Journal of Behavioral Decision Making
J. Behav. Dec. Making, **23**: 383–403 (2010)
Published online 27 August 2009 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/bdm.663

Aggregating Conclusive and Inconclusive Information: Data and a Model Based on the Assessment of Threat

JOSEPH V. BARANSKI[1*,†] and WILLIAM M. PETRUSIC[2]
[1]*Defence Research and Development Canada—Toronto,Toronto, Ontario, Canada*
[2]*Department of Psychology, Carleton University, Ottawa, Ontario, Canada*

ABSTRACT

This study examined the process of combining conclusive and inconclusive information using a Naval threat assessment simulation. On each of 36 trials, participants interrogated 10 pieces of information (e.g., speed, direction, bearing, etc.) about ''targets'' in a simulated radar space. The number of hostile, peaceful, and inconclusive cues was factorially varied across targets. Three models were developed to understand how inconclusive information is used in the judgment of threat. According to one model, inconclusive information is ignored and the judgment of threat is based only on the conclusive information. According to a second model, the amount of dominant conclusive information is normalized by all of the available information. Finally, according to a third model, inconclusive information is partitioned under the assumption that it equally represents both dominant and non-dominant evidence. In Experiment 1, the data of novices (i.e., civilians) were best described by a model that assumes a partitioning of inconclusive evidence. This result was replicated in a second experiment involving variation of the global threat context. In a third experiment involving experts (i.e., Canadian Navy officers), the data of half of the participants were best described by the partitioning model and the data of the other half were best described by the normalizing model. In Experiments 1 and 2, the presence of inconclusive information produced a ''dilution effect'', whereby hostile (peaceful) targets were judged as less hostile (peaceful) than the predictions of the Partitioning model. The dilution effect was not evident in the judgments of the Navy officers. © 2009 Crown in the right of Canada.

KEY WORDS inconclusive information; threat assessment; subjective probability; confidence; balance of evidence; dilution effect

INTRODUCTION

The present research is concerned with a specific type of judgment that is based on a known and fixed set of informational cues. Some of the cues may favor a particular decision (e.g., a particular diagnosis, a guilty

* Correspondence to: Joseph V. Baranski, Defence Research and Development Canada—Toronto, 1133 Sheppard Ave. West, P.O. Box 2000, Toronto, Ontario M3M 3B9, Canada. E-mail: joe.baranski@drdc-rddc.gc.ca
†Chief Scientist.

verdict, or a ''hostile'' military target), and some of the cues may favor the alternative decision (e.g., an alternative diagnosis, a not-guilty verdict, or a ''peaceful'' military target). Importantly, in all of these contexts, some of the cues that are relevant to the judgment or assessment may be missing, unavailable, or, more generally, inconclusive[1].

In the absence of normative or prescriptive theory, the broader research topic of human judgment when confronted with missing or inconclusive information has to date been addressed in different ways. For example, several studies have shown that with relatively small cue sets, people will routinely infer missing cue values based on provided (Jagacinski, 1991, 1994, 1995; Körner, Gertzen, Bettinger, & Albert, 2007; Kühberger & Huber, 1998), learned (White & Koehler, 2004), or assumed (e.g., Ebenbach & Moore, 2000; Johnson, 1987) relationships between the missing and available cues. Importantly, in the present studies, missing or inconclusive information is distinguished from related research examining the impact of useless (Bastardi & Shafir, 1998; Windschitl & Chambers, 2004) or irrelevant information (Gaeth & Shanteau, 1984) on judgments, although there is a definite connection here. Specifically, it has been shown that the presence of non-diagnostic information can reduce the impact of diagnostic information, a phenomenon now widely referred to as the ''dilution effect'' (e.g., Nisbett, Zukier, & Lemley, 1981; but see Troutman & Shanteau, 1977, for an earlier demonstration dubbed the ''non-diagnostic effect''). To date, the dilution effect has been demonstrated in a wide range of judgment tasks, including the classic ''beads and jars'' task (LaBella & Koehler, 2004; Troutman & Shanteau, 1977), social judgments (Kemmelmeier, 2004; Nisbett et al., 1981; Peters & Rothbart, 2000; Tetlock & Boettger, 1989; Tetlock, Lerner, & Boettger, 1996), and even expert judgments made by professional auditors (Hackenbrack, 1992; Shelton, 1999; Waller & Zimbelman, 2003).

The problem of how to deal with missing or inconclusive information is a ubiquitous concern in numerous real-world judgment and decision making environments. One domain where this problem is routinely encountered is military threat assessment. Indeed, a major and ongoing concern for military commanders is to assess the potential threat of all entities in their area of operations. Automated decision support systems and data fusion technologies that aid in such tasks seek to provide the decision maker with greater situation awareness and thus some degree of confidence that the overall information package favors either ''hostile'' or ''friendly'' forces (see e.g., Bryant & Webb, 1999; Hutchins, 1996; Roy, Breton, & Paradis, 2001). Inevitably, however, some of the information that is relevant to the judgment of threat may be missing or unavailable at the moment of decision and thus a real challenge for command decision makers is how to deal with such inconclusive information.

A real-world operational context that highlights this judgmental challenge is Naval–Air defense and threat assessment (Liebhaber & Feher, 2002; Liebhaber, Kobus, & Feher, 2002). Scientific interest in this area increased dramatically following the now infamous ''Vincennes tragedy'', when, on 3 July 1988, the USS Vincennes shot down an Iranian airliner, killing all 290 people aboard (see Klein, 1996; Roberts & Dotterway, 1995). Shortly thereafter, the US Office of Naval Research sponsored an ambitious scientific research and development program called Tactical Decision Making Under Stress (TADMUS), which explored various human factors issues related to individual and team decision making under conditions of stress and uncertainty (for a review see Cannon-Bowers & Salas, 1998). Numerous research directions on judgment under uncertainty were identified and advanced by the TADMUS program (see Collyer & Malecki, 1998). However, one aspect of the broader problem that remains in need of further research is the impact of inconclusive information on the assessment of threat. The main objective of the present research was to obtain data on this problem and to develop and test alternative models of combining conclusive and inconclusive information.

[1]Following Slow And Fast Guessing Theory (SFGT: Baranski & Petrusic, 2003; Petrusic, 1992; Petrusic & Jamieson, 1978; Petrusic & Baranski, 1989), we conceptualize ''inconclusive'' information as a non-diagnostic sample of evidence that at the moment of decision favors neither the dominant nor the non-dominant response alternative. Accordingly, throughout this paper we use the terms ''missing'', ''non-diagnostic'', and ''inconclusive'' information, synonymously.

DOI: 10.1002/bdm

MODELS OF USING INCONCLUSIVE INFORMATION

In this section we develop three hypotheses concerning how inconclusive information may be used when forming subjective probability (SP) threat assessments (Baranski & Petrusic, 1999, 2000). In each case, we consider variations on a basic ''*strength*'' model, or ratio rule, which forms the foundation of well-known conceptualizations of SP; i.e., Luce's (1959) choice model, Anderson's (1981) Information Integration theory, and Support theory (Koehler, 1996; Tversky & Koehler, 1994). To illustrate the predictions of the various models, we will consider a hypothetical situation in which a military commander must make a threat assessment for a particular target that has appeared on her radar display. She has access to exactly ten cues[2] about this target (e.g., direction, signal strength, maneuver pattern, etc.). Cues can favor a hostile (H) target, a peaceful (P) target, or cues can be inconclusive (I) (e.g., the data is unavailable, or its value does not clearly favor either hostile or peaceful). Table 1 provides a matrix of cue configurations in which we factorially vary the number (n) of peaceful [n(P)], hostile [n(H)], and inconclusive [n(I)] cues. In this matrix, which provides the stimulus set for the experiments to be reported, the total number of cues is held constant at 10, but the number of peaceful, hostile, and inconclusive cues is systematically varied, producing 36 unique ''targets''.

The ''Ignore'' hypothesis

According to the ''Ignore'' hypothesis, missing or inconclusive information is not considered when forming SP threat assessments. In general terms, this approach would be most consistent with a more formal judgment analysis, such as Bayes' theorem (Edwards, Lindman, & Phillips, 1965; and see Schum & Martin, 1968, for a formal treatment of inconclusive evidence in the context of Bayes' theorem), although the present judgment task does not correspond to traditional probability revision. To illustrate the properties of the Ignore hypothesis in the present context, consider the target in Table 1 comprised of six hostile cues, two peaceful

Table 1. The Number of Peaceful (P), Hostile (H), and Inconclusive (I) cues for the 36 targets in Experiment 1

P	I	H	P	I	H	P	I	H	P	I	H	P	I	H	P	I	H
10	0	0	8	2	0	6	4	0	4	6	0	2	8	0	**0**	**10**	**0**
9	0	1	7	2	1	5	4	1	3	6	1	**1**	**8**	**1**			
8	0	2	6	2	2	4	4	2	**2**	**6**	**2**	0	8	2			
7	0	3	5	2	3	**3**	**4**	**3**	1	6	3						
6	0	4	**4**	**2**	**4**	2	4	4	0	6	4						
5	**0**	**5**	3	2	5	1	4	5									
4	0	6	2	2	6	0	4	6									
3	0	7	1	2	7												
2	0	8	0	2	8												
1	0	9															
0	0	10															

Relatively Peaceful Target
Relatively Hostile Target

Note: We highlight one relatively ''Peaceful'' target comprised of 6 peaceful cues, 2 hostile cues, and 2 inconclusive cues, as well as one relatively ''Hostile'' target comprised of 6 hostile cues, 2 peaceful cues, and 2 inconclusive cues.
Bold values indicate neutral target.

