Defence Research and Development Canada | Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

Heuristic Classification in a Simulated Air Threat Task

David J. Bryant

Defence R&D Canada – Toronto

Technical Report

DRDC Toronto TR 2005-191

June 2005

Canada

Heuristic Classification in a Simulated Air Threat Task

David J. Bryant

Defence R&D Canada – Toronto

Technical Report

DRDC Toronto TR 2005-191

June 2005

Author

David J. Bryant

Approved by

Carol McCann

Head, Command Effectiveness and Behaviour Section

Approved for release by

K.M. Sutton

Chair, Document Review and Library Committee

Abstract

Fast and frugal heuristics have proved effective models of human judgment for certain kinds of problems. To further explore the value of this approach, two experiments investigated the decision procedures used by human subjects to perform a cue-based classification task in a simulated air threat assessment task. Threat assessment is the classification of aircraft on the basis of sensor data that can be likened to probabilistic cues. Subjects learned to classify simulated aircraft using four probabilistic cues then classified test sets designed to contrast predictions of several heuristics, including the Take-the-Best-for-Classification (TTB-C) and Pros Rule developed specifically for the threat classification task. Results indicated that a proportion of subjects could be classified using TTB-C and another significant proportion as using the less frugal Pros Rule. No subject was observed to respond as predicted by a Bayesian strategy. Despite predictions that time pressure and perceived uncertainty of cues would affect how many subjects employed TTB-C, no effect of these variables was observed. These results suggest that it is possible to model multi-attribute decision tasks like threat assessment with fast and frugal heuristics but no single heuristic is a general model for the simulated threat assessment task.

Résumé

L'heuristique simple et rapide a fait la preuve de modèles efficaces de jugement humain pour certains types de problèmes. Dans le but d'étudier plus à fond la valeur de cette approche, deux expériences ont porté sur les procédures décisionnelles utilisées par les sujets humains afin d'exécuter une tâche de classification fondée sur les repères dans le cadre d'une simulation de tâche d'évaluation des menaces aériennes. L'évaluation des menaces est la classification des aéronefs fondée sur des données de capteurs comparables à des repères probabilistes. Les sujets apprenaient à classer des aéronefs classifiés au moyen de quatre repères probabilistes, puis subissaient des séries de tests classifiés conçus pour mettre en contraste les prédictions de plusieurs heuristiques, dont « ne garder que le meilleur en vue de la classification » (TTB-C) et la règle des « pour » élaborée spécifiquement pour la tâche de classification des menaces. Les résultats laissent voir qu'une proportion des sujets pouvait être classée au moyen de l'heuristique TTB-C, tandis qu'une autre proportion importante pouvait être classée au moyen de la règle des « pour » qui est moins simple. Aucun sujet observé n'a réagi comme prévu à une stratégie bayesienne. Malgré les prédictions voulant que les contraintes de temps et l'incertitude perçue aient un effet sur le nombre de sujets recourant à la TTB-C, on n'a noté aucun effet sur ces variables. Les résultats suggèrent qu'il est possible de modéliser des tâches décisionnelles à attributs multiples comme l'évaluation des menaces au moyen d'heuristiques simples et rapides, mais qu'aucune heuristique ne constitue un modèle général pour la simulation de tâche d'évaluation des menaces.

This page intentionally left blank.

Executive summary

To aid naval operators to perform the situation assessment process, researchers have been working to develop various forms of decision support systems, such as more informative displays and data analysis tools. A constant challenge in this domain, however, is to understand exactly what kinds of information operators need to build more accurate pictures. Key to the situation assessment process are threat assessment judgments, in which shipboard operators monitor sensors (e.g., radar, sonar) that provide information about the aircraft, surface vessels, and underwater vessels (called "contacts" by operators) in the area around the ship. Using this information, operators attempt to detect all craft around them, then classify them (i.e. what type of craft) and identify them if possible (i.e. what nation does the craft belong to and from where did it come?). Unfortunately, empirical data regarding how people actually weigh and combine information is not complete, particularly with respect to the kinds of tasks performed in situation assessment.

Fast and frugal heuristics have been proposed as possible models for human judgment in a range of tasks. A fast and frugal heuristic is a computationally simple procedure for making judgments with limited information, consisting of a search rule, stopping rule, and heuristic decision principle. Such a heuristic is assumed to be fast because it involves relatively few processing steps and frugal because it requires little information. Although they have proved effective in solving certain problems, evidence that fast and frugal heuristics provide plausible models of human judgement is equivocal. For this reason, two experiments explored whether a fast and frugal heuristic could successfully model subjects' behaviour in a simulated air threat assessment task.

In Experiment 1, participants learned to classify contacts in a simulated naval warfare environment and then were tested, under low or high time pressure, on sets of contacts that contrasted predictions of several heuristics, including the Take-the-Best-for-Classification (TTB-C) and Pros Rule developed specifically for the threat classification task. Results indicated that a proportion of subjects could be classified using TTB-C and another significant proportion as using the less frugal Pros Rule, regardless of time pressure. No subject was observed to respond as predicted by a Bayesian strategy.

Experiment 2 examined the effect of information uncertainty on participants' use of heuristics. Information uncertainty has been suggested as a factor that promotes use of simple decision heuristics and is an important concern in operational settings (e.g., the HALIFAX class frigate). As in Experiment 1, a proportion of subjects could be classified using TTB-C and another significant proportion as using the less frugal Pros Rule but no subject was observed to respond as predicted by a Bayesian strategy. The manipulation of perceived cue uncertainty had no effect on participants' use of fast and frugal heuristics.

The results of the two suggest that it is possible to model multi-attribute decision tasks like threat assessment with fast and frugal heuristics but no single heuristic is a general model for the simulated threat assessment task. Further study is needed to determine what factors might govern how people select a heuristic to use in the threat assessment task.

Bryant, D. J. 2005. Heuristic Classification in a Simulated Air Threat Task. DRDC Toronto TR 2005-191. Defence R&D Canada - Toronto.

Sommaire

Afin d'aider les opérateurs navals à exécuter le processus d'évaluation de la situation, des chercheurs ont élaboré divers types de systèmes de soutien décisionnel, notamment des affichages plus informatifs ainsi que des outils d'analyse des données. La compréhension exacte du genre de renseignements dont les opérateurs ont besoin pour brosser des tableaux exacts demeure cependant un enjeu permanent dans ce domaine. Les jugements établis à la suite de l'évaluation des menaces sont essentiels au processus d'évaluation de la situation. Pour ce faire, les opérateurs embarqués surveillent les capteurs (p. ex., les radars, les sonars) qui fournissent des renseignements sur les aéronefs, les navires de surface et les sous-marins (appelés « contacts » par les opérateurs) dans la zone entourant le navire. Se servant de ces renseignements, les opérateurs tentent de détecter toutes les embarcations qui se trouvent autour d'eux, de les classifier (c.-à-d. de trouver le genre d'embarcation) et, si possible, de les identifier (c.-à-d. à quel pays l'embarcation appartient-elle et d'où vient-elle?). Malheureusement, les données empiriques sur la façon dont les gens apprécient et regroupent les renseignements sont incomplètes, particulièrement en ce qui a trait aux genres de tâches effectuées lors de l'évaluation de la situation.

On a proposé de recourir à une heuristique simple et rapide comme modèle éventuel de jugement humain pouvant être appliqué à diverses tâches. L'heuristique simple et rapide est une méthode de calcul simple permettant d'établir des jugements même avec des renseignements limités. Elle comprend une règle de recherche, une règle d'arrêt et un principe de décision heuristique. On dit de l'heuristique qu'elle est rapide parce qu'elle implique relativement peu d'étapes de traitement et qu'elle est simple parce qu'elle requiert peu de renseignements. Même si elle s'est avérée efficace pour résoudre certains problèmes, la preuve que l'heuristique simple et rapide produit des modèles plausibles de jugement humain demeure équivoque. Voilà pourquoi on a réalisé deux expériences visant à déterminer si l'heuristique simple et rapide pouvait réussir à modéliser les comportements des sujets lors d'une simulation de tâche d'évaluation des menaces aériennes.

Lors de l'expérience n° 1, les participants ont appris à classifier les contacts dans un contexte simulé de guerre navale. Ils ont ensuite été mis à l'épreuve, dans le cadre de contraintes de temps mineures ou importantes, sur des ensembles de contacts mettant en contraste les prédictions de plusieurs approches heuristiques, dont « ne garder que le meilleur en vue de la classification » (TTB-C) et la règle des « pour » élaborée particulièrement pour la tâche de classification des menaces. Les résultats laissent voir qu'une proportion des sujets pouvait être classée au moyen de l'heuristique TTB-C, tandis qu'une autre proportion importante pouvait l'être au moyen de la règle des « pour » qui est moins simple, peu importe la contrainte de temps. Aucun sujet observé n'a réagi comme prévu à une stratégie bayesienne.

L'expérience n° 2 examinait les effets de l'incertitude relative aux renseignements sur l'utilisation que les participants faisaient de l'heuristique. On a déjà suggéré que l'incertitude relative aux renseignements était un facteur qui favorise le recours à l'approche heuristique pour prendre des décisions simples et que cela constitue une source de préoccupation importante dans les installations opérationnelles (p. ex., la frégate de classe HALIFAX). Comme dans le cas de l'expérience n° 1, une proportion des sujets a pu être classée au moyen

de l’heuristique TTB-C, tandis qu’une autre proportion importante l’a été au moyen de la règle des « pour », qui est moins simple, sans qu’aucun sujet observé ne réagisse comme prévu à une stratégie bayesienne. La manipulation de l’incertitude perçue relative aux repères n’a eu aucune incidence sur l’utilisation que les participants faisaient de l’heuristique simple et rapide.

Les résultats des deux expériences laissent entrevoir qu’il est possible de modéliser des tâches décisionnelles à attributs multiples comme l’évaluation des menaces au moyen d’heuristiques simples et rapides, mais qu’aucune heuristique ne constitue un modèle général pour la simulation d’une tâche d’évaluation des menaces. Il faudra procéder à d’autres études pour déterminer les facteurs qui régissent la façon dont les individus choisissent l’heuristique à utiliser dans le cadre de la tâche d’évaluation des menaces.

Bryant, D. J. 2005. Heuristic Classification in a Simulated Air Threat Task. R & D pour la défense Canada - Toronto TR 2005-191. R & D pour la défense Canada - Toronto.

This page intentionally left blank.

Table of contents

List of figures

List of tables

Acknowledgements

The author would like to express his deep gratitude to Andrea Hawton, Tonya Hendriks, Elaine Maceda, and Wendy Sullivan-Kwantes for their invaluable assistance in conducting the experiments reported here.

This page intentionally left blank.

Introduction

Background

Gigerenzer, Todd, and the Adaptive Behavior and Cognition Group [1] have proposed that *fast and frugal heuristics* can serve as models for human judgment. A fast and frugal heuristic is a computationally simple procedure for making judgments with limited information, consisting of a search rule, stopping rule, and heuristic decision principle [2]. Such a heuristic is assumed to be fast because it involves relatively few processing steps and frugal because it requires little information

A variety of studies have shown fast and frugal heuristics to be accurate and efficient solutions to certain judgment tasks. The Recognition Heuristic, for example, has been shown to perform as accurately as more complex statistical analyses on problems in which a decision maker must select one option from just two possibilities [1, Ch. 2] [3]. Another example of a fast and frugal heuristic is Gigerenzer and Goldstein's [3] "Take the Best" (TTB) heuristic in which a decision alternative is chosen on the basis of a single cue that is probabilistically associated with the dimension on which the choice is made [1, pp. 77-81] [3].

