Influence of Meta-Information on Decision-Making: Lessons Learned from Four Case Studies

Maria Riveiro, Tove Helldin and Göran Falkman
Informatics Research Centre, University of Skövde, 541 28 Skövde, Sweden
Email: {maria.riveiro, tove.helldin, goran.falkman}@his.se

Abstract—This paper discusses the results of four empirical evaluations that assess the effects that visualizing system meta-information have on decision-making, particularly on confidence, trust, workload, time and performance. These four case studies correspond to the analysis of (1) the effects that visualizing uncertainty associated to sensor values (position, speed, altitude, etc. and track quality) have on decision-making on a ground to air defense scenario; (2) the effects that the visualization of the car’s certainty on its own capability of driving autonomously have on drivers’ trust and performance; (3) the influence that the visualization of various qualifiers associated to the proposals given by the support system have on air traffic operators carrying out identification tasks and (4) the effects that the presentation of different abstraction levels of information have on classification tasks carried out by fighter pilots. We summarize the results of these four case studies and discuss lessons learned for the design of future computerized support systems regarding the visualization of meta-information.

Keywords: system meta-information, uncertainty, decision-making, trust, situation awareness, decision support.

I. INTRODUCTION

Humans constantly analyze information to make decisions and, in many cases, computer-based systems are used to support users in this analysis in order to reach effective and optimal solutions. However, data is normally imperfect, i.e., has some degree of error, quality or uncertainty associated. According to Endsley [1], a certain level of situation awareness must be reached in order to make a complex decision. This understanding normally involves managing complex information at various levels, not only domain related information but also qualifiers associated to it (meta-information). These qualifiers contextualize the information and can critically influence how a decision-maker processes, understands and acts on it [2].

As discussed by Guarino et al. [2] and Pfautz et al. [3], decision-makers reason using both information and meta-information, and how successful the decision-makers are on deciding upon effective courses of action depends on their skills and experience in processing and understanding this information. There are a number of different sources and types of meta-information that can have a large impact on an operator’s decision-making. These include, for example, characteristics of the information source, the uncertainty and ambiguity of the collected information, the information context, the reliability and credibility of the source content, as well as temporal qualifiers (see [2] for more examples). However, as stated by Pfautz et al. [3] and Bisantz et al. [4], it is not common that decision support systems (DSSs) incorporate such information qualifiers into the primary displays. This is further discussed by McGuirl and Sarter [5] who argue that support systems often present users with a diagnosis or solution to a problem with little or no explanation or qualification, which places analysts in a position where they must fully accept the systems’ advice or perform the entire decision-making process on their own, often under time pressure. Most of the research discussions of meta-information relate to the effects of uncertainty visualization on decision-making [6]. Lipshitz and Strauss [7] state that uncertainty constitutes a major obstacle to decision-making. Analysts are often challenged to recognize, understand and manage uncertainty, and, even when they do, it is difficult to identify the best decision [8].

In order to understand the effects that presenting meta-information, as for example, (un)certainty, has on decision-making, this paper presents a summary of the results obtained from four empirical studies where different types of system meta-information were presented to the participants. During these studies, visualizations of sensor uncertainty, automation performance (un)certainty, decision reliability and different abstraction levels of information were presented to the operators/participants. The effects of these visualizations on the operators’ confidence, response times, workload, performance and trust in the automated aid were evaluated. We present a discussion of the results in light of relevant literature, outlining implications and lessons learned from our findings.

This paper is organized as follows. Section II provides a brief review of related works, mainly covering the areas of meta-information visualization. The four empirical investigations are summarized in section III while section IV discusses their results. Finally, section V provides some conclusions.