[2]Discussions between the first author and several senior Canadian Naval officers suggested that the actual number of cues that are used in this assessment varies from 8 to 12, depending on the situation. This estimate was empirically validated by Liebhaber et al. (2002) who showed that most US Naval officers in their study used between 8 and 11 cues out of a possible 18 identified as potentially relevant.

DOI: 10.1002/bdm

cues, and two unknown, or inconclusive cues. According to a ratio or strength model, the SP that the target is hostile as opposed to peaceful would be

$$\mathrm{SP}(\mathrm{H} > \mathrm{P}) = \frac{n(\mathrm{H})}{n(\mathrm{H}) + n(\mathrm{P})} \tag{1}$$

Hence after eliminating the inconclusive cues, $\mathrm{SP}(\mathrm{H} > \mathrm{P}) = 6/8 = 0.75$. Panel A of Figure 1 provides a view of the predictions of the ''Ignore'' hypothesis after plotting the predicted levels of $\mathrm{SP}(\mathrm{H} > \mathrm{P})$ for each of the 36 targets in Table 1. In the figure, model predictions are plotted as a function of the number of hostile (x-axis) and inconclusive cues. Note that this view predicts a distinctive SP curve for each level of inconclusive information.

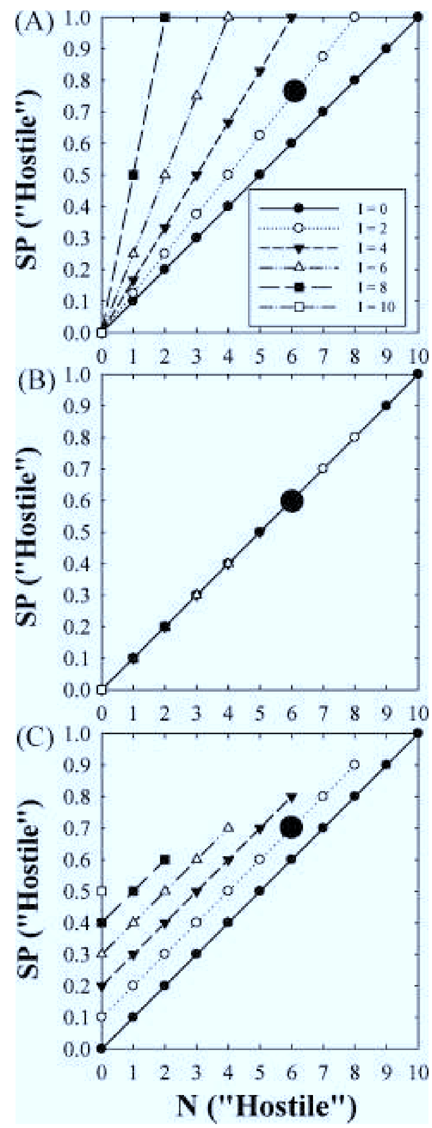

Figure 1. Predictions of the ''Ignore'' (Panel A), ''Normalize'' (Panel B), and ''Partition'' (Panel C) hypotheses. In each case, the subjective probability that a target is judged ''Hostile'' is plotted as a function of the number of hostile cues (x-axis), separately for each level of inconclusive information. The large filled circles denotes the predictions for the same target with 6 Hostile cues, 2 Peaceful cues, and 2 Inconclusive cues (see text for details)

DOI: 10.1002/bdm

The ''Normalize'' hypothesis

According to the ''Normalize'' hypothesis, SP(H > P) is based on a ratio of dominant (majority) cues to all cues, including inconclusive ones. That is

$$\mathrm{SP(H > P)} = \frac{n(\mathrm{H})}{n(\mathrm{H}) + n(\mathrm{P}) + n(\mathrm{I})} = \frac{n(\mathrm{H})}{n(\mathrm{total})} \tag{2}$$

Accordingly, for the same target discussed above, $\mathrm{SP(H>P)} = 6/6+2+2 = 0.60$. Figure 1B provides a view of the predictions of the ''Normalize'' hypothesis after plotting out the predicted levels of SP(H > P) for all 36 targets in Table 1. Note that this model predicts an identical SP curve for each level of inconclusive information.

The ''Partition'' hypothesis

According to the ''Partition'' hypothesis, SP(H > P) is based on a ratio of dominant information to all information, but here we make the additional assumption that the inconclusive cues are interpreted literally as equal-valued information elements and thus are equitably partitioned into peaceful and hostile cues. That is

$$\mathrm{SP(H > P)} = \frac{n(\mathrm{H}) + \frac{1}{2}n(\mathrm{I})}{n(\mathrm{H}) + \frac{1}{2}n(\mathrm{I}) + n(\mathrm{P}) + \frac{1}{2}n(\mathrm{I})} = \frac{n(\mathrm{H}) + \frac{1}{2}n(\mathrm{I})}{n(\mathrm{total})} \tag{3}$$

On this view, for example, if presented with two inconclusive cues, the participant will assign one to ''Hostile'' and one to ''Peaceful''. Accordingly, for the same target discussed above, $\mathrm{SP(H>P)} = 7/10 = 0.70$. Figure 1C provides a view of the predictions of the ''Partition'' hypothesis after plotting out the predicted levels of SP(H > P) for the 36 targets in Table 1. Note that this view again predicts a distinctive SP curve for each level of inconclusive information, but the pattern is quite different from the ''Ignore'' hypothesis.

In summary, the three hypotheses developed above can be summarized on the basis of what happens to the slopes and intercepts across values of $n(\mathrm{I})$: The Ignore hypothesis predicts equal intercepts and differing slopes; the normalize hypothesis predicts equal intercepts and equal slopes; and the partition hypothesis predicts differing intercepts and equal slopes. The objective of Experiment 1 was to empirically test the predictions of the three hypotheses.

EXPERIMENT 1

Method

Participants

Thirty-two adults participated in Experiment 1. Participants were Toronto area university students, students conducting undergraduate and graduate thesis research at DRDC—Toronto, and civilian employees of DRDC—Toronto. All participants were recruited by poster advertisement, were paid approximately $12.00 (Canadian) for participation, and were naive with respect to the nature and aims of the experiment. Each session took approximately 1 hour to complete. The study was approved by the DRDC—Toronto ethics review board for research involving human participants.

Apparatus

The task was presented on an AST Vision 7L 17-inch flat screen color monitor. An AST Pentium processor controlled event sequencing, randomization, and the recording of responses. A PC ''mouse'' was used as an input device.

DOI: 10.1002/bdm

Threat assessment task

The experimental platform employed in the present study is called Team and Individual Threat Assessment Network (TITAN). TITAN is a computer-based simulation of a Naval shipboard surveillance and threat assessment task designed for studying individual and team judgment and decision making performance (see also Baranski et al., 2007; Bryant, 2007; McCann, Baranski, Thompson, & Pigeau, 2000). In the current instantiation of TITAN, participants were presented with a radar-like display on which 36 contacts (symbolized by large asterisks) were shown surrounding a symbol representing ‘‘own-ship’’. The contacts represented surface, sub-surface and air traffic that has been identified by the ship’s radar system. Each contact must be processed based on information available from a drop-down menu. Ten pieces of information were potentially available for each contact (e.g., maneuver pattern, signal strength, direction, or origin, etc.) and participants were informed that each piece of information should be weighted equally toward the overall threat assessment. An example of a trial display is provided in Figure 2.

The processing of an individual track involves a specific sequence of events. First, a contact is selected (i.e., ‘‘hooked’’) by pointing and clicking on the target with a PC mouse. The track is then highlighted and a