In simulation studies, TTB performs a choice task, such as choosing the larger of two cities roughly as accurately as more computationally intensive linear regression models [3], a finding that has been replicated with 19 other data sets drawn from psychology, economics, and other fields and involving comparison along a variety of dimensions [4]. In addition to achieving comparable accuracy, the TTB consistently exhibits a clear advantage over linear procedures in terms of frugality, consulting, on average, fewer cues and performing fewer computations than linear procedures. A non-compensatory heuristic like TTB generally performs well when the task environment is itself structured such that the validity of cues falls off dramatically in a non-compensatory fashion [5].

Although fast and frugal heuristics can solve certain problems well, it remains to be convincingly demonstrated that they provide plausible models of human decision making. Several studies have examined subjects' choice behaviour in tasks in which subjects are required to use probabilistically predictive cues to select an alternative (e.g., [6] [7] [8] [9]). Typically, these studies report that a proportion of subjects can be classified as employing TTB to make choices on the basis of their patterns of responses to a series of items but that a significant, often majority, proportion of subjects seem to use more complex, compensatory procedures.

The propensity of subjects to employ TTB is affected by a range of factors, such as costs imposed on obtaining cues. Broder [6] found 40% of subjects used TTB when the cost of cue information was relatively low and 60% when the cost was relatively high. Newell and Shanks [9] have also replicated the finding that a greater proportion of subjects employ TTB when cue costs are relatively higher.

People also seem to be sensitive to the payoff structure of the task environment. Employing a simulated stock market task, Broder [10] varied not only the cost of purchasing cues but the way cues were associated to appropriate choices. In one case, cues were designed to create a compensatory environment in which no single cue outweighed the combination of all others. In the other case, the environment was compensatory with one highly weighted cue. Roughly half of the subjects were classified as TTB-users when the underlying task structure was non-compensatory but only a quarter of subjects were so classified when the structure was compensatory. Broder [10] notes that purchasing more than one cue in the compensatory condition could enhance subjects' predictive power whereas the cost would outweigh the benefit in the non-compensatory condition. The implication of such findings is that TTB is a strategy *available* to decision makers but task conditions, such as the cost of examining cues and the cue structure, play a large role in determining whether people actually employ it.

One thing immediately stands out from the preceding discussion. It is typically the case that only a subset (albeit a majority in some cases) of subjects can be classified as using TTB. Even under favourable conditions, a significant proportion of subjects used more complex decision procedures. In other words, the very evidence that supports TTB as a potential model of choice also calls into question its generality as a model. Subjects have frequently been observed to deviate from the principles of fast and frugal heuristics. Newell and Shanks [9], for example, found that the order in which subjects search for cues, in particular, often deviates substantially from what would be expected if decision makers were employing a fast and frugal heuristic (see also [8]).

Not only is it the case that individuals differ in the kind of decision strategy they appear to use in experiments, conditions typically have to be contrived to make TTB an attractive strategy [10]. Costs to data acquisition, for example, must be imposed and the underlying task structure must be such that TTB yields a high probability of success. Moreover, these manipulations, it seems, must be fairly extreme to produce a significant proportion of subjects using TTB [9]. Even simplifying the task environment and emphasizing the non-compensatory structure of a task does not dramatically increase the proportion of subjects using TTB [8]. Adaptive use of TTB also likely depends on individual factors, such as intelligence [10].

Purpose of the Current Study

Little work has been done on applying the fast and frugal heuristic approach to military decision making, even though the limitations of time and available information associated with command and control suggest that heuristics could prove very useful in the military context (see [11]). The fast and frugal heuristic approach is consistent with naturalistic or recognition-based theories of decision making, which have been applied to military decision making (e.g., [12] [13]). In fact, fast and frugal heuristics could augment naturalistic theories by providing computational models for various decision tasks [14]. Thus, the aim of this paper is to explore whether a fast and frugal heuristic can be developed and validated for a basic military task.

Threat assessment is a complex task performed by multiple systems and human operators that entails detection and classification of aircraft (called "contacts" by operators) according to their degree of threat to one's self and allied units.[1] Modern warships carry multiple sensors (e.g., radar, electronic emission detection systems, Interrogate-Friend-Foe or IFF interrogation systems) and computer systems that take in and process data in ways designed to refine sensor signals and assist in the identification process. Human operators receive data via a combat information system that displays contacts (also called tracks) and associated data. The task can be difficult because data are typically distributed across several systems and operators (e.g., [15] [16]). Operators must also integrate sensor data with other kinds of information (e.g., geography, intelligence reports) and their operational knowledge (e.g., tactics, doctrine, rules of engagement) [17]. All of this occurs in complex, dynamic operational environments and operators often work under extreme time pressure and stress [18].

Formally, threat assessment can be viewed as categorization based on uncertain cues, which is not a straightforward task because no cue provides certain classification. Operators cannot rely on just the identification of a contact's type and seem to judge the threat of a contact on the basis of the extent to which its behavior deviates from expectations [19]. It is however, not well understood exactly how operators use cue information to make threat assessments [18].

Given concerns that military decision making is vulnerable to problems of information overload and uncertainty, the fast and frugal heuristic approach may provide a useful framework in which to study time- and information-stressed decision making. Consistent with this possibility, it appears operators do not consider or weigh all available data equally and that they employ decision making procedures that differ from those previously assumed [18].

Take-the-Best-for-Classification (TTB-C)

A variant of TTB, called Take-the-Best-for-Classification heuristic (TTB-C), was devised to perform the threat classification task.[2] Illustrated in Figure 1, TTB-C is based on the premise that the single most valid cue can be used to make accurate threat classification judgments in a task environment in which that cue is highly predictive. Unlike TTB, which chooses between two objects along a single dimension, TTB-C places a single object into one of two categories. Thus, TTB-C is simpler in some respects than TTB but it takes from TTB the basic search concept of locating the single best cue to make its decision.

---

[1] In this paper, we focus on air threat assessment in the naval context in which units are generally surface vessels operating in areas with a wide range of military and commercial air traffic.

[2] TTB-C is also derivable from the Lexicographic heuristic for two-alternative choice, which is a generalization of Take-the-Best [22, p. 143].

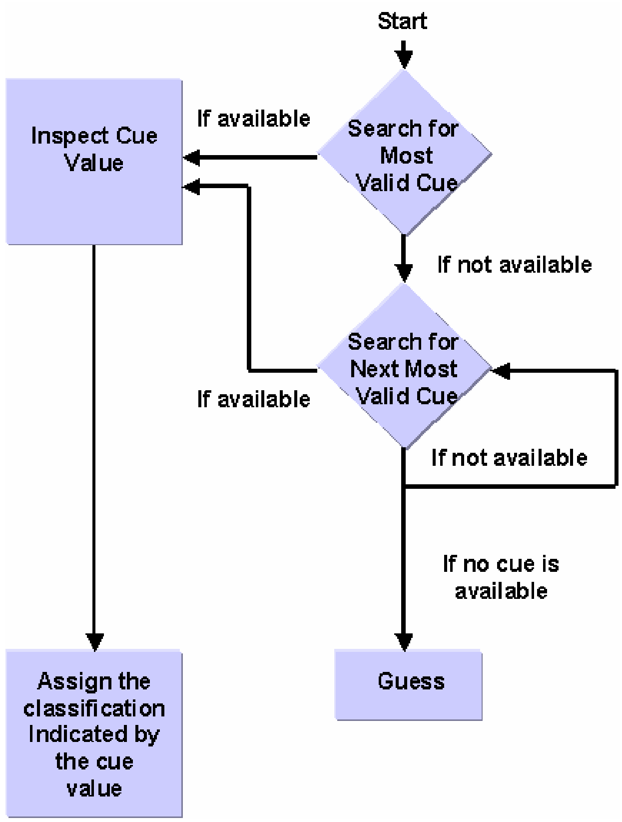

Figure 1: *The Take-the-Best-for-Classification (TTB-C) Heuristic*

Given an as-yet-unclassified contact, TTB-C begins by searching for the single most valid cue to serve as the basis for classification. In the experiment described in this report, all cues associated with contacts will be available, so the most valid cue should always be inspected. When the most valid available cue is located, the heuristic assesses which threat class has the greater probability of being true given the value of that cue and makes that threat class the output of the heuristic. The heuristic will be applied here to an experimental task in which subjects make a simplified two-category choice (friend or foe) but the heuristic could apply to threat classification with a larger set of threat classes. With the contact classified, the heuristic terminates. Should no valid cue be found, the heuristic can only guess.

TTB-C, as illustrated here, assumes that there exist one or more cues that have some non-random association to the threat class of contacts and that all, or some subset, of these cues can be inspected by the decision maker. Moreover, the decision maker must have acquired, through experience or training, knowledge of the relative validities of these cues. These, of course, are not minor assumptions but there is sufficient evidence that people can learn cue validities, even if their learning is imperfect [20] [21].

Compensatory Heuristics

Just as TTB-C is an adaptation of the TTB heuristic to the single-choice classification problem, other two-alternative choice decision strategies can be adapted. Among the decision strategies that have been examined are Franklin's Rule and Dawes' Rule. Franklin's Rule is a procedure by which a decision maker calculates the sum of cue values weighted by the corresponding cue validities for each alternative and selects the alternative with the highest score [22]. Dawes' rule is similar and calculates the sum of un-weighted cue values and selects the alternative with the highest score. Because both Franklin's and Dawes' Rules count up bits of evidence, or "pros," for an alternative, they can be termed Pros Rules (for the sake of clarity, these rules will be referred to as the Weighted Pros and Unweighted Pros rules throughout this report). Both are compensatory, meaning that they employ all available cues although they do not compute probabilities to reach a decision.

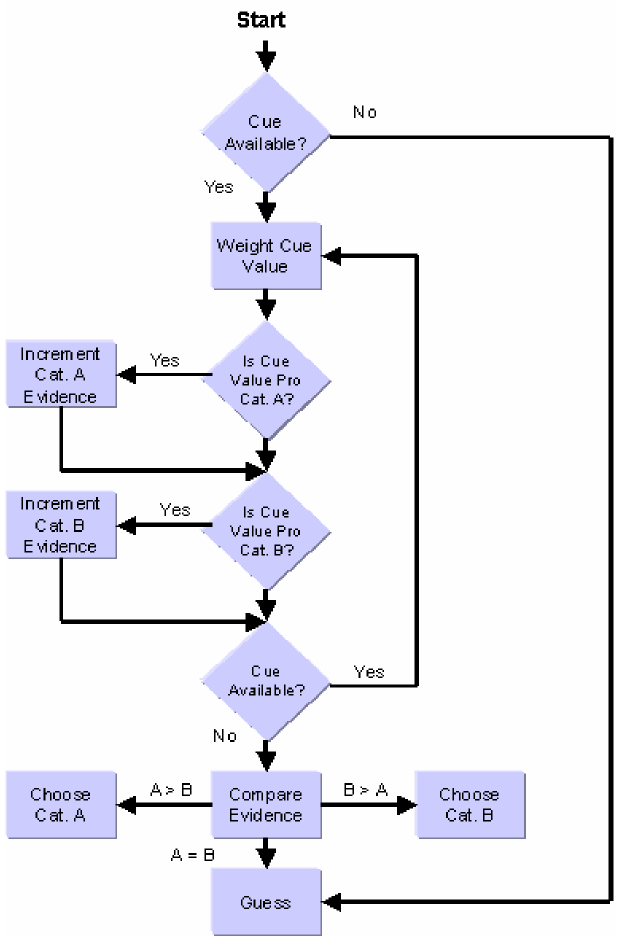

Figure 2: *Weighted Version of the Pros Rule for Classification*

Versions of the Pros Rules were formulated for the threat classification task. Unlike their progenitors, they do not compare cue values for two alternatives but rather examine each cue value and assign evidence toward either friend or foe classification, depending on the associations of cue values to threat class. A running sum is maintained and, after all available cues have been inspected, used to place the contact in the friend or foe category. Figure 2 contains an illustration of the Weighted Pros Rule, which weights cues by their validity, adapted for threat classification. A classification version of the Unweighted Pros Rule is performed just as illustrated in Figure 2 but without the weighting step following the selection of a cue. These rules use more information than TTB-C but are more generally useful because their accuracy is not limited to cases where a single cue is highly predictive.

