

Risk Game: Capturing impact of information quality on human belief assessment and decision making

Anne-Laure Joussemme¹, Giuliana Pallotta², Jonathan Locke¹

¹ NATO STO Centre for Maritime Research and Experimentation (CMRE)

Viale San Bartolomeo, 400, La Spezia, Italy

{Anne-Laure.Joussemme,Jonathan.Locke}@cmre.nato.int

² Lawrence Livermore National Laboratory (LLNL)

7000 East Avenue, Livermore, CA 94550, USA, pallotta2@llnl.gov

Abstract

This paper presents the Risk Game, a general methodology to elicit experts' knowledge and know-how, in their ability to deal with information provided by different types of sources (sensors or humans) of variable quality, to take into account the information quality and to reason about concurrent events. It is a contrived technique capturing data expressing human reasoning features during a specific task of situation assessment. The information is abstracted by cards and its quality, which varies along the three dimensions of uncertainty, imprecision and falseness, is randomly selected by dice roll. The game has been played by experts of maritime surveillance, mostly marine officers from several nations. The Risk Game is domain-independent and can be designed for any specific application involving reasoning with multi-sources. The preliminary results obtained are promising and allow validating the efficiency of the elicitation method in capturing the link between information quality and human belief assessment. Besides the positive feedback collected from the players and their perceived effectiveness of the method, the data effectively capture the impact some specific information quality dimensions on belief assessment. We highlight, for instance, that the relevance of information perceived by the players may differ from the effective information relevance, that a high ratio of false information increases the uncertainty of the player before decision and may lead to wrong decisions, or that the context has a high impact on the decision made. Future extensions of the Risk Game are finally sketched.

Keywords: Gamification, Information Quality, Uncertainty, Information Fusion, Belief Assessment, Decision Making.

1 Introduction

Reaching an adequate level of Situation Awareness (SAW) for an informed and confident decision not only requires processing information of various types (numerical values, natural language statements, objective or subjective assessments, etc) and dealing with imperfect information (uncertain, imprecise, conflicting, doubtful, ambiguous, etc) but requires also understanding and taking advantage of the operational context. Moreover, it requires processing information across different levels of semantic content, ranging from lower levels of processing (e.g., contact and target detection, target tracking, classification) to higher levels of processing (e.g., activity recognition, behavior analysis, target identification, threat and impact assessment), while considering user's needs and context [1]. Processing information which can be irrelevant and thus distracting is highly demanding for the human

operator who is often under stress due to time constraints. All these factors have a measurable impact on the decision in terms of timeliness, confidence or correctness.

The impact of information quality on decision quality has been studied in different domains such as finance [2] or enterprise performance [3],[4]. Within these works, different information quality dimensions are considered and their impact on the decision maker is highlighted. For instance, the quality dimensions of *accessibility*, *accuracy*, *completeness*, *interpretability* are used in [4]. Only a few works study the impact of information quality on situation and threat assessment in the maritime domain, but it is worth mentioning the *Data Fusion Levels Two and Three Workshop* sponsored by the Office of Naval Research (ONR) and held in Arlington, VA in 2005. The summary paper [12] describes “how information pedigree is used to support and enhance situation and threat assessment”, considering dimensions such as *reliability*, *confidence*, *trust*, and their integration in automation of sensor-data processing. In [4], the authors follow Wang’s Total Data Quality Management (TDQM) cycle [14], which implements four components: (1) Identification of information quality dimensions, (2) Measurement, providing information quality metrics, (3) Analysis, identifying causes for problem and computing the impacts of poor information quality, and (4) Improvement, providing techniques for improving information quality. The approach proposed in this paper follows the steps of [14] and is comparable to the one proposed in [3], while the impact of information quality on decision quality is studied along the quality dimensions of *falseness*, *imprecision*, *uncertainty* and *relevance*.

Such studies rely on different methodologies: for instance in [13], semi-structured interviews of Marine Protected Areas (MPAs) experts are conducted and their interpretation of different attributes is analysed. Hoffman *et al.* [15] distinguish between three categories of elicitation techniques: “(1) analysis of the tasks that experts usually perform, (2) various types of interviews and (3) contrived tasks which reveal an expert’s reasoning processes without necessarily asking about these processes”. The Risk Game presented here falls under the third category. Using games for Knowledge Elicitation (KE) has the advantage, over structured or semi-structured interviews (e.g., [13]), of being more engaging for the experts and less time consuming [15], [16]. The effectiveness of serious games to enhance the risk management process has been recently reported in [5]. A study published in 2018 [5] describes a serious game for natural risk assessment where an interactive interface is designed and some indicators (such as motivation and immersion) are used to assess the players’ feedback. A very recent survey [7] provides an interesting discussion on serious games role in learning disaster risk management. Serious games are also an efficient means to highlight reasoning biases as reported for instance in [8] and [9]. Besides the advantages of interactive-based game approaches reported in the literature (e.g., effectiveness in learning performance and satisfaction [10], immediate feedback, active participation [11]), this kind of approach coupled with facilitation allows a rich lively verbal exchange with experts very keen to share their knowledge and experience, and in the case of the Risk Game, explaining intermediate belief states (doubts, certainties) or justifying information queries.

The main research question addressed in this study is whether serious games are an efficient means to capture the impact of information quality on human belief assessment and decision making. A secondary question is which information quality dimension impacts belief assessment. To this end, we hypothesise that information quality can be reduced to the three basic information quality dimensions of uncertainty, imprecision and falseness. We also assume that human belief assessment is impacted by the type of the source of information and the nature of information (i.e. attribute), while the decision depends on the operational context and associated risk.

The Risk Game presented in this paper is a gaming approach to elicit experts’ knowledge and know-how in processing heterogeneous information (from sensor measurements to human statements), considering information and source quality and reasoning about concurrent events. It is a contrived technique in the sense of [15], analysing the quality dimensions of uncertainty, imprecision and falseness. The Reliability Game presented in [18] targets the specific aspect of source quality, leaving the information quality dimension fixed. Contrary, the Risk Game explicitly varies the information quality dimensions and the source

quality is a latent variable. Further analysis of the structured data gathered during experiments would contribute to the development of automated algorithms for an improved synergy with the human operator. Taking a “game with a purpose” approach [17], the Risk Game is aimed at capturing data expressing human reasoning capabilities while performing a specific task. The scenario-based design approach makes the Risk Game application independent and easy to develop several versions of the game suited to particular applications. In this paper, the maritime surveillance problem is tackled, but intelligence applications can be considered as well, in support to the intelligence analyst hypotheses evaluation task, medical diagnosis or crisis management.

The paper is organised as follows: In Section 2 we present the Risk Game design and detail the methodology developed. Section 3 presents the players’ feedback gathered through a survey at the end of the exercise. The experimental setting involving the players together with some exploratory data analysis is presented in Section 4. Section 5 summarises some findings and sketches some research avenues to be further explored.

2 Game design

The Risk Game has been designed to validate the hypothesis that Information Quality (IQ) has an impact on Belief Assessment (BA), Risk Assessment and Decision Quality. We present in this section a methodology which allows (1) to elicit experts’ knowledge and know-how about belief and risk assessment and decision making, in an entertaining way, (2) to gather data which, once analysed, would support or refute our hypothesis and (3) to characterise and measure such an impact.

2.1 Gameplay

A set of players has first been selected, mostly experts in maritime surveillance and threat assessment, who played individually and successively (see Figure 1) the designed Risk Game following the chronology of the gameplay:

- The player plays the role of the Officer of the Watch (OoW) who has to assess a vessel behaviour whose track has been lost one hour ago;
- The player has two candidate tracks (corresponding to vessel A and vessel B) that could correspond to the unknown lost vessel, with approximate positional information: One track is located in the Area Of Responsibility (AOR), the other one is in the country of origin of the lost vessel; the player **does not know** that that in the scenario of the game, Vessel A is the missing vessel and thus should be assessed as a threat. Vessel B is a fishing vessel from outside the AOR which is going back to its port. The added difficulty in the threat assessment is that the two vessels are of the same type.
- The player should follow the following general reasoning: If the track in the AOR corresponds to the lost vessel, then it means that the lost vessel crossed the border without authorisation and should be considered as a potential threat; however, if the other track, not in the AOR, corresponds to the lost vessel, then that means that the lost vessel is still in its area and there is no threat;
- The player successively selects pieces of information (PoIs) of randomised quality, about the two observed and unknown tracks to assess which track corresponds to the lost vessel;
- After discovering each PoI, the player is asked to assess which of the two tracks corresponds to the lost vessel, under the form of a belief degree, and fill the corresponding assessment form accordingly;
- At any time during the Risk Game (i.e., after having examined an arbitrary PoIs), the player can take the final resolution to either send or not to send a patrol aircraft to inspect the vessel. This decision marks the end of the game. The ultimate goal of the

Risk Game is to take the best decision about a threat by using the minimum number of PoIs.