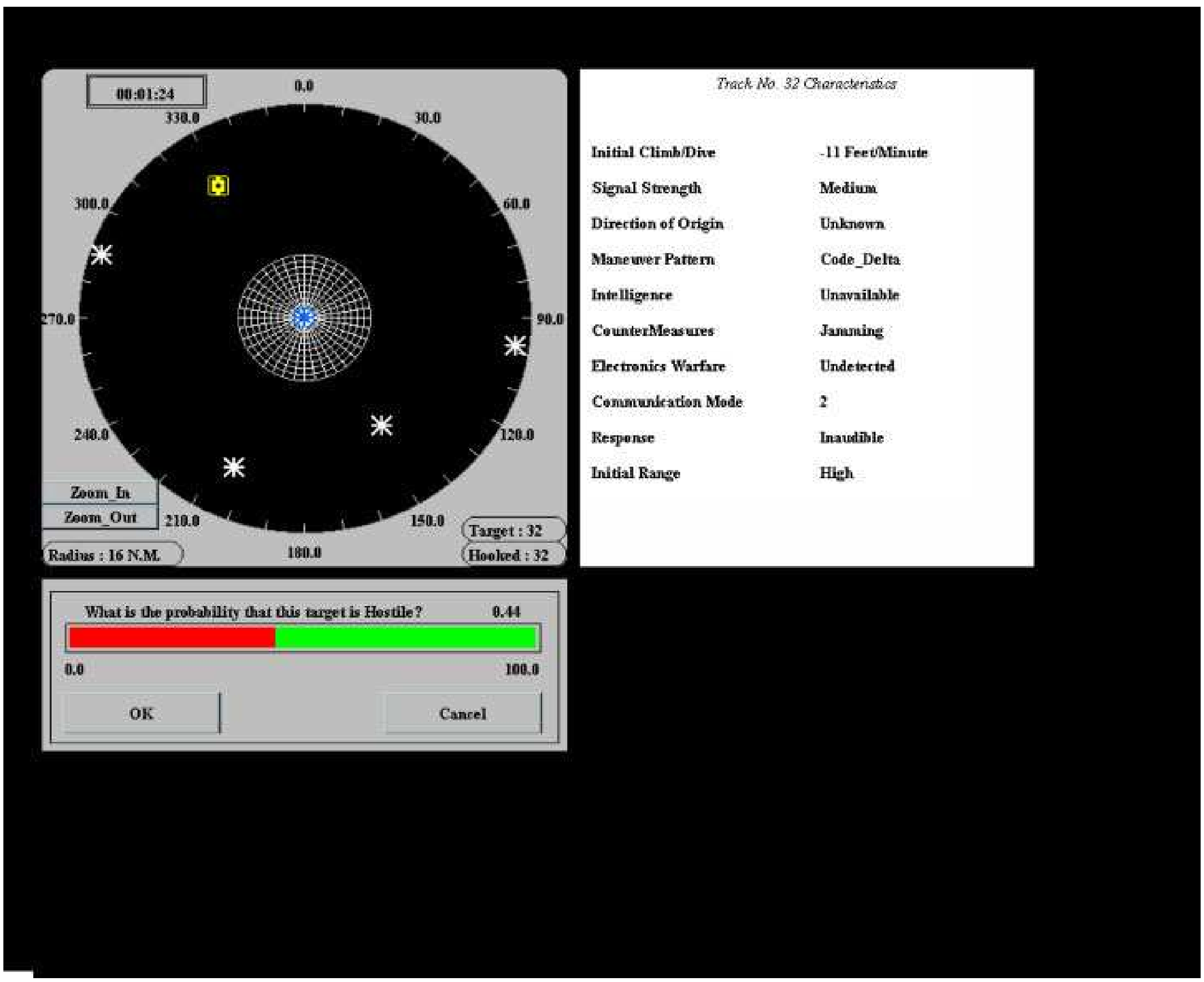

Figure 2. Screen shot of the TITAN task showing the radar display, the threat assessment scale, and the ten cue characteristics for a selected target (highlighted)

DOI: 10.1002/bdm

menu appears with 10 information elements, or cues. These values (e.g., signal strength = High) are then cross-referenced with a decision aid matrix, which allows the participant to categorize each cue value as ''peaceful'', ''hostile'', or ''unknown''. The decision aid matrix was positioned on a table stand directly beside the computer monitor in order to facilitate cue classification. Participants were required to review the information provided for each track and to submit their threat assessment to the system using a sliding visual analog probability scale (from 0 to 1.0). In sum, the task for the participant was effectively a multi-level decision task where ten cues are first classified as signalling ''hostile'', ''peaceful'', or ''unknown'' and then the participant integrates these cues to render a final ''probability of threat'' judgment. Participants were given detailed instructions on the use of the SP threat rating scale but were given no instructions on how to combine the cue information. Once the threat assessment was entered and confirmed, the contact disappeared from the screen and the participant selected the next contact. The task continued until all 36 targets were assessed (approximately 30–40 minutes). No feedback was provided on the accuracy of the assessments, precisely because our objective was to determine the way inconclusive information is naturally used by each participant.

Stimuli

Each participant processed the same set of 36 contacts but in a different random order. The 36 targets were comprised of a factorial combination of Peaceful (P), Hostile (H), and Inconclusive (I) cues. Table 1 provides a view of the cue composition of the 36 targets. Thus, for example, a target comprised of 6 Peaceful cues, 2 Hostile cues, and 2 Inconclusive cues would be viewed as relatively Peaceful; conversely, a target comprised of 6 Hostile cues, 2 Peaceful cues, and 2 Inconclusive cues would be viewed as relatively Hostile. Note that as such the entire stimulus set was symmetrical and balanced with respect to overall threat levels.

Design and procedure

In order to confirm that hostile and peaceful targets are judged symmetrically, the 32 participants were randomly assigned to one of two groups; 16 participants were required to judge the targets on a dimension of ''hostileness''; the other 16 participants were required to judge the same set of targets on a dimension of ''peacefullness''. In each case, ''0'' and ''1.0'' on the probability scale were to denote the lowest and highest levels, respectively.

Results and discussion

Figure 3A provides a plot of the mean probability threat assessments for the 16 participants in the ''hostile judgment'' condition as a function of the number of hostile cues, separately for each level of inconclusive information; Figure 3B provides the data for the 16 participants in the ''peaceful judgment'' condition. In each case, the data display a pattern that is consistent with the predictions of the Partition hypothesis. Indeed, Figure 3C shows a strong correspondence between the overall data[3] set and the predictions of the Partition hypothesis.

Figure 4 shows the correspondence between the three models and the complete data set for Experiment 1; i.e., 36 targets × 32 participants = 1152 observations. Again, the Partition hypothesis ($r^2 = .824$) provides the best account, followed by the Ignore ($r^2 = .687$) and Normalize ($r^2 = .666$) hypotheses, respectively. Finally,

[3]In order to combine the data from the two groups, we assumed that, for each target, SP[H > P] = 1 − SP[P > H], under the assumption of additivity of the probability judgments. Indeed a plot of SP[H > P] versus SP[P > H] for each target confirmed this assumption ($r^2 = 0.97$, $n = 36$).

DOI: 10.1002/bdm

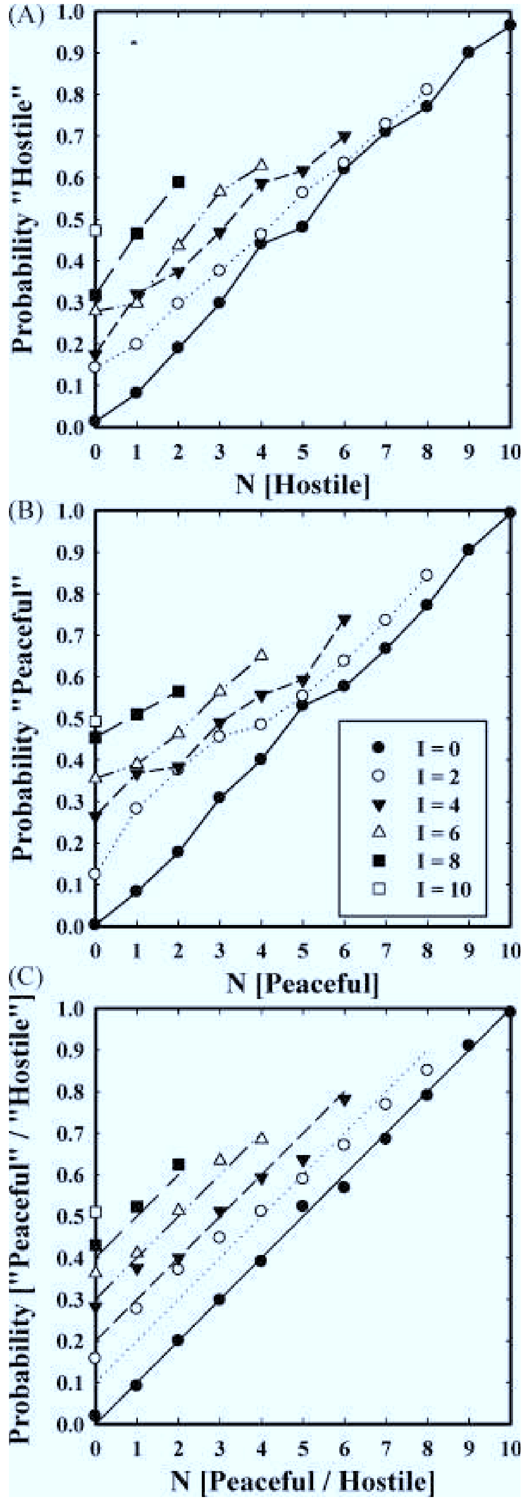

Figure 3. Probability ''Hostile'' (''Peaceful'') plotted as a function of the number of Hostile (Peaceful) cues, separately for each level of inconclusive information in Experiment 1. Data are reported separately for the judge ''Hostile'' group (Panel A), the judge ''Peaceful'' group (Panel B), and for the combined data (Panel C), together with the predictions of the Partition hypothesis

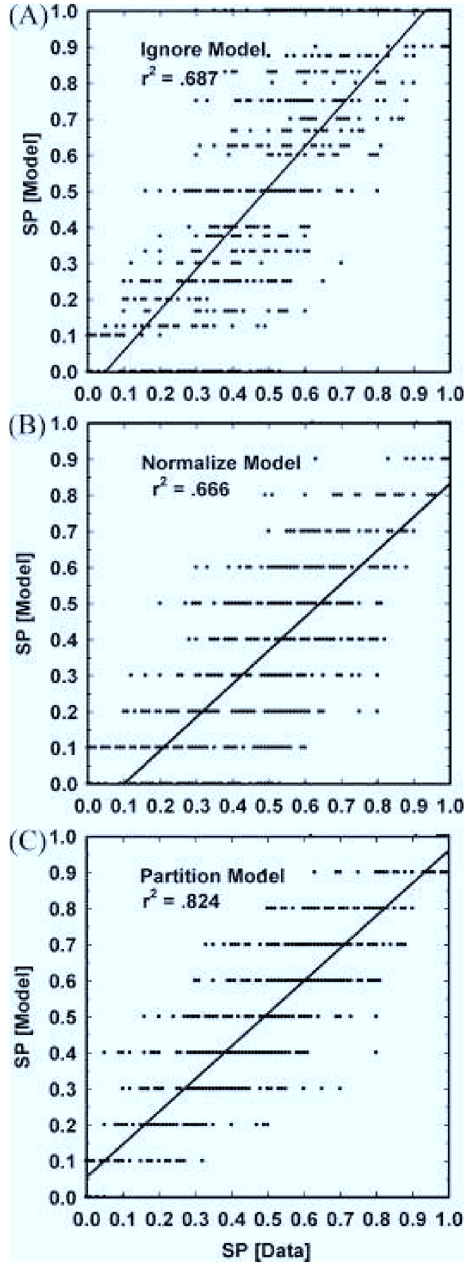

Figure 4. Correspondence between the three hypotheses and data for Experiment 1 ($n = 1152$): Ignore (Panel A), Normalize (Panel B), and Partition (Panel C)

Table 2 provides the correspondence of the three models to the data of each of the 32 participants in Experiment 1; in each case the Partition model provides the best account of the data.