A question of particular interest to effective threat assessment is whether operators employ compensatory or non-compensatory strategies. Compensatory and non-compensatory procedures achieve different trade-offs of accuracy and costs of cognitive resources. Examining whether humans tend to employ compensatory or non-compensatory procedures will provide insight into the ways operators are likely to respond to operational constraints on time, data, and computational power.

Experiment 1

The experimental task was framed in a simulated threat assessment environment. Various "contacts" (simulated aircraft) were presented on a simulated radar screen for subjects to classify as either *friend* or *foe* based on the values of four cues. The availability of just four cues is unrealistic, as was the fact that all the cues were strictly binary, but the task was not intended to accurately describe an operational context. Instead, the experimental task allowed precise variations of the relationships of cues to contact classification. Each cue value had a specific probability of being associated with friend and foe contacts, with these probabilities determining the cue's validity in classifying contacts.

Subjects were required to learn the associations between cue values and threat classification. Numerous studies (e.g., [20] [21]) have demonstrated that people can learn to weigh probabilistic cues according to the extent those cues are correlated with a criterion. In cases where cues are linearly related to a criterion, peoples' judgments of the criterion are consistent with a weighted averaging model [23] [24], suggesting that people can internalize some representation of the probabilistic relationships of cues to criterion. Probabilistic cue learning by trial and error, however, is difficult and people are far from perfect in their ultimate performance (e.g., [25] [26]). Indeed, Klayman [27] has noted that many factors strongly impair learning non-linear relationships of cues to criterion, including the number and abstractness of cues and the nature of feedback provided during learning (e.g., [28] [29]).

Results of a pilot study employing a similar procedure showed that subjects exhibited steady improvement in classification accuracy in the training session and provided fairly accurate estimates of cue validity for all cues [30]. The pilot study also made clear that it can be difficult to distinguish hypothesized decision procedures because the procedures predict the same responses to many items. In the pilot study, for example, only two cue patterns elicited different predicted responses of TTB-C and a strategy based on the Bayesian conditional probabilities. Moreover, the patterns occurred only rarely in the test set.[3] Thus, although a proportion of subjects seemed to respond to those items as though using TTB-C, it was not possible to conclusively determine whether subjects' responses actually did conform to the predictions of that heuristic.

To better examine subjects' decision procedures, the current experiment employed a test set of contacts designed to offer maximal contrast between hypothesized cognitive procedures. This test set contained an artificially high number of items for which TTB-C, the Pros Rules, and an integrative Bayesian strategy made different predictions. The Bayesian strategy was assumed to compute the conditional probabilities of friend and foe classifications given the particular pattern of cue values for a contact and select the alternative with the higher probability of being the correct classification. By examining subjects' responses to these items, it was hoped to identify the cognitive procedure used for each subject.

[3] Given that the cues used to describe contacts varied in cue validity and contacts were generated randomly according to these validities, the 16 possible cue patterns varied in their likelihood of occurring in the test set of contacts. The most common patterns tended to elicit the same prediction from decision procedures.

A second objective of the experiment was to examine the effects of time pressure on decision strategy and cue use. Subjects performed the threat classification task under two time limits, a relatively high time pressure (8 seconds) and a relatively low time pressure (14 seconds).

Method

Subjects

Subjects were 24 men and women who were employees of Defence Research and Development Canada - Toronto (DRDC Toronto), students conducting research at DRDC Toronto, or individuals recruited from local universities. All subjects were aged 18 and older, had normal or corrected-to-normal vision, and were unfamiliar with the specific hypotheses and stimulus configurations of the experiments. All received payment in exchange for participation.

Materials

The experiment was conducted using the Team and Individual Threat Assessment Network (TITAN) experimental platform, which is a low fidelity threat assessment simulator. The interface presents a radar screen on which "contacts" are presented by asterisk symbols. Each contact corresponds to a single entity around the subject's "own ship," which is indicated by a blue circle at the center of the radar screen. Using the computer mouse, the subject can click on ("hook") a contact, which activates a set of buttons that allow access to information about that contact. This information consists of four characteristics, such as speed, altitude, and so on. The interface can be customized to allow subjects to view all of the contact's characteristics simultaneously or to restrict subjects to viewing one characteristic at a time. By clicking another button, subjects call up a box in which two possible classifications, "friend" and "foe," are indicated. Radio buttons under each classification allow the subject to indicate a classification judgment. Further windows open to allow subjects to indicate a confidence judgment and receive feedback concerning their classification accuracy. TITAN was run on Pentium PC computers.

Subjects performed two conditions, one involving low time pressure during test and the other high time pressure. Consequently, two sets of 300 contacts (150 friend and 150 foe) were created for the training sessions and two sets of 50 contacts (25 friend and 25 foe) were created for the test sessions.

Design

Three variables were manipulated in this experiment. The first, varied within subjects, was the Cue Validity of each cue used to describe contacts in the training stimuli sets. To vary Cue Validity, each possible value of a cue (values 1 and 2) was probabilistically associated to friend and foe classifications such that each cue differed in diagnosticity. Thus, for one cue each possible value was paired with the friend or foe classification 90% of the time, for another cue 80% of the time and so on. Table 1 indicates the proportions of friend and foe contacts possessing each cue value for the four cues in the two training sets.

Table 1: Relative Frequencies of Cue Values for Friend and Foe Contacts (Experiment 1)

	SET 1							
	Cue 1 (Signal Strength)		***Cue 2 (Direction of Origin)***		***Cue 3 (Counter Measures)***		***Cue 4 (Response)***	
	Value 1 (Medium)	***Value 2 (High)***	***Value 1 (B. Lag.)***	***Value 2 (Red Sea)***	***Value 1 (None)***	***Value 2 (Jamming)***	***Value 1 (No Resp)***	***Value 2 (Given)***
Friend	10%	90%	40%	60%	70%	30%	80%	20%
Foe	90%	10%	60%	40%	30%	70%	20%	80%
	SET 2							
	Cue 1 (Intelligence)		***Cue 2 (Manoeuvre Pattern)***		***Cue 3 (Comm. Mode)***		***Cue 4 (Electronic Warfare)***	
	Value 1 (Private)	***Value 2 (Platform)***	***Value 1 (Delta)***	***Value 2 (Foxtrot)***	***Value 1 (1)***	***Value 2 (3)***	***Value 1 (Big B)***	***Value 2 (None)***
Friend	10%	90%	40%	60%	70%	30%	80%	20%
Foe	90%	10%	60%	40%	30%	70%	20%	80%

Note: B. Lag. = Blue Lagoon; Big B = Big Bulge

The second variable manipulated was the Contact Type in the test stimuli. Each test set was made up of patterns that offered contrasting predictions of the three contending classification strategies discussed previously; namely TTB-C, the Bayesian strategy, and the Pros Rule.

Table 2: *Predicted Responses to Contact Types by Hypothesized Strategies (Experiment 1)*

	Predicted Response of Strategy			**Contact Types**	
Cue Pattern	**TTB-C**	**Bayesian**	**Pros Rule**	**Foe**	**Friend**
1,2,1,1	Foe	Friend	Friend	B	A
2,1,2,2	Friend	Foe	Foe	A	B
1,1,1,1	Foe	Foe	Guess	D	C
1,2,1,2	Foe	Foe	Guess	D	C
1,2,2,1	Foe	Foe	Guess	D	C
2,1,1,2	Friend	Friend	Guess	C	D
2,1,2,1	Friend	Friend	Guess	C	D
2,2,2,2	Friend	Friend	Guess	C	D

Note: Cue pattern indicates the value (as 1 or 2) for each cue in order of cues listed in Table 1.

Eight cue patterns were identified for which at least one strategy offered a differing response than that predicted by the other strategies. These are presented in Table 2. The patterns listed in Table 2 refer to the sequence of cue values for the four cues used to describe contacts. Thus, the pattern 1,2,1,1 indicates that Cue 1 possessed Value 1 (see Table 1 above), Cue 2 possessed Value 2, and so on. The predictions of the three strategies are indicated for each pattern in Table 2. The Pros Rule (unweighted version) produced no clear preference for Friend or Foe in six patterns in which two cues indicated Friend and two cues indicated Foe and so this strategy can only guess the correct classification.[4]

In the test set, each of the critical patterns was paired an equal number of times with friend and foe contacts. From these contacts, we created four Contact Types (A, B, C, and D), listed for each test item in Table 2. The types distinguish the predicted accuracy of the possible strategies for each item type based on their predicted responses. Thus Type A and B items, which were the same patterns but differing in the threat class to which they were associated, elicit opposing predictions from TTB-C and the Bayesian strategy. Where TTB-C would predict that these patterns indicate a friend, the Bayesian strategy would predict they indicate a foe, and vice versa. Type C and D patterns elicit the same predictions from TTB-C and the Bayesian strategy but force the Unweighted Pros Rule to guess because equal numbers of cues suggest friend and foe classifications. The predicted accuracy levels in Table 3 form hypotheses for each decision strategy that can be used to identify which, if any strategy, describes subjects' responses to the four Contact Types.

Table 3: *Predicted Accuracy Levels by Contact Type (Experiment 1)*

	Contact Type*			
Heuristic	**A**	**B**	**C**	**D**
TTB-C	0%	100%	0%	100%
Bayesian	100%	0%	0%	100%
Pros Rule	100%	0%	Guess	Guess

* See Table 2

The final variable, also varied within subject, was the Time Pressure applied during the test session (Low vs. High). In the low time pressure condition, subjects had 14 seconds to inspect cues and make a judgment, whereas in the high time pressure condition they had 8 seconds. The low time pressure limit is roughly equal to the average response time observed in the pilot experiment, in which subjects were under no time pressure. The high time pressure limit is roughly equal to two standard deviations less than the mean response time of the pilot experiment.

Because subjects performed the experimental task in both a High and Low Time Pressure condition, two contacts sets were created using different cue labels and cue values but the underlying cue validity structures of the two sets were the same. The two sets were fully

[4] Experiment 2 examines differential predictions of the weighted and unweighted versions of the Pros Rule for this class of item.

counterbalanced with the two Time Pressure conditions and the order in which they were performed.

Procedure

The experiment was divided into two sessions for the High and Low Time Pressure conditions, each with a training and test phase. In the training phase, subjects received 300 contacts, of which 150 were friends and 150 foes. The training set was representative of all possible cue patterns. Given the structure of cue information, some patterns were more likely to occur than others and the training set reflected this. All contacts were presented simultaneously in random positions on the radar screen, although the subject was required to use “zoom in” and “zoom out” buttons to view all of the contacts. Each contact had four cues associated with it, specifying cue values generated according to the probability matrix shown in Table 1. Subjects selected one contact at a time in whatever order they wished and accessed that contact’s cue values. All four values were available on the screen at the same time but the order in which cues were listed was random from contact to contact. Subjects then made a classification judgment, under no time pressure, indicating that the contact is either friend or foe. After this, subjects received accuracy feedback on their classification judgment in the form of a message indicating whether they were correct or incorrect and provision of the correct classification. Subjects received no initial information concerning the predictiveness of cues. In the both the training and test phases, contacts disappeared from the screen immediately after subjects made a judgment to prevent subjects from revisiting the same contact more than once.