Figure 1. *A player assessing the threat from an unknown track in the vicinity of the port, based on information previously queried*

The game mechanics are articulated around on three main elements of (1) a scenario describing a story and settling the player in a specific role with specific task in a specific context, (2) information cards abstracting pieces of information which are familiar to the domain experts and are usually displayed on screens or provided by operators, (3) a belief gathering method to capture experts' sequentially updated belief degrees. The independent variables cover a wide range of information quality dimensions among which three (falseness, uncertainty and imprecision) are controlled and randomized by dice roll. Their impact on experts' belief degrees and decision made (dependent variables) are captured.

The board version of the Risk Game presented in this paper is attractive and playful thanks to (1) a colored board and a set of colored cards which allow a nice visual discovery phase of the game, (2) the actions from the player of rolling the dice, selecting cards and rating the belief degrees which keep the player engaged, (3) the sometimes challenging reasoning task which forces the player to maintain the focus on the task and finally (4) an interactive facilitation which eases the dialogue with the scientist and makes pleasant and efficient the sharing of experience and knowledge.

2.2 Scenario-based game design

A short scenario description with some context is given to the players, who will be actors in the story, with the role of decision maker. The vignette developed is part of the larger general scenario which takes place in the port of Herosé (Figure 2(a)) at the border between CentreLand (the AOR) and RightLand (to which the lost fishing vessel belongs), after a period of crisis and the Harbour Protection Level is set to TWO (over a scale of three levels). Information about the lost vessel is provided to the player as it would be the case from typical vessel databases: Name, Type, Subtype, Length, Width, Flag together with its last Location, Heading and Speed are provided (Figure 2(b)).

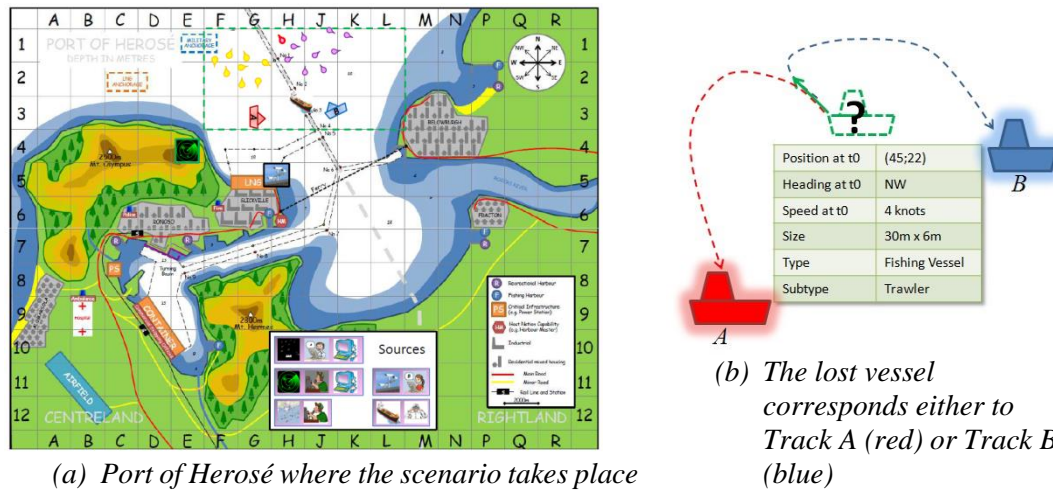


Figure 2. Board and scenario for the Risk Game

2.2.1 Underlying reasoning

Let us denote by *A* the event (or proposition) “The lost vessel corresponds to Track A” (in CentreLand) and by *B* the event “The lost vessel corresponds to Track B” (in RightLand). The two events are exclusive since the lost vessel cannot be in both CentreLand (event *A*) and RightLand (event *B*). Thus:

$$A \wedge B \vdash \perp \tag{1}$$

where \wedge denotes the logical conjunction, and \perp denotes the logical contradiction. Equation (1) reads: “Event *A* AND Event *B* cannot be both true”. We however do not exclude the possibility that neither *A* nor *B* is true meaning that the vessel may have left the area under surveillance. The exhaustivity of the events is referred in [19] as the *open-world* assumption and writes:

$$A \vee B \not\vdash \top \tag{2}$$

where \top is the logical tautology.

Any evidence for *A* should either at worst decrease the belief for *B*, i.e. provide some evidence that the lost vessel is Vessel *B*, or at least should not provide any information. Equivalently, any evidence toward *B* should either at worst decrease the belief for *A*, or at least should not provide any information. For instance, if the player learns (based on some information the player is provided with) that Track *B* does not correspond to a fishing vessel, the player’s belief that Track *A* is the lost vessel should increase. Let us denote by *Bel(A)* the belief degree assigned by the player towards the event *A*. It is known that human belief assessment does not necessarily satisfy probabilistic coherence (see for instance [20]), meaning that the belief degrees towards events *A* and *B* may not sum up to 1: They can be lower than, higher than or equal to 1. In the case the player follows a probabilistic assessment, the coherence principle under the *closed-world* assumption imposes that *Bel(A)+Bel(B)=1*. Under the open-world assumption, the same principle would be stated as *Bel(A)+Bel(B)≠1* meaning that another event not considered here (i.e. the vessel is outside the area) would be assigned the complementary degree to 1.

To respect the natural non-additive belief assessment of humans while considering both open- and closed-world assumptions, we allow the player a complete freedom in belief degree assessments about the two events. The belief degrees elicited about events *A* and *B* are thus free from any probabilistic interpretation and thus do not need to sum up to 1.

2.3 A domain-independent methodology

The scenario-based game design allows the flexibility to customise the game to suit any specific application and domain. The key design steps are the following:

1. Design of the scenario with help of domain experts, to target a relevant problem and ensure the realism.
2. Identification of the meaningful sources of information together with their characteristics.
3. Definition of the different pieces of information under the specific format of the game (see Section 2.6) with suitable quality variations.
4. Printing of the board and cards.
5. Selection of a representative set of experts.
6. Play the Risk Game.

The Risk Game methodology can thus be applied to any domain where the problem of aggregating (i.e. fusing) pieces of information under uncertainty appears challenging. For instance, in intelligence applications the aggregation of experts opinions expressed with uncertainty is rather known as “multi-INT fusion”. Some methods such as the Analysis of Competing Hypotheses (ACH) are aimed at supporting the intelligence analyst in evaluating multiple hypotheses provided by different sources, considering some quality dimensions such as credibility or relevance. Other applications involving decision making under uncertainty such as medical diagnosis or crisis management can be targeted by a Risk Game and benefit from the outcomes of the data analysis.

2.4 Attributes

Each vessel is described by a series of five attributes: LOCATION, SPEED, HEADING, TYPE, SIZE. These attributes have been selected to cover different types of scales (continuous, discrete, nominal, etc). Moreover, some are independent (e.g., LOCATION and SPEED) while others are dependent (e.g., SIZE and TYPE).

We define the a priori effective relevance¹ of information regarding the attribute only, ending up with the qualitative ranking and further categorise the attributes as behavioural: LOCATION, SPEED, HEADING and classification: TYPE, SIZE, SPEED. The task of the player is to discover “where” the suspect vessel is, or other said which track it corresponds to. It is thus essentially a classification problem where the player needs to *match* queried information with the corresponding attributes know about the lost vessel. The behavioural attributes are thus less relevant to this task, and the most relevant (i.e., discriminating) attribute is the SIZE (both length and width).