The relation between the ''Partitioning'' and ''Balance of Evidence'' hypotheses

Recall that according to the Partitioning hypothesis

$$\mathrm{SP(H > P)} = \frac{n(\mathrm{H}) + \frac{1}{2}n(\mathrm{I})}{n(\mathrm{H}) + n(\mathrm{P}) + n(\mathrm{I})} \tag{4}$$

DOI: 10.1002/bdm

Table 2. Model fit indices for each participant (P) in Experiment 1

	Hostile							PEACEFUL					
	Ignore		Normalize		Partition			Ignore		Normalize		Partition	
P	R^2	RMSE	R^2	RMSE	R^2	RMSE	P	R^2	RMSE	R^2	RMSE	R^2	RMSE
2	0.684	0.191	0.619	0.174	*0.783	0.116	1	0.562	0.225	0.603	0.178	*0.793	0.113
4	0.663	0.198	0.714	0.151	*0.873	0.089	3	0.694	0.188	0.621	0.174	*0.816	0.107
6	0.743	0.173	0.788	0.130	*0.815	0.107	5	0.767	0.164	0.724	0.148	*0.900	0.078
8	0.863	0.126	0.773	0.134	*0.886	0.084	7	0.861	0.127	0.721	0.149	*0.970	0.043
10	0.576	0.222	0.685	0.158	*0.767	0.120	9	0.727	0.178	0.652	0.166	*0.895	0.080
12	0.831	0.140	0.696	0.156	*0.894	0.081	11	0.782	0.159	0.792	0.129	*0.955	0.053
14	0.762	0.166	0.711	0.152	*0.890	0.083	13	0.782	0.198	0.758	0.139	*0.880	0.086
16	0.802	0.152	0.659	0.165	*0.815	0.107	15	0.866	0.125	0.762	0.138	*0.947	0.057
18	0.837	0.137	0.793	0.128	*0.974	0.040	17	0.515	0.237	0.634	0.171	*0.707	0.135
20	0.759	0.167	0.706	0.153	*0.822	0.105	19	0.392	0.265	0.428	0.213	*0.679	0.141
22	0.670	0.196	0.600	0.178	*0.743	0.126	21	0.609	0.213	0.533	0.193	*0.764	0.121
24	0.657	0.199	0.689	0.361	*0.746	0.125	23	0.810	0.148	0.595	0.179	*0.902	0.078
26	0.711	0.183	0.606	0.177	*0.852	0.096	25	0.457	0.251	0.595	0.179	*0.702	0.136
28	0.721	0.180	0.799	0.126	*0.919	0.071	27	0.706	0.184	0.676	0.160	*0.780	0.117
30	0.702	0.186	0.796	0.127	*0.889	0.083	29	0.750	0.170	0.771	0.135	*0.940	0.061
32	0.840	0.136	0.783	0.131	*0.981	0.034	31	0.578	0.221	0.604	0.178	*0.783	0.116
Mean	**0.739**	**0.172**	**0.714**	**0.163**	**0.853**	**0.091**	Mean	**0.679**	**0.191**	**0.654**	**0.164**	**0.838**	**0.095**
SD	**0.081**	**0.027**	**0.070**	**0.056**	**0.074**	**0.027**	*SD*	**0.144**	**0.042**	**0.098**	**0.023**	**0.097**	**0.032**

Note: Best fits are indicated by an asterisk.

In the present study, $n(\mathrm{H}) + n(\mathrm{P}) + n(\mathrm{I}) = 10$, and so we can substitute for $n(\mathrm{I})$. With some cancellation and rearranging of terms we obtain

$$\mathrm{SP}(\mathrm{H} > \mathrm{P}) = \frac{n(\mathrm{H}) - n(\mathrm{P}) + 10}{20} \tag{5}$$

Note that $n(\mathrm{H}) - n(\mathrm{P})$ is on a scale from -10 to $+10$ and thus the constants simply convert the equation to a SP scale; i.e., 0–1. Therefore,

$$\mathrm{SP}(\mathrm{H} > \mathrm{P}) = F[n(\mathrm{H}) - n(\mathrm{P})] \tag{6}$$

where F is a monotonic function that converts the cue configuration to a SP scale. Note importantly that this representation is precisely the Vickers (1979) ‘‘balance of evidence’’ hypothesis for the basis of confidence in human judgment (see also Vickers & Packer, 1982; Vickers et al., 1985b); i.e., where confidence is a monotonic function of the difference in the evidence totals favoring the dominant and non-dominant response alternatives. A critical prediction of this result is that SP, according to the partitioning model, is a direct linear function of the ‘‘balance of evidence’’ (i.e., the difference between the number of hostile and peaceful cues) and thus must be independent of the amount of inconclusive information.

Figure 5A provides a view of the predicted SP threat assessments (‘‘Hostile’’ and ‘‘Peaceful’’) as a function of the absolute value of the ‘‘balance of evidence,’’ separately for each level of inconclusive information. Thus, if participants are adhering strictly to a partitioning of inconclusive evidence, then, as is evident in Figure 5A, the predicted pattern of judgments will be based only on the balance of diagnostic cues and thus should be uninfluenced by the number of inconclusive cues. Figure 5B provides a plot of the observed data from Experiment 1. The pattern of responses was virtually identical for the judged ‘‘hostile’’ and ‘‘peaceful’’ conditions, so we combined the data for the two groups. As is evident in Figure 5B, and

DOI: 10.1002/bdm

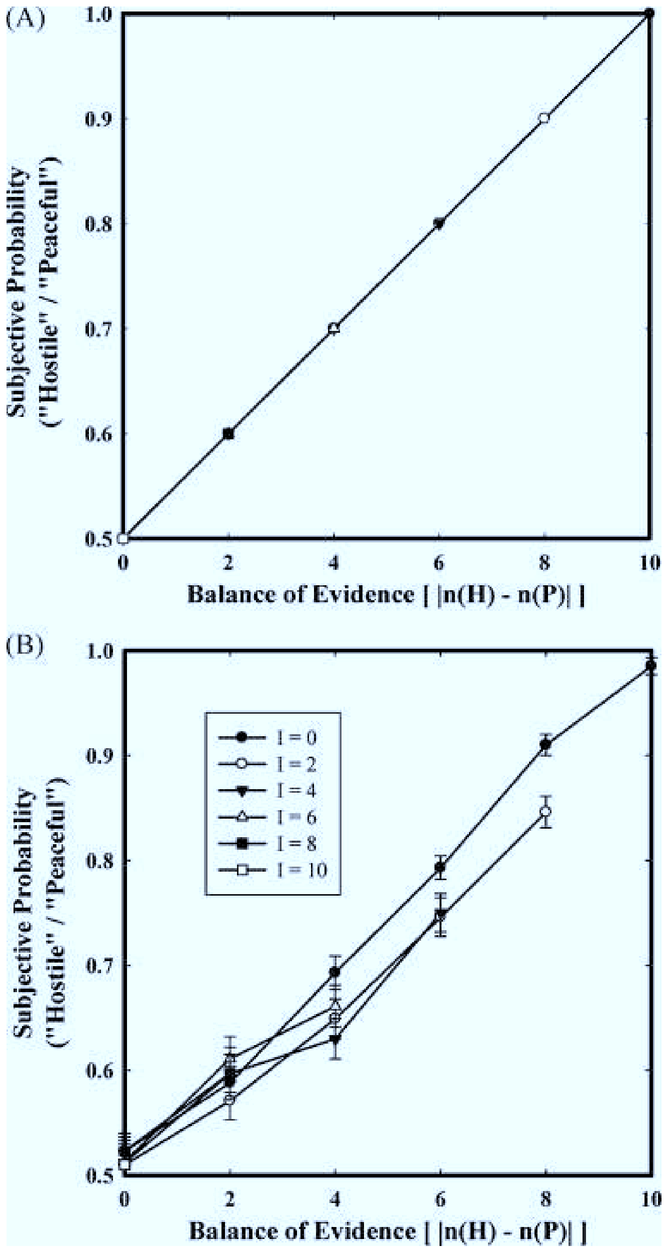

Figure 5. Panel A: predictions of the "Balance of Evidence" hypothesis, plotted separately for each level of inconclusive information. Panel B: data from Experiment 1 illustrating the "dilution effect". Error bars denote the standard error of the mean across participants

despite the apparent success of the partitioning model vis-à-vis the other models, SP threat assessments are clearly affected by the presence of inconclusive information. The interpretation of the results is that the presence of inconclusive information reduces threat levels for hostile contacts and increases threat levels for peaceful contacts, a general form of regression to the mean or central tendency, consistent with the well-known "dilution effect" (e.g., LaBella & Koehler, 2004; Nisbett et al., 1981; Waller & Zimbelman, 2003).