II. RELATED WORK

Meta-information has been defined as characteristics or qualifiers of information that affect a human’s decision-making and behavior [2]. Classifications of types and sources of meta-information can be found in [3] and [2]. Examples of meta-information are uncertainty, ambiguity, reliability of source, relevance, characteristics of the information source, level of abstraction of information, etc. Attempts to represent meta-information have been found in the literature, most of which concern the visualization of uncertainty [3], [4]. Research that concerns the visualization of uncertainty is quite extensive (see, e.g., [9], [10]). Nevertheless, there is a lack of empirical results indicating the effectiveness of these visualization...
techniques in terms of how well they are perceived, understood and accepted by the users [11]. An equally important question is how the visualization of uncertainty influences reasoning and decision-making within problem contexts for which uncertainty matters [10].

Besides the visualization of uncertainty, there are few examples of visualizations or representations of other types of meta-information. Two exceptions are [12] (that proposes the visualization of data aiding molecular biologists identify certain features of genes and to detect inconsistencies in the data) and [13], [14] (that visualizes the reliability of the automatically generated identity of a detected object in a command and control application).

The general agreement resulting from these studies highlights the importance of visualizing relevant meta-information to the operators in order to enable them to make better decisions. However, such presentations might not only lead to better, more well-informed decisions. It might also result in improved operator-automation cooperation where the operator is informed of the performance of the automated support system and the reliability of the automatically generated data and information analyses, decisions and actions. Such automation transparency may lead to better calibrated operator trust in the automated DSS, where known human-automation problems such as automation misuse and disuse may be avoided (see, e.g., [15], [16] for more information).

Which meta-information to present depends on the context in which the operators are to make their decisions, as well as the characteristics of the specific information displays used by the operators [4]. Further, it also depends on the expected cooperation between the operator and the automated technologies used. According to Parasuraman et al. [17], automated functions can aid operators in four different ways: by collecting data, by analyzing this data, by providing recommendations of decisions and actions, as well as by implementing these decisions into actions. The higher the automation, the more information regarding the performance of the automation and the quality of the recommendations generated should be presented in order for the operator to make good, effective decisions as well as to be able to more appropriately calibrate his/her trust in the system. However, as stated by [18], there is surprisingly little systematic guidance of how to improve the operators’ trust calibration process when interacting with advanced technological systems. Further, there is also a lack of empirical results indicating the effectiveness of presenting meta-information in terms of how well they are perceived, understood and accepted by the operators [4], [11]. This is further discussed by Fortenbery et al. [19], who argue that more research is needed to investigate ways of effectively providing meta-information to decision-makers as well as to evaluate the effects of such provision on the operators’ understanding or decision-making at the task level.

### III. Case studies

The four case studies presented in this manuscript are typical research problems in information fusion research\(^1\). Three of these case studies are within military and civil security research, and one of them (autonomous driving) relates to traffic safety. In these scenarios, operators, pilots and drivers need to analyze imperfect sensor data, operate imperfect recommender systems and have limited time to make a decision that could have severe consequences if such decisions are late or wrong. Information about the qualifiers of the data analyzed (i.e., meta-information) might help users to reach optimal solutions. Therefore, we design four empirical studies were participants need to analyze data and meta-information (at different abstraction level, i.e., meta-data and meta-information) using a recommender system and carry out known tasks where decisions need to be taken under certain time constraints. Table 1 presents a summary of the four empirical studies presented, the meta-information considered in each of them and a brief summary of the results obtained in relation to the effects of meta-information visualization on the participants’ confidence, trust, workload, performance and time to carry out their tasks. Figures 1 to 4 illustrate some of the materials used during the experiments.

#### Case study nr. 1. Uncertainty visualization

Military personnel normally carry out tasks related to the recognition, identification and prioritization of objects in an environment of interest. The execution of these tasks is often challenging due to the need to synthesize large amounts of often disparate data into a meaningful whole as well as considering the

\(^1\)“Information fusion encompasses theory, techniques and tools conceived and employed for exploiting the synergy in the information acquired from multiple sources [...] such that the resulting decision or action is in some sense better [20].”