Following the training phase, subjects were allowed a short break and then performed the test phase. The test phase followed the same procedure as the training phase with a few important differences. Subjects could no longer access all cue information simultaneously. Instead, each cue was represented by an individual button that subjects pressed to view the value of that cue. The order of the buttons was randomized from contact to contact. Subjects were given no specific instructions concerning how many cues to select; they were told to view whatever cue information they wanted before making their classification judgment. In addition, subjects were presented with only 50 contacts (10 type A, 10 type B, 15 type C and 15 type D) and they did not receive feedback on the accuracy of their judgments. Subjects were under either Low or High Time Pressure when making their judgments. A clock display counted down the seconds remaining so subjects could monitor how much time remained to make each judgment. If a subject failed to respond within the time limit, the subject’s non-response was counted as an error.

Following the training and test phases, subjects were asked to indicate the diagnosticities of each cue as a predictor of contact classification by judging the proportions of cue values associated with friend and foe contacts. Subjects also were asked to describe what type of strategy they employed.

Results

Subjects’ performance in the training session and responses to the post-experiment survey were analysed first to determine how well subjects learned to classify contacts. Then subjects' performance and cue use in the test session were analysed to examine subjects' decision strategies.

Training Session

The contacts presented during the training session were divided into six blocks of 50 contacts each, based on the order of presentation (i.e., the first 50 contacts, the next 50, etc.). Accuracy scores (the percentage of contacts correctly classified as friend or foe) were calculated for each block for each subject to create mean accuracy scores, which are shown broken down by Time Pressure condition in Figure 3. A two-way, within-subjects Analysis of Variance (ANOVA) revealed a significant effect of Training Block [$F(5,115) = 49.54$, $MSe = 0.30$, $p < .01$] but no significant main effect of Time Pressure [$F(1,23) = 1.28$, $MSe = 0.058$, *n.s.*], which is to be expected because the training was identical for High and Low Pressure conditions. There was also no significant interaction effect between the two factors [$F(5,115) = 0.15$, $MSe = 0.0005$, *n.s.*].

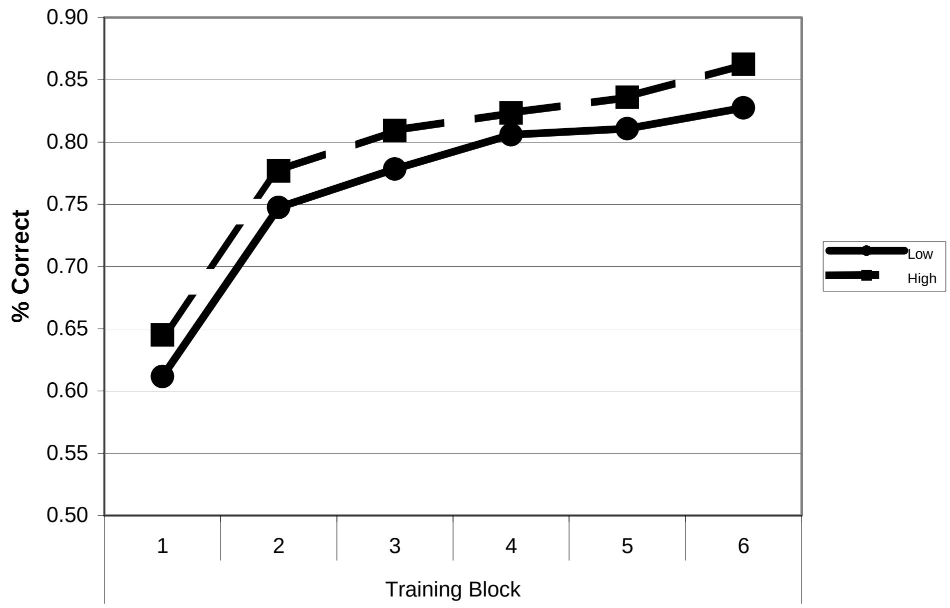

Figure 3: *Classification Accuracy by Block in the Training Session (Experiment 1)*

As can be seen in Figure 3, accuracy generally increased over Trial Block and subjects' performance in the final block was near, but not quite, optimal with respect to both the Bayesian and TTB-C strategies. The mean Block 6 accuracy of .83 in the Low Time Pressure condition was significantly less than the optimum of .94 predicted by the Bayesian strategy [$t(23) = 3.73$, $p<.01$], as was the mean accuracy of .86 in the High Time Pressure condition [$t(23) = 3.44$, $p<.01$]. The mean Block 6 accuracy of the Low Time Pressure condition did not significantly differ from the optimum of .90 predicted by TTB-C [$t(23) = 1.66$, *n.s.*]. The mean Block 6 accuracy in the High Time Pressure condition did, however, differ significantly from .90 [$t(23) = 2.40$, $p<.01$].

Another measure of subjects' degree of learning comes from their responses to the post-experiment survey. Subjects were asked to estimate, for each cue, the percentages of friends and foes possessing each of the two possible cue values (the actual percentages are presented in Table 1). Error scores were calculated for each subject by subtracting the actual cue value percentage from the subject's estimate for each cue value and computing the average difference or error score across the four cues. These mean error scores are shown in Table 4 and indicate the average amount by which subjects over- or underestimated the percentages of friends and foes possessing each cue value as well as the standard deviations of subjects' error scores, which indicate the variability of subjects' error scores.

Table 4: *Mean Error and Mean Standard Deviation of the Error for Perceived Cue Uncertainty (Experiment 1)*

Time Pressure	Mean Error (%)	SD of Error (%)
Low	-0.08	21.22
High	-3.85	21.57

Note: Negative error scores indicate underestimation of true cue value association probabilities.

Although the absolute mean error was somewhat greater in the High than Low Time Pressure condition, this effect was not significant [$t(23) = 1.48$, *n.s.*]. Again, this is to be expected because the training was the same for each condition. Similarly, there was no reliable difference in the standard deviations of the error [$t(23) = 0.16$, *n.s.*]. More revealingly, the overall mean error score (combined for Low and High Pressure conditions) was quite low and did not differ significantly from zero [$t(23) = 1.38$, *n.s.*]. Inspection of the standard deviations of mean error scores, however, revealed that subjects' estimates of cue validities were highly variable and the overall mean standard deviation of the error was significantly greater than zero [$t(23) = 14.11$, $p < .01$]. Thus, although subjects' errors averaged to near zero, their individual judgments of cue validities often exhibited substantial error. This result suggests that subjects had learned that for each cue value that was more frequently associated with friend or foe, its complementary value was as frequently associated with the opposite contact class. Subjects balanced overestimations of cue associations to one threat class by corresponding underestimation in to the other class but did not necessarily learn the precise values of cue validities.

Overall, subjects clearly learned to classify contacts to a high degree of accuracy, which indicates learning of individual cue validities. Subjects' explicit expressions of cue validities, however, were not so accurate. Subjects also exhibited less than perfect ability to list the cues in order of cue validity [$t(23) = 13.97$, $p < .01$], which indicates that subjects had not completely internalized the cue structure of the stimuli set.

Classification Strategy

The test set was made up entirely of patterns that offered contrasting predictions of the three contending classification strategies (see Table 2). Type A and B items, for example, elicit opposing predictions from TTB-C and the Bayesian strategy. Type C and D patterns elicit the same predictions from TTB-C and the Bayesian strategy but force the (unweighted) Pros Rule to guess because equal numbers of cues suggest friend and foe classifications. Thus, it should be possible to identify post hoc the type of classification strategy employed by subjects by examining their responses to the different patterns.

To assess strategy use by subjects, we examined each subject's proportion of correct responses for each pattern type (A, B, C, and D) and compared this to the predicted accuracy levels listed in Table 3. A subject was considered to have performed with 100% accuracy for a pattern type if that subject responded correctly to that pattern more often than would be expected by chance, the assumed criterion of which was 5%. For example, the test set contained 10 items with pattern A. Assuming random responding according to a binomial distribution, the probability of selecting seven correct responses is 11.7% whereas the probability of selecting eight correct responses is 4.4%. Consequently, subjects scoring 8, 9, or 10 correct for items with pattern A were scored as showing "100%" performance for those items. Although subjects exhibiting only 8 or 9 correct for these items did not actually score 100%, we concluded that these subjects were consistently applying a strategy that predicted correct response for those items and that any incorrect responses could be attributed to response error. Similarly, we determined that accuracy of 2 items was only 4.4% likely by chance and scored subjects as showing "0%" performance if they correctly responded to 0, 1, or 2 items with pattern A. Similar criteria were computed for all pattern types. A subject was considered to be responding at "50%"(guessing rate) if they responded accurately to a number of items consistent with a binomial probability greater than .05.

Table 5 presents subjects' accuracy scores for each contact type and the decision strategy assigned to each subject. Table 6 presents the number of subjects classified as using a given decision strategy. Looking at each individual's patterns of accuracy over the critical cue patterns, no subjects exhibited the pattern predicted by the Bayesian strategy (i.e. consistently incorrect for patterns B and C and consistently accurate for patterns A and D). The majority of subjects' patterns matched predictions of either the TTB-C (i.e. consistently incorrect for patterns A and C and consistently accurate for patterns B and D) or Pros Rule (i.e. consistently incorrect for pattern B, consistently accurate for pattern A, and indistinguishable from random choice for patterns C and D), although a sizeable portion seemed to be guessing or using some unclassified strategy.

Table 6: *Number of Subjects Classified as Using Hypothesized Decision Strategies (Experiment 1)*

	Decision Heuristic			
Time Pressure	**TTB-C**	**Pros Rule**	**Guessing**	**Unclassifiable**
Low	9	7	2	6
High	10	6	6	2

* N = 24 at each level of Time Pressure

The proportions of subjects using TTB-C, the Pros Rule, and either guessing or unclassifiable were very similar in the High and Low Time Pressure conditions. A Pearson Chi-Square test revealed no significant difference between the two conditions [$\chi2$ =4.51, df = 3, n.s.].

Table 5: *Response Accuracies by Contact Type and Decision Strategy Classification (Experiment 1)*

	Low Time Pressure									High Time Pressure								
	True Accuracy				Interpolated Accuracy				Decision Strategy	True Accuracy				Interpolated Accuracy				Decision Strategy
Sub.	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**		**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**	
1	0.00	1.00	0.00	1.00	0	1	0	1	TTB-C	0.10	0.90	0.06	0.86	0	1	0	1	TTB-C
2	0.30	0.90	0.50	0.50	G	1	G	G	X	0.50	0.40	0.64	0.44	G	G	G	G	Guess
3	0.90	0.10	0.56	0.43	1	0	G	G	Pros	0.00	0.90	0.14	0.81	0	1	0	1	TTB-C
4	0.90	0.00	0.71	0.44	1	0	G	G	Pros	0.80	0.20	0.56	0.29	1	0	G	G	Pros
5	0.00	1.00	0.36	0.81	0	1	G	1	X	0.00	0.80	0.13	1.00	0	1	0	1	TTB-C
6	0.30	0.60	0.31	0.29	G	G	G	G	Guess	0.40	0.60	0.50	0.56	G	G	G	G	Guess
7	0.00	0.90	0.19	0.71	0	1	0	G	X	0.10	0.90	0.43	0.81	0	1	G	1	X
8	0.00	1.00	0.21	0.75	0	1	0	1	TTB-C	0.00	1.00	0.13	0.71	0	1	0	G	X
9	0.00	1.00	0.00	0.69	0	1	0	G	X	1.00	0.00	0.31	0.64	1	0	G	G	Pros
10	0.00	0.90	0.13	1.00	0	1	0	1	TTB-C	0.00	1.00	0.00	0.94	0	1	0	1	TTB-C
11	1.00	0.00	0.38	0.57	1	0	G	G	Pros	0.00	1.00	0.07	0.88	0	1	0	1	TTB-C
12	0.50	0.50	0.36	0.69	G	G	G	G	Guess	1.00	0.00	0.63	0.29	1	0	G	G	Pros
13	0.00	1.00	0.07	1.00	0	1	0	1	TTB-C	0.50	0.50	0.25	0.64	G	G	0	G	Guess
14	0.90	0.00	0.25	0.57	1	0	0	G	Pros	0.60	0.60	0.64	0.44	G	G	G	G	Guess
15	0.00	0.90	0.06	0.86	0	1	0	1	TTB-C	1.00	0.00	0.29	0.69	1	0	G	G	Pros
16	0.00	1.00	0.00	1.00	0	1	0	1	TTB-C	0.00	1.00	0.00	1.00	0	1	0	1	TTB-C
17	0.00	1.00	0.14	0.94	0	1	0	1	TTB-C	0.00	1.00	0.00	1.00	0	1	0	1	TTB-C
18	1.00	0.00	0.69	0.29	1	0	G	G	Pros	0.90	0.20	0.71	0.38	1	0	G	G	Pros
19	0.00	1.00	0.19	0.79	0	1	0	1	TTB-C	0.00	0.90	0.14	0.88	0	1	0	1	TTB-C
20	0.30	0.80	0.64	0.38	G	1	G	G	X	0.60	0.70	0.31	0.36	G	G	G	G	Guess
21	0.00	0.90	0.21	0.81	0	1	0	1	TTB-C	0.00	1.00	0.00	1.00	0	1	0	1	TTB-C
22	0.90	0.00	0.31	0.57	1	0	G	G	Pros	1.00	0.00	0.64	0.63	1	0	G	G	Pros
23	1.00	0.00	0.38	0.71	1	0	G	G	Pros	0.50	0.60	0.50	0.75	G	G	G	1	Guess
24	0.56	0.80	0.43	0.81	G	1	G	1	X	0.00	1.00	0.00	1.00	0	1	0	1	TTB-C