Separate Analyses of Variance (ANOVAs) were conducted at each level of Balance of Evidence, comparing threat assessments across each level of inconclusive information ("I"). Briefly, the results confirmed what is evident in Figure 5B; the dilution effect is significant when the balance of evidence exceeds 2 and is of approximately constant magnitude across the higher levels of inconclusive information.

EXPERIMENT 2

Experiment 1 provided support for the Partitioning hypothesis but the finding of a dilution effect suggests that inconclusive information can affect threat assessments in a regressive manner. Experiment 2 sought to replicate and extend these findings in a study that varied the global threat context of the judgment task. Twenty-four participants were recruited from the same population as in Experiment 1. Because Experiment 1 found that judgments on the dimensions of ‘‘hostileness’’ and ‘‘peacefulness’’ were symmetrical, all participants in Experiment 2 rated target threat level on the dimension of ‘‘hostileness’’ only. Each participant participated in two sessions involving the TITAN game described in Experiment 1. The sessions were identical except for variations in a scenario context that depicted a land claim dispute involving two fictitious countries, Mongoba and Islandia. One version involved escalating hostilities and one version involved de-escalating hostilities (see Appendix). Half of the participants were exposed to the escalating (de-escalating) scenario in Session 1 and the de-escalating (escalating) scenario in Session 2. These scenarios were read by the participants prior to commencing each session, which was conducted on separate days. The objective of the study was to examine the effect of the scenario context on the weighting of inconclusive cues. We hypothesized that in the hostile context, inconclusive information would be weighted as more hostile, particularly for hostile targets; in the peaceful context, inconclusive information would be weighted as more peaceful, particularly for peaceful targets.

Results and discussion

A preliminary analysis revealed that the contextual scenario manipulation had no effect on the weighting of inconclusive cues. Accordingly, Figure 6 provides a plot of the mean probability assessments for the 24 participants after combining the data over the two sessions and two context conditions. Once again the data are well represented by the Partition model. As in Experiment 1, we examined the correspondence between each of the models and the entire data set; i.e., 36 targets $\times$ 24 participants $\times$ 2 sessions $=$ 1728 observations. Again, the Partition hypothesis ($r^2 = .829$) provided the best account of the data, followed by the Ignore ($r^2 = .684$) and Normalize ($r^2 = .664$) hypotheses, respectively. In addition, Table 3 shows the correspondence between the three models and the data of each of the 24 participants in Experiment 2. The partitioning model provided the best account of the data for 23/24 participants; for Participant 19 the Normalize model provided a slightly better account.

Finally, Figure 7 provides a view of the Partition hypothesis as reflected by the Balance of Evidence. Once again there is clear evidence of a dilution effect, such that the presence of inconclusive evidence produces a regression effect resulting in hostile targets being judged slightly less hostile than predicted by the model, and peaceful targets being judged slightly less peaceful than predicted by the model. ANOVAs paralleling those conducted in Experiment 1 were conducted at each level of balance of evidence. Once again the results confirmed what is evident in Figure 7; the dilution effect is clearly evident when the balance of evidence exceeds 2 and is of approximately constant magnitude across the higher levels of inconclusive information.

EXPERIMENT 3

The findings of Experiments 1 and 2 provide strong support for the Partition hypothesis for the use of inconclusive information in threat assessment. The objective of Experiment 3 was to examine the Partition

DOI: 10.1002/bdm

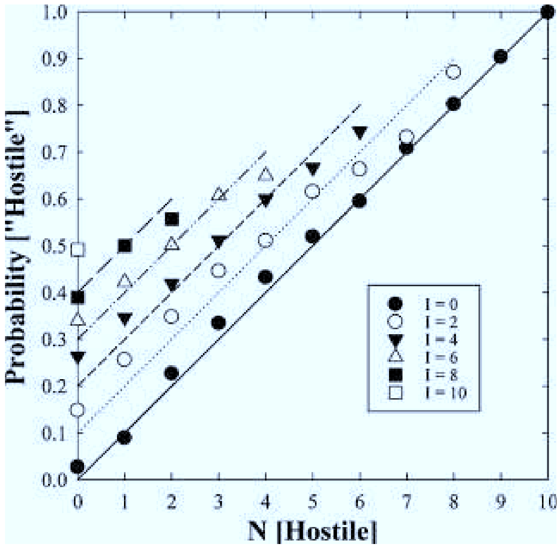

Figure 6. Combined data of 24 participants in Experiment 2, together with the predictions of the Partitioning hypothesis

Table 3. Model fit indices for each participant (P) in Experiment 2

	Ignore		Normalize		Partition	
P	R^2	RMSE	R^2	RMSE	R^2	RMSE
1	0.787	0.155	0.693	0.154	*0.898	0.078
2	0.736	0.172	0.630	0.169	*0.848	0.095
3	0.648	0.199	0.714	0.149	*0.745	0.124
4	0.849	0.130	0.722	0.147	*0.962	0.048
5	0.691	0.186	0.669	0.160	*0.823	0.103
6	0.855	0.128	0.766	0.134	*0.995	0.017
7	0.526	0.231	0.640	0.167	*0.751	0.122
8	0.696	0.185	0.528	0.191	*0.839	0.098
9	0.496	0.238	0.543	0.188	*0.620	0.151
10	0.734	0.173	0.730	0.145	*0.909	0.074
11	0.746	0.169	0.706	0.151	*0.912	0.073
12	0.823	0.141	0.765	0.135	*0.976	0.038
13	0.611	0.209	0.713	0.149	*0.796	0.111
14	0.838	0.135	0.758	0.137	*0.977	0.037
15	0.715	0.179	0.731	0.144	*0.892	0.080
16	0.617	0.208	0.582	0.180	*0.764	0.119
17	0.593	0.214	0.610	0.174	*0.769	0.118
18	0.854	0.128	0.708	0.150	*0.902	0.077
19	0.629	0.204	*0.872	0.099	0.823	0.103
20	0.592	0.214	0.732	0.144	*0.837	0.099
21	0.668	0.193	0.687	0.155	*0.871	0.088
22	0.606	0.211	0.519	0.193	*0.765	0.119
23	0.751	0.167	0.738	0.142	*0.931	0.064
24	0.896	0.108	0.718	0.148	*0.954	0.052
Mean	**0.707**	**0.178**	**0.686**	**0.154**	**0.857**	**0.087**
SD	**0.111**	**0.036**	**0.084**	**0.021**	**0.092**	**0.033**

Note: Best fits are indicated by an asterisk.

DOI: 10.1002/bdm

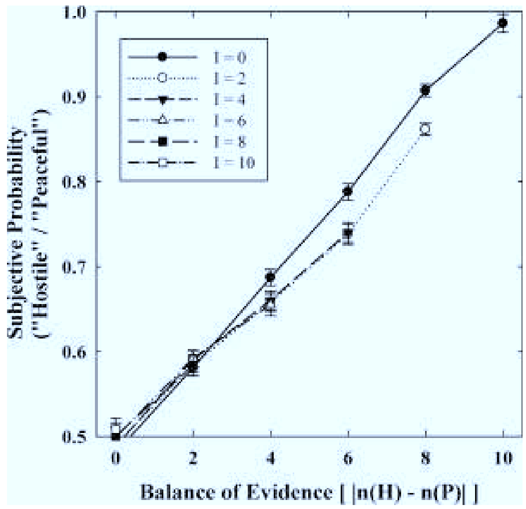

Figure 7. Balance of evidence plot for Experiment 2 illustrating the "dilution effect". Error bars denote the standard error of the mean across participants

hypothesis and the dilution effect using a group of eight, experienced Operations Room Officers (OROs) from the Canadian Navy. The study was identical to Experiment 1, except that participants judged the targets only on the dimension of "hostileness."

Results and discussion

As in the previous experiments, we examined the correspondence between each model and the entire data set; i.e., 36 targets $\times$ 8 participants $=$ 256 observations. In comparison to Experiments 1 and 2, the Partition hypothesis provides a slightly poorer ($r^2 = .733$) overall account than the Normalize hypothesis ($r^2 = .736$); as in Experiments 1 and 2, the account of the Ignore model was not as good ($r^2 = .633$). Table 4 shows the correspondence of the three models to the data of the eight participants in Experiment 3. In this case,

Table 4. Model fit indices for each participant (P) in Experiment 3

	Ignore		Normalize		Partition	
P	R^2	RMSE	R^2	RMSE	R^2	RMSE
1	0.597	0.216	0.654	0.166	*0.686	0.139
2	0.700	0.186	*0.880	0.098	0.814	0.107
3	0.567	0.224	*0.820	0.120	0.640	0.149
4	0.591	0.218	*0.902	0.088	0.707	0.134
5	0.695	0.188	0.632	0.171	*0.760	0.122
6	0.753	0.169	*0.848	0.110	0.775	0.118
7	0.603	0.214	0.628	0.172	*0.826	0.104
8	0.835	0.138	0.827	0.117	*0.970	0.043
Mean	**0.667**	**0.194**	**0.774**	**0.130**	**0.772**	**0.114**
SD	**0.094**	**0.030**	**0.116**	**0.034**	**0.102**	**0.033**

Note: Best fits are indicated by an asterisk.