<table>
<thead>
<tr>
<th>Case study</th>
<th>Type of meta-information</th>
<th>Confidence</th>
<th>Trust</th>
<th>Workload</th>
<th>Performance</th>
<th>Time</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Uncertainty visualization</td>
<td>Uncertainty in sensor readings and tracks</td>
<td>↑</td>
<td>Not eval.</td>
<td>↓ or =</td>
<td>=</td>
<td>↑</td>
<td>Less false positives</td>
</tr>
<tr>
<td>2. Autonomous driving</td>
<td>Car (uncertainty)</td>
<td>↓</td>
<td>Not eval.</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>More time doing other things</td>
</tr>
<tr>
<td>3. System inner workings</td>
<td>Reliability and source conflict on identity proposed by system</td>
<td>Not sig. diff.</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>Less attempts to make decision</td>
</tr>
<tr>
<td>4. Different levels of information abstraction</td>
<td>Various abstraction levels: from overview to detail</td>
<td>Not sig. diff.</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>More false positives</td>
</tr>
</tbody>
</table>

Table 1: summary of the four user studies (between subjects) and the results obtained. The results show the comparison between the test group (with meta-information) and the control group. Example (1st row): confidence by test group was ↑ (higher) than the control group. Not eval. = not evaluated. Not sig. diff. = no statistically significant difference found between the groups.
Figure 1: case study nr. 1. Threat evaluation prototype used during the experiments: (1) the uncertainty associated to the sensor readings (error) is proportional to the thickness of the graphical lines that show the values for speed, altitude, distance, etc. for the selected object; (2) the uncertainty associated to the track is represented by a track quality value between 0-1; (3) the uncertainty associated to the position of the objects is displayed using semi-transparent circles.

prevailing political situation, time constraints, possible decision alternatives and their consequences. In this case study, we selected an air defense scenario, where expert operators have a few minutes to make a well-informed decision based on uncertain sensor data regarding the identity of an object and where the consequences of a late or wrong decision are severe. A proof of concept prototype designed and implemented for carrying out the empirical investigations allow the operators to analyze multivariate spatio-temporal sensor data (see fig. 1). Uncertainty associated to the sensor readings is visualized using the thickness of the lines in the graphical representation of the sensor values. Semi-transparent filled circles represent the uncertain position of the objects in the environment, while a track quality value between 0 and 1 accounts for the quality of the estimated track for each target. Eight experienced air traffic operators were divided into two groups (with and without uncertainty visualization) and carried out identification and prioritization tasks using the prototype. Figure 1 illustrates the prototype, the uncertainty visualization and the scenario.

Meta-information: Uncertainty associated to the sensor readings (in the form of error or precision), to the quality of the track estimated by the tracking system as well as the geographical uncertainty associated to the position of the objects.

Results: The group of experts aided by visualizations of uncertainty assigned significantly higher confidence values to their identifications, while presenting similar performance levels and slightly higher average identification time.

Case study nr. 2. Autonomous driving: To investigate the impact of visualizing car uncertainty on drivers’ behavior and trust during an automated driving scenario, a simulator study was conducted (a detailed description of our investigations can be read in [21]). A between-group design experiment with 59 Swedish drivers was carried out at the Human Machine Interaction (HMI) laboratory at Volvo Car Corporation, Gothenburg. The laboratory contained both a driving simulator and a fully functioning cockpit. A continuous representation of the uncertainty of the car’s ability to drive autonomously during snow conditions was displayed to one of the groups, whereas omitted for the control group. Figure 2 illustrates the interface and the graphical representation of the car’s certainty used during the experiments. At a certain point the car could no longer drive autonomously (due to adverse weather), and the driver should take over the control of the car.

Meta-information: (Un)certainty associated to the car’s ability to drive autonomously.

Results: On average, the group of drivers who were provided with the uncertainty representation took control of the car faster when needed, while they were, at the same time, the ones who spend more time looking at other things than on the road ahead. Thus, drivers provided with the uncertainty information could, to a higher degree, perform tasks other than driving without compromising with driving safety. The analysis of trust shows that the participants who were provided with the uncertainty information trusted the automated system less than those who did not receive such information, which indicates a more proper trust calibration than the control group.