Note: G = Guess, X = Unclassifiable

Response Time

Response times were measured from the time at which the subject pressed the "Set Leader" button (which called up the Friend/Foe judgment box) to the time he/she indicated a threat classification and pressed the Return key on the computer keyboard. Note that this time includes the time to inspect cues (which was limited by the Time Pressure condition) and the time to indicate a response. Although no predictions concerning response times were drawn from the decision strategies under consideration, mean response times were computed for subjects. Generally, subjects took a fair amount of time, on the order of 5.5 to 7.5 seconds, to indicate their decisions.

A two-factor within-subjects ANOVA revealed no significant effect of Item Type [$F(1,23) = 0.95$, MSe = 0.14, *n.s.*], but subjects did respond significantly faster on average in the High Time Pressure than Low Time Pressure condition [$F(1,23) = 15.09$, MSe = 91.14, $p<.05$], indicating that the Time Pressure manipulation did affect how subjects performed the task. The interaction of Item Type and Time Pressure proved not significant [$F(3,69) = 1.63$, MSe = 0.27, *n.s.*].

Cue Selections

When a subject clicked on a cue button in the test session to inspect the value of that cue, that action was recorded as a "cue selection." All such cue selections were recorded for every test item to determine which cues subjects inspected and the order in which they inspected them. The hypothesized strategies make different predictions concerning the order in which subjects should select cues. TTB-C predicts that subjects will select the highest validity cue first and rarely, if ever, inspect other cues. Although the strategy requires no further selections, subsequent selections might occur in descending order of cue validity. The Bayesian and Pros Rule inspect all cues so there is no need to inspect cues in a particular order. Given that the order of cues on the display was random for each contact, these strategies predict essentially random cue selection patterns, although they do not rule out the use of a consistent search order by individuals.

Table 7: *Average Number of Cue Selections by Decision Strategy (Experiment 1)*

		Decision Strategy Assigned	
Time Pressure	**Overall**	**TTB-C**	**Pros**
Low	3.93	3.72	4.01
High	3.24	2.92	3.48

Table 7 shows the average number of cue selections made by subjects overall and for the subsets of subjects assigned to the TTB-C and Pros Rule. Although each contact possessed only four cues, subjects were able to inspect each cue multiple times, so that for any given contact there could be more than four cue selections (i.e. one or more cues was repeated). Subjects inspected slightly fewer cues on average in the High than Low Time Pressure conditions [$t(23) = 2.19$, $p<.05$]. In both conditions, however, the mean number of selections was well above the single selection predicted by the TTB-C strategy.

To determine whether the subset of subjects identified as using TTB-C did, in fact, select a single cue as predicted, separate mean number of cue selections was calculated for conditions in which subjects were inferred to have employed TTB-C or the Pros Rule on then basis of subjects' responses to test items. In both the Low [$t(23) = 1.67$, *n.s.*] and High [$t(23) = 1.75$, *n.s.*] Time Pressure conditions, mean number of selections did not differ between TTB-C and Pros Rule users. Even those subjects classified as users of TTB-C according to their responses generally selected more than one cue to inspect. Thus, these subjects did not use TTB-C in the strict sense that cue search terminates after a single cue is found. They did, however, use the one-reason decision rule that is the basis of TTB-C, that the single best cue determines the threat class assigned regardless of what other cues indicate.

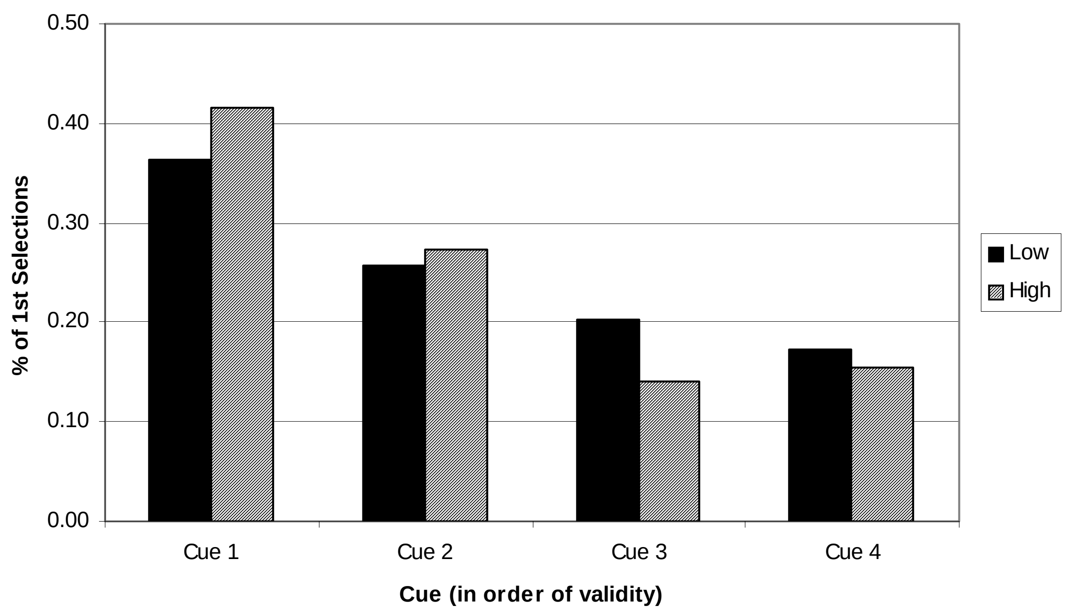

Figure 4: *Average Number of Cues Selected by Sequence (Experiment 1)*

Subjects were free to inspect cues in any order they chose but all subjects exhibited a tendency to inspect the most valid cue first. To examine this, we identified for each item which cue was inspected first (First Selection). Figure 4 shows the proportion of items for which the subject selected each cue first, where the cues are indicated in descending order of validity. As can be seen, subjects selected the most predictive cue first more frequently than the next most predictive cue. A repeated measures ANOVA revealed a significant effect of cue validity on First Selections [$F(3,138) = 7.77$, MSe = 2.13, $p<.01$] but no effects of either Contact Type (A, B, C, or D) [$F(3,138) = 0.77$, MSe = 0.001, *n.s.*] or Time Pressure [$F(3,138) = 0.73$, MSe = 0.003, *n.s.*]. No interaction effect was significant. Thus, in all cases subjects were most likely to select the most predictive cue first, followed by the second most valid cue. There was no appreciable difference in the proportion of First Selections that were the third- and fourth-most valid cues.

Discussion

The results provide some evidence of the use of fast and frugal heuristics in the threat classification task. A subset of subjects appeared to use TTB-C, as evidenced by their responses to test items, although they did not engage in frugal search, even in the high time pressure condition. They did, however, show a tendency to select the best cue first. Other subjects used the Pros Rule, which employs all cues and is not as frugal as TTB-C, suggesting that some subjects were willing to give up some speed and frugality to make use of potentially important cue information. No subject, however, exhibited the pattern of responses predicted by the Bayesian strategy.

There were no apparent differences in heuristic use between the high and low time pressure conditions. It may have been that the time pressure was not severe enough; subjects did have enough time to inspect multiple cues in the high time pressure condition. A more extreme manipulation may result in greater use of TTB-C. That said, making the time pressure too severe undermines the value of contrasting different heuristics. If time pressure is made so severe that subjects could only inspect one cue in the time given, then they would be forced into using TTB-C rather than exhibiting a preferred strategy.

Experiment 2

The purpose of the second experiment was, first, to determine whether the use of heuristics could be replicated. Thus, subjects performed the same task under similar conditions. To enhance the power of the methodology to discriminate the kind of heuristic used by subjects, two additional contact types were added to the test sets to distinguish predictions of the Weighted and Unweighted versions of the Pros Rule.

A second objective of the experiment was to examine the effects of information uncertainty on decision strategy and cue use. In this context, information uncertainty refers to the probability that available cue information does not correctly reflect the true value of the cue. Manipulations of cue uncertainty affect the value of cue information, which has been shown to affect subjects' use of a fast and frugal heuristic [10]. More specifically, when uncertainty is higher, there is less value in inspecting and using multiple cues. The greater the uncertainty of information, the greater is the incentive to use a simple heuristic that bases decisions on a limited subset of cue information, hence limiting the extent to which inaccurate information affects the decision. Subjects performed the threat classification task in the test session under two conditions of uncertainty, in which the uncertainty associated with a cue increased each time the subject selected a cue to inspect for a given contact. In one case, uncertainty followed a linear pattern that gradually increased uncertainty with each cue selection, whereas in the other case it followed an exponential pattern that increased uncertainty very little for the first few cue selections but dramatically for the fourth cue selection.

Method

Subjects

Subjects were 24 men and women who received payment in exchange for participation and met all the criteria indicated in Experiment 1.

Materials

The experiment was conducted using the TITAN platform. Two sets of 300 contacts (150 friend and 150 foe) were created for the training sessions and two sets of 50 contacts were created for the test sessions. These sets were counterbalanced with the Uncertainty condition.

Design

We manipulated three variables in this experiment. The first, varied within subjects, was the Cue Validity as in Experiment 1. Table 8 indicates the proportions of friend and foe contacts possessing each cue value for the four cues in the two training sets.

Table 8: *Relative Frequencies of Cue Values for Friend and Foe Contacts (Experiment 2)*

	SET 1							
	Cue 1 (Signal Strength)		***Cue 2 (Direction of Origin)***		***Cue 3 (Counter Measures)***		***Cue 4 (Response)***	
	Value 1 (Medium)	***Value 2 (High)***	***Value 1 (B. Lag.)***	***Value 2 (Red Sea)***	***Value 1 (None)***	***Value 2 (Jamming)***	***Value 1 (No Resp)***	***Value 2 (Given)***
Friend	60%	40%	30%	70%	20%	80%	90%	10%
Foe	40%	60%	70%	30%	80%	20%	10%	90%
	SET 2							
	Cue 1 (Intelligence)		***Cue 2 (Manoeuvre Pattern)***		***Cue 3 (Comm. Mode)***		***Cue 4 (Electronic Warfare)***	
	Value 1 (Private)	***Value 2 (Platform)***	***Value 1 (Delta)***	***Value 2 (Foxtrot)***	***Value 1 (1)***	***Value 2 (3)***	***Value 1 (Big B)***	***Value 2 (None)***
Friend	60%	40%	30%	70%	20%	80%	90%	10%
Foe	40%	60%	70%	30%	80%	20%	10%	90%

Note: B. Lag. = Blue Lagoon; Big B = Big Bulge

The second variable, Contact Type in the test stimuli, was manipulated within subjects. As in Experiment 1, the test sets contained patterns that offered contrasting predictions of contending classification strategies. This time, however, both sets also contained items intended to distinguish between the Weighted Pros Rule and Unweighted Pros Rule versions.