DOI: 10.1002/bdm

the Normalize hypothesis provides the best account of the data for Participants 2, 3, 4, and 6, whereas the Partition hypothesis provides the best account of the data for the remaining four participants. In Figure 8 we group these two patterns as a function of the number of inconclusive cues. To a reasonable approximation, and considering the limited number of participants representing each model, the respective patterns match the predictions of the two models quite well. Finally, Figure 9 provides a view of the two patterns of responding as a function of the balance of evidence. In contrast to Experiments 1 and 2, no dilution effect is evident in the data of the experienced naval officers, and this is particularly evident for the participants whose data were best accounted for by the Partition model (Figure 9B). For the participants whose data were best accounted for by the Normalize model (Figure 9A), the data were identical, except for the response to a zero balance of evidence. In this case, the SP "Hostile" increases monotonically with increases in the number of inconclusive cues.

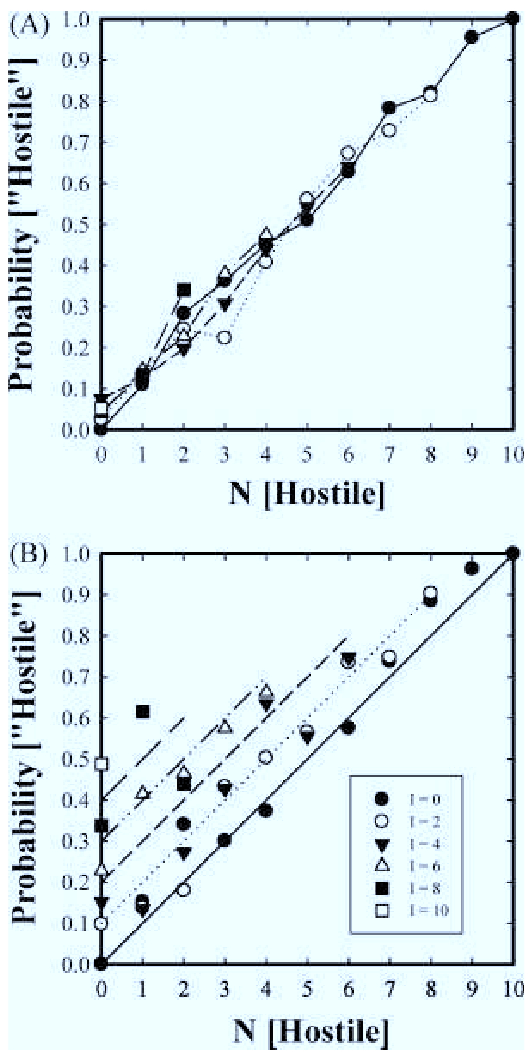

Figure 8. Data from Experiment 3. Panel A provides the data for the four Navy officers whose data were best accounted for by the Normalize hypothesis; Panel B provides the data for the four Navy officers whose data were best accounted for by the Partition hypothesis

DOI: 10.1002/bdm

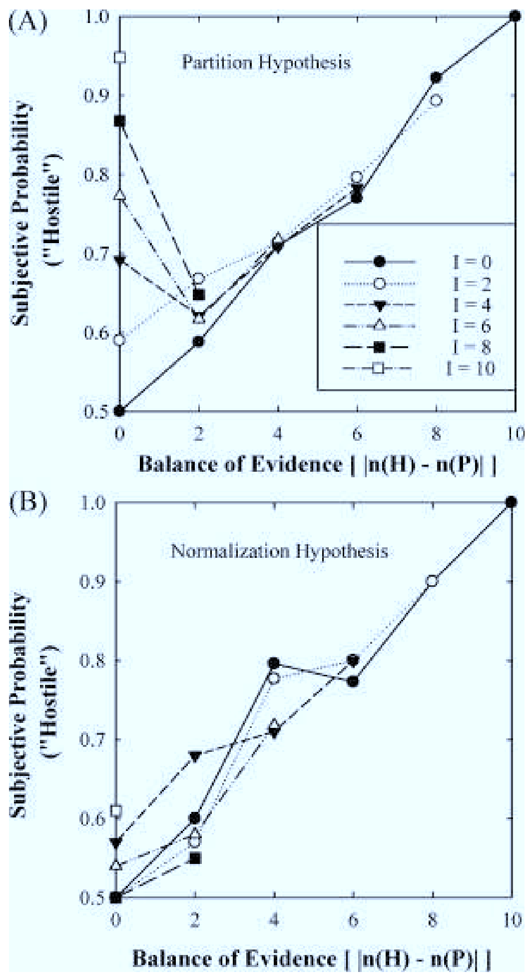

Figure 9. Balance of evidence plots for Experiment 3. Panel A provides the data for the four Navy officers whose data were best accounted for by the Normalize hypothesis; Panel B provides the data for the four Navy officers whose data were best accounted for by the Partition hypothesis. In each case there is no evidence of a dilution effect

GENERAL DISCUSSION

The present research developed and tested three deterministic models for the use of inconclusive information in human judgment. The evidence provided in support of the Partition model suggests a novel, yet reasonable conceptualization of inconclusive evidence; it is literally evidence that favors *both* response alternatives. This is provided in contrast to the predictions of the Ignore model, in which inconclusive evidence literally favors *neither* of the response alternatives and thus should be eliminated from the judgment process. Of course this is not to say that the latter model is generally incorrect. In contrast, it would be an important theoretical advance in our understanding of the processing of inconclusive information if judgment contexts could be determined in which the different models could be favored; for example a distinction between inconclusive and irrelevant evidence discussed in the introduction. In the present context of threat assessment however, on balance, the results overwhelmingly support the Partition model.

The present findings may have implications for the development and testing of quantitative models of confidence in human judgment. Arguably the most successful formulation to date has been the Balance of Evidence hypothesis developed in the context of Vickers' (1979) accumulator-based decision model.

DOI: 10.1002/bdm

Although originally developed with application to two-choice perceptual comparative judgments, the Balance of Evidence view was intended to describe a very general basis for confidence in human judgment. More recently, Baranski and Petrusic (1998) proposed that confidence might be regulated by the balance of evidence and a scaling of the amount of inconclusive information, a hypothesis developed in the context of the Slow and Fast-Guessing decision model (Baranski & Petrusic, 2003; Petrusic, 1992; Petrusic & Baranski, 1989; Petrusic & Jamieson, 1978).

In the present study we systematically varied the amount of inconclusive information in a simple, discrete information simulation with a fixed data set. Our findings showed that subjective probabilities were well described by a balance of evidence but that inconclusive information clearly affects SP confidence judgments. Accordingly, a natural extension of this work would be to examine the effect of both discrete and continuous information elements in both perceptual and non-perceptual tasks. Methodologically the challenge is to quantify and standardize informational cues *a priori* and then allow the cues to be presented in a controlled manner.

One approach, again developed by Vickers (1985) and his associates (Vickers et al., 1985; 1989) involves the so-called expanded judgment task which permits the study of the time course, or evolution, of the decision making process. Another approach could involve an extension of the present task to category learning (Sieck & Yates, 2001) or multiple-cue probability learning (MCPL) tasks (White & Koehler, 2004), involving probabilistic as opposed to discrete informational cues. Indeed, White and Koehler (2004) developed and tested alternative models for dealing with missing information in a MCPL medical diagnosis task. Their findings revealed that in the judgment phase, participants ‘‘fill in’’ missing information with the mean cue value presented during the learning phase. Interestingly, the present findings in support of the partition model are consistent with this view on the assumption that an inconclusive cue in the threat assessment task is weighted as 0.5.

From an applied perspective, the present findings can inform the development of future decision support technologies. Ultimately, if these technologies are to be trusted and used, they should complement naturalistic decision-making processes while reducing the likelihood of errors and biases. For example, Experiment 3 highlighted clear individual differences in how experts process inconclusive information with the implication that advanced decision support technologies should dynamically adapt to the information processing strategies of the user. Similarly, the present study employed a fixed and constrained cue set with equal weighting of the cues. However, as illustrated in Figure 9A, some experts weighted inconclusive cues as more ‘‘hostile’’ when the balance of evidence was low. Of course, in real world tasks the cues weights would vary (see Liebhaber et al., 2002) depending on contextual factors and cue contingencies (e.g., ‘‘direction of origin’’ may have a higher weighting if ‘‘signal strength’’ exceeds some criterion). Importantly, the models developed in the present study were able to capture these general approaches to how individuals weigh and process inconclusive information.

Our final discussion point concerns the dilution effect which, in the present studies appeared as a regression effect, resulting in more conservative judgments (i.e., less extreme) than predicted by a strict balance of evidence. Interestingly, this regression effect was evident in the judgments of the civilian participants (Experiments 1 and 2), but not in the data of the experienced naval officers (Experiment 3). This lack of a dilution effect in the judgments of the Naval officers is also interesting because, as mentioned at the outset, the dilution effect has been reported in other expert domains, most notably professional auditors (Hackenbrack, 1992; Shelton, 1999; Waller & Zimbelman, 2003). In their examination of archival data of professional audits, Waller and Zimbelman posit that personal (accountability) and organizational (i.e., firm policy) motivations may be consistent with conservative judgment and thus the dilution effect in that context (see, e.g., Tetlock & Boettger, 1989). For military commanders the conflict is obvious; they cannot be conservative if the threat is real, and they must be conservative if, as with the Vincennes tragedy, the threat is not real. The lack of a dilution effect in the judgments of the naval officers may therefore reflect their highly trained personal and organizational motivations to defend against potential adversaries. Of course this

DOI: 10.1002/bdm

suggestion is speculative and will require examination in a more naturalistic, high fidelity training environment, perhaps comparing verbal protocol (e.g., Burke, 1995; Kühberger & Huber, 1998), process tracing (Körner et al., 2007), and direct assessment methodologies.