2.5 Sources of information

Six sources of information have been selected to cover the diversity of commonly available maritime surveillance sources, ranging from hard (physical sensors) to soft (humans) sources, providing either subjective or objective information, either in a numerical or qualitative (using natural language) format. Moreover, sources are a combination of an *information container* and an *information provider* [21], which allows distinguishing between automatic processors (such as trackers or classifiers) and human analysts processing the same initial

¹ Basing the relevance assessment on the attributes only does not cover the information content. Indeed, a piece of information of very low informational content is not relevant as it would barely impact the previous belief state.

signal or image. Sources are further characterised according to their expertise defined by the list of attributes about which they are able to provide information [21], among the list of the five basic attributes. For each track, each attribute is reported by either two or three distinct sources, providing the player with both some redundancy and complementarity in information received. Table 1 lists the six sources of information, their expertise (in terms of the attributes about which they are able to provide information) and their range (in terms of the track covered, e.g., VA, VB or both). No specific information about the sources' quality (in terms of reliability) is provided but the player is expected to rely on prior own knowledge (or perception) about this aspect and to process the information accordingly.

Table 1. Sources coverage and expertise

Sources		Range	Expertise				
Container	Provider		LOCATION	HEADING	SPEED	SIZE	TYPE
Radar	Operator C	VA and VB		X		X	
	Tracker/ATR		X		X		X
SAR	Analyst A		X	X			
	ATR				X		
Camera	Analyst B	VA only	X	X	X	X	X
Cargo	Captain	VB only	X	X	X	X	X

2.6 Information quality dimensions

The comprehensive consideration of all Information Quality Dimensions (IQDs) (see for instance [22], [23], [24]) is beyond the purpose of the Risk Game, and we focus on the three dimensions below assumed to have a possible impact on the reasoning and decision processes and that we consider as basic and independent:

- IDQ1. **Trueness/Falseness** - Refers to the “true” (or reference) value a vessel attribute and is a synonym of “correctness”;
- IDQ2. **Certainty/Uncertainty** - Refers to a degree of confidence or belief assigned to a specific (or set of) value to be “true”. Its cause can be either a lack of knowledge (epistemic uncertainty) or the random variability of the underlying process (aleatory uncertainty). When assigned by the source itself it may be called “self-confidence”;
- IDQ3. **Precision/Imprecision** - Refers a set of possible values: The smaller the cardinality of the set (or the length of the interval), the higher the precision. It reflects the inability of the source to provide a single value or to discriminate between several values and is a synonym of “nonspecificity”.

On the one hand, imprecision (or precision) and uncertainty are opposed [23]: “I’m certain that the speed of the vessel is between 3 and 6 knots” (imprecise but certain statement) versus “I’m not certain that the speed of the vessel is 5 knots” (precise but uncertain statement). On the other hand, precision and trueness are often associated in performance assessments of systems, gathered under the notion of accuracy. These three IQDs will be our basic independent variables of the Risk Game.

In addition to the three above-selected IQDs, the following four dimensions will also be considered, assuming they could also have an impact on the reasoning and decision processes. They are latent variables as they result from some variations of the observed (and controlled) variables IDQ 1 to IDQ 3:

- IDQ4. **Conflict** - Relative to some inconsistency between the different pieces of information. Indeed, because some pieces of information may be **false**, they may conflict with others;
- IDQ5. **Relevance** - Relative to the output of the reasoning process in general. It can be relative to the goal (here the decision to be made), or to the belief state of the

player. Understood as informational content, relevance directly arises from **imprecision** and **uncertainty**. Relevant information impacts previous belief, or is helpful to make the decision;

IDQ6. **Reliability** - Relative to how much a source's statement can be relied on, which could be derived from the ability of the source to provide **correct (true)** information (previously assessed or known);

IDQ7. **Independence** - Relative to the correlation of the sources (i.e., if the output of one source influences the output of another one), and relative to the correlation of attributes (e.g., TYPE and SIZE are dependent **attributes**).

The relevance of the input pieces of information will not explicitly vary, but will vary implicitly with the type of attribute and the uncertainty or imprecision of the piece of information. For instance, information about the location is less relevant to identify the vessel than information about its type. Similarly, the sources' reliability will not be explicitly specified but due to their diversity, the sources exhibit different reliability degrees: For instance, the camera analyst may be considered more reliable than the cargo captain, or the radar providing the location may be considered more reliable than the analyst of the SAR imagery.

Moreover, some of the attributes reported are not independent from each other. In particular, the instant components of LOCATION and HEADING can reasonably be considered independent from each other and from the other attributes while SPEED, SIZE and TYPE are dependent. Figure 3 displays the different variable of the Risk Game and their dependencies.

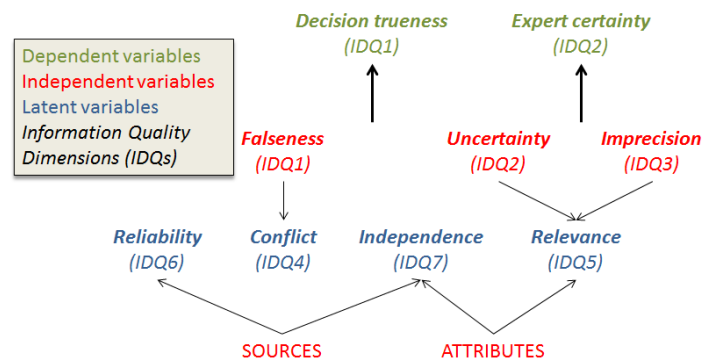


Figure 3. *Independent, latent and dependent variables of the Risk Game. Seven quality dimensions (italic) are considered, three are independent variables (red, italic) controlled randomly through dice roll, four are latent (blue, italic), derived through other two independent variables (Sources and Attributes) controlled by the players' queries, and two are the dependent variables (green, italic) which are measured during the experiment*

Note that a suitable formalisation of the reasoning process supported by a dedicated mathematical framework to reasoning under uncertainty would allow the measurement of these quality dimensions. This is however out of the scope of the current paper and will be addressed in future work.

2.7 Levels of information quality

For the sake of simplicity (i.e., to keep the number of combined quality levels tractable), we consider only two possible states for the three basic IQD variables: Pieces of information are either perfectly true, precise, certain, or moderately false, imprecise, uncertain, leading to a set of eight possible combinations:

- **Falseness** is a shift from the true value but is always in the range of acceptable mistakes so that it is never obvious that the PoI is actually false. For instance, no source reports that the observed vessel is of length 150 meters or that it is a Ferry;
- **Uncertainty** scale follows the Standardised Lexicon used by the National Intelligence Council in US, and we consider only two values, *likely* and *almost certain*, to which we assign the two corresponding numerical degrees of 0.6 and 1: Soft sources (humans) report phrases and hard sources (sensors or algorithms) report numerical values for an equivalent meaning. For instance, “likely” exactly means “confidence of 0.6”;
- **Imprecision** is represented either by intervals for the numerical attributes (SPEED, LENGTH, WIDTH) or sets of values for nominal and discrete attributes (LOCATION, HEADING, TYPE). Moreover, some soft sources report fuzzy statements (Small, Medium, Large for the SIZE for instance).

Table 2. *Eight information quality levels and corresponding randomisation*

False	Imprecise	Uncertain	Randomisation
0	0	0	0.11
0	0	1	0.22
0	1	0	0.22
0	1	1	0.11
1	0	0	0.06
1	0	1	0,11
1	1	0	0.11
1	1	1	0.06

The information quality the player obtains is randomised over the eight versions of the same piece of information through a dice roll, as displayed in Table 2. The best and worst quality levels are assigned a lower probability. Also, a higher probability is assigned to the true pieces of information either only imprecise or only uncertain. This probability distribution is certainly a bit pessimistic regarding sources’ performance in general. However, the global low quality of information is explained by (1) the rough sea state and (2) the high distance between the sources and the vessels, relative to their range.

2.8 Information cards and fusion

A piece of information (PoI) is denoted by f and is represented by a tuple:

$$g = \langle s, v, a, x, i, u, f \rangle \quad (3)$$

where s is the source providing the piece of information, v is the vessel (or track), a is the attribute, x is the estimation or measurement of a for v by s , i is the imprecision of g , u is its uncertainty and f is its falseness. For instance,

$$g = \langle (\text{Radar}, \text{Tracker}), VA, \text{SPEED}, 4\text{knots}, 0, 1, 0 \rangle \quad (4)$$

reads: “The radar with its associated tracker estimates the speed of Vessel A at 4 knots with a confidence of 0.6”, and this information is true. The last element of trueness/falseness is obviously not provided to the player and known only to the game facilitator.