Case study nr. 3. System inner workings: As seen in case study nr. 1, military decision makers identify and prioritize objects and situations in the environment of interest before they take action. Different support systems have been implemented that are able to provide recommendations regarding the identity and threat of the objects present in the environment and give recommendations for actions. However, most models for
identifying objects and evaluating threats are complex and difficult to understand, which might result in inappropriate levels of operator trust and other performance related problems. To limit the negative effects of such inadequate cooperation, we have studied the effects that visualizing the inner workings of a identification system have on decision-making. An empirical evaluation was carried out in order to assess if explanations of the inner workings of the complex identification model and visualizing its reliability and possible sources of conflict affect the expert operators’ performance, trust and workload when carrying out identification and prioritization tasks with real-world data. A modified version of the prototype used in case study nr. 1 was used during the experiments (see figure 3). Eight expert air defense operators participated in the study.

**Meta-information:** Information associated to the identity of the object predicted by the system was presented (identity might be hostile, suspect, friend, neutral, unknown and pending), along with the probability of belonging to one of these classes over the time, its reliability and if there were conflicts between the sources of information used for the prediction.

**Results:** The presentation of system meta-information indeed positively affected the expert operators’ trust in the support system, a result which however must be weighed against the slight increase of the operators’ perceived workload.

**Case study nr. 4. Different levels of information abstraction:** Fighter pilots need to constantly discriminate between hostile and friendly targets during their missions. When inconclusive or uncertain data regarding detected targets is available, they have to manually fuse target values such as speed, altitude and g-force, as well as investigate if onboard or team systems have been able to determine which radar or motor is used by the target. Based on such information, the pilots use their experience and knowledge of known identity characteristics to determine the most probable identity or class of an object. In this context, we investigated the effects of automating this process on the pilots’ performance and initial trust in the automated combat classification system. We further analyzed the influence of visualizing different levels of transparency of the classification fusion process on the pilots’ performance and subjective confidence in their classification decisions made. These levels consisted of (1) only presenting the text-based information regarding the specified object (without any automated support), (2) accompanying the text-based information with an automatically generated object class suggestion, and (3) adding the incorporated sensor values with associated (uncertain) historic values in graphical form (see figure 4). A simple classification system based on rules was designed and applied to a set of 11 scenarios including a fixed arrangement of objects and associated characteristics. Each scenario contained the 3 display conditions. As such, 33 momentary images were prepared and presented to the pilots participating in the study.

**Display condition 1** uses display 1, **condition 2** uses displays 1+2 while **condition 3** shows a detailed view of the included parameters with associated historical sensor values. Condition 1 uses display 1, condition 2 uses displays 1+2 while condition 3 shows displays 1+2+3.

**Results:** The pilots needed more time to make a classification decision when being provided with display conditions 2 and 3 than display condition 1 (significant difference). The results also show that the number of incorrect classifications was the lowest during display condition 3 (no significant difference). Comparing the reported trust ratings from display conditions 2 and 3, higher ratings can be found in display condition 3 (significant difference). However, no difference in decision confidence was reported during the three display conditions. The pilots’ reported slightly increased workload measures when using the aid, but also the general opinion that such support system would aid them make confident decisions faster, having trained and learned the system beforehand.

**IV. Discussion**

Before discussing the findings of the user studies in light of relevant literature, we present a summary of the results.

**Confidence:** The visualization of meta-information, particularly, uncertainty associated to the sensor readings and the track quality had a significant positive effect on the confidence given to the classifications made by the experts (case nr. 1). However, no significantly difference on confidence has been noted in case nr. 3 and 4 (confidence was not measured in the autonomous driving case study).
Trust: Trust in the systems’ recommendations was significantly higher in those cases in which more information was given to the participants regarding the system inner workings and explanations (i.e., case nr 3 and 4). In case nr. 1, trust in the system was not applicable (since the system did not provide any recommendation), and in case nr. 2 it was found that the drivers with aid reported lower trust values (but they were able to calibrate their trust in the system better than the control group).