ACKNOWLEDGEMENTS

Portions of this work were presented at the 40th annual meeting of the Psychonomic Society, Los Angeles, 1999 and the joint meeting of the Canadian Society for Brain, Behavior and Cognitive Science and the Experimental Psychology Society, Cambridge, UK, 2000. The authors thank Tonya Hendriks, Andrea Hawton, Heather Devine, and Heidi Roesler-Mulroney for technical support with this project.

APPENDIX—THREAT CONTEXT SCENARIOS FOR EXPERIMENT 2

Hostile Scenario

Imagine yourself as a naval officer aboard a ship. You are part of a UN task force deployed to monitor activities off the coasts of Islandia and Mongoba. Mongoba is a small coastal country governed by an aggressive military dictatorship. It recently threatened to claim, by force if necessary, 400 miles of prime coastline from the peaceful, neighboring, and democratic country of Islandia. Islandia immediately appealed to the UN for assistance. Accordingly, the situation in Islandia is under UN military observation while UN-mediated negotiations proceed.

Mongoba's leadership is openly critical of the bargaining strategies employed by the UN delegation. Their leadership, backed by a strong right-wing terrorist faction, issued open threats against the UN through CNN. In addition to directly threatening the UN, they also threatened to disrupt civilian shipping traffic in the waters off the coast of Islandia and Mongoba. Intelligence reports confirm that Mongoba possesses a potent air strike capability and is practicing attack maneuvers on land, air, and sea. Furthermore, over the last few weeks Mongoba has been preparing troops for occupation of Islandia and are reportedly stockpiling weapons and artillery.

The situation has changed

Heavy gun battles erupted overnight after talks broke down between Mongoba and Islandia. Islandia refuses to negotiate under the threat of violence; however, Mongoba insists that the violence will continue until a satisfactory settlement is reached on their land claim. The Mongoban military is currently marshalling troops and equipment to proceed with their plan to occupy Islandia.

Your mission

You are positioned 1500 nautical miles off the coast of Islandia. Your primary mission is to assess the potential threat posed to your task force by all traffic (i.e., contacts) in your radar space. Specifically, you are required to review information for the contacts on your radar display and, with the aid of your decision matrix, decide how hostile each of those contacts is. Remember that the contacts you are evaluating could be Mongoban, civilian, or your own.

Peaceful Scenario

Imagine yourself as a naval officer aboard a ship. You are part of a UN task force deployed to monitor activities off the coasts of Islandia and Mongoba. Mongoba is a small coastal country governed by an aggressive military dictatorship. It recently threatened to claim, by force if necessary, 400 miles of prime coastline from the peaceful, neighboring, and democratic country of Islandia. Islandia immediately appealed to the UN for assistance. Accordingly, the situation in Islandia is under UN military observation while UN-mediated negotiations proceed.

Mongoba's leadership is openly critical of the bargaining strategies employed by the UN delegation. Their leadership, backed by a strong right-wing terrorist faction, issued open threats against the UN through CNN. In addition to directly threatening the UN, they also threatened to disrupt civilian shipping traffic in the waters off the coast of Islandia and Mongoba. Intelligence reports confirm that Mongoba possesses a potent air strike capability and is practicing attack maneuvers on land, air, and sea. Furthermore, over the last few weeks Mongoba has been preparing troops for occupation of Islandia and are reportedly stockpiling weapons and artillery.

The situation has changed

Today, during hours of intense negotiations the leaders of Mongoba and Islandia made significant progress on key issues of the land claim dispute. As a show of good faith the Mongoban leadership retracted their public condemnation of the UN. Notwithstanding the friendly nature of this gesture the leaders remain unwilling to withdraw their threat to occupy Islandia. In fact, the Mongoban terrorist faction continues to conduct military exercises and organize troops in preparation for occupation. Needless to say, their stance has tempered Islandia's optimism following the latest round of talks.

Your mission

You are positioned 1500 nautical miles off the coast of Islandia. Your primary mission is to assess the potential threat posed to your task force by all traffic (aka. Contacts) in your radar space. Specifically, you are required to review information for the contacts on your radar display and, with the aid of your decision matrix, decide how hostile each of those contacts is. Remember that the contacts you are evaluating could be Mongoban, civilian, or your own.

REFERENCES

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 929–945.

Baranski, J. V., & Petrusic, W. M. (1999). On the use of inconclusive information in the generation of subjective probabilities. Paper presented at the 40th annual meeting of the Psychonomic Society, Los Angeles, CA. *Abstracts of the Psychonomic Society, 4* 54.

Baranski, J. V., & Petrusic, W. M. (2000). Subjective probability in the assessment of threat: Comparing expert vs. novice use of inconclusive information. Paper presented at a symposium on "Confidence in human judgment" at the Joint Meeting of the Experimental Psychology Society and the Canadian Society for Brain, Behavior, and Cognitive Science, Cambridge, UK. *Abstract p.24*.

Baranski, J. V., & Petrusic, W. M. (2003). Adaptive decision processes in perceptual comparisons: Effects of changes in the global difficulty context. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 658–674.

Baranski, J. V., Thompson, M. M., Lichacz, F. M. J., McCann, C., Gil, V., Pasto, L., et al. (2007). Effects of sleep loss on team decision making: Motivational loss or motivational gain? *Human Factors*, *49*, 646–660.

Bastardi, A., & Shafir, E. (1998). On the pursuit and misuse of useless information. *Journal of Personality and Social Psychology*, *75*, 19–32.

Bryant, D. J. (2007). Classifying simulated air threats with fast and frugal heuristics. *Journal of Behavioral Decision Making*, *20*, 37–64.

Bryant, D. J., & Webb, R. D. (1999). Literature survey for issues in naval decision support: Phase II. Defence Research and Development Canada Toronto, *Contract Report CR-1999-014: Toronto*. 210 pages.

DOI: 10.1002/bdm

Burke, S. J. (1995). The dimensional effects of missing information on choice processing. *Journal of Behavioral Decision Making*, *8*, 223–244.

Cannon-Bowers, J. A., & Salas, E. (1998). *Making decisions under stress: Implications for individual and team training*. Washington, DC: APA Press.

Collyer, S. C., & Malecki, G. S. (1998). Tactical decision making under stress: History and overview. In J. Cannon-Bowers , & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training*. Washington, DC: APA Press. (pp. 3–15).

Ebenbach, D. H., & Moore, C. F. (2000). Incomplete information, inferences, and individual differences: The case of environmental judgments. *Organizational Behavior and Human Decision Processes*, *81*, 1–27.

Edwards, W., Lindman, H., & Phillips, L. D. (1965). Emerging technologies for making decisions. In T. M. Newcomb (Ed.), *New directions in psychology II* (pp. 259–325). New York: Holt, Rinehart, and Winston, Inc.

Gaeth, G., & Shanteau, J. (1984). Reducing the influence of irrelevant information on experienced decision makers. *Organizational Behavior and Human Performance*, *33*, 263–282.

Hackenbrack, K. (1992). Implications of seemingly irrelevant evidence in audit judgment. *Journal of Accounting Research*, *30*, 126–136.

Hutchins, S. G. (1996). Principles for intelligent decision aiding. *US Office of Naval Research Technical Report 1718*, San Diego, CA. 34 pages.

Jagacinski, C. M. (1991). Personnel decision making: The impact of missing information. *Journal of Applied Psychology*, *76*, 19–30.

Jagacinski, C. M. (1994). Evaluation of job candidates with missing information: Effects of attribute importance and interattribute correlation. *Journal of Behavioral Decision Making*, *7*, 25–42.

Jagacinski, C. M. (1995). Distinguishing adding and averaging models in a personnel selection task: When missing information matters. *Organizational Behavior and Human Decision Processes*, *61*, 1–15.

Johnson, R. D. (1987). Making judgments when information is missing: Inferences, biases, and framing effects. *Acta Psychologica*, *66*, 69–82.

Kemmelmeier, M. (2004). Separating the wheat from the chaff: Does discriminating between diagnostic and non-diagnostic information eliminate the dilution effect? *Journal of Behavioral Decision Making*, *17*, 231–243.

Klein, G. (1996). The effect of acute stressors on decision making. In J. E. Driskell , & E. Salas (Eds.), *Stress and human performance* (pp. 49–88). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Koehler, D. J. (1996). A strength model of probability judgments for tournaments. *Organizational Behavior and Human Decision Processes*, *66*, 16–21.

Körner, C., Gertzen, H., Bettinger, C., & Albert, D. (2007). Comparative judgments with missing information: A regression and process tracing analysis. *Acta Psychologica*, *125*, 66–84.

Kühberger, A., & Huber, O. (1998). Decision making with missing information: A verbal protocol study. *European Journal of Cognitive Psychology*, *10*, 269–290.

LaBella, C., & Koehler, D. J. (2004). Dilution and confirmation of probability judgments based on nondiagnostic evidence. *Memory & Cognition*, *32*, 1076–1089.