Pieces of information are abstracted through *information cards* available for query to the player. Each card is uniquely identified by a coded ID.

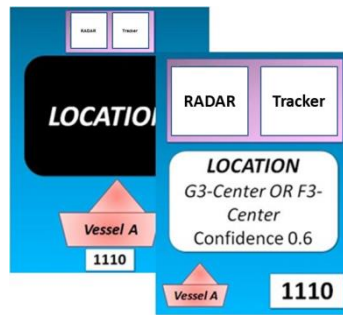


Figure 4. Example of information card abstracting a piece of information (PoI) about the location of Vessel A provided by the tracker processing the radar signal

Figure 4 is an example of an information card abstracting a piece of information about the location of Vessel A provided by the tracker processing the radar signal.

First, only the back of the card (black) is visible to the player. Then, the player selects (1) the track to query, (2) the attribute about which to obtain information and (3) the source providing information. Thus, at each round, the player has 3 degrees of freedom to select the piece of information (i.e. the card): The track (or vessel), the attribute and the source. These point towards a single pile of (8) cards from which information quality varies. After rolling two dice, the player is given the card of corresponding quality by the game facilitator and can read the information content (white part). The player is now ready to rate the belief degree towards events A and B.

The information processing to be performed by the player to reach suitable situation awareness can be formulated as fusion task:

$$g = \oplus_n(g_n) \tag{5}$$

where g_P would represent information aggregated by the player P from the individual pieces of information g_n queried and \oplus denotes some aggregation operation used by the player to derive g_P from the set of g_n s. g_n is the piece of information under the form of Eq. (4). g_P can be further synthesised as a pair (Bel(A); Bel(B)) of belief degrees towards events A and B respectively.

2.9 Gathering belief states and decision making

In practice, the Officer of the Watch would have all this information (i.e., the five attributes values for the two vessels) available at the same time, possibly displayed on several screens. In the Risk Game we would like to decompose the reasoning process to the granularity level of individual pieces of information, in order to track the successive belief states of the player while discovering and processing the different pieces of information. It is thus explained to the player that time is fixed, the situation does not evolve, and the time taken for processing information does not matter.

Player:												
	Belief that MV is Vessel A					Belief that MV is Vessel B						
# POI	0	0.2	0.4	0.6	0.8	1	0	0.2	0.4	0.6	0.8	1
1327		X								X		
2415			X									
1427					X		X					
2113					X			X				
Decision	Send the aircraft					X	Do not send the aircraft					

Figure 5. Belief assessment form to be filled in by the player after discovering each piece of information. “MV” stands for “Missing Vessel” (or lost vessel)

Once the player is given the information card by the facilitator and has read the message, the belief can be assessed about events *A* and *B* through the form displayed in Figure 5. A scale between 0 and 1 with steps of 0.2 is proposed to the player. The player is expected to rate both events at each time round, regardless the vessel queried, but is not reminded to do so which may result in missing data. One row corresponds to one game round: the first column corresponds to the ID of the card which captures the track, attribute, source and the information quality (independent variables), the red section captures the player’s belief degree about event *A* and the blue part about event *B*. A cross indicates the player’s belief state, the dependent variable. The player can select as many cards desired (up to the 26 cards available) and when ready to decide, can check the corresponding box at the bottom of the form and the game is over. The set of forms filled by the players and collected during the Risk Game constitutes the dataset to be analysed in Section 4.

3 Methodology assessment

A rehearsal session with some experts allowed validating the scenario, the set of sources as well as the set of pieces of information, making the game quite realistic. The players of the rehearsal were generally happy with the “draft” design of the game, quickly caught the rules, and validated the story and sources reports. They were committed at the beginning of the game and quite excited during the play.

A total of 32 players (different from the rehearsal) selected as experts in maritime surveillance and threat assessment played the Risk Game during a Table Top eXercise (TTX) held at the NATO Centre for Maritime Research and Experimentation (CMRE). A survey was conducted in which the players were asked to assess the Risk Game according to the criteria of *Realism*, *Operational relevance*, *Understandability*, *Engaging ability*, *Elicitation efficiency* and *Facilitation*. Figure 6 displays the players’ feedback, the darker the purple color, the better the feedback. In the light of this survey, the players generally found the Risk Game either *extremely realistic* or *very realistic* (60%) and either *extremely* or *very operationally relevant* (80%). They all found the elicitation method *effective* while 72% of them found it either *extremely* or *very effective*. They also all understood the purpose of the game only 20% moderately understood it. Most of them (96%) found the game at least *moderately engaging*, while 88% found it either *extremely* (44%) or *very engaging* (44%). Finally, they generally appreciated the facilitation (92% found the game either *extremely* or *very well facilitated*).

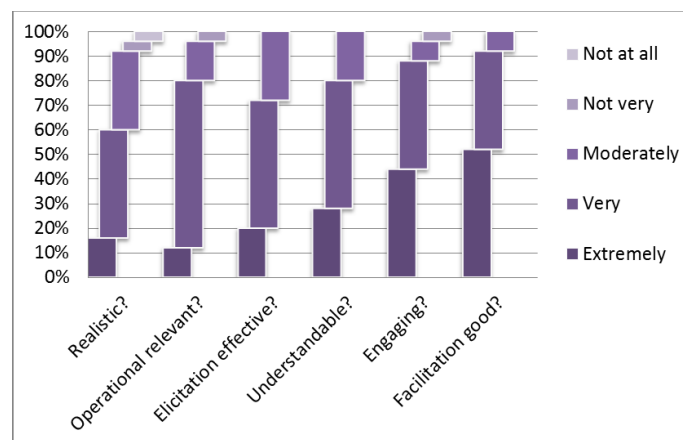


Figure 6. Players’ feedback after the Risk Game

The outcomes of this survey demonstrate first, that the players generally greatly enjoyed playing the Risk Game and had a very general positive impression afterward. In this respect, it makes the Risk Game “efficient” as the engagement of experts is required for gathering meaningful data. Second, this survey also shows that the players positively perceived the efficiency of the elicitation method. Although this feedback is encouraging, this perception needs to be validated by the data gathered, which is the purpose of the next section.

4 Exploratory data analysis

We provide below some results from the exploratory analysis of the data collected during the game, in the aim of demonstrating the ability of the Risk Game to capture links between some information quality dimensions and human belief assessment and decision making. Due to the high number of variables of the game for information queries, information quality and decision quality, we provide below examples of the kind of analyses that can be derived from the Risk Game, while an exhaustive analysis of the data is left for future work. The results and analyses presented here are thus a subset only of the interesting and relevant ones. Nevertheless, they cover the main aspects of information quality and its impact on decision making.

We first investigate the queries from the players which illustrate their information needs (Section 4.1), followed by the final belief states and decision made (Section 4.2) and finally information quality (Section 4.3).

4.1 Analysis of queries

The players had four degrees of freedom to query the information: The track (or vessel) A or B, the attribute (LOCATION, HEADING, SPEED, TYPE and SIZE), the source (Radar, SAR, Camera and Cargo), and the number of queries. Also, they had a complete freedom in the order of their queries. The analysis of the players’ query behaviour is meaningful to highlight their information needs, their perception (or a priori knowledge) of sources’ quality and attribute relevance to the problem, as well as their reasoning strategy.

4.2 Information needs

Each player selected a variable number of cards as displayed by the histogram of Figure 7. The average number of POIs is 13, with a standard deviation of 6 queries. There is considerable variability in the number of queries, some players selecting a very few POIs while other selected all available ones. If we compute artificially the time of reaction to the number of POIs retrieved and processed, it is interesting to note that the “quickest” and the “slowest” players made the same decision that is to send the patrol aircraft, by querying 3 and 26 POIs respectively. The decision to send the patrol aircraft dominates the entire range of numbers of POIs.

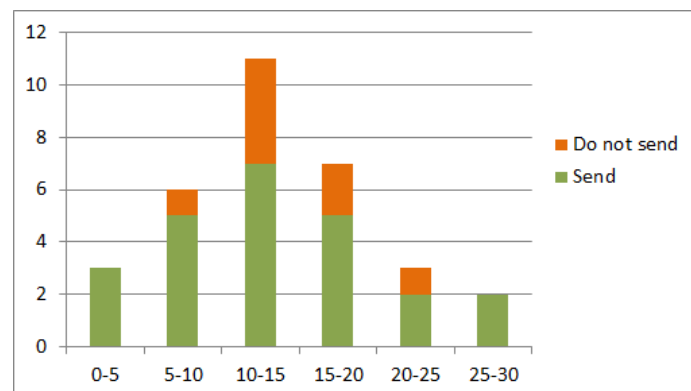


Figure 7. *Number of queried pieces of information based on the final decision made*

Table 3 shows the ratio of the players which queried either Vessel A or Vessel B first (left column) or most (right column).

