and the sensors report the presence of ATK at the enemy outpost. In your judgement, what is the probability of ATK being present at enemy outpost?

17. Both the Recce Gp and the sensors DO NOT report the presence of ATK at the enemy outpost. In your judgement, what is the probability of ATK being present at enemy outpost?

18. The Recce Gp report is unavailable, but the sensor report indicates the presence of ATK at the enemy outpost. In your judgement, what is the probability of ATK being present at the enemy outpost?

ANNEX B: Response sheet for an experiment into subjective value elicitation: A comparison of direct rankings, likelihood values, and conditional probabilities for use in decision support systems.

Scenario background

This scenario is based on the events leading up to the start of hostilities of the Yom Kippur War in 1973. Within the scenario you will be asked to provide a series of answers, some of which will require you to consider the situation from the perspective of an Israeli cabinet member. When providing your answers throughout this experiment please bear in mind the historical relationship between the Arab coalition (Egypt and Syria) and Israel.

Prior to the commencement of hostilities the Arab coalition specifically created a series of ‘false alarms’ within the Israeli military through the concentration and movement of their forces. At this time, the Israeli government was preparing to fight a general election and so did not want to be seen mobilising reserves for no reason. The false alarms created by the Arab coalition de-sensitised the Israeli military to such events – seeing troop movements and concentrations became less unexpected. As such, it became harder for the Israeli military to accept and realise that the events occurring before them now actually represented a hostile attack and were not another planned exercise or manoeuvre.

Scenario hypotheses

The experiment will consider two collectively exhaustive and mutually exclusive hypotheses considered to represent the possible status of the Egyptian and Syrian coalition at various times from April to October 1973.

Hypothesis A = The Egyptian and Syrian coalition are capable of a hostile attack against Israel within the next few months. (Imminently hostile)

Hypothesis B = The Egyptian and Syrian coalition is NOT planning a hostile attack against Israel within the next few months. (Not imminently hostile)

Elicitation of rank ordering of nodes.

The aim of this first section of the experiment is to determine in your opinion the order of importance of the potential information sources available to the Egyptian-Syrian coalition:

- Mobilisation of reserves
- Concentration of forces
- Soviet citizens leaving Egypt
- Development of an air defence capability

1. In your opinion which of the 4 potential information source is the most important?

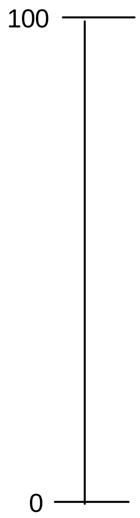
2. In your opinion which of the 4 potential information sources is the least important?

The chart below show a vertical line starting at zero and ending at 100. By zero write the potential information source you considered to be the least important. Conversely, by 100 write the potential information source you considered to be the most important. Now answer the following two questions:

3. In your opinion which of the 4 potential information sources is the second most important?

4. In your opinion which of the 4 potential information sources is the next to least important?

If zero is the least important and 100 the most important, place two marks on the chart representing the positions of the remaining 2 potential information sources and label them. You need to be comfortable that the gaps between the marks represent in your opinion the increase in importance of each potential information source from one to the next. Finally, place numbers by the marks between 0 and 100 to represent the importance of the two marks you have made on the (0,100) scale.



Elicitation of likelihood values

As previously stated, the two hypotheses being considered are:

Hypothesis A = The Egyptian and Syrian coalition are capable of a hostile attack against Israel within the next few months. (Imminently hostile)

Hypothesis B = The Egyptian and Syrian coalition is NOT planning a hostile attack against Israel within the next few months. (Not imminently hostile)

Based on the events in Table 1 below, you are asked in this second part of the experiment to provide a series of assessments, called likelihood ratios. These ratios relate the relative probabilities of an event being reported when hostilities are either imminent or not imminent and when ALL the preceding (i.e. lower-numbered) events have been reported.

For example, consider event two: Intelligence report received that Egypt has ordered a partial mobilisation of reserve forces. You are asked to provide the ratio of the probability of this occurring, knowing that there is an intelligence report that a SAM missile installment has been received, if hostilities were imminent compared to event two occurring if no hostilities were planned. A ratio of 1 implies you believe that the event is equally likely to occur whether the Arab coalition's intention is imminently hostile or not, given all previous events which have taken place. A ratio greater than 1 means that you believe the event is more likely to occur when hostilities are imminent. Conversely, a ratio less than 1 represents a belief that the event being considered is more likely to occur when hostilities are not imminent. The value of the ratio expressed how many times more likely the event being considered is given all previous events when hostilities are imminent compared to when they are not imminent.

Now look at Table 1. You are requested to give your answers as ratios based on the method explained above. Work through the table top to bottom, a row at time, From the second row on, assume that you know all of the previous (lower-numbered) events have already occurred.

Event #	Description:	Likelihood ratio
1	Intelligence report received that Syria and Egypt have developed a limited air defence capability through a 1st instalment of SAM missiles	
2	Intelligence report received that Egypt has ordered a partial mobilisation of reserve forces	
3	Intelligence report received that Syria and Egypt have developed an extensive air defence capability through a 2nd instalment of SAM missiles	
4	Special air reconnaissance over Siani reports a concentration of Egyptian forces.	
5	Intelligence report received that Soviet citizens in Egypt have started leaving.	