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

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

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

McCann, C., Baranski, J. V., Thompson, M. M., & Pigeau, R. (2000). On the utility of experiential cross-training for team decision making under time stress. *Ergonomics*, *43*, 1095–1110.

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, *13*, 248–277.

Peters, E., & Rothbart, M. (2000). Typicality can create, eliminate, and reverse the dilution effect. *Personality and Social Psychology Bulletin*, *26*, 177–187.

Petrusic, W. M. (1992). Semantic congruity effects and theories of the comparison process. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 962–986.

Petrusic, W. M., & Baranski, J. V. (1989). Context, context shifts, and semantic congruity effects in comparative judgments. In D. Vickers , & P. Smith (Eds.), *Human information processing: Measures, mechanisms, and models* (pp. 231–251). Amsterdam, North-Holland: Elsevier Science Publishers.

Petrusic, W. M., & Jamieson, D. G. (1978). The relation between probability of preferential choice and the time to choose changes with practice. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 471–482.

Roberts, N. C., & Dotterway, K. A. (1995). The Vincennes incident: Another player on the stage. *Defence Analysis*, *11*, 31–45.

DOI: 10.1002/bdm

Roy, J., Breton, R., & Paradis, S. (2001). Human-computer interface for the study of information fusion concepts in situation analysis and command decision support systems. SPIE Proceedings, Vol 4380, Signal Processing, Sensor Fusion, and Target Recognition. Orlando, FL.

Sieck, W. R., & Yates, J. F. (2001). Overconfidence effects in category learning: A comparison of connectionist and exemplar memory models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 1003–1021.

Shelton, S. W. (1999). The effect of experience on the use of irrelevant evidence in auditor judgment. *The Accounting Review*, *74*, 217–224.

Schum, D. A., & Martin, D. W. (1968). Human processing of inconclusive evidence from multinomial probability distributions. *Organizational Behavior and Human Performance*, *3*, 353–365.

Tetlock, P. E., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology*, *57*, 388–398.

Tetlock, P. E., Lerner, J. S., & Boettger, R. (1996). The dilution effect: Judgmental bias, conversational convention, or a bit of both? *European Journal of Social Psychology*, *26*, 915–934.

Troutman, C. M., & Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organizational Behavior and Human Performance*, *19*, 43–55.

Tversky, A., & Koehler, D. J. (1994). Support theory: A non-extensional representation of subjective probability. *Psychological Review*, *101*, 547–567.

Vickers, D. (1979). *Decision processes in visual perception*. New York: Academic Press.

Vickers, D. (1985). Antagonistic influences on performance change in detection and discrimination tasks. In G. d'Ydewalle (Ed.), *Cognition, information processing and motivation: Proceedings of the XXIII international congress of psychology (* Vol. 3, pp. 79–115). Amsterdam: North Holland.

Vickers, D., Burt, J., Smith, P., & Brown, M. (1985). Experimental paradigms emphasizing state or process limitations: I. Effects on speed accuracy trade-offs. *Acta Psychologica*, *59*, 129–161.

Vickers, D., Foreman, E. A., Nicholls, M. E. R., Innes, N. J., & Gott, R. E. (1989). Some experimental tests of the application of an interval of uncertainty model to a time-limited expanded judgment task. In D. Vickers , & P. Smith (Eds.), *Human information processing: Measures, mechanisms, and models* (pp. 253–265). Amsterdam, North-Holland: Elsevier Science Publishers.

Vickers, D., & Packer, J. S. (1982). Effects of alternating set for speed or accuracy on response time, accuracy, and confidence in a unidimensional discrimination task. *Acta Psychologica*, *50*, 179–197.

Vickers, D., Smith, P., Burt, J., & Brown, M. (1985). Experimental paradigms emphasizing state or process limitations: II. Effects on confidence. *Acta Psychologica*, *59*, 163–193.

Waller, W. S., & Zimbelman, M. F. (2003). A cognitive footprint in archival data: Generalizing the dilution effect from the laboratory to field settings. *Organizational Behavior and Human Decision Processes*, *91*, 254–268.

White, C. M., & Koehler, D. J. (2004). Missing information in multiple-cue probability learning. *Memory & Cognition*, *32*, 1007–1018.

Windschitl, P. D., & Chambers, J. R. (2004). The dud-alternative effect in likelihood judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 198–215.

Authors' biographies:

Joseph V. Baranski is Chief Scientist at Defence Research & Development Canada—Toronto. His research interests include individual and team decision-making under stress, confidence in judgment, and the ability to self-monitor aspects of cognitive performance. He received his PhD in psychology in 1992 from Carleton University in Ottawa, Canada.

William M. Petrusic is Emeritus Distinguished Research Professor in the Department of Psychology at Carleton University in Ottawa. His research interests include theories of response time and comparative judgment, perceptual and memory psychophysics, and confidence in judgment. He received his PhD in psychology in 1968 from the University of Michigan.

Authors' addresses:

Joseph V. Baranski, Defense Research and Development Canada—Toronto, 1133 sheppard Ave. West, P.O. Box 2000, Toronto, Ontario M3M 3B9, Canada.

William M. Petrusic, Department of Psychology, Carleton University, Ottawa, Ontario, Canada.

DOI: 10.1002/bdm

UNCLASSIFIED

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 document, 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)

Aggregating Conclusive and Inconclusive Information: Data and a Model Based on the Assessment of Threat (U)
(U)

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

Joseph V Baranski; William M. Petrusic

5. DATE OF PUBLICATION (Month and year of publication of document.)

July 2008

6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)

21

6b. NO. OF REFS (Total cited in document.)

50

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

Scientific Literature Published in: Journal of Behavioral Decision Making 23: 383–403 (2010)

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 SL 2008−139

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 AVAILABILITY (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

UNCLASSIFIED

UNCLASSIFIED

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) This study examined the process of combining conclusive and inconclusive information using a Naval threat assessment simulation. In the present context, inconclusive information refers to data that is relevant but does not clearly support a choice alternative at the moment of decision. On each of 36 trials, participants interrogated 10 pieces of information (e.g., speed, direction, bearing, etc…) about ‘targets’ in a simulated radar space. The number of hostile [n(H)], peaceful [n(P)], and inconclusive [n(I)] cues was factorially varied across targets. Three models were developed to understand how inconclusive information is used in the judgment of threat. According to one model, inconclusive information is ignored and the judgment of threat is based only on the conclusive information. According to a second model, the amount of dominant conclusive information is normalized by all of the available information. Finally, according to a third model, inconclusive information is partitioned under the assumption that it equally represents both dominant and non−dominant evidence. In Experiment 1, the data of novices (i.e., civilians) were best fit by a model that assumes a partitioning of inconclusive evidence. This result was replicated in a second experiment involving variation of the global threat context of the scenario. In a third experiment involving experts (i.e., Canadian Navy officers), the data of half of the participants were best fit by the partitioning model and the data of the other half were best fit by the normalizing model. In Experiments 1 and 2, the presence of inconclusive information produced a “dilution effect”, whereby hostile targets were judged as less hostile and peaceful targets were judged as less peaceful, than the predictions of the Partitioning model. The dilution effect was not evident in the judgments of the Navy officers.

(U) L’étude se penche sur le processus de recoupement de renseignements probants et non probants au moyen d’une simulation de l’évaluation de la menace navale. Ici, « renseignements non probants » s’entend de données qui sont pertinentes mais qui n’appuient pas clairement une option ou une autre dans l’immédiat. Lors de chacun des 36 essais, les participants se sont penchés sur 10 éléments d’information (p.ex. vitesse, direction, cap) concernant des « cibles » repérées sur un espace radar simulé. Le nombre de signaux hostile [n(H)], de signaux pacifiques [n(P)], et de signaux non probants [n(I)] ont fait l’objet d’une variation factorielle. Trois modèles ont été élaborés pour tenter de comprendre dans quelle mesure des renseignements non probants sont utilisés pour se prononcer sur une menace potentielle. Suivant le premier modèle, les renseignements non probants ne sont pas pris en compte, c’est à dire que le jugement posé concernant la menace ne se fonde que sur les renseignements probants. Suivant le deuxième modèle, le nombre de renseignements de la catégorie dominante est normalisé par rapport à l’ensemble des renseignements reçus. Enfin, suivant le troisième modèle, on prend pour hypothèse que les renseignements non probants sont répartis également entre les deux autres types de renseignements (lesquels peuvent être dominants ou non). Lors de l’expérience 1, les données de novices (civils) montraient un meilleur ajustement avec le modèle reposant sur la répartition des éléments non probants. On a obtenu le même résultat lors d’une deuxième expérience, où on avait modifié le contexte global de la menace dans le scénario. Lors d’une troisième expérience, à laquelle participaient des experts (officiers de la Marine canadienne), les données de la moitié des participants cadraient mieux avec le modèle reposant sur la répartition des renseignements non probants, alors que les données de l’autre moitié des participants montraient un meilleur ajustement avec le modèle reposant sur la normalisation. Dans le cadre des expériences

1 et 2, la présence de renseignements non probants a produit un « effet de dilution » : les cibles hostiles étaient jugées moins hostiles et les cibles pacifiques étaient jugées moins pacifiques, par rapport aux prévisions du modèle reposant sur la répartition des renseignements. Cet effet de dilution n’est pas ressorti des jugements posés par les officiers.

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) Inconclusive information; threat assessment; subjective probability; confidence; balance of evidence; dilution effect