Table 3. *Ratio of vessels queried by the players*

	First	Most
Vessel A	72%	84%
Vessel B	28%	12%

Unsurprisingly, Vessel A appears to be the priority vessel to the players (first queried for 72% of the players, and most queried for 84%), as indeed it is the one in the area of responsibility.

The relevance of information perceived by the players can be estimated through the analysis of the attribute queries. The ratio of the first and most queried attribute by the players regardless the vessel is displayed in Table 4. The number of queries has been normalised to the number of available PoIs, either 3 or 2 (see Table 1).

Table 4. *Ratio of attributes queried by the players*

	First	Most
LOCATION	81%	53%
SPEED	3%	6%
HEADING	13%	6%
TYPE	3%	28%
SIZE	0%	6%

We observe that the location was perceived as the most relevant attribute by the players (both first and most queried), although it was the least relevant to solve the identification problem (see Section 2.2). Even if the players had initially a rough idea of the location of each vessel, it seems that a more precise location was needed, maybe to better visualise the scene, to better appreciate the closeness of the vessel to the coast. We do not have so far data to support any of these hypotheses but only the feedback of the players after the game to help clarify this behaviour. Table 5 displays the ratio of queried PoIs of the two categories of attributes: The category “behavioural” corresponds to the information about LOCATION, HEADING and SPEED which help clarifying the vessel behaviour, while the classification category includes SPEED, SIZE and TYPE.

Table 5. *Ratio of attribute categories queried by the players*

	Most
Behavioural	60%
Classification	40%

This table relates the effective relevance of the attributes (as defined in Section 2.2) and the perceived relevance if we consider that the players would query more frequently the information perceived as relevant. Most queries focused on some attributes relevant to anomaly detection and threat assessment in general (i.e. behavioural attributes) but less relevant to the purpose of the game, which was rather a classification problem (i.e. the classification attributes would have been more helpful to discriminate between the two vessels).

We assume that the expected reliability of sources can be estimated by the first and most queried source among Radar, SAR, Camera and Cargo, as displayed in Table 6. These results provide an idea of the trust or confidence of the players in the different sources. The radar appears as the most trusted source (first and most queried). The cargo captain and the camera analyst received significant interest as well, as it appears that human information was very valuable to the players.

Table 6. *Ratio of sources queried by the players*

	First	Most
Radar	63%	72%
SAR	13%	22%
Camera or Captain	26%	6%

However, these results must be analysed carefully since the source query is not independent from the attribute query (i.e., not all the sources provided information about all the attributes). Indeed, the study could be refined to consider couples (source, attribute) as it became clear during the game that the players rely on a source to provide a given attribute and on another for another one.

4.3 Query strategies

Based on the way the players queried the information, we observed different reasoning strategies (without being able to judge about their respective quality, though). As an example, Figure 8 shows the ratio of switches in the queries between Vessels *A* and *B*, either from *A* to *B* or from *B* to *A*, relatively to the maximum number of possible switches. A null ratio means that the player queried a single vessel (either *A* or *B*), thus never switched. A low ratio means that the player mostly queried one vessel and then the other one. A high ratio means that the player systematically queried one vessel and right after the other one, demonstrating a reasoning strategy by comparison. This later group of players probably took advantage of the evidences against one event to increase their belief for the other one. It is still very difficult however to confirm this hypothesis in the light of the data gathered.

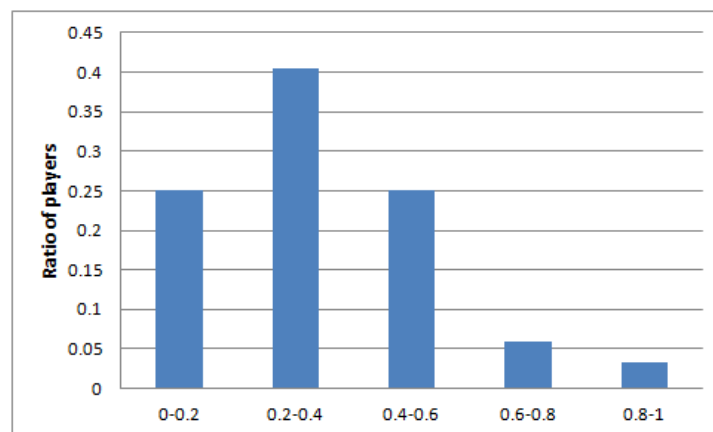


Figure 8. *Ratio of switches in queried information between Vessel A and Vessel B*

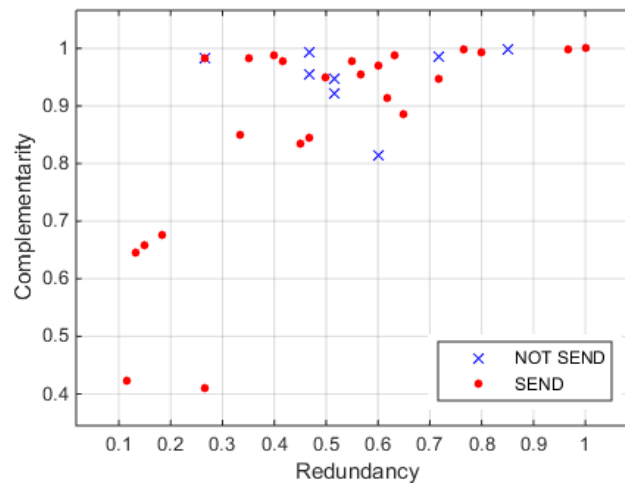


Figure 9. *The redundancy/complementarity values grouped based on the decision made*

Because all attributes are available from either two or three sources, the players have some control on the degree of *redundancy* and *complementarity* of the information gathered. An empirical way to estimate the need in redundant information is to count the number of queries toward the same attribute from different sources (for a given vessel). Analogously, an empirical way to measure the level of the complementarity in information was derived by computing the number queries on different attributes (for a given vessel). These two dimensions of redundancy and complementarity are jointly plotted in Figure 9. Each dot represents a player, together with the final decision made (blue cross or red bullet). The need for redundancy however may be explained by a “bad” piece of information, conflicting with previous ones, which requires further investigation. On the other hand, complementarity is needed to cover the diversity of attributes and that is what most of the players were looking for. We observe that most of the players adopted a strategy based on a high level of complementarity (i.e. (0.8, 1)) and a medium level of redundancy, giving priority to multi-attribute investigation.

4.4 *Belief states and decision*

Two kinds of output were recorded from the players: (1) Their sequential belief states relative to both events *A* and *B* and assessed after each step *n* of information discovery and (2) the decision to send or not to send the patrol aircraft once they felt they knew enough about the situation. Figure 10 provides a sample of the data recorded for two players. Player #17 (Figure 10(a)) systematically assessed both events (no missing assessment), symmetrically rated events *A* and *B*, and while initially generally leaned toward event *A* (with a step of high uncertainty), finally changed mind to reach a high belief degree toward *B*. Player #8 (Figure 10(b)) generally assessed only one of the two events, went through a phase of high confidence for event *B* to finally reached a state of high uncertainty.