Table 1: Collation of likelihood values

Elicitation of conditional probabilities

In this third and final part of the experiment you will be asked to provide a series of probabilities as percentages. Please remember when deciding upon your values that there are no right or wrong answers, we are simply interested in your opinion of the situation. **You are asked not to use values of 0 or 100 since these imply things which are either impossible or certain and, given the perspective we ask you to adopt in this experiment, none of the events or hypotheses considered here fall into either of those categories.**

When completing the following questions please answer them as though you were an Israeli cabinet minister being presented with reports from the Israeli intelligence service. It is important to remember that the reports you be accurate, however they may also be mistaken in their interpretation of events or there may be non-hostile explanations for them such as military exercises

1. **Please note that the sum of your answers to this question must equal 100**. Considering the background to the conflict, but BEFORE you know any of the events shown in Table 1 on the last page occurred what would be your degree of belief, expressed as a percentage probability, that the Arab intent was:
    a. Imminently hostile (i.e. planning an attack within the next few months)

    b. Not imminently hostile (i.e. not planning an attack within the next few months)

2. **Please note that the sum of your answers in EACH ROW must equal 100.**
   Considering the scenario, please enter the following probabilities

When Arab intent is	The probability of mobilizing reserves		
	No mobilisation	Partial mobilisiation	Full mobilisation
Imminently hostile			
Not imminently hostile			

3. **Please note that the sum of your answers in EACH ROW must equal 100.**
   . Considering the scenario, please enter the following probabilities

When Arab intent is	The probability of forces being concentrated	
	Concentrated	Not concentrated
Imminently hostile		
Not imminently hostile		

4. **Please note that the sum of your answers in EACH ROW must equal 100.**
   Considering the scenario, please enter the following probabilities

When Arab intent is	The probability of Soviet citizens leaving	
	Leaving	Staying
Imminently hostile		
Not imminently hostile		

5. **Please note that the sum of your answers in EACH ROW must equal 100.** Considering the scenario, please enter the following probabilities

When Arab intent is	**The probability of an air defence capability being developed**		
	No such capability	**Limited AD capability**	**Extensive AD capability**
Imminently hostile			
Not imminently hostile			

For questions 6 to 9 you need to bear in mind the difference between an event actually occurring and an apparent event being reported by an imperfect source such as an intelligence agency. These questions require you to judge how likely you would consider an intelligence agency to correctly detect and report an event of the type indicated given that it has occurred.

You also need to judge the chance of them mistakenly reporting something that is not true, e.g. mistaking a full mobilisiation of reserves for a partial mobilisiation of reserves or even no mobilisation at all. Logically, the higher the reliability of the intelligence agency the more often its reports should correspond with the event that has actually occurred. However, even a very reliable agency is not perfect and can make mistakes. Of course, the propensity to make such mistakes might be different for different types of events – that is something for you to judge.

6. **Please note that the sum of your answers in EACH ROW must equal 100.**

   Considering the scenario, please enter the following probabilities

When the actual mobilization of reserves is	**The probability that the type of mobilisiation reported by the intelligence agency is**		
	None	**Partial**	**Full**
None			
Partial			
Full			

7. **Please note that the sum of your answers in EACH ROW must equal 100.**

   Considering the scenario, please enter the following probabilities

When the actual enemy force deployment is	**The probability that the enemy force deployment reported by the intelligence agency is**	
	Concentrated	**Not concentrated**
Concentrated		
Not concentrated		

8. **Please note that the sum of your answers in EACH ROW must equal 100.**

   Considering the scenario, please enter the following probabilities

When Soviet citizens in Egypt (diplomats, etc) are actually	**The probability that Soviet citizens in Egypt are reported by the intelligence agency to be**	
	Leaving	**Staying**
Leaving		
Staying		

9. **Please note that the sum of your answers in EACH ROW must equal 100.**
Considering the scenario, please enter the following probabilities

When the development of an Egyptian-Syrian air defence capability is	**The probability that the Egyptian-Syrian air defence capability will be reported by the intelligence agency to be**		
	None	**Limited**	**Extensive**
None			
Limited			
Extensive			

10. **Please note that the sum of your answers in EACH ROW must equal 100.**
For each part of question 10 you are asked to express your agree of belief in each hypothesis as a percentage probability given that you have received ONLY the intelligence report(s) specified in the first column of the table from your intelligence agency.

Hypothesis A = The Egyptian and Syrian coalition are capable of a hostile attack against Israel within the next few months. (Imminently hostile)

Hypothesis B = The Egyptian and Syrian coalition is NOT planning a hostile attack against Israel within the next few months. (Not imminently hostile)

Intelligence reports available	Probability that the hypothesis is true	
	A (imminently hostile)	**B (Not imminently hostile)**
Egypt has extensively built up its air defence capability		
Egypt has ordered a partial mobilisation of reserves		
Egypt has ordered a concentration of its forces		
Egypt has ordered the removal of Soviet Citizens		
Egypt has extensively built up its air defence capability and ordered a partial mobilisiation of reserves		
Egypt has extensively built up its air defence capability, ordered a partial mobilisiation of reserves and concentrated its forces		
Egypt has extensively built up its air defence capability, ordered a partial mobilisiation of reserves, concentrated its forces and Soviet Citizens are leaving		