Workload: An significant increase of workload has been seen in case nr. 3 and 4, while the same workload or in some cases, less, was reported in case nr. 1.

Performance: Significantly better or at least equal performance between the control and test groups has been seen in the presence of meta-information in all the cases.

Time: Mixed results have been collected regarding time to carry out the tasks under investigation. Time needed to complete the identification tasks by participants in the presence of meta-information was reported higher in cases nr. 1 and 4, and was significantly lower in cases nr. 2 and 3. This can be partly explained by the differences among the tasks and application areas of the four case studies. In the autonomous driving scenario, more information about the car made the drivers more aware of the situation, and they could take over the control faster, i.e. less time than the control group. In case study nr. 1 and 4 the meta-information provided seems to increase the response time, perhaps because more information needed to be analyzed and taken into account.

Andre and Cluter [22] found out that if uncertainties are not presented to operators, they tend to adopt risky behavior, and that operators have a tendency to ignore implicit or hidden decision criteria when they are not visually presented. As we have seen in case nr. 1, the participants without visualization of uncertainty performed equally good as those with uncertainty considering the true positives values (correct threatening classifications), however, they did mark more targets as dangerous than the test group, i.e., more false positives. In this context, this is an indication of ‘riskier behavior’, as reported in [22], since without information about uncertainty experts became more inaccurate, and reported more targets as dangerous when there were not (possible inappropriate actions towards non threatening targets could have been taken). We would like to highlight that, even if the performance of both groups was similar taking into account the correct identifications (true positives), the confidence reported by the group with uncertainty visualization was significantly higher and the false positives rates significantly lower. These positive measures
came with a cost, a slightly higher response time.

In a particular application domain where a software for water balance is presented, Cliburn et al. [23] highlight the importance that the DSS display both information and an estimation of the uncertainty associated. After carrying out a empirical study with such software, they concluded that decision-makers do not appreciate to see uncertainty without a related advice that presents an appropriate action when uncertainty is present. Similar results were presented by Thill et al. [24], where a navigation aid with an arrow and text line justifying the choice shown in a driving scenario was perceived as the most intelligent of the three choices provided.

The authors argue that a more intelligent aid can increase the driver’s awareness of the whole driving situation. These are interesting results, that could guide our future work: would the visualization of course of actions when uncertainty is present improve the confidence, the performance and on the other hand, mitigate the adverse effects on time seen in our experiments?

The trust values reported by the participants were higher in the study cases 3 and 4, i.e., participants trusted more the system when meta-information (system reliability associated to identity proposal) and explanations (higher degree of detailed information) were given to the participants. These results are in line with those reported in [14], that claims that to engender appropriate reliance on combat identification systems, users should be made aware of the system reliability. These positive effects which came with the cost of higher workload, fulfilled, as well, our preliminary hypothesis. However, the trust values reported in the autonomous driving case were more challenging to analyze, i.e., the visualization of the car’s (un)certainty on its own capability for driving autonomously did not lead to higher trust values.

The finding is in contrast to Beller et al. [25] and Seppelt and Lee [26], that recommend that providing drivers with continuous information about automation is preferable to providing warnings, and that information about automation increase trust and acceptance. A possible explanation for interpreting our results can be found in [15] and [27]. Dzindolet et al. [15], where the role of trust in automation reliance is studied, suggest that participants initially considered automated decision aids trustworthy and reliable, but, after observing the automated aid make errors, participants distrusted even reliable aids, unless an explanation was provided regarding why the aid might err. Knowing why the aid might err increased trust in the decision aid and increased automation reliance, even when the trust was unwarranted.