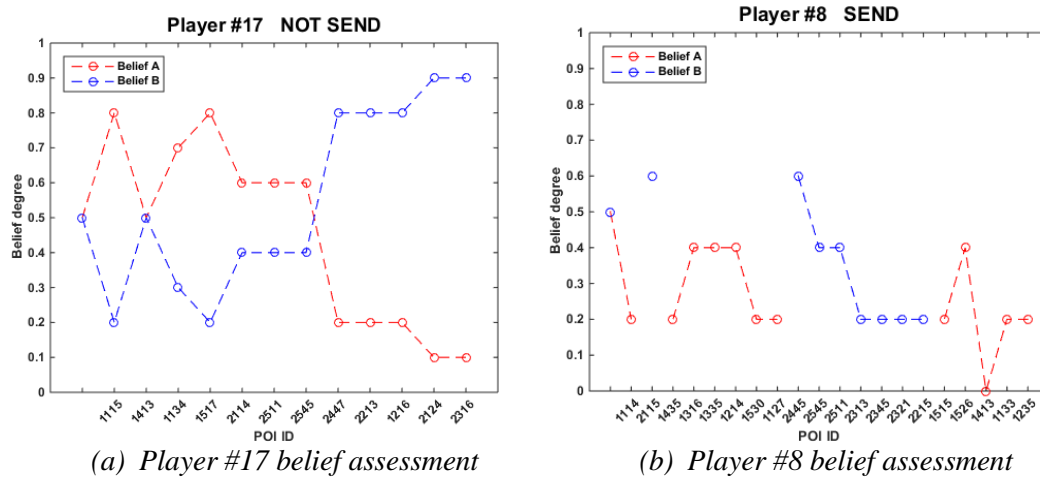


Figure 10. Two examples of sequential belief assessments for both events A (red) and B (blue) for two players. The ID of the card is display on the x-axis.

Because of the binary problem (either A or B), we will consider only trueness and certainty (ignoring imprecision possibly expressed by the player). Trueness is simply estimated relative to the known correct track, while certainty is assessed through the last belief state before decision. If we denote by $Bel_n^{(i)}(A)$ and $Bel_n^{(i)}(B)$ the belief degrees of Player i at step n toward events A and B respectively, assuming these belief degrees range from 0 to 1, the certainty degree of the player will be defined by:

$$c^{(i)}(n) = |Bel_n^{(i)}(A) - Bel_n^{(i)}(B)| \quad (6)$$

so that $c^{(i)}(n) = 1$ means that the player is totally certain about one event or the other, and $c^{(i)}(n) = 0$ means that the player is totally uncertain (does not lean toward an event or the other). Figure 11(a) shows the final belief state (at step N) of each player. We partitioned then the set of players according to their final belief leading to three groups (see Figure 11(b)):

- **Group B**, denoted as G_B : Strong belief for event B, for $Bel_N^{(i)}(A) - Bel_N^{(i)}(B) \leq -0.4$. A negative value means that the player has a stronger belief for event B before deciding, i.e. low uncertainty toward B;
- **Group A**, denoted as G_A : Strong belief for event A, for $Bel_N^{(i)}(A) - Bel_N^{(i)}(B) > 0.4$. A positive value means that the player has a stronger belief for event A before deciding, i.e. low uncertainty toward A;
- **Group 0**, denoted as G_0 : High uncertainty regarding A and B, for $-0.4 < Bel_N^{(i)}(A) - Bel_N^{(i)}(B) \leq 0.4$. A null value means that the player has equal belief for both A and B, i.e. a high uncertainty.

As briefly described in Section 2, in our scenario, Vessel A was indeed the missing vessel and thus should have been assessed as a threat. Vessel B was another fishing vessel from RightLand of the same type but slightly smaller than the missing vessel, and it was going back to its port.

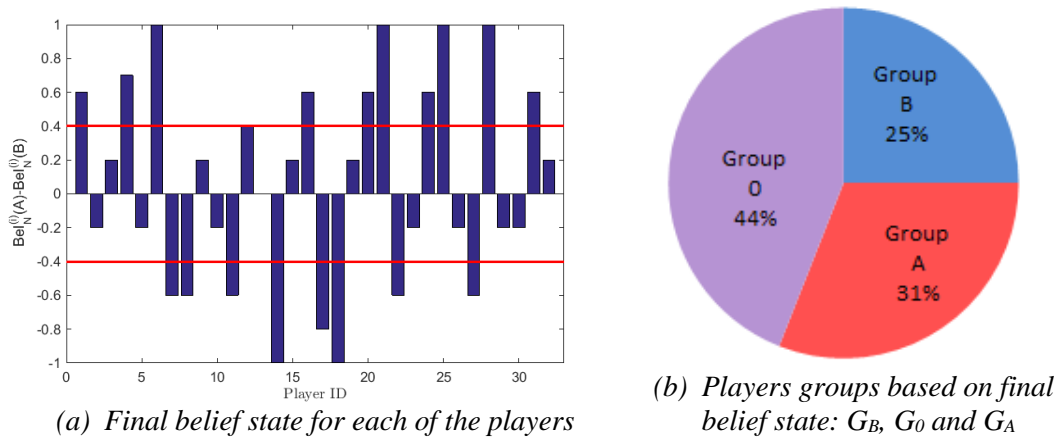


Figure 11. Final belief state before decision

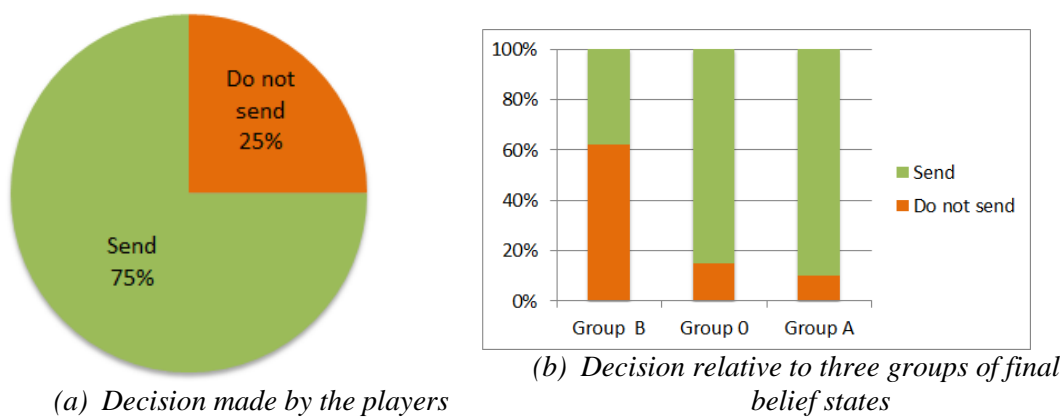


Figure 12. Decision and relation to final belief state

Figure 12 provides a summary of the decision made by the players (green for decision “Send” and orange for decision “Do not send”). Figure 12(a) shows that the majority (3/4) decided to send the patrol aircraft. This result is not surprising due to the asymmetry in the two vessels’ risk level: Only Vessel A was possibly at risk since Vessel B was far and not in the AOR. Figure 12(b) refines the decision made within the three groups defined above. We first observe that while we would expect that a strong belief toward A (group G_A) would lead automatically the players to decide to send the patrol a small percentage actually did not. When asked about a justification for such a decision, some players answered that they still had time to send the patrol and delayed their decision at that time due to its cost. We can make a symmetrical observation about players who strongly believed that the missing vessel was still in its area (event B) but who actually decided to send the patrol. Their decision was justified by the contextually risky environment which implied a cautious decision to avoid missing any suspicious event.

Finally, we observe that of the players with high uncertainty before decision (group G_0) a large majority decided to send the patrol aircraft. This highlights the strong influence of the context as noticed at the beginning of this analysis section: Because the Harbour Protection Level is still TWO the decision favoured was to send the patrol.

4.5 The effect of information quality

This section contains the core outcomes of the study as we show that the Risk Game allows not only to capture the impact of information quality on belief assessment, but also to identify the impact of specific dimensions. We thus analysed the effect of information quality on both

the decision and final belief state with two different approaches. First, we performed an analysis player by player considering the final belief state only. Then, we analysed each piece of information individually regarding its impact on previous beliefs, regardless the player.

4.6 Falseness

In Figure 13 the impact of the ratio of false information (over the number of pieces of information picked-up) on the final belief state: The darker, the higher the ratio of false information. We observe that the players who leaned toward event *B* (group *B*) indeed received a higher ratio of false pieces of information than the ones who were highly uncertain (or confused, group 0), or than the ones who leaned toward event *A* (group *A*) This result confirms our expectations that a high ratio of false pieces of information increases the confusion in the decision maker mind (as group 0 players were more uncertain than groups *A* and *B*), up to “miss-assessing” the situation (as group *B* players were rather certain of the wrong hypothesis).