Table 9: *Predicted Responses to Contact Types by Hypothesized Strategies (Experiment 2)*

	Predicted Response of Strategy				**Contact Types**	
Cue Pattern	**TTB-C**	**Bayesian**	**Franklin's**	**Dawes'**	**Foe**	**Friend**
1,2,2,2	Foe	Friend	Friend	Friend	B	A
2,1,1,1	Friend	Foe	Foe	Foe	A	B
1,1,1,1	Foe	Foe	Guess	Guess	D	C
2,2,2,2	Friend	Friend	Guess	Guess	C	D
1,1,2,2	Friend	Friend	Friend	Guess	F	E
1,2,1,2	Friend	Friend	Friend	Guess	F	E
2,1,2,1	Foe	Foe	Foe	Guess	E	F
2,2,1,1	Foe	Foe	Foe	Guess	E	F

Note: Cue pattern indicates the value (as 1 or 2) for each cue in order of cues listed in Table 8

The critical test patterns are listed in Table 9 with the predicted responses of the four decision rules. In the test set, each of the critical patterns was paired an equal number of times with friend and foe contacts.

From these contacts, we created six Contact Types (A, B, C, D, E, and F) and predicted the levels of accuracy predicted by the hypothesized decision procedures, shown in Table 10. Types E and F contacts distinguish the two forms of the Pros strategies.

Table 10: *Predicted Accuracy Levels by Contact Type (Experiment 2)*

	Contact Type*					
Heuristic	**A**	**B**	**C**	**D**	**E**	**F**
TTB-C	0%	100%	0%	100%	0%	100%
Bayesian	100%	0%	0%	100%	0%	100%
Weighted Pros Rule	100%	0%	Guess	Guess	0%	100%
Unweighted Pros Rule	100%	0%	Guess	Guess	Guess	Guess

* See Table 9

The final variable, also varied within subject, was the Perceived Cue Uncertainty during the test session (Linear vs. Exponential). In both conditions, subjects were told that each time a cue selection was made there would be a chance that, instead of simply showing the true cue value, the interface would randomly select value 1 or 2 for that cue (in which case the subject would be receiving no useful diagnostic information). Moreover, subjects were told that the probability of this would increase with each cue selection made. Subjects were shown a matrix (see Table 11) that indicated the probabilities of randomizing the cue value shown as a function of cue selection sequence. The Linear Uncertainty condition featured steadily increasing probability of random error in cue value, whereas the Exponential Uncertainty condition featured probability of random error that increased very little initially, then increased dramatically for the fourth selection. The progressive nature of the error matrix (i.e. that the probability of error increased with each cue selection) was intended to reduce the value of each subsequent cue inspected, which should reduce subjects' incentive to continue a cue search.

Table 11: *Uncertainty Matrix Shown to Subjects Prior to Test Session*

	ORDER OF CUE SELECTION			
CONDITION	**1st**	**2nd**	**3rd**	**4th**
Linear Uncertainty	10%	15%	20%	25%
Exponential Uncertainty	3%	6%	12%	48%

Note: In both conditions, average error is roughly 17.5%

Note that the random error matrix was used only to instill in subjects a belief in the likelihood that cue information is uncertain; no randomization was actually applied to cues. To assess the decision strategies used by subjects it is necessary that the cue patterns remain stable. Without feedback, subjects had no way of determining whether any random error had actually been applied.

Subjects were required to make judgments within a time limit of 14 seconds that corresponds to the low time pressure condition of Experiment 1.

Procedure

The experiment was divided into two sessions, one for the Linear Uncertainty and the other for the Exponential Uncertainty conditions, each with a training and test phase. In the training phase, subjects made classification judgments for 300 contacts (150 friends and 150 foes) presented in the same manner as Experiment 1. Each contact had four cues associated with it, specifying cue values generated according to the probability matrix shown in Table 7. Subjects selected contacts, inspected cue values, and made threat classification judgments in the same way as the first experiment.

Subjects received 62 contacts in the test sessions (10 each of types A, B, C, D, E, and F plus two non-critical contacts not considered in the data analysis) and they did not receive feedback on the accuracy of their judgments. Subjects were under either Linear or Exponential Cue Uncertainty instructions when making their judgments. A clock display counted down the seconds remaining so subjects could monitor how much time remained to make each judgment.

Following the training and test phases, subjects were asked to indicate the diagnosticities of each cue as a predictor of contact classification by judging the proportions of cue values associated with friend and foe contacts and to list the cues in order of their diagnosticity.

Results

Training Session

The contacts presented during the training session were divided into six blocks of 50 contacts each, based on the order of presentation, and mean accuracy scores were calculated (see Figure 5). A two-way, within-subjects design Analysis of Variance (ANOVA) revealed a significant effect of Training Block [$F(5,115) = 5.05$, MSe = 0.003, $p < .05$] but no significant main effect of Perceived Cue Uncertainty [$F(1,23) = 0.03$, MSe = 0.01, *n.s.*]. There was also no significant interaction effect between the two factors [$F(5,115) = 1.03$, MSe = 0.003, *n.s.*]. As in Experiment 1, subjects learned how to classify friends and foes more accurately through trial and error training. Subjects' performance in the final block was near, but not quite, optimal with respect to both the Bayesian and TTB-C strategies. The mean Block 6 accuracy, combined across Linear and Exponential conditions, was .87 and significantly less than both the optimum of .90 predicted by TTB-C [$t(23) = 4.04$, $p<.01$] and the optimum of .94 predicted by the Bayesian strategy [$t(23) = 9.52$, $p<.01$].

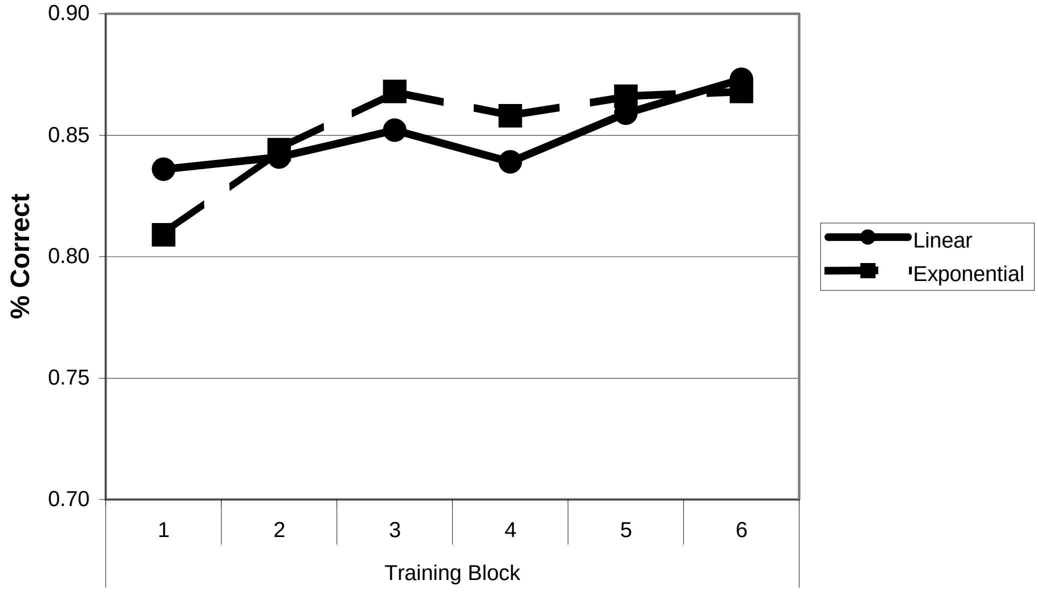

__Figure 5:__ Classification Accuracy by Block in the Training Session (Experiment 2)

Subjects' responses to the post-experimental survey were again used to assess subjects' learning of cue diagnosticities. Error scores, calculated by subtracting the actual cue value percentage from each subject's estimate for each cue value and computing the average difference or error score across the four cues, are shown in Table 12. The absolute mean error in the Linear and Exponential Cue Uncertainty conditions, which shared the same training procedures, did not differ significantly [$t(23) = 0.25$, *n.s.*]. Similarly, there was no reliable difference in the standard deviations of the error [$t(23) = 0.14$, *n.s.*]. The overall mean error score (combined for Linear and Exponential Cue Uncertainty Pressure conditions) was quite low and did not differ significantly from zero [$t(23) = 1.20$, *n.s.*]. Subjects' estimates of cue validities were highly variable and the overall mean standard deviation of the error was significantly greater than zero [$t(23) = 12.63$, $p < .01$]. As in Experiment 1, subjects' errors averaged to near-zero but their individual judgments of cue validities often exhibited substantial error. Subjects' error in listing the cues in order of cue validity was greater than zero [$t(23) = 5.57$, $p < .01$].

__Table 12:__ Mean Error and Mean Standard Deviation of the Error of Cue Validity Estimates (Experiment 2)

Perceived Cue Uncertainty	Mean Error (%)	SD of Error (%)
Linear	-1.91	22.86
Exponential	-2.48	22.49

Note: Negative error scores indicate underestimation of true cue value association probabilities.

Classification Strategy

To assess strategy use by subjects, we examined each subject's proportion of correct responses for each pattern type (A, B, C, D, E, and F) and compared this to the predicted accuracy levels listed in Table 10. As before, a subject was considered to have performed with 100% accuracy for a pattern type if that subject responded correctly to that pattern more often than would be expected by chance (assumed to be 5%). Table 13 presents subjects' accuracy scores for each contact type and the decision strategy assigned to each subject. Subjects exhibited a high degree of consistency in strategy use.

Table 14: *Number of Subjects Classified as Using Hypothesized Decision Strategies (Experiment 2)*

Perceived Cue Uncertainty	Decision Heuristic				
	TTB-C	**Dawes'**	**Franklin's**	**Guessing**	**Unclassifiable**
Linear	11	6	1	0	6
Exponential	11	4	0	4	5

* N = 24 at each level of Perceived Cue Uncertainty

Table 14 presents the number of subjects classified as using a given decision strategy. As in Experiment 1, no subject exhibited the pattern predicted by the Bayesian strategy. Rather, the majority of subjects' patterns matched predictions of either the TTB-C or a Pros Rule, although a sizeable portion seemed to be guessing or using some unclassified strategy. In all but one case where subjects consistently employed a Pros Rule, the rule used was the Unweighted Pros Rule. Unlike Experiment 1, however, more subjects employed TTB-C than a Pros Rule.

Exactly the same number of subjects were classified as TTB-C users in the Linear and Exponential uncertainty conditions. Three more subjects used a Pros Rule in the Linear than Exponential condition, although four subjects seemed to be guessing in the Exponential condition compared to none in the Linear condition. A Pearson Chi-Square test revealed no significant difference between the two conditions [$\chi 2 = 5.49$, df = 4, n.s.].