A known negative consequence of too much trust in an automated generated solution is the automation bias, where humans do not seek disconfirming evidence and fully trust the system’s recommendations (this effect has been seen in, for example, C2 systems [27]). Over-trust can lead to serious problems in many settings, and we should not necessarily desire the highest trust rating in an automated system, especially when the underlying automation that drives the system is flawed [27]. Therefore, the effect of lower trust values reported by the drivers aided by the visualization of the car’s (un)certainty can be easily explained if we consider that those that could see the car’s (un)certainty of its capabilities could adjust their trust levels better than the group who was not provided with this information. We believe that if we show this type of meta-information we support trust calibration, and, based on our experiences, we agree with Cummings et al.’s [27] recommendation, i.e., designers of intelligent systems should seek some median level of acceptable trust, and that high and low ratings are both equally problematic.

However, it should be further investigated if the visual representation of the car uncertainty used in our study should be complemented with additional information regarding why the level of uncertainty was high. In general, we could say that the drivers with (un)certainty information can calibrate better their trust in the automatic driving system, and this finding is in line with the work presented by McGuril and Sarter [5] where the participants who were informed of the system confidence were better able to more appropriately calibrate their trust in the decision aid.

The benefits and costs of introducing automated technologies have been well documented across a variety of domains. In relation to military identification and threat evaluation tasks, studies have indicated that the availability of an identification decision aid either had no effect on the operators’ performance [15] or improved performance in only the most difficult conditions [28]. This findings are in line with our empirical tests in military contexts, either the performance between the groups was very similar or increase slightly (case nr. 3). The positive effects were greater in other metrics, such as confidence, trust or less false positives rates. In [38], it was further concluded that the use of an automatic identification system had no adverse effect on combat effectiveness as well as that there was no significant increase in the number of missed opportunities to engage hostile targets, nor a significant increase of the time needed to reach such decisions. The participants of our experiments, regardless of the information provided to them, were able to successfully identify the hidden threats, within a reasonable time frame. It would be interesting to investigate why there is not a larger difference between the performance of the groups, which role plays training and experience carrying out these tasks?

From a methodological point of view, we highlight that designing empirical tests assessing the influence of meta-information within DSSs on decision-making is a challenging task that requires holistic approaches (as recommended by Va- chon et al. [29]), since many competing constraints influence the processes, and if one decides to carry out more controlled experiments the naturalistic decision-making context is lost.

As seen in the these four case studies, the visualization of meta-information might lead to positive results in terms of confidence, trust or even response time, without reducing the performance. These positive results need to be weighted against their costs, for example, increase of perceived workload. In each application domain, there are great differences among the constraints around a decision situation that define
the optimal way of visualizing and interacting with the information provided by a computerized support system. Even if some authors have argued that, no matter how aspects such as uncertainty, risk, stress or error have been quantified, their full effect are hard to understand outside the context of the individual [30], the lessons learned from these case studies might help us to design future computerized support systems, capitalizing on the positive effects of visualizing meta-information and mitigating their possible adverse ones.

V. Conclusions

Human decision-making in dynamic and data rich environments is a complex task, that is increasingly being supported by computerized systems. Decision-makers need to manage not only the situation-related information, but also the qualifiers or meta-information that contextualizes it.

This paper summarizes four empirical studies carried out to investigate the effects that visualizing different types of meta-information have on decision-making, especially on confidence, trust, workload, response time and performance. The results show that, even if there are differences among the various studies due to the various constraints and tasks, the visualization of meta-information had positive effects on confidence and response times, without reducing the performance. Moreover, participants aided by meta-information visualization calibrated better their trust. On the other hand, an increase in perceived workload was reported.

Lessons learned from these case studies may help us to design future DSSs, benefiting from the positive effects of visualizing meta-information and finding means of palliating the adverse ones.

Acknowledgment

This research has been supported by the Swedish Knowledge Foundation under grant 2010/0320 (UMIF). Vinnova through the National Aviation Engineering Research Program (NFFP5-2009-01315), Saab AB and the University of Skövde.

References