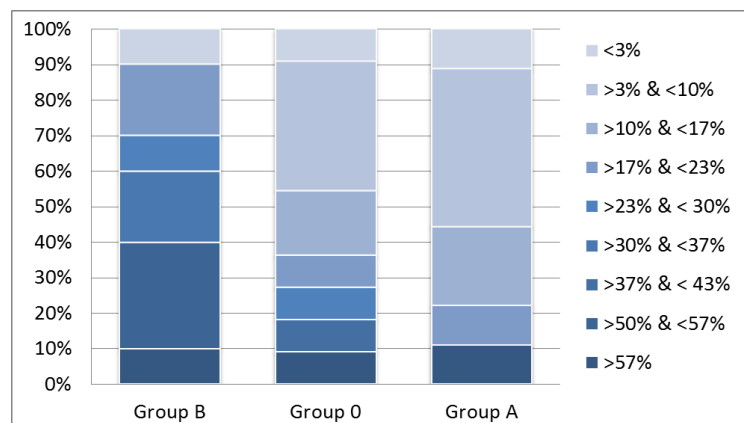


Figure 13. Impact of false information ratio on the final belief state

4.6.1 Certainty and precision

Figure 14 displays the impact of information content only (i.e. regardless the falseness of the information) on the belief change in the player’s mind it produced (event *A* only is shown). We considered three categories of belief change:

- **Null belief change** when the belief about *A* did not change at all, i.e. $|Bel_n^{(i)}(A) - Bel_{n-1}^{(i)}(A)| = 0$, where $Bel_n^{(i)}(A)$ is the belief if the player *i* at the step *n* about Vessel *A* and $Bel_{n-1}^{(i)}(A)$ is the belief if the player *i* at the previous step *n-1* about Vessel *A*.
- **Low/Medium belief change** when the belief about *A* changed but remained in favour of the same event, i.e. $|Bel_n^{(i)}(A) - Bel_{n-1}^{(i)}(A)| < 0.5$;
- **High belief change** when the belief about *A* changed so that the player changed his mind either from *A* to *B* or from *B* to *A*, i.e. $|Bel_n^{(i)}(A) - Bel_{n-1}^{(i)}(A)| \geq 0.5$.

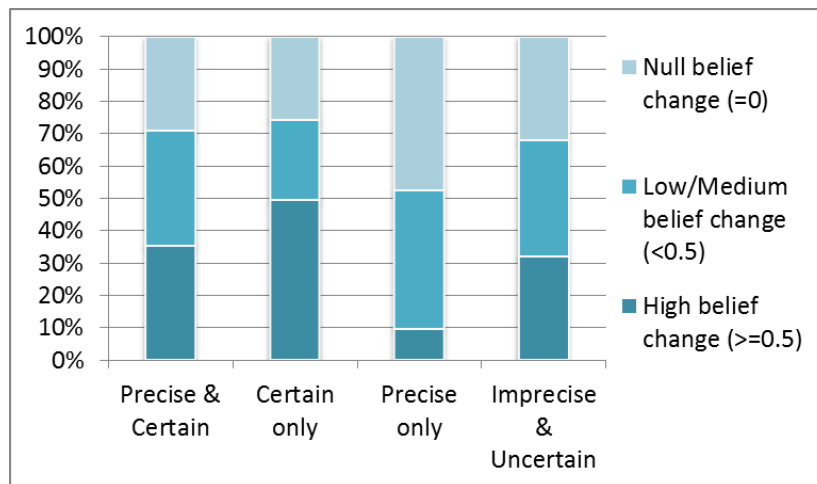


Figure 14. *Impact of information content only on the belief change (about event A)*

The relative ratio of pieces of information which produced either no belief change or a low/medium belief change respects more or less the available information quality (see Table 2). More interesting however is the group of pieces of information which made the players switching from a belief toward A to a belief toward B (or reversely), darker blue in Figure 14. It appears that this group contains a very high proportion of *certain* pieces of information. In other words, the set of pieces of information that were certain only (i.e. not imprecise), represented by the second column in the figure, were more likely to make the players change their belief from one event to the other one. Although further investigations are required, it seems that information expressed with high certainty has a high impact on the players’ belief change.

4.7 Relevance

We interpret a frequent query of a specific attribute as having some level of *perceived relevance* by the player to the problem. We highlighted that although the behavioural attributes were generally frequently queried, they did not contain many relevant discriminant elements which were rather carried by the classification attributes (TYPE and SIZE) since the two vessels differed mainly in their length and width.

We thus analysed the impact of attributes on the belief change and especially investigated which attributes provoked the highest belief change such that the players changed their mind from one event to another (either A to B or B to A). Figure 15 displays the ratio of attributes which indeed made the players changing their minds. It is interesting to observe that unsurprisingly the size has the most impact followed by the type. What is surprising though, is that the location had a similar impact to the size while it in fact did not contain any information which would have helped the player discovering where the missing vessel was.

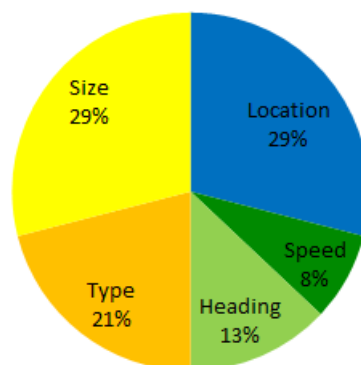


Figure 15. *Impact of attribute on belief change from A to B or B to A*

4.8 Possible biases and limitations

Although we demonstrated the ability of the Risk Game to capture the impact of information quality (along different dimensions) on human belief assessment, the results presented herein should be interpreted in the light of possible biases induced by the game design. In particular:

- **Operational context:** the global Harbour Protection Level (TWO) strongly influenced both the belief assessment and decision toward event *A* and “Send the patrol” respectively. Further study would change the operational context only (for a Search and Rescue one for instance) where the events are symmetric in terms of risk assessment, in such a way that the decision and belief assessment would be guided only by the information at hand and the concern of a timely and accurate intervention;
- **Belief assessment:** a few players might have misunderstood the rules or the form to be filled and possibly confounded the instant belief assessment with the assessment based on accumulated pieces of information. This may be an important bias that we could have observed for one player whose belief assessments did not exhibit an evolution.
- **Complexity:** the number of information quality dimensions is probably too high so it became tedious to find interesting results in the data gathered. These became obvious after partitioning the data. We could expect that a simpler design, with a reduced number of variables would simplify the analysis and lead to more direct link exhibiting the impact of information quality of human belief assessment.
- **Limited number of players:** the board version of the game has the advantage of a close interaction with the experts who additionally provide explanations of their reasoning path and justification for their decision. This aspect of the facilitation is extremely beneficial to understand the operational concerns of the players. However, the small amount of samples gathered limits strongly the scope of the conclusions drawn. A digital version of the game would allow a deeper study and more convincing results supported by meaningful statistics.

5 Conclusions

We presented the Risk Game, a methodology to elicit experts’ knowledge and know-how in threat assessment and decision making in risky environments, with imperfect information. The Risk Game was designed to highlight in particular the impact of information quality on belief and threat assessment and decision making. We selected the three quality dimensions of *falseness*, *uncertainty* and *imprecision* that we made vary independently. The other dimensions of *relevance*, *source reliability* and *conflict* were also studied as naturally derived from the three basic ones.

Thanks to the involvement of the 32 players, we have been able to validate (1) the risk game as an entertaining elicitation method, (2) the efficiency of the approach to gather data demonstrating the impact of different aspects of information quality on belief assessment and decision making and (3) the framework for reasoning analysis establishing formal links between different information and source quality dimensions, as well as between context, threat assessment and decision. We validated our main research question and conclude that serious games are an efficient means to capture the impact of information quality on human belief assessment and decision making.

In the light of the data gathered and analysed, we have been able to effectively address the secondary research question about which information quality dimensions impact belief assessment, and to highlight some impact patterns about *relevance*, *complementarity*, *redundancy*, *uncertainty*, *imprecision*, and *falseness*. First, the analysis of the players queries

through information cards about source, attribute and vessel dimensions demonstrates the distinction between *information needs* (which information the players think is useful to solve the problem) and the *effective relevance* of information (which information actually helps solve the problem). It appears that while positional information (location, heading and speed to some extent) was less useful than the classification information (size and type), it was the most queried by the players. We also highlighted the information quality dimensions of *complementarity* and *redundancy*, as a natural basis of the human fusion process. It was shown that most of the players advocated for a medium redundancy by high complementarity. From the queries analysis, we set up the basis to underline different reasoning strategies, i.e. *comparative reasoning* versus *single event focus*.