Response Time

Response times were again measured from the time at which the subject pressed the "Set Leader" button to the time he/she indicated a threat classification and pressed the Return key. Generally, subjects took a fair amount of time, on the order of 7 to 7.75 seconds, to indicate their decisions. A two-factor within-subjects ANOVA revealed no significant effect Perceived Cue Uncertainty [$F(1,23) = 0.45$, MSe = 4.12, *n.s.*]. The analysis did reveal, however, significant effects of Item Type [$F(5,115) = 2.94$, MSe = 0.43, $p<.05$], which seems to reflect somewhat lower response times for item type F and perhaps C and E than all other item types. There is really no reason to expect such a difference and the effect is likely an artefact. The ANOVA revealed no significant interaction of Item Type and Perceived Cue Uncertainty [$F(5,115) = 1.67$, MSe = 0.45, *n.s.*].

Table 13: *Response Accuracies by Contact Type and Decision Strategy Classification (Experiment 2)*

	Perceived Linear Error													**Perceived Exponential Error**													
	True Accuracy						**Interpolated Accuracy**						**Decision Strategy**	**True Accuracy**						**Interpolated Accuracy**						**Decision Strategy**
Sub.	**A**	**B**	**C**	**D**	**E**	**F**	**A**	**B**	**C**	**D**	**E**	**F**		**A**	**B**	**C**	**D**	**E**	**F**	**A**	**B**	**C**	**D**	**E**	**F**	
1	0.1	0.9	0.1	1	0.3	0.8	0	1	0	1	G	1	X	0	0.9	0	1	0	1	0	1	0	1	0	1	TTB-C
2	1	0	0.7	0.3	0.6	0.4	1	0	G	G	G	G	DAWES	0.9	0.4	0	1	0.6	0.2	1	G	0	1	G	0	X
3	1	0	0.4	0.7	0.3	0.7	1	0	G	G	G	G	DAWES	0.9	0.1	0.5	0.4	0.3	0.7	1	0	G	G	G	G	DAWES
4	0.8	0.2	0.5	0.5	0.3	0.7	1	0	G	G	G	G	DAWES	0.5	0.9	0.8	0.1	0	1	G	1	1	0	0	1	X
5	1	0	1	0	0.4	0.4	1	0	1	0	G	G	X	0	1	0	1	0	1	0	1	0	1	0	1	TTB-C
6	0.1	1	0	1	0	1	0	1	0	1	0	1	TTB-C	1	0	0.4	0.4	0.5	0.5	1	0	G	G	G	G	DAWES
7	0	0.8	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C
8	0.6	0.4	0.6	0.2	0.2	0.8	G	G	G	0	0	1	X	0.5	0.7	0.3	0.9	0	0.7	G	G	G	1	0	G	X
9	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C	0	1	0	1	0	1	0	1	0	1	0	1	TTB-C
10	0	0.6	0.1	0.9	0.1	0.9	0	G	0	1	0	1	X	0.7	0.4	0.4	0.7	0.3	0.6	G	G	G	G	G	G	GUESS
11	0.9	0.1	0.4	0.7	0.3	0.6	1	0	G	G	G	G	DAWES	0.2	0.8	0.1	1	0.2	0.8	0	1	0	1	0	1	TTB-C
12	0.2	0.9	0.1	1	0.2	0.8	0	1	0	1	0	1	TTB-C	0.6	0.3	0.4	0.6	0.2	0.7	G	G	G	G	0	G	GUESS
13	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C	0	1	0	1	0	1	0	1	0	1	0	1	TTB-C
14	1	0.1	0.6	0.5	0.3	0.8	1	0	G	G	G	1	DAWES	0.3	0.5	0	1	0.5	0.8	G	G	0	1	G	1	X
15	0	1	0.1	1	0.1	1	0	1	0	1	0	1	TTB-C	0.2	0.7	0	1	0.4	0.7	0	G	0	1	G	G	X
16	1	0	0.6	0.3	0	1	1	0	G	G	0	1	FRANK	1	0	0.4	0.6	0.5	0.5	1	0	G	G	G	G	DAWES
17	0.1	0.9	0.3	0.9	0.3	0.7	0	1	G	1	G	G	X	0.1	0.8	0	1	0.2	0.9	0	1	0	1	0	1	TTB-C
18	0.9	0.1	1	0.1	0	1	1	0	1	0	0	1	X	1	0.1	0.7	0.4	0.3	0.6	1	0	G	G	G	G	DAWES
19	1	0	0.6	0.4	0.6	0.5	1	0	G	G	G	G	DAWES	0.3	0.4	0.3	0.6	0.3	0.7	G	G	G	G	G	G	GUESS
20	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C	0.6	0.4	0.3	0.6	0.4	0.6	G	G	G	G	G	G	GUESS
21	0	1	0	0.9	0.2	0.8	0	1	0	1	0	1	TTB-C	0.1	1	0	0.9	0	1	0	1	0	1	0	1	TTB-C
22	0	1	0	1	0	1	0	1	0	1	0	1	TTB-C	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C
23	0.2	0.8	0	1	0.2	0.9	0	1	0	1	0	1	TTB-C	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C
24	0	1	0	1	0.2	0.8	0	1	0	1	0	1	TTB-C	0.1	0.9	0	1	0	0.9	0	1	0	1	0	1	TTB-C

Note: G = Guess, X = Unclassifiable

Cue Selections

Table 15 shows the average number of cue selections made by subjects overall and for the subsets of subjects assigned to the TTB-C and a Pros Rule (all but one of the subjects in this subgroup were classified as using the Unweighted Pros Rule). Subjects inspected slightly fewer cues on average in the Exponential than Linear Cue Uncertainty condition but this difference did not achieve statistical significance [$t(23) = 1.95$, *n.s.*]. In both conditions, the mean number of selections was well above the single selection predicted by the TTB-C strategy. Subjects classified as TTB-C users selected fewer cues overall than those classified as Unweighted Pros Rule or Weighted Pros Rule users but these differences did not achieve statistical significance in either the Linear [$t(16) = 1.55$, *n.s.*] or Exponential [$t(13) = 2.08$, *n.s.*] Cue Uncertainty conditions. Even those subjects who seemed to use TTB-C according to their responses generally selected more than one cue to inspect.

Table 15: *Average Number of Cue Selections by Decision Strategy (Experiment 2)*

		Decision Strategy Assigned	
Perceive Cue Uncertainty	**Overall**	**TTB-C**	**Pros**
Linear	4.16	3.40	4.66
Exponential	3.75	2.70	4.54

As in the previous experiment, subjects exhibited a tendency to inspect the most valid cue first. Figure 6 shows the proportion of items for which the subject selected each cue first, where the cues are indicated in descending order of validity. A repeated measures ANOVA revealed a significant effect of cue validity on First Selections [$F(3,69) = 20.80$, $MSe = 0.64$, $p<.01$], confirming that subjects were reliably more likely to selected the most valid cue first than any other. Unlike the finding of Experiment 1, the ANOVA of the current data revealed a significant interaction of Cue Validity with Item Type (A, B, C, D, E, F) [$F(3,138) = 0.77$, $MSe = 0.001$, *n.s.*] or Time Pressure [$F(15,345) = 1.70$, $MSe = 0.03$, $p<.05$]. There was also a significant three-way interaction of Cue Validity, Item Type, and Perceived Cue Uncertainty [$F(15,345) = 2.20$, $MSe = 0.03$, $p<.01$]. These interaction effects reflect variation in the extent to which the most valid cue was preferred as a first selection. For some combinations of the three factors, subjects exhibited a relatively smaller or larger percentage of first selections that were the most valid cue. In all cases, however, the most valid cue was selected first by a wide margin over all other cues. No other effects were found to be significant.

Discussion

As in the first experiment, most subjects used either TTB-C or a Pros Rule to perform the threat assessment task. Similarly, no subject was classified as using the Bayesian strategy. All but one subject classified as a user of the Pros Rule were classified as using the Unweighted Pros Rule. Subjects' preference for the simpler version may have been due to their less than perfect learning of cue validities, which would have undermined the accuracy of any weighted procedure. Like those in Experiment 1, subjects who were classified as users of TTB-C on the basis of their responses to test items tended to select more cues to inspect than needed by that heuristic.

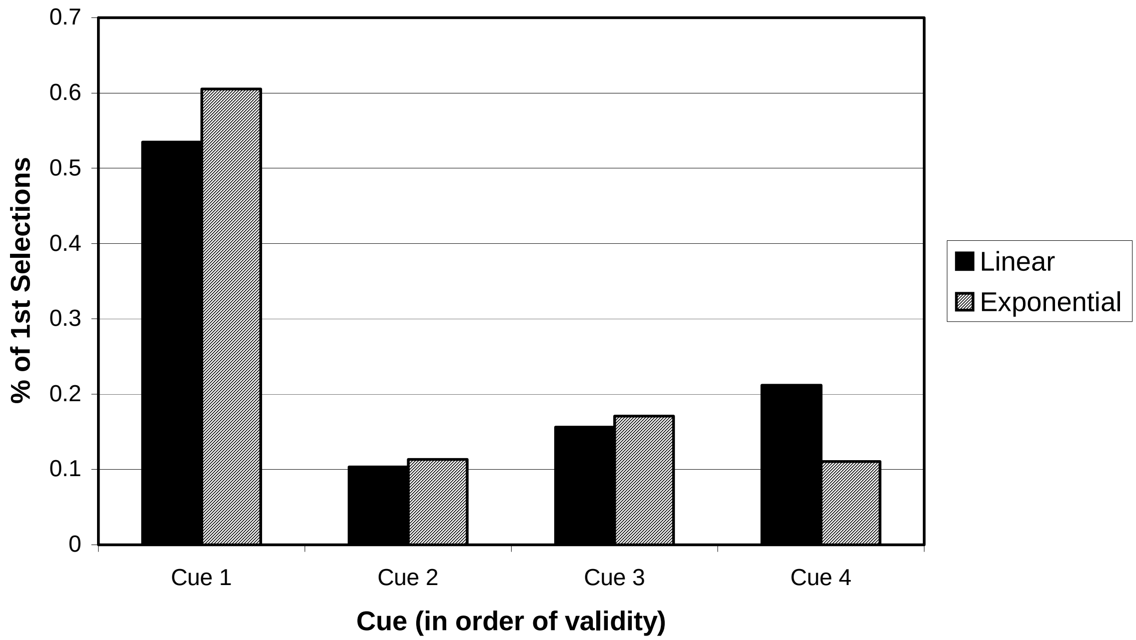

Figure 6: *Average Number of Cues Selected by Sequence (Experiment 2)*

There was no apparent effect of the cue uncertainty manipulation. It was expected that introducing uncertainty would induce subjects to inspect fewer cues and increase the use of TTB-C. Nevertheless, the same proportions of subjects used TTB-C and The Unweighted Pros Rule in both conditions and the numbers of TTB-C users were not dramatically larger than those seen in Experiment 1. Again, the manipulation may not have been extreme enough to induce subjects to move toward one or another procedure. It may be, however, that individual differences may play a more important role in selecting a preferred strategy for performing the task (see [10]).

Conclusion

A number of studies have found that some subjects employ fast and frugal heuristics to perform choice tasks (e.g., [6] [7] [8] [9]). The results of two experiments extended these findings to the threat classification task. Threat classification differs from previous choice tasks in that subjects learn to use cues to assign an entity to one of two classes rather than chose an alternative that is superior along some associated dimension. In both cases, however, subjects must deal with uncertainty and employ some procedure to assess what is most likely the correct response.

As in previous studies, only a subset of subjects employed the simplest possible heuristic, which in the case of threat assessment was dubbed TTB-C. Like its progenitor, "Take-the-Best," used in choice tasks, TTB-C relies on a single cue to make a decision and is, thus, easy to use. In the design of the current task, TTB-C was also fairly effective, capable of achieving a level of performance only somewhat lower than a strategy based on calculating the conditional probabilities of friend or foe. Nevertheless, a large proportion of subjects chose to employ a somewhat more complicated compensatory heuristic, the Unweighted Pros Rule. This heuristic uses all available cues to accumulate evidence of the most likely classification. It is not as frugal as TTB-C but does not require the computation of the Bayesian strategy. Interestingly, it is also generally a little less accurate overall than TTB-C for the task structure used in these experiments [30].