Second, we analysed the impact of the information quality along the three dimensions of *falsehood*, *uncertainty* and *imprecision*, on the uncertainty state of the player before decision and upon making the decision. The decision to send the patrol was made by 75% of the players, which actually was the correct decision given our scenario. We explained this asymmetry by the context (i.e., Harbour Protection Level of TWO) which made the players favour this safer decision in case of high uncertainty, as supported by the results. We also clearly observed that the higher amount of false information, the higher the uncertainty before decision.

We finally characterised the impact of the individual pieces of information quality on the belief change they induced in the player's mind. It appeared that information expressed with *certainty* by the sources had a high impact on the players' instantaneous belief.

Also, we confirmed that the most effective relevant attributes (size and type) had very high impact in belief change, together with the location. This last result is surprising since the location cards did not contain any effectively relevant information which would have helped the player discriminate between the two vessels. This maybe stresses the need for humans to visualise the scene to better grasp the situation. Certainly, further investigation would be required to be able to draw any valid conclusion.

In future work, we will develop the methodology to focus on other specific aspects of information quality (e.g., source quality, inconsistency). We will analyse the reasoning profiles of the players by comparing their assessment with automatic reasoners. Finally, a formalisation of the process will lead to deeper and quantitative analysis of information quality on belief assessment.

References

- [1] A.N., Steinberg, C.L., Bowman, Revisions to the JDL data fusion model, in: D., L., Hall, J., Llinas (Eds.), Handbook of Multisensor Data Fusion, The Electrical Engineering and Applied Signal Processing Series, CRC Press, Boca Raton, 2001, Ch. 2, pp. 2–1 to 2–19. <https://doi.org/10.1201/9781420038545.ch2>
- [2] A., Borek, A.K., Parlikad, P., Woodall, M., Tomasella, A risk based model for quantifying the impact of information quality, Computers in Industry 65 (2014) 354–366. <https://doi.org/10.1016/j.compind.2013.12.004>
- [3] S., Raghunathan, Impact of information quality and decision-maker quality on decision quality: a theoretical model and simulation analysis, Decision Support Systems 26 (1999) 275–286. [https://doi.org/10.1016/S0167-9236\(99\)00060-3](https://doi.org/10.1016/S0167-9236(99)00060-3)
- [4] H.T.M., Moges, K.M., Dejaeger, W.M., Lemahieu, B.M., Baesens, A multidimensional analysis of data quality for credit risk management: New insights and challenges, Information & Management 50 (2013) 43–58. <https://doi.org/10.1016/j.im.2012.10.001>
- [5] A., Schönbohm and A., Jülich, On the Effectiveness of Gamified Risk Management Workshops: Evidence from German SMEs, International Journal of Serious Games, 3(2), 2016. <https://doi.org/10.17083/ijsg.v3i2.117>
- [6] F., Taillandier and C., Adam, Games Ready to Use: A Serious Game for Teaching Natural Risk Management. Simulation & Gaming 49(4) (2018), 441–470. <https://doi.org/10.1177/1046878118770217>

- [7] A., Solinska-Nowak, P., Magnuszewski, M. Curl, A., French, A., Keating, J., Mochizuki, W., Liu, R., Mechler, M., Kulakowska and L., Jarzabek, An overview of serious games for disaster risk management – Prospects and limitations for informing actions to arrest increasing risk, *International Journal of Disaster Risk Reduction* 31 1013-1029, 2018. <https://doi.org/10.1016/j.ijdr.2018.09.001>
- [8] K., Seo, H., Ryu, and J., Kim, Can Serious Games Assess Decision-Making Biases? Comparing Gaming Performance, Questionnaires, and Interviews, *European Journal of Psychological Assessment*, Sep., 2018, doi:10.1027/1015-5759/a000485.
- [9] N.E., Dunbar, C. H., Miller, B.J., Adame, J., Elizondo, S., N., Wilson, S., G., Schartel, B., Lane, A., A., Kauffman, S., Straub, J., K., Burgoon, J., Valicich, E., Bessarabova, M.L., Jensen, J., Jenkins, J., Zhang, and D., Morrison, Mitigating Cognitive Bias through the Use of Serious Games: Effects of Feedback. In: Spagnolli A., Chittaro L., Gamberini L., (eds) *Persuasive Technology. PERSUASIVE 2014. Lecture Notes in Computer Science*, vol 8462. Springer, Cham. https://doi.org/10.1007/978-3-319-07127-5_9
- [10] J., Fernandes, D. Duarte, C., Ribeiro, C., Farinha, J., Madeiras Pereira, M., Mira da Silva, iThink : A game-based approach towards improving collaboration and participation in requirement elicitation, in *Virtual Worlds for Serious Applications (VS-GAMES'12)*, *Procedia Computer Science* 15, 66 – 77, 2012. <https://doi.org/10.1016/j.procs.2012.10.059>
- [11] N., Parde, M., Papakostas, K., Tsiakas, M., Dagioglou, V., Karkaletsis, R., Nielsen, I Spy: An Interactive Game-Based Approach to Multimodal Robot Learning, in *Proc. of the Knowledge, Skill, and Behavior Transfer in Autonomous Robots AAAI Workshop*, 2015.
- [12] M., Ceruti, S., Das, A., Ashenfelter, G., Raven, R., Brooks, M., Sudit, G., Chen, E., Wright, Pedigree information for enhanced situation and threat assessment, in: *Proceedings of the 9th International conference on Information Fusion*, Florence, Italy, 2006. <https://doi.org/10.1109/ICIF.2006.301797>
- [13] E.C., Carballo-Cardenas, A.P., Mol, H., Tobi, Information systems for marine protected areas: How do users interpret desirable data attributes?, *Environmental Modelling & Software* 41 185–198, 2013. <https://doi.org/10.1016/j.envsoft.2012.11.016>
- [14] R., Y., Wang, A product perspective on total data quality management, *Communications on the ACM* 41 (2). <https://doi.org/10.1145/269012.269022>
- [15] R.R., Hoffman, N.R., Shadbolt, A.M., Burton, G., Klein, Eliciting knowledge for experts: A methodological analysis, *Organizational* 62 (2) 129–158, 1995. <https://doi.org/10.1006/obhd.1995.1039>
- [16] Y., Cao, Ludic elicitation: Using games for knowledge elicitation, Ph.D. thesis, The Pennsylvania State University, 2014.
- [17] L., von Ahn, L., Dabbish, Designing games with a purpose, *Communication of the ACM* 51 (8) 58–67, 2008. <https://doi.org/10.1145/1378704.1378719>
- [18] F., De Rosa, A.L., Jusselme and A., De Gloria, A Reliability Game for Source Factors and Situational Awareness Experimentation, *International Journal of Serious Games*, 5(2), 2018, DOI 10.17083/ijsg.v5i2.243.
- [19] P., Smets, R., Kennes, The transferable belief model, *Artificial Intelligence* 66 (1994) 191–234. [https://doi.org/10.1016/0004-3702\(94\)90026-4](https://doi.org/10.1016/0004-3702(94)90026-4)
- [20] C., W., Karvetski, K., C., Olson, D.R., Mandel, C.R., Twardy, Probabilistic coherence weighting for optimizing expert forecasts, *Decision Analysis* 10 (4) 305–326, 2013. <https://doi.org/10.1287/deca.2013.0279>
- [21] A.L., Jusselme, A.C., Boury-Brisset, B., Debaque, D., Prevost, Characterization of hard and soft sources of information: A practical illustration, in: *Proceedings of the International Conference of Information Fusion*, Salamanca, Spain, 2014.
- [22] R.Y., Wang, D., M., Strong, Beyond accuracy: What data quality means to data consumers, *Journal of Management Information Systems* 12 (4) 5–34, 1996. <https://doi.org/10.1080/07421222.1996.11518099>
- [23] P., Smets, Imperfect information: Imprecision - uncertainty, in: A., Motro, P., Smets (Eds.), *Uncertainty Management in Information Systems. From Needs to Solutions*, Kluwer Academic Publishers, pp. 225–254, 1997. https://doi.org/10.1007/978-1-4615-6245-0_8
- [24] G.J., Klir, B., Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall International, Upper Saddle River, NJ, 1995.
- [25] ISO 5725-1, Accuracy (trueness and precision) of measurement methods and results - part 1: General principles and definitions, Tech. rep., ISO International Standardization, 1994.