Neither a manipulation of time pressure nor of perceived cue uncertainty had any apparent effect on subjects' heuristic use. For example, essentially the same proportions of subjects used TTB-C in the High and Low time pressure conditions, although time pressure is one factor often cited as a rationale for fast and frugal heuristics (e.g., [22]). It is possible that the manipulations employed here were not strong enough to affect subjects. It is also possible that individual factors play a dominant role in determining which procedure individuals use. Broder [10] has observed that some individual factors such as intelligence do affect the use of TTB in choice tasks.

The two experiments introduced a new approach to identifying individual subjects' decision strategies. Assigning a decision strategy to a given subject on the basis of his or her behavior can be problematic because it is the invisible, underlying process itself that is of most interest rather than the accuracy of the outcome. In tasks such as two-item choice and threat assessment, multiple procedures predict the same response for a large proportion of items. Other researchers have analysed subjects' selection of cue information (e.g., [9]) and developed statistical methods to categorize subjects' patterns of responding [7]. The approach used in the two experiments was to create a test set that was over-representative of relatively rare items that elicited contrasting predictions from hypothesized decision procedures. By assessing the consistency with which subjects responded one way or another to items, it was possible to assign a decision procedure.

A potential problem of this methodology is that it assumes that inconsistency in responding to a given item type derives from random error. It is not known for certain whether subjects might ever chose one procedure initially then intentionally employ a different procedure.

Certainly, subjects could switch decision strategy from item to item within the test set, which would mask the nature of subjects' decision procedures. Another problem arises when assessing heuristics that predict guessing for certain items (i.e. the Pros Rule). It was assumed that guessing would not yield any systematic preference for either classification. To operationalize this assumption for individual patterns of responding, it was necessary to set an arbitrary probability level for assessing whether a subject's pattern of responding to a type of item consistently favoured one or the other classification or not. Setting that level at .05 seemed reasonable but leaves a wide range of individual response patterns that are consistent with a guessing strategy. Moreover, the presence of a consistent tendency to respond one way for a type of item does not absolutely rule out the possibility of guessing. A subject could randomly select a response for the first instance of a test item then stick with that response for all subsequent instances. Despite guessing, the subject would appear to exhibit a consistent preference for the initial response.

The results of the two experiments suggest that it is possible to model multi-attribute decision tasks like threat assessment with fast and frugal heuristics. It is clear, however, that no single fast and frugal heuristic has proven to be a general model for the simulated threat assessment task, as there are individual differences in the choice of decision strategy as well as the consistency with which individuals employ a strategy. TTB-C, which was chosen as the simplest starting point, is one possible heuristic but a compensatory heuristic is also applicable and may be more valid in real-world settings (cf. [18]). The next step in this research is to characterize the underlying task ecology of threat assessment in operational settings and characterize the efficiency demands most relevant to the task.

References

1. Gigerenzer, G., Todd, P. M. & The ABC Research Group. (1999). Simple heuristics that make us smart. Oxford, UK: Oxford University Press.

2. Gigerenzer, G. & Selten, R. (2001). Rethinking rationality. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox*, pp. 1-12. Cambridge, MA: The MIT Press.

3. Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650-669.

4. Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart*, pp. 97-118. Oxford, UK: Oxford University Press.

5. Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart*, pp. 119-140. Oxford, UK: Oxford University Press.

6. Broder, A. (2000). Assessing the empirical validity of the "Take the Best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1332-1346.

7. Broder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making, 16*, 193-213.

8. Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes, 91*, 82-96.

9. Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 53-65.

10. Broder, A. (2003). Decision making with the "Adaptive Toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 611-625.

11. Blackmond Laskey, K., D'Ambrosio, B., Levitt, T. S., & Mahoney, S. (2000). Limited rationality in action: Decision Support for military situation assessment. *Minds and Machines*, *10,* 53-77.

12. Klein, G. (1992). *Decision making in complex military environments*. Klein Associates, Fairborn, OH: Final Contract Summary Report, prepared for Naval Command, Control and Ocean Surveillance Center, San Diego, CA.

13. Klein, G. (1997). An overview of naturalistic decision making applications. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making* (pp. 49-59). Lawrence Erlbaum Associates Inc., Mahwah, NJ.

14. Todd, P. M., & Gigerenzer G. (2000). Putting naturalistic decision making into the Adaptive Toolbox. *Journal of Behavioral Decision Making, 14*, 353-384.

15. Chalmers, B. A. (1998). On the design of computer-based C2 decision support for a modern frigate. *Proceedings of the 4th International Command and Control Research and Technology Symposium*. U.S. Department of Defense, C4ISR Cooperative Research Program. Nasby Park, Sweden.

16. Chalmers, B. A., Easter, J. R., & Potter, S. S. (2000). Decision-centered visualizations for tactical decision support on a modern frigate. *Proceedings of 2000 Command and Control Research and Technology Symposium* [CD-ROM]. Naval Postgraduate School, Monterey, CA.

17. Liebhaber, M. J., & Feher, B. A. (2002). *Surface warfare threat assessment: Requirements definition*. Technical Report 1887. SPAWAR Systems Center, San Diego, CA.

18. Liebhaber, M. J., Kobus, D. A., & Feher, B. A. (2002). *Studies of U.S. Navy air defense threat assessment: Cues, information order, and impact of conflicting data*. Technical Report 1888. SPAWAR Systems Center, San Diego, CA.

19. Liebhaber, M. J., & Smith, C. A. P. (1999). Tactical Decision-Making Under Stress: Baseline results." Cited in Liebhaber, Kobus, & Feher (2002).

20. Brehmer, B. (1973a). Single-cue probability learning as a function of the sign and magnitude of the correlation between cue and criterion. *Organizational Behavior and Human Decision Processes, 9*, 377-395.

21. Himmelfarb, S. (1970). Effects of cue validity differences in weighting information. *Journal of Mathematical Psychology*, *7*, 531-539.

22. Reiskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell? In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart*, pp. 141-167. Oxford, UK: Oxford University Press.

23. Brehmer, B. (1973b). Effects of cue validity on interpersonal learning of inference tasks with linear and nonlinear cues. *American Journal of Psychology, 86*, 29-48.

24. Castellan, N. J. Jr. (1992). Relations between linear models: Implications for the Lens Model. *Organizational Behavior and Human Decision Processes, 51*, 364-381.

25. Brehmer, B., & Johansson, R. (1977). Subjects' ability to use information about cue validity in multiple-cue judgment tasks. (Umea Psychological Reports No. 118). University of Umea.

26. Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1083-1119.

27. Klayman, J. (1984). Learning from feedback in probabilistic environments. *Acta Psychologica, 56*, 81-92.

28. Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty and experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 317-330.

29. Lindberg, L-A., & Brehmer, B. (1977). Effects of task information and active feedback control in inductive inference. (Umea Psychological Reports No. 123). University of Umea, Umea, Sweden.

30. Bryant, D. J. (2003). *Cue validity learning in threat classification judgements*. DRDC Toronto TR 2003-041, Defence R&D Canada – Toronto, Toronto, Canada.

List of symbols/abbreviations/acronyms/initialisms

$\chi 2$	Pearson Chi-Square Test
ANOVA	Analysis of Variance
DRDC	Defence Research and Development Canada
IFF	Interrogate Friend-or-Foe
MS_e	Mean-Squared Error
TITAN	Team and Individual Threat Assessment Network
TTB	Take-the-Best Heuristic
TTB-C	Take-the-Best-for-Classification Heuristic

Distribution list

INTERNAL DISTRIBUTION

DRDC Toronto TR 2005-xxx

1	-	Director General
1	-	Deputy Director General
1	-	Chief Scientist
6	-	Document Library
1	-	David Bryant (author)
1	-	Joseph Baranski
1	-	Carol McCann
1	-	Megan Thompson
1	-	David Smith
1	-	Angela Febbraro
1	-	Donna Pickering
1	-	Anne-Rene Blais
1	-	Don McCreary
1	-	Keith Hendy
1	-	Justin Hollands
1	-	Lochlan Magee
1	-	Stuart Grant
1	-	Peter Kwantes
1	-	Walter Dyck
1	-	Jocelyn Keillor
1	-	Sharon McFadden

EXTERNAL DISTRIBUTION

DRDC Toronto TR 2003-105

1	-	DRD KIM
1	-	DRD KIM (unbound)
1	-	DRDC
3	-	DGOR
4	-	DGJFD
1	-	ST HP
2	-	D ST CCIS
2	-	D ST L
2	-	D ST A
2	-	D ST M
2	-	D MRS
1	-	D MRS 6
1	-	D MRS 6-2
2	-	DAT
2	-	DAD
2	-	DLSC
2	-	DLPS
2	-	D DCEI
1	-	D LCSPM
1	-	DRDB
1	-	DMSS
1	-	DMSS – 6
1	-	DMSS – 8
1	-	DMPPD
1	-	DMPPD 6
1	-	DMPPD 6 –2
1	-	DSTCCIS
1	-	DSTA
1	-	DSTM
1	-	DSTM – 2

2	-	DLR
2	-	CFEC
2	-	DRDC Atlantic Attn: B. McArthur B. Chalmers
5	-	DRDC Valcartier Attn: R. Breton M. Belanger S. Paradis E. Bosse J.M.J. Roy P. Labbe
1	-	Canadian Forces Command and Staff College Toronto Attn: Commanding Officer
1	-	CF Maritime Warfare School CFB Halifax Halifax, Nova Scotia Attn: Commanding Officer

DOCUMENT CONTROL DATA

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document, Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Publishing: DRDC Toronto Performing: DRDC Toronto Monitoring: Contracting:	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification is indicated by the appropriate abbreviation (S, C, R, or U) in parenthesis at the end of the title)

Heuristic Classification in a Simulated Air Threat Task (U)

4. AUTHORS (First name, middle initial and last name. If military, show rank, e.g. Maj. John E. Doe.)

David J. Bryant

5. DATE OF PUBLICATION (Month and year of publication of document.) June 2005	6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 52	6b. NO. OF REFS (Total cited in document.) 30

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The names of the department project office or laboratory sponsoring the research and development − include address.)

Sponsoring:

Tasking:

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant under which the document was written. Please specify whether project or grant.) 16kf	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document) DRDC Toronto TR 2005−191	10b. OTHER DOCUMENT NO(s). (Any other numbers under which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILIY (Any limitations on the dissemination of the document, other than those imposed by security classification.)

Unlimited distribution

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11), However, when further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited announcement

DOCUMENT CONTROL DATA

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

(U) Fast and frugal heuristics have proved effective models of human judgment for certain kinds of problems. To further explore the value of this approach, two experiments investigated the decision procedures used by human subjects to perform a cue−based classification task in a simulated air threat assessment task. Threat assessment is the classification of aircraft on the basis of sensor data that can be likened to probabilistic cues. Subjects learned to classify simulated aircraft using four probabilistic cues then classified test sets designed to contrast predictions of several heuristics, including the Take−the−Best−for−Classification (TTB−C) and Pros Rule developed specifically for the threat classification task. Results indicated that a proportion of subjects could be classified using TTB−C and another significant proportion as using the less frugal Pros Rule. No subject was observed to respond as predicted by a Bayesian strategy. Despite predictions that time pressure and perceived uncertainty of cues would affect how many subjects employed TTB−C, no effect of these variables was observed. These results suggest that it is possible to model multi−attribute decision tasks like threat assessment with fast and frugal heuristics but no single heuristic is a general model for the simulated threat assessment task.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

(U) heuristics; threat assessment; probabilistic cues

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca