Visualizing Uncertainty Information in Engineering Design Processes to Assist Individual and Team Decision Making

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Xiao Dong

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Caroline Hayes

September 2012

Acknowledgements

A truely thank you is given to my advisor, Prof Caroline Hayes, for her guidance and encouragement over the course of my time at the University of Minnesota. Your expertise and support will carry on wherever life leads me.

Appreciation is also offered to Dr. Thomas Stoffregen, Dr. Karen LaBat, and Dr. Sant Arora for serving on my examination committee.

Dedication

This dissertation is dedicated to my Mom and Dad, who taught me so many important lessons in life.

Abstract

This dissertation work describes a decision support system (DSS), two experiments, and the resulting insights regarding the role and influence of an uncertainty visualization in assisting decision makers to gauge more realistically whether they have enough information to make decisions. The primary goal of this work is to improve the scientific understanding of decision making in situations where the best design options are ambiguous. Decision makers are often faced with the task of identifying the “best” option from a set, where “best” is defined by multiple criteria came out by the decision maker. However, uncertainty and lack of information can make the task hard.

Uncertainty is inherent in all real work contexts; it creates ambiguities that make decision making difficult. To help decision makers recognize and manage ambiguity the author developed and evaluated a domain-independent decision support tool (DSS), the Uncertainty DSS. It is designed to help decision makers recognize situations in which uncertainty creates ambiguity in their choices, and to identify information which can help reduce that ambiguity. It does so by providing a simple graphical representation of the relative value ranges of multiple options. The aim is to help decision makers visually recognize when options overlap in their possible values, in which case it is difficult to identify the “best” option with information available. In order to evaluate the impact of the Uncertainty DSS, the author created a pared-down version, the Certainty DSS, which provides no uncertainty visualizations. Two experiments were designed and conducted with engineering designers in the context

of both individual and team decision making. Participants carried out real decision making task in the context of real design projects in which they had been engaged for at least a month. The results of the experiment with individual decision makers showed that without the visualizations, engineering designers did not distinguish between ambiguous and unambiguous sets of options despite being aware in a general sense that there was much uncertainty in their design alternatives. However, with Uncertainty DSS, participants exhibited a significantly improved ability to recognize ambiguous decision situations, and expressed appropriately reduced confidence in ambiguous situations. Additionally, Uncertainty DSS increased the likelihood that participants would form plans to seek clarifying information on critical, uncertain parameters. The results of the experiments with team decision makers were similar to the individual decision makers. Moreover, teams using the Uncertainty DSS communicated more within the team and developed better shared situation awareness. These results suggest that a relatively simple visualization of uncertainty can benefit both individual decision makers, and teams of decision makers by assisting them in: 1) identifying when they do not have enough information to make an unambiguous choice, 2) identifying what additional information might reduce uncertainty and 3) providing a common structure is which teams can discuss uncertainty and its impact on the design decisions.

While there is still room for improvement to fine tune the DSS measurements, and training given to designers to enable them to think about how uncertainty and information (a lack of it) impact design decisions. However, this work represents a first step demonstrating that uncertainty visualization can change the way in which designers think about uncertainty—New guidance for tools for decision makers.

Keywords: **Decision support systems, decision making under uncertainty, team decision making, situation awareness**

Table of Contents

List of Tables

List of Figures

Chapter 1 -- Introduction

Decision Support Systems (DSSs) are becoming increasingly popular in various domains as a way to aid decision makers in making better decisions in a more efficient and effective manner. From economics to medical problems and aircraft navigation to military operations, the implementation of DSSs is growing. However, there is no guarantee that a DSS will improve decision making or problem solving performance. In fact, it is a widely held belief in this area of study that all DSSs will exhibit brittle behavior at one time or another (Larson & Hayes, 2005; Smith et al., 1997). Thus it is important to carefully design DSS to support decision maker's natural needs and behavior.

There are many challenges inherent in complex decision making domains such as military operations, economics, business, and medical operations. Decision makers in many of these areas are presented with high stress levels, potential information overload or lack of information, various sizes and social dynamics of teams or groups involved in the decision making process, dynamic operational environments, and vital objectives that must be met. These factors can drastically impact the quality and types of decisions made and hence, there is a need to research decision making in complex environments to determine if, and how, decision support systems may be useful to aid decision makers.

For this dissertation, among all the complex decision making scenarios, It is especially interesting to look at engineering design domain. Engineering design as a domain has yet been studied thoroughly. It is mostly left unclear of the challenges that decision makers face, and the behavior that they adopt. Engineering design is an excellent example of a complex task that is not well modeled as a decision event. Design options are constantly being added, modified, or refined, as are design goals. Much information is simply unknown or hard to obtain. Moreover, there are usually time considerations because of the iteration of the design, the sheer number of decisions means that most must be made very rapidly. Finally, many important design activities are not captured by a decision event, for example, the information seeking behaviors that precede the selection of an alternative, are critical to product design especially when designers are of little experience.

This dissertation adopts a user centered approach to identify the needs of engineering designer, design a novice decision support system for engineering designers and evaluate the system with real design tasks. Various forms of studies are conducted such as understanding engineering design context and designer's behavior with ethnographic observation and longitudinal study, evaluating proposed DSS with two controlled lab studies. The most critical information that ethnographic observation revealed is that designers could benefit from incorporating the sense of uncertainty in DSS when comparing design alternatives. This is because uncertainty could better represent the ill-defined problem and the lack of information at the stage of making decisions. Also uncertainty could be harnessed to guide the search for information that supports a decision making task.

The DSS designed by the author aims to present an uncertainty visualization to increase decision makers' "situation awareness" with respect to uncertainty. The visualization was implemented in a domain-independent decision support system (DSS) which was called the Uncertainty DSS. This DSS allows decision makers to specify multiple alternatives, and to specify criteria and values which characterize those alternatives. Decision makers can specify uncertain values as ranges, which are visually represented by bars. Most importantly, the DSS provides a simple visualization in which decision makers can simultaneously compare the relative uncertainty of multiple alternatives. If the bars for two alternatives overlap, then it may not be possible to determine which is the better alternative given the current information. The decision maker can then decide whether to make a choice anyway knowing that the decision may be suboptimal (when deadlines make a choice necessary), or seek more information to resolve the ambiguity.

To evaluate the visualization, a second version of the DSS was created with no uncertainty visualizations which was named the Certainty DSS. Because real decision making is always situated in a task, often a complex task, then it was important to observe the effect of the DSS in a real and complex task. The author conducted two controlled lab studies with real design tasks manipulating different mediating factors such as domain expertise and individual vs. team use of Uncertainty DSS to understand the impact of uncertainty in engineering design decision making. The results showed that, in both individual and team context, the Uncertainty DSS did improve participants' situation awareness in that it helped them to recognize when they did not have enough information to make an informed choice (e.g. when uncertainty interfered with decision making). Furthermore, the preference, workload and usability results demonstrated the benefit of Uncertainty DSS had over Certainty DSS. Finally there was evidence that domain expertise can be a mediator when decision makers utilize uncertainty visualization in DSS.

Generally speaking, this piece of research can increase the community's scientific understanding of decision making in engineering design context. In additional to that, a computer based DSS (Uncertainty DSS) that can visualize uncertainty when comparing decision options so that decision makers can benefit from understanding whether there is enough information and where to dig more information if needed. It is likely that researches for other domains are able to draw lessons from the implementation of this DSS.

Chapter 2 --Literature Review

This chapter includes all the relevant domain researches relevant to this dissertation. The first three sections: “Decision support systems”, “Decision making models” and “The nature of engineering design processes” focus on the general background and domain knowledge introduction of this field. The remaining sections “Domain expertise in decision making”, “Information seeking in design”, “Information seeking and the relationship to uncertainty” and “Team decision making” are more specific topics in decision making and more relevant to the hypothesis and contribution of the dissertation.

2.1 Decision support systems

Decision support systems are gaining an increased popularity in various domains, including business, engineering, the military, and medicine (Marek & Roger 2002). They are especially valuable in situations in which the amount of available information is prohibitive for the intuition of an unaided human decision maker and in which precision and optimality are of importance. Decision support systems can aid human cognitive deficiencies by integrating various sources of information, providing intelligent access to relevant knowledge, and aiding the process of structuring decisions. They can also support choice among well-defined alternatives and build on formal approaches, such as the methods of engineering economics, operations research, statistics, and decision theory. They can also employ artificial intelligence methods to address heuristically problems that are intractable by formal techniques. Proper application of decision-making tools increases productivity, efficiency, and effectiveness and gives many businesses a comparative advantage over their competitors, allowing them to make optimal choices for technological processes and their parameters, planning business operations, logistics, or investments.

According to Carlisle (2003) and O'Connor (2004), the primary role of decision support systems is to help people make informed decisions by providing and managing information, clarifying preferences and presenting the tradeoffs involved in various possible choices. Marek and Roger (2002) envision decision support systems as interactive, computer-based systems that aid users in judgment and choose activities. They pointed out the typical functions of decision support systems are framing decision, modeling decision process, and problem solving. Clyde and Andrew (1996) stated that decision support systems could offer other functions such as: alerting on decision opportunity, facilitating and enhancing user's ability to process knowledge, offering advice, stimulating user's creativity, and coordinating interactions among participants in decision- making process. Carlisle (2003) and O'Connor (2004) also pointed out that providing unbiased information to decision maker is an important function of decision support systems.

2.2 Decision making models

It is important to note that the field of decision making support does not have a universally accepted model, meaning that there are many theories vying for supremacy in this broad field. The decision making models can be grouped into rational decision making models and naturalistic decision making models. Rational models assume decision makers are rational human beings, who make decisions following a logical order of a cognitive process. The comparison between options is also rational and often performed by filling out forms or charts utilizing decision makers' knowledge. Multi-criteria decision making method (MCDM) is a popular method within this category.

2.2.1 Multi-criteria decision making methods

There are a broad range of mathematically-based "rational decision making" approaches aimed at improving human decision making by providing more systematic and logical methods. The work described in this paper focuses on a particular class of mathematically-based decision making process called multi-criteria decision making (MCDM) techniques. MCDM techniques are a broad family of mathematical methods that compare alternatives in a set, using multiple criteria (Klein, 1993). Decision makers can use MCDM techniques to choose from available alternatives, characterized by multiple qualitative or quantitative criteria (Saaty, 1980). For example, a prospective car buyer might compare his or her car choices using criteria such as fuel efficiency, cost, and comfort. The criteria used may vary from buyer to buyer depending on what is most important to that particular person.

2.2.1.1Weighted Sums Model

One common MCDM approach is the weighted sum model in which the value of each term in the sum represents the degree to which an alternative satisfies a given criterion. The term's weight represents that criterion's importance to the decision maker (Triantaphyllou, et al. 1995). Variants of the weighted sum method are popular because they are relatively easy to understand and use.

2.2.1.2 Deterministic and Non-deterministic MCDM models

One can further divide MCDM methods into deterministic and non-deterministic methods. Deterministic decision making methods are those that do not explicitly incorporate a representation of uncertainty, for example, the cost of an alternative may be represented as a specific number, or "point" value (Triantaphyllou, 1997). In contrast, non-deterministic decision making methods are those that incorporate some explicit representation of uncertainty or unknowns. For example, uncertainty in the cost of an option may be represented as a range of possible costs, or as a function describing the likelihood of various costs. In this dissertation, the impact of a deterministic and a non-determinist MCDM method were compared on decision makers' awareness of uncertainty.

2.2.1.3 Drawbacks of MCDM techniques

Some challenges in use of MCDM techniques include the difficulty of deciding how to set the weights on the criteria (Triantaphyllou, 1997), and the relatively high effort required on the part of the decision maker to create the model of the decision situation and entering the input information (Hayes, 1981; Law, 1996). Furthermore, Hayes &Akhavi (2008) found that it took users more time to compare alternatives with a non-deterministic than a deterministic MCDM method, mainly because of the additional data entry required. Thus, in the work reported here, one of the objectives was to minimize data entry effort, particularly for the Uncertainty DSS. So that users would not feel the DSS was too burdensome and refuse to use it (when not in a laboratory setting). Thus, while there were many ways to represent uncertainty, for example as numbers (Erevet. al1993), distribution curves, or fuzzy sets (Bellman & Zadeh, 1970; Thurston & Carnahan, 1992), the author chose to represent uncertainty as a range of values, presented visually as a bar with two sliders to set the upper and lower ends of the range. Users set the range by moving the upper and lower sliders to create a range of values. This representation will be described in more detail in the system description, below.

2.2.2Naturalistic decision making models

On the other side of the spectrum of decision making models, naturalistic decision making methods focus on the descriptive view of how people make decisions in actual settings that often feature unstructured problems imbedded within complex and dynamic systems (Klein, Orasanu, Calderwood, & Zsambok, 1993; Zsambok & Klein 1997). Decision making in these settings tends to differ significantly from the analytic style inferred from structured laboratory decision tasks that form the basis for traditional decision theory research. Orasanu and Connolly (1993) stated that naturalistic decision making models attempted to address how people can make decisions in situations where the conditions are changing over time, where information is ambiguous, and where the plausibility of potential goals and courses of action is shifting over time. Also pointed out by Orasanu and Connolly(1993) and Klein (1989), naturalistic decision research often concerns real-life situations in which there are limited time to act, and in which the decision maker are domain experts. A growing body of research indicates that under realistic conditions experts make decisions using a holistic process involving situation recognition and pattern matching to memory structures to make rapid decisions (Dreyfus, 1981; Klein, 1989, 1993; Klein, Calderwood, & Clinton-Cirocco, 1986).

Although naturalistic decision making has been studied in a wide range of domains from traffic management and mission analysis, to health care and space exploration (Salas and Klein, 2001; Schraagen, Militello, Ormerod, and Lipshitz, 2008), engineering design has not yet been much studied through naturalistic decision making approaches. Engineering design is an excellent example of a complex task that is not well modeled as a decision event. Alternatives are constantly being added, modified, or refined, as are design goals. Much information is simply unknown or hard to obtain. Additionally, there are practical time considerations, the sheer number of decisions means that most must be made very rapidly. Finally, many important design activities are not captured by a decision event, for example, the information seeking behaviors that precede the selection of an alternative, are critical to product design especially when designers are of little experience. Thus one of the contributions of this work is to bring naturalistic decision making studies to engineering design process.

The approach in this research is trying to take advantage of the benefits by marrying rational models and naturalistic models. Multi-criteria decision methods (MCDM) was chosen as a foundation but expanded in a way intended to support and augment decision makers' natural and typical approaches to decision making. By doing so, the decision support systems can be used in decision makers' daily work, having the benefit of the combined strengths of human and mathematical decision approaches.

2.3 The nature of engineering design processes

It is pointed out by Kemper et al. (2006) that decision making is integrated to the engineering design process and is an important element in nearly all phases of design, from defining the problem, synthesizing alternatives, evaluating what is acceptable and what is not, identifying which design elements to work on first, specifying what information is needed and by whom, selecting which alternatives are worth further investigation and finally configuring the optimal design.

To better understand the decision making in engineering design process, it is important to be aware of the salient properties and structure of design context.

Uncertainty in design. Uncertainty is present in all designs (Aughernbaugh & Paredis, 2006), from hand-held computer devices to space station systems. Even when a design is considered to be complete, there may still be uncertainty concerning issues such as the performance of the design under all the conditions to which it may be exposed in its working life, its manufacturing feasibility, or the final cost. Uncertainty is most prevalent in the early stages of design, also known as conceptual design, when the alternatives under consideration may be little more than quick sketches or brief outlines. One approach to somewhat resolve the uncertainty during design process is the presentation of probability information. Probabilistic information such as forecast uncertainty can be communicated accurately and efficiently to both professional users as well as those with less background knowledge. Participants were able to learn and make excellent use of completely novel displays with very little training. It is important however, that the presentation format be matched to the task at hand. In a study done by LimorNadav-Greenberg et al., it is clear that uncertainty products can be used to good advantage, if these factors are taken into consideration during the development of such products (LimorNadav-Greenberg et al., 2007).

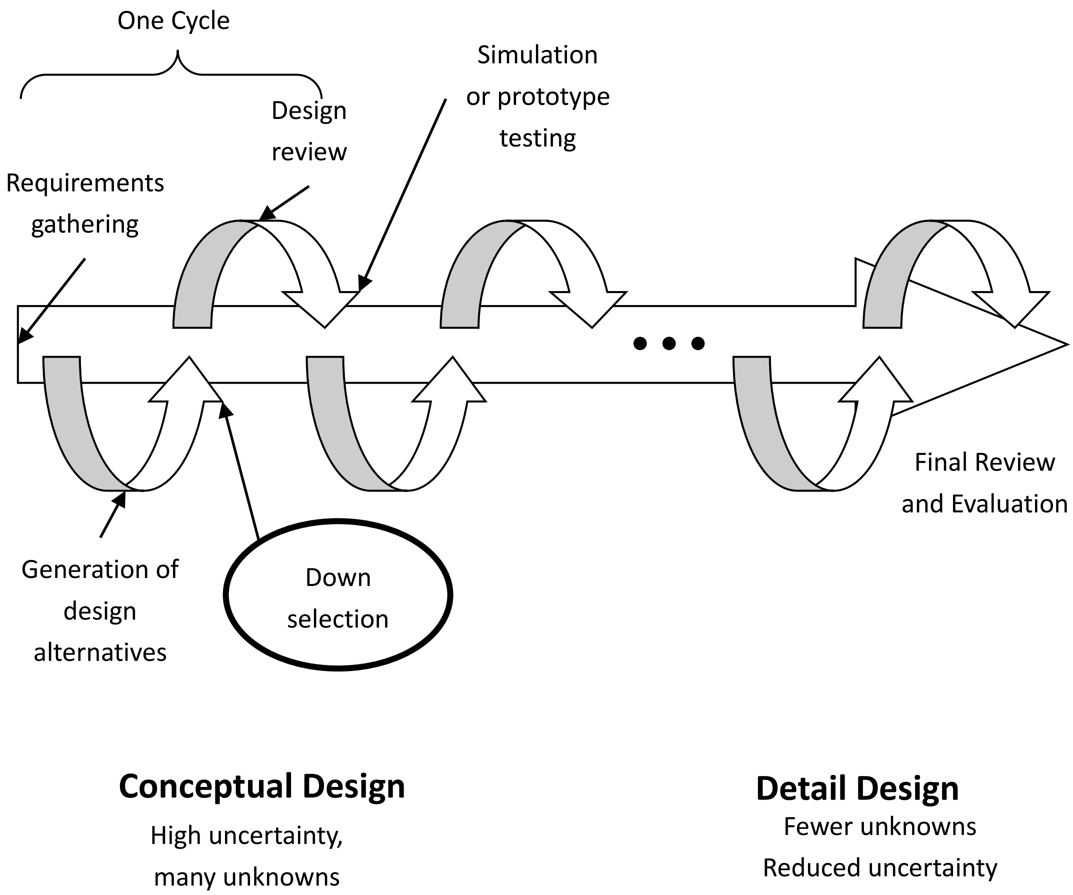


Figure 1. An iterative model of the design process (Hayes and Akhavi, 2008)

<u>Descriptions of the design process.</u> At all stages of the design process, designers must repeatedly choose the most promising alternatives for further development. This winnowing of alternatives is known as the down selection process; this is the step that most people think of as "decision making" because alternatives must be compared and

the best ones chosen. Conceptual design, down selection, and their relationship to the overall design process are shown in Figure 1.

What makes the design process complex is design problems are always ill-structured (Newell, 1973; Simon, 1973) meaning that the goal specification is not completely known and needs to be elaborated and determined as the design problem is being solved. Rittel and Webber (1973) found similar properties when they studied planning problems. They call this class of problems ‘‘wicked’’. Buchanan (1992) first attempted to answer the question why the design problems are wicked or, in Simon’s terminology, ill-structured. He concluded that “the answer to the question lies in something rarely considered: the peculiar nature of design subject matter. Design problems are wicked because design has no special subject matter of its own apart of what a designer conceives it to be.”

The designer therefore must use his/her judgment and experience in order to decompose and specify both the problem and solution. Designer’s commitment to the problem-solving process brings in additional domain specific knowledge that transforms the original ill-structured problem toward a well-structured one. There is a need to minimize the time taken to reach a satisfactory design solution whilst avoiding design errors that could have significant cost implications if taken forward into the product.

Normative and empirical design methods. How the design process is undertaken can be divided into those that come from a normative strain or empirical strain of design research (Stempfle & Badke-Schaub, 2002).Those belonged to the normative strain provide a prescriptive methodology of how design should be conducted. Research from the empirical strain reflects what designers really do in practice. Research demonstrates that design practice often differs from prescriptive methodologies of how design should be undertaken. Ball et al. (1994) found that designers sometimes focus on early solution paths that lead to satisfactory though often sub-optimal solutions to the design problem. Davies (1991) analyzed the design behavior of computer programmers. He found that although much of their design activity conforms to a top-down decomposition strategy, they occasionally make opportunistic shifts to lower level components of the design. Ehrlenspiel (1999) claims that deviations from a prescriptive methodology are often a result of economic, time and other constraints on the real life design process. This suggests that although designers do often conform to a normative design strategy, deviations do occur, often due to strategic knowledge or constraints and there is therefore a limit to what extent a detailed design methodology can be imposed on designers. There is therefore a need for any methodological design support to be sufficiently flexible not to prohibit variations in the design process.

2.4 Domain expertise in decision making

Expertise impacts the decision process in several ways. First, experienced decision makers exhibit high competence within their domain and have accumulated a vast repertoire of instances or cases they can draw on (e.g., Chi, Feltovich & Glaser, 1981; Chi, Glaser, &Farr, 1988; Klein, 1993a, 1993b; Patel & Groen, 1991; Simon

&Chase, 1973). Supporting the relationship between domain expertise and high-quality outcomes, Charness and Tuffiash (2008) noted that experts' superior performance in domains such as sports, medicine, and aviation is directly dependent on their domain-specific knowledge and that this knowledge enables experts to anticipate future actions and prepare for them more efficiently.

Second, experts see and process information differently than novices do. It is important to note that in naturalistic environments experts are able to quickly identify the most ecologically valid cues—that is, the subset of information most critical to accurate situation assessment—and to match these cues against patterns in their experience base. For expert firefighters, for example, cues such as the sponginess of the floor predict the path of a fire. Experienced pilots match the shape, color, and size of clouds with patterns in memory to determine whether they should fly through them or around them. The ability to use pattern-matching processes in naturalistic environments enables quick diagnosis and decision making; experts are able to comprehend quickly what may be at the core of a problem and what actions should be taken (Cellier, Eyrolle, & Marine, 1997; Orasanu & Connolly, 1993). This is a critical factor in time-constrained situations.

Experts are also sensitive to constantly changing values of information and adapt their mental models accordingly (Waag & Bell, 1997). Often, decisions are made incrementally and iteratively as decision makers use feedback from the environment to adjust their actions and are able to change course in the middle of a situation when a decision is not working effectively (Connolly, 1988; Lipshitz, 1993). In the medical domain, physicians often monitor results of a treatment to refine their diagnoses (Orasanu & Connolly, 1993).Situation assessment is characterized by action/feedback loops, as experts use an iterative process to incorporate changes that result from incremental decisions.

In addition, experts can employ strategies that enable them to cope with ambiguity, dynamic conditions, and time pressure. They are proactive, anticipating potential failures, risks, or conflicts, and prepare for them. They are aware of time constraints and know how to "buy time." For instance, pilots may request a holding pattern to gather additional decision relevant information (Orasanu, 1990). To manage time wisely, experts anticipate developments, make contingency plans, prioritize tasks, and use low-workload periods to prepare for upcoming events (Fischer, Orasanu, & Montalvo, 1993; Orasanu, 1990; Xiao, Milgram & Doyle, 1997a). Xiao, Milgram, and Doyle (1997b) found that expert anesthetists planned points for consideration that enabled them to activate knowledge and rules for problem regulation, action preparation, and arrangement of material assistance. In other work, expert physicians were found to activate a significantly larger proportion of knowledgebase control for regulating conflict—problems involving contradictions between surgical goals and patient status and properties—than did intermediates and residents (Morineau et al., 2009).

2.5 Information seeking in design

Choosing an alternative in the down select process mentioned previously at the "nature of engineering design processes" section is tightly tied with information seeking. At many points in the design process, designers lacked sufficient information to make informed comparisons between alternatives, particularly during conceptual design stage. There are many ways that designers can seek information. Sometimes they create the information themselves by developing more detailed drawings of targeted areas of a design or by building and testing prototypes. Sometimes information is produced through analytical methods such as calculations and simulations. And sometimes it is collected from external sources. Some information seeking activities require significant effort, knowledge, and cost. Designers must make judgments about when the cost of information seeking is likely to pay-off in the final product. An issue pointed out by Hayes and Akhavi (2008) was that for student designer teams (senior undergraduate), they did not always know when to seek more information or when to stop. Meeting time constraints was very important in forcing them to think critically about what information was the most important and to focus their information seeking efforts. Furthermore, the study also shows that student designers were far more likely to seek information in areas where they felt knowledgeable and comfortable and to avoid seeking it in areas unfamiliar to them. For example, they were very comfortable elaborating physical, three-dimensional details and conducting mathematical analyses of specific aspects such as stress and torque, but they were far less comfortable developing cost and manufacturability estimates. They did not know where to look for cost information or who to call or consult. They also often grossly underestimated costs and appeared completely unaware of the degree of uncertainty in their estimates. This avoidance is simply a hesitancy to engage in an effort of unknown magnitude for an unknown benefit. (Hayes, C. C. and Akhavi, F. 2008)

Bradley and Agogino (1994) also described information seeking as an important part of the process through which design alternatives are selected. They studied this process in the context of automated selection of design component choices from a catalog. They described a mathematical formulation which can be used to decide when it is worthwhile to expend the cost and effort required to gather additional information. However, the method required the designer to put in effort to collect the input data for the method, which they may not be willing to do if they believe they can fare almost as well without using any special analysis for information seeking decisions.

In designing DSSs, understanding when and how to seek information may be as important, if not more important than comparing alternatives, because good comparisons are predicted based on appropriate and sufficient information. Weick (1995) and Heuer (1999)'s research shows that accuracy increases with data elements up to a point (perhaps 8-10 data elements) and then asymptotes while confidence continues to increase (Oskamp, 1965). Lanir (1991) studied cases on intelligence failure, such as the Yom Kippur war, and found that in each case the decision makers already had sufficient data. Their difficulty was in analyzing these data. Gathering additional data would not necessarily have helped them. In a recent naturalistic observation of twenty one Navy forecasters at the Naval Pacific Meteorological and Oceanographic (NPMOC) facility in San Diego (Smallman & Hegarty, 2007), they found that when performing a weather forecasting task, forecasters accessed weather maps that were more complex than they needed, displaying variables that were extraneous to their task. Omodei, Wearing, McLennan, Elliott, and Clancy (2004) showed that at some point increasing data may lead to worse decision quality. To make the complicated problem even more complex, domain novice most likely lack the resource and knowledge in terms of when and where to seek additional information. They could benefit from guided information seeking plan much more than domain experts.

This dissertation can expand the previous literature by investigating how expertise impact design decision making. More specifically how decision makers with different levels of domain expertise may react differently on uncertainty visualization and formulate information seeking plans.

2.6 Information seeking and the relationship to uncertainty

In information science, the idea of uncertainty underlies all aspects of information seeking and searching. Mignerey, Rubin, and Gorden (1995) show the relationship between uncertainty and information seeking in the case of the newcomer to an organization, and Borgers et al., (1993) show the relationship between experienced uncertainty and the information-seeking behavior of cancer patients, Belkin's (1980) built a concept of the Anomalous State of Knowledge which is "An anomalous state of knowledge is the one in which the searcher recognizes a gap in the state of knowledge. The awareness of this state leads to information searching." Kuhlthau's (1993a) linked affective states to information-seeking stages. Kuhlthau's

(1993a) six-stage model (Initiation, Selection, Exploration, Formulation, Collection, Search closure) of the information-seeking process links uncertainty to various stages in that process, specifically in the Initiation stage, when a person first becomes aware of a need for information, and during the Exploration stage, when the individual is seeking to establish the general field of the problem. The six stages and their brief introduction can be found in Table 1.Kuhlthau (1993b) has also proposed uncertainty as a basic principle for information seeking, defining uncertainty as: “a cognitive state which commonly causes affective symptoms of anxiety and lack of confidence.” and, drawing upon her research, noting that: “Uncertainty and anxiety can be expected in the early stages of the information search process. Uncertainty due to a lack of understanding, a gap in meaning, or a limited construct initiates the process of information seeking.”

This collection of researches informed the author that in order to help decision maker understand what information is worthwhile to pursue can be achieved by letting decision makers realize more easily and accurately where they have uncertainty in decisions.

Table 1.Kuhlthau’s (1993a) six-stage model of information-seeking process

Stage 1: Initiation	During the first stage, the information seeker recognizes the need for new information to complete an assignment. As they think more about the topic, they may discuss the topic with others and brainstorm the topic further. This stage of the information seeking process is filled with feelings of apprehension and uncertainty
Stage 2: Selection	In the second stage, the individual begins to decide what topic will be investigated and how to proceed. Some information retrieval may occur at this point. The uncertainty associated with the first stage often fades with the selection of a topic and is replaced with a sense of optimism.

Stage 3: Exploration	In the third stage, information on the topic is gathered and a new personal knowledge is created. Individuals endeavor to locate new information and situate it within their previous understanding of the topic. In this stage, feelings of anxiety may return if the information seeker finds inconsistent or incompatible information.
Stage 4: Formulation	During the fourth stage, the information seeker starts to evaluate the information that has been gathered. At this point, a focused perspective begins to form and there is not as much confusion and uncertainty as in earlier stages. Formulation is considered to be the most important stage of the process. The information seeker will formulate a personalized construction of the topic from the general information gathered in the exploration phase.
Stage 5: Collection	During the fifth stage, the information seeker knows what is needed to support the focus. Now presented with a clearly focused, personalized topic, the information seeker will experience greater interest, increased confidence, and more successful searching.
Stage 6: Search closure	In the sixth and final stage, the individual has completed the information search. Now the information seeker will summarize and report on the information that was found through the process. The information seeker will experience a sense of relief and, depending on the fruits of their search, either satisfaction or disappointment.

2.7 Team decision making

In many professional settings decisions are carried out by teams of experts. For instance, medical decisions involved in patient diagnosis and care often require the input of different specialists and the collaboration of physician and nurse teams. Similarly, decisions during military or firefighting operations and other high-risk, dynamically changing and time-pressured tasks are the product of a team effort even though one individual may ultimately bear the responsibility for the decision. In the same domain with this dissertation, decision making by engineering design teams involves time and effort management, conflict resolution, and collaborative problem solving. Decision making in these contexts is a complex process, as information from many sources has to be integrated into a coherent representation of the situation and errors may be fatal. Because of this, merely putting several individuals to work ensures neither that they function as a team nor that their decision making will be better. Research noted the unwillingness of team members to collaborate with others (Perry & Wears, 2011) or to correct inaccurate diagnoses and decisions (Orasanu et al., 1998; Tschan et al., 2009), suggesting the presence of social phenomena, such as "silo" thinking (individualism), conformity (matching own behaviors to the norm of the society or social group), and confirmation bias (a tendency to favor information that confirms own preconception and hypothesis regardless of whether the information is true).

On the other hand, the benefits of decision making as a team are the collective knowledge, team communication and the consensus on decisions. When working on a project alone, you can only do little in the sense of brainstorming. By function as a team, individuals often find themselves inspired by other people's creativities. The ability to brainstorm with people makes working as team effective both in terms of time management and the end result. A few researches in product development domain pointed out that team can reduce development costs (Brown & Eisenhardt, 1995; Kessler & Chakrabarti, 1996). Also team can shorten the time from idea to commercialization, especially if the project activities are conducted simultaneously (Brown & Eisenhardt, 1995). Asan example, Cooper and Klein schmidt (1994) concluded that a cross functional and dedicated product development team was the most important factor associated with project timeliness.

By analyzing cases in which teams have failed and by comparing successful and unsuccessful teams, researchers have identified components of effective teamwork. First, there needs to be a sound shared cognition or situation awareness between team members (Orasanu, 1994; Rouse et al., 1992; Salas, Cooke, & Rosen, 2008; Salas, Sims, &Burke, 2005). Second, communication is essential to team decision making. Whether team members are inches apart or hundreds of miles distant, they need to share task-critical information and to ensure mutual understanding. They need to call attention to problems in the task environment or concerning their teamwork, to direct the actions of others, and to bring their expertise to their team's decision making.

2.7.1 Shared Cognition and Situation Awareness in Team

Coordinated action by team members requires that they have critical domain knowledge in common, in particular knowledge concerning the operational environment, equipment, standard procedures, practices, and strategies. In addition, they must have a common understanding of their team's mission: what their objectives are, their plans, their individual roles and responsibilities, and how they as a team are to interact (Cannon-Bowers & Salas, 2001).

Shared task and teamwork knowledge and, to a lesser degree, common domain knowledge were found to foster effective team processes (i.e., extent of planning, collaboration, and communication), which in turn increased task performance (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Smith-Jentsch, Cannon-Bowers, Tannenbaum & Salas, 2008). Shared task and teamwork knowledge may also increase team members' sensitivity to the information needs of others, resulting in implicit coordination—that is, team members volunteer task-relevant information before it is requested (Entin & Serfaty, 1999; MacMillan, Entin, & Serfaty, 2004; Stout, Cannon-Bowers, Salas, & Milanovich, 1999).

For successful task performance, team members must also have a shared understanding of current task conditions and shared expectations of future events(Mosier & Fisher, 2010).. As teams operate in dynamically changing environments, plans that were appropriate at some point in time may become inadequate. Thus, accurate situation assessment is critical. Team members must monitor their task environment for cues and information that are inconsistent with their original assumptions and may call for a change of action. Orasanu (1994) employed the term shared situation models to capture the dynamic nature of team members' situational awareness.

2.7.2 Team communication

Although the importance of team communication is critical, its successful execution is anything but guaranteed. Failures and breakdowns in team communication have been implicated in many accidents and incidents, in commercial aviation (Cushing, 1994; Fischer & Orasanu, 2000; Orasanu, Fischer, & Davison, 1997), offshore oil drilling operations (Flin, O'Connor, & Crichton, 2008), and health care (Leonard, Graham, & Bonacum, 2004; Reader, Flin, & Cuthbertson, 2008).Effective and efficient team communication is especially critical in task environments in which conditions are dynamically changing and decisions have to be made under considerable time pressure and in the face of ambiguity concerning the nature of the problem(Mosier & Fisher, 2010).It has been reported that successful team performance is associated with explicit communication. Research showed that well-performing teams shared more information about the problem they faced and talked more about their response to it than did poorly performing teams (Krifka, Martens, & Schwarz, 2004;Mazzocco et al., 2009; Sexton & Helmreich, 2000); in particular, high-performing teams were more likely to articulate new plans and changes in task allocation and expectations about future events (Gillan, 2003; Orasanu & Fischer, 1992).Members of successful teams were also more explicit about their reasoning and their intentions and tended to contextualize information (Johannesen, 2008; Tschan et al., 2009).

In this dissertation, to measure the effectiveness of team work, shared situation awareness and team communication are taken into consideration. Moreover, this dissertation also expends the understanding of team decision making to the naturalistic engineering design environment when the decision makers are real engineering designers and the tasks are real design tasks.

Chapter 3 -- Methodology

This research is following a user-centered approach to (1) understand decision making in the context of product design, (2) identify struggle point of engineering designers that needs to be supported, (3) inform the design of decision support system to facilitate decision making in product design and other domains, and (4) evaluate the system in different scenarios. In order to accomplish the research goals, four methods were used in the work: (1) ethnographic observations, (2) interviews, (3) field study and (4) controlled laboratory studies.

The author did a preliminary study to inform the design of a DSS and to develop hypotheses on how the DSS might change the decision making process. The studies provided a starting point for understanding designers' needs and design processes. Two experiments, one studying the impact of the DSS on individual decision makers, and another studying the impact on decision making teams were conducted in serial.

3.1 Methods

Ethnographic observations are observations of work as it is carried out in a normal setting. Ethnographic studies are particularly useful for understanding the needs and constraints imposed by an actual work environment. In this research, the author cooperated with a senior mechanical design team, whose weekly meeting was observed. In results, required functionalities were identified based on designer's needs. This process makes the decision support tool, which is part of the outcome of this research, useful to the designers. Detailed procedure and discussion can be found in the "Preliminary works by author" section.

Laboratory study is more controlled than ethnographic observations. While the situation in laboratory studies may be somewhat artificial, they allow measurement and quantification of phenomena in a way that ethnographic studies cannot. In this research, there will be one lab study with controlled variables to test hypothesis which will be discussed in the “experiment design of lab study” section.

Longitudinal field/naturalistic study is a research study that involves repeated observations of the same people over long periods of time. In a longitudinal study participants are followed over time with continuous or repeated monitoring. As a result, it provides data about the same individual at different points in time allowing the researcher to track change at the individual level. In this research, this longitudinal study comes in hand with observation of designer’s naturalistic decision making process. Designer’s weekly meeting is observed and activities are identified in a weekly base. It is anticipated that this study will take 4 months.

Each of these study types, ethnographic, laboratory studies, and longitudinal field/naturalistic studies can provide different views of the complex phenomena associated with product design processes. Together they provide a mix of qualitative and quantitative data that allow construction of a richer overall picture than is possible using any one method alone.

In the remaining part of this dissertation, topics will be covered in the following order: Chapter 4 talks about using ethnographic observation and interviews to understand the engineering design context and designer’s behavior and needs. Chapter

5 describes how the needs transform into DSS requirement and the implementation of the Uncertainty DSS. Chapter 6 discusses the first experiment with engineering designers using the Uncertainty DSS individually. Chapter 7 talks about the rationale of studying team environment and changes to the Uncertainty DSS after the first experiment. Chapter 8 discusses the second experiment with engineering designers using the Uncertainty DSS as a team. Chapter 9 summarizes the work of this dissertation.

Chapter 4 -- An Ethnographic Study of Engineering Designers

In order to get a high-level picture of what engineering designers do during a project and what the major challengers they are facing, the author worked cooperatively (conducted interviews and observations) with 5 mechanical designers (from ME 4054: Design Projects) to study their design process throughout the semester. It is worth mentioning that this study is a fairly informal study combined with observations and unstructured interviews. Observed activities and challenges were based on the author's subjective interpretation.

4.1 Methods

In Spring 2009, in order to understand the context of engineering design and identify engineering designers' requirements of making design decisions, the author worked cooperatively with 5 mechanical designers (from ME 4054: Design Projects) to study their design process throughout the semester.

The author observed the designers' 12 weekly design meetings and took notes. All five designers were interviewed separately at week 4. Each interview lasted roughly 20 minutes. At week 4 they had worked long enough to become familiar with the task, but they had not yet made a specific choice as to which idea to develop further. In the early weeks of the process, designers explored many possible ways of addressing the problem, but at week 5 they were required by the class instructor to choose a single, most promising idea for further development. Otherwise, many teams would be inclined to exploring possible ideas until it was too late to actually build and test any one idea effectively.

4.2 The Designers' Task

The designers' objective was to design a robotic arm for use by a specific quadriplegic person to support some daily activities. The robot arm must be capable of manipulating a variety of lightweight objects found in the man's home and office environment such as paper, small books, compact discs, and soft drink cans. It must have a control interface that a quadriplegic person can manipulate and be powered by the on-board battery system of his electrically powered wheel chair. Additionally, it must be simple for an assistant to mount and un-mount from the arm of the wheel chair. The electronics and motors must be reasonably weatherproof, light weight, and inexpensive. Finally, the student designers must build and test their best design. The student designers are a good resource for author to do research with. The author will observe design processes, and try out decision support tools together with student designers. Figure 2 demonstrates a previously created robot arm in action.

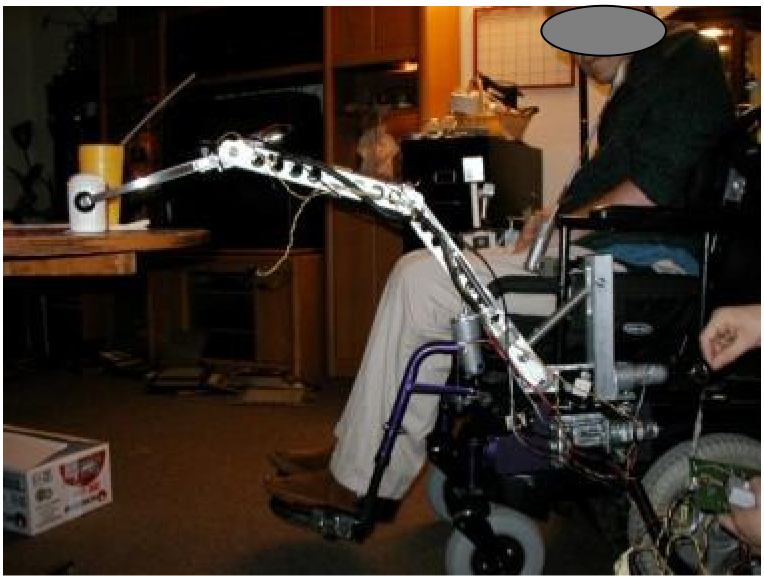


Figure 2. Robotic arm prototype for a quadriplegic man.

4.3Major Observations

- Designers tend to draw sketches on paper and whiteboard. The visual representation of design ideas helps designers communicate as a team and compare design alternatives.
- Designers seek information frequently at initial design stage (e.g., concept abstraction, compare design alternatives).At this stage, more information benefits designers' decision making. Information can be obtained by many methods including searching the web, by further development and detailing of design alternatives, mathematical analysis, simulation of performance, or building and testing prototypes. One of the observations was that entry-level designers lack the knowledge of how to seek information effectively; they were easily distracted and often lost focus. Furthermore, they are not always even aware when they are lacking information necessary to make a choice (as verified in the subsequent experiment described below).
- When comparing design alternatives, student designers draw them on cards and post cards on the wall. They use fast elimination to get rid of the less favorable ones. They also use decision matrix formed in Excel to quantitatively compare design alternatives.
- Using decision matrix and comparing design alternatives not only help designers find a good design but also help them think thoroughly about design goals and each alternatives.

- Designers use a lot of intuition and heuristic experience to solve problems, such as how much should the arm weigh, how much does the material cost, how fast the gripper should close, etc.
- Designers constantly refine their design alternatives when more information is available (e.g., talking to target subject, online information seeking, etc.) Most information seeking behavior took place in the initial design stage.
- Decision matrix (multi-criteria, weighted sums decision making process) was used to compare design alternatives, discuss the merits and draw backs of each, and negotiate to select a best option. Forming decision matrix was a team activity, and each score was assigned with a team consensus. They implemented it in an Excel spreadsheet and projected it on a white board in the lab where their meetings were held. They could then use this decision matrix as a discussion tool to facilitate team negotiation and design making. (This is likely a fairly common practice, not unique to this design team). Design concepts were listed on one axis (e.g. two fingered gripper lined with rubber, three fingered gripper, air suction device, etc.) and criteria (cost, manufacturability, usability, etc.) were listed on the other. In team discussions, the design team negotiated the criteria, created weights for criteria, etc. If there is confliction on the score, the team resolved confliction by spend more time discussing different ideas. In some cases one idea was convinced by another idea, and in the others they postponed decision making because of lack of critical information.

4.4 Major Challenges Based on Observation

- Decision support tool should be easy to use. Data input should be less cumbersome and the procedure of comparison should be intuitive. Decision makers are likely to move away from the tool if it requires a lot of effort to learn.
- Sometimes assigning specific scores to designs is neither helpful nor useful because designers do not have enough information to decide the value of score.
- In many cases, how to spend limited time on valuable topics is challenging. This is partly because bad time management and partly because valuable topics are not necessary clear to engineering designers.
- When decision makers feel there is a tie between designs, they had hard time browse through pros and cons of designs.
- Designers had difficulty agreeing on specific individual scores to assign for each alternative and criteria because there is usually much uncertainty associated with each design alternative, especially in the early design stages. This observation is similar to the results of other studies (Aughernbaugh & Paredis, 2006; Hayes & Akhavi, 2008; Orasanu & Connolly, 1993).This is an inherent challenge faced by all designers. On the positive side, it has been witnessed that going through multi-criteria decision process facilitated team decision making interactions, structure discussions, and articulate design goals and properties.

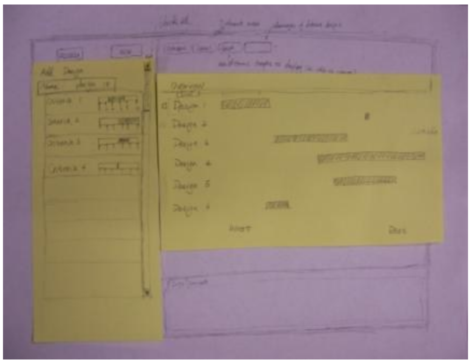
Based on the observed activities and challenges from designers, the author proposed a DSS to address their needs for flexibility, guided information seeking and awareness of uncertainty. The DSS took the format of decision matrix since the general concept is familiar to many designers thus it would provide a good, basic framework for the DSS. The DSS added flexibility by granting users the capability to express uncertainty as a range, eliminating the cumbersome of assign single values in the matrix. Finally, the uncertainty in the overall score for each design is expected to help designers understand when they have enough information to make decisions, as will be described below.

Chapter 5 -- Development of Uncertainty Decision Support System for Engineering Design

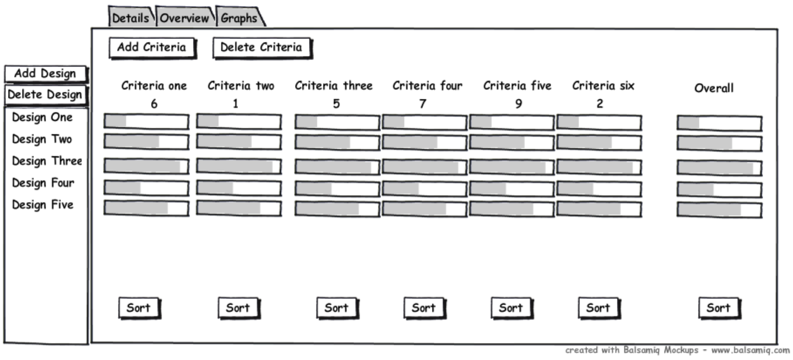
Based on the findings of the ethnographic field study, a DSS was proposed with the aim to address the needs of engineering designers (e.g. to compare design alternatives, express uncertainty about the parameters for each option, and identify when they have enough information to make a choice between alternatives). The envisioned DSS takes the format of decision matrix with the aim of limiting the effort needed to input data while visualizing uncertainty information to inform information seeking activities. The uncertainty DSS allows designers to quickly identify areas where seeking more information may be beneficial, but to use their own judgment and experience in deciding by what methods to gather information, and if the benefits are likely to be worth the costs of getting that information. Thus it provides them with some useful but approximate guidance, without a large data entry burden of complex, experience-based knowledge.

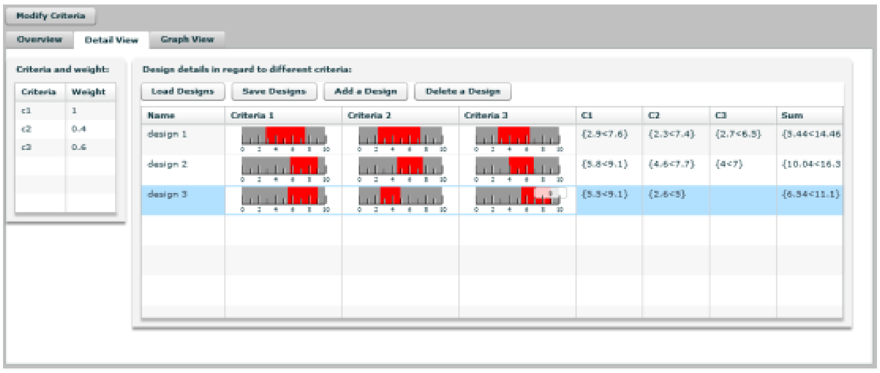
Drawings and paper prototypes can be found in Appendix A. After several rounds of design ideation, usability and functionally testing (Figure 3 shows the evolution of the interface), the first finalized version of the DSS is illustrated in Figure 6. One thing worth mentioning is that to compare the impact of uncertainty in DSS, after finalizing Uncertainty DSS, a Certainty DSS which is shown in Figure 5 was created.

The DSS is created from scratch and is implemented in Adobe Flex SDK. Adobe Flex is a software development kit released by Adobe Systems for the development and deployment of cross-platform rich Internet applications based on the Adobe Flash platform. It is challenging in traditional programming language to support rich graphical user interfaces and modern styles of user interaction. Flex seeks to minimize this problem by using MXML, an XML-based markup language, offers a way to build and layout graphic user interfaces. Interactivity is achieved through the use of ActionScript, the core language of Flash Player



.





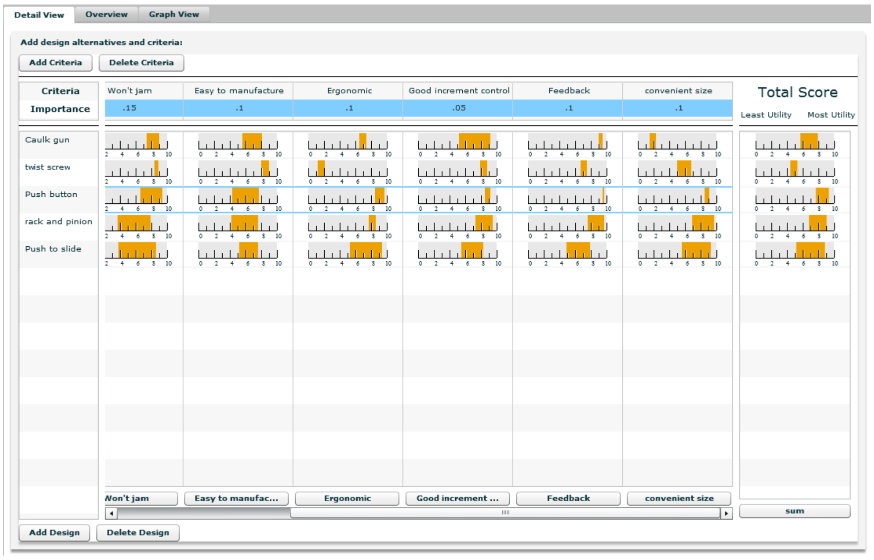


Figure 3. The evolution of uncertainty DSS interface

Both DSSs are based on the decision matrix structure of MCDM methods. This has the advantage that decision matrices are familiar concepts to many designers. The DSSs are not specific to any particular domain; they provide general purpose structures in providing decision makers with the flexibility to define their own alternatives and decision criteria to reflect their specific task. For the sake of this section, Certainty DSS will be introduced first and then Uncertainty DSS. The two DSSs differ in the following ways. Uncertainty DSS provides users with a simple approach for representing and visualize uncertainty. Most importantly to this discussion, the Uncertainty DSS allows users to visualize the extent to which the uncertainty in alternatives overlaps with each other. This concept was named "uncertainty overlap". It will be explained through detailed examples below, and it is used to help decision makers identify situations in which critical information is lacking.

5.1 The Certainty DSS

Figure 5 shows the interface for the certainty DSS. It implements a standard, weighted sums MCDM method for comparing alternatives. Its functions are described below.

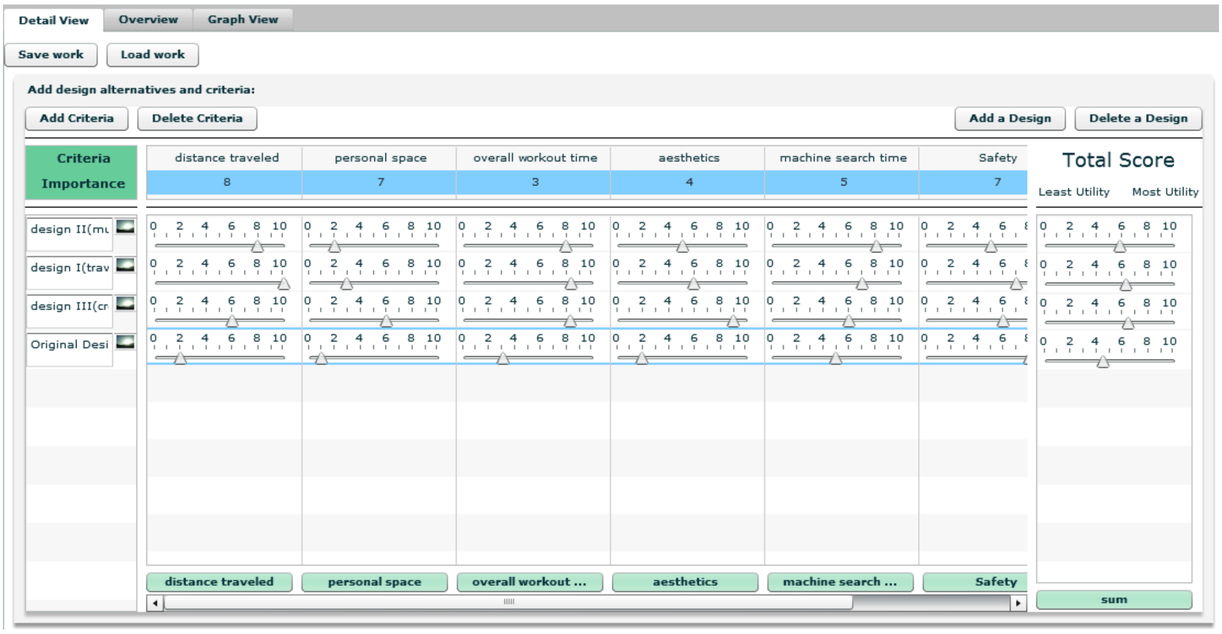


Figure 4.The interface for the Certainty DSS.

5.1.1 Rows

The name of each design alternative can be entered at the left end of each row. The names can be anything that is meaningful to the designer. For example, the names used for the alternatives for the end effecter on the device for the paraplegic person could be: “rubber tipped finger,” “multi-fingered gripper” and “suction device.”

5.1.2 Columns

The names of the criteria which are relevant to their project can be entered at the top of each column. Criteria for the end effecter could include: “material cost,” "manufacturing cost," "reliability," etc. A number between 0 and 10 representing the "importance weight" must be entered immediately under each criterion. Larger weights indicate more important criteria, and smaller weights are less important criteria. Decision makers determine the weights based on their perceptions of the relative importance of each criterion. For example, cost is often given a large weight. There is no right or wrong answer for what the weights should be, but they should accurately reflect the values of the stake holder in decision.

5.1.3 Matrix Cells

Each cell in the matrix created by the rows and columns represents a value indicating the degree to which that design alternative fulfills that criterion on a relative scale between 0 and 10 (where 0 indicates that the design alternative does not fulfill the criteria at all, and 10 indicates it does so completely). The Certainty DSS allows users to enter only one value for each design parameter. This value is chosen by the decision maker. If they do not yet know the precise value they must pick a single value to enter. For example, if the designer does not yet know the precise material cost for a vacuum device he or she might enter an estimate of the likely cost.

5.1.4 Total Score

The Certainty DSS calculates a Total Score for each alternative as the weighted sum of all values in the row. Thus, the Total Score for alternative_i is:

$\text{Total Score}_i = (\text{value}_{i1} * \text{weight}_1) + (\text{value}_{i2} * \text{weigh}_2) + \ldots (\text{value}_{in} * \text{weight}_n)$

The Total Score is displayed as single value, as shown in the right-most column of the interface in Figure 5. Total Score represents the overall value of an alternative to the decision maker. It can also be viewed as summary of the decision maker's opinions on the degree to which the alternative meets the project goals.

5.1.5 Adding, Deleting, Sorting

Users can add or delete additional columns as needed. If the rows or columns cover an area larger than the screen, users can use the scroll bars at the side and bottom to scroll through them. However, alternative names, criteria, weights and Total Scores stay in fixed locations on the screen. Users can also sort the criteria from most to least important, and they can sort the alternatives by Total Score, from largest to smallest.

5.2 The Uncertainty DSS

The top-level interface for the Uncertainty DSS is shown in Figure 5.

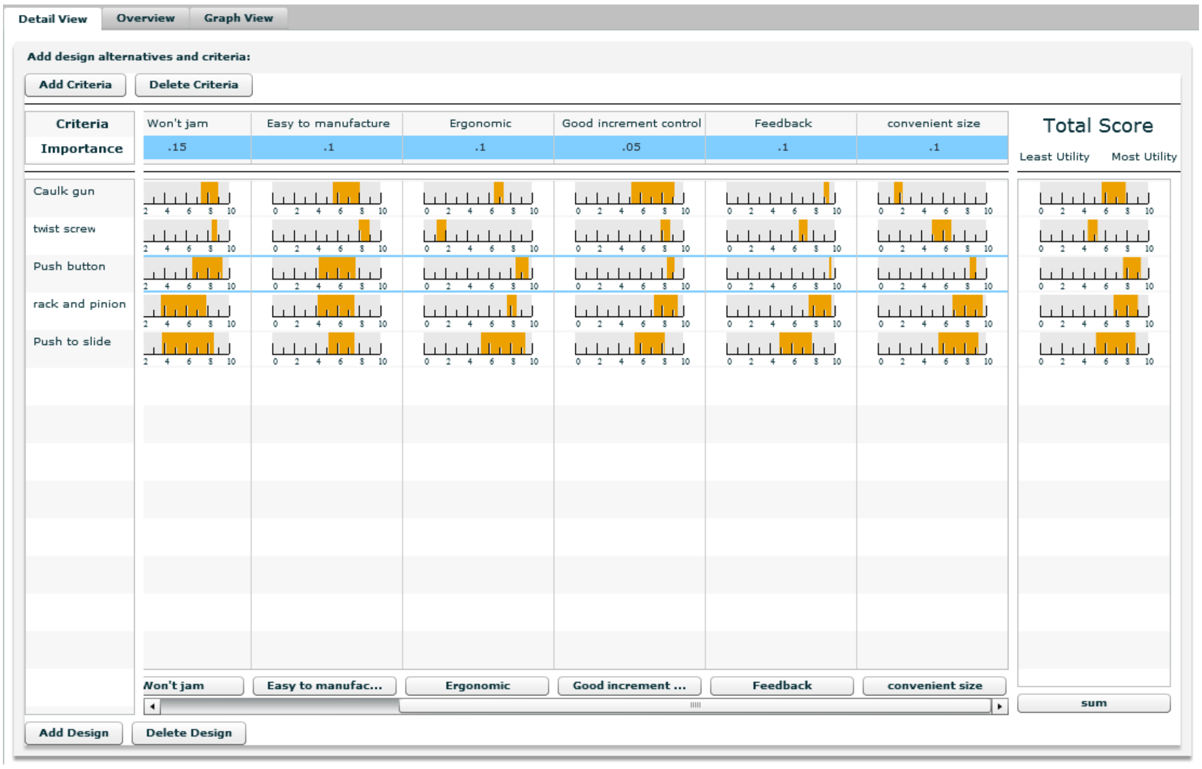


Figure 5. The top-level interface for the Uncertainty DSS.

This interface is the same as the Certainty DSS in most respects, but with several important differences. It:

- Allows designers to enter the *values* in cells as ranges,
- Displays *Total Score* as a range,
- Provides a simple visualization to assist users in identification of *uncertainty overlap* between alternatives.

The latter is used to help identify pairs, or teams of alternatives with overlapping ranges which may require more research to determine which is "best," where best is defined as highest total score.

5.2.1 Rows

The names of each alternative are entered at the right side of each row, just as in the Certainty DSS.

5.2.2 Columns

The names of the criteria and their weights are entered at the top of each column.

5.2.3 Matrix Cells

Instead of entering single values in each cell, values are entered as ranges using a pair of sliders. Thus, if the designer is uncertain about what exact value to enter, such as the manufacturing cost, he or she can enter a range of values. The distance between the sliders (e.g. the width of the bar) indicates how much uncertainty is associated with that value. Initially, the default state for a cell is total uncertainty: the two ends of the bar are set at the minimum and maximum ends of the range so that the bar extends across the whole width of the cell. If the designer has no information about a value, it can be left at the default. The default feature was actually an important part of the design because one of the drawbacks of MCDM methods is the time required to fill in all the values in the matrix (Hayes, 1981; Law, 1996; Hayes & Akhavi, 2008). The availability of a neutral default means that decision makers do not need to fill them all in. The range can be narrowed later if more information is obtained.

In designing this interface, a few complex visualizations of uncertainty were considered but dropped simply because it was more appropriate to start with a very simple representation. A simple representation is easy for users to learn and use, and creates the least visual clutter when many alternatives are displayed.

5.2.4 Total Score

The DSS computes a "Total Score" for each alternative which is displayed in the right-hand column as a bar representing a range of values (See Figure 2). The lower end of the range for "Total Score" is a computed as weighted sum of all the lowest values in each cell for that row. The upper end is a weighted sum of all the upper values in each cell. Thus, if lower $total_i$ and upper $total_i$ are the values at the upper and lower ends of the Total Score bar for $alternative_i$, and lv_{ij} and uv_{ij} are the lower and upper ends of the bar in $cell_{ij}$, then:

$$\text{Lower total}_i = (lv_{i1} * \text{weight}_1) + (lv_{i2} * \text{weigh}_2) + \ldots (lv_{in} * \text{weight}_n)$$

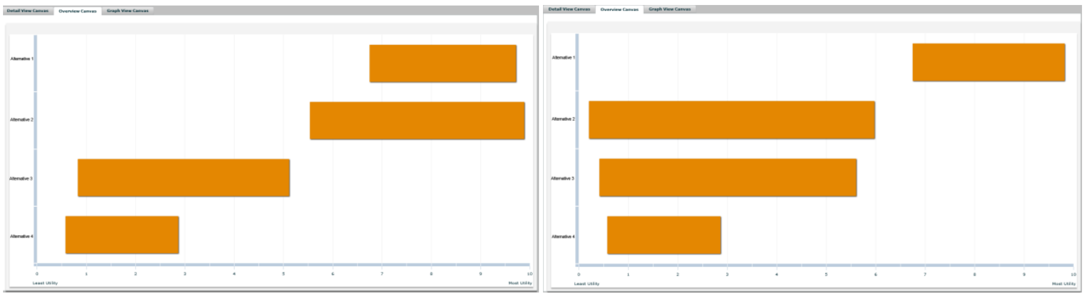
$$\text{Upper total}_i = (uv_{i1} * \text{weight}_1) + (uv_{i2} * \text{weigh}_2) + \ldots (uv_{in} * \text{weight}_n)$$

5.2.5 Adding, Deleting, Sorting

Rows (alternatives) and columns (criteria) can be added and deleted as in the Certainty Interface. Users can sort the criteria by weight, and alternatives by the Total Score.

5.2.6 A Visualization for Information Sufficiency

The displays in Figure 6a and b show expanded views of the Total Score column. Figure 6a shows an uncertainty visualization of a situation in which there is sufficient information to make a choice. Figure 6b shows the one in which there is insufficient information, as will be explained below.



a. b.

Figure 6.Visualizations of overlapping uncertainty ranges for the Total Scores of multiple alternatives.a. shows a situation in which there is sufficient information to make a choice. b. shows a situation in which there is sufficient information to make a choice.

<u>No overlap among top alternatives.</u> Sometimes, one alternative stands far above the others, as in Figure 7b. The range of its Total Score does not overlap with the ranges for the other three alternatives. A decision maker can feel reasonably confident choosing this option as the best of the set because in the expected worst case the first option will have a better Total Score than the others. However, there are no guarantees; the result depends on the accuracy of the information entered. Note that the bottom 3 alternatives in Figure 7b completely overlap with each other. This means that the decision maker lacks sufficient information to definitively identify which is the second best option (or which is the worst). However, this is not

a problem if one only needs to decide on the top option; the current information is sufficient to identify the top option.

Overlap among top alternatives. However, it frequently happens that several "best" alternatives overlap in their overall value, as shown in Figure 7a. While the average Total Score of the first alternative is better than the average Total Score of the second, the second may eventually prove to be better than the first. As more details are determined, and more uncertainty is resolved, the ranges of each of these alternatives will narrow, and they may no longer overlap.

Note that overlap between alternatives and information sufficiency are functions of the individual decision maker and what he or she knows about the set of alternatives. For example, Joe may lack sufficient knowledge and information to determine whether alternative A is better than alternative B, while Mary has information that allows her to say A is definitely better than B. Thus, given what Joe knows, A and B overlap; but from Mary's perspective they do not overlap and she has sufficient information to make an informed decision.

5.2.7 Seeking more information

When the top alternatives overlap, the decision maker may follow several courses of action. The designer may simply choose one alternative at random if both are good enough, or the designer may wish to seek more information in order to resolve the overlap. However, not all information is equally useful in informing the choice between alternatives. What decision maker need the most is information that will most reduce the width of the Total Score bars associated with the top alternatives.

Thus, the decision maker needs to find criteria of high importance (with large weights) that also have relatively wide bars.

As mentioned earlier, the design of Uncertainty DSS was to give engineering designers more flexibility to compare decision options, to guide information seeking plans and to support team decision making. Two studies were conducted to test the Uncertainty DSS. The focus of the first study was to test the individual use of the Uncertainty DSS. It is compelling to learn whether designers like the concept of uncertainty in DSS, whether the uncertainty can be understood and used properly by them and whether domain experience is a moderating factor. The goal of the second study was to test the Uncertainty DSS in a team using environment. In additional to the metrics in the first study, team situation awareness and system usability were measured.

5.3 Reduce effort and support iterative use

It is important to create interfaces that make it easy and fast for users to enter the information about their design alternatives. After all, if software is not easy to use, it is hard to be adopted by users even though it may be powerful once you learnt it. In preliminary research carried out by Hayes and Akhavi (2008), one observation is that designers regard data entry as painful task to do because they need to switch from mouse to keyboard and then require hand-eye coordination to input the data. Based on that, the author proposes a "click and drag" solution to input data. For each design in each criterion, mouse clicking is programmed as a way to input data. The mouse dragging can specify how certain users think the inputted data is. Now keyboard is not needed any more, because the only control to perform all the data entry is mouse. The process is easy and straightforward, all you need to do is click to anchor a start point then drag to the end point. The range will reflect how much uncertainty users think the data entry.

Another way to mitigate the effort to use decision support tools is to allow users to avoid data entry where possible. For example, when not using structured decision aids which required designers to enter a complete matrix of values for all criteria and all alternatives, designers were observed to take many "short cuts" to speed up comparison and avoid unnecessary data entry. These short-cuts involve avoiding detailed consideration of many values, alternatives and criteria, dropping out clearly inferior designs to a few "top" contenders. From that point on, further criteria were considered only for those top few alternatives, avoiding consideration of much data pertaining to the alternatives that dropped out.

The DSS allows designers to add or subtract criteria and alternatives; adjust weights and values at any time. This flexibility is important for allowing them to explore design problem, and to come re-cast and reconsider a major choice as their understanding of the problem deepens. The author’s decision support philosophy is that one should not try to support all activities but only a critical few where support can have greatest impact. Simple, "light weight" decision support tools that augment the skills of humans, often provide far more benefit than complex tools that attempt to replace or replicate complex human activities.

Chapter 6--Experiment One

The author conducted an evaluation of the Uncertainty DSS in which 22 engineering designers were recruited to compare sets of design alternatives using three comparison methods: no DSS (the control), the Certainty DSS and the Uncertainty DSS. The sets of design alternatives were drawn from on-going multi-month design projects in the participants' areas of expertise. This was an important aspect of the evaluation because it allowed the researcher to understand the impact of the DSS in a practical context.

6.1 Research questions

In this study, the following questions were investigated:

- Can uncertainty visualizations help designers understand when a lack of information interferes with the ability to make a choice?
- Can visualization of uncertainty overlap help designers to calibrate their decision confidence (i.e. to be more confident when there is sufficient information, and less confident when there is not enough information to make a clear choice),
- Can uncertainty visualizations encourage designers to seek clarifying information when appropriate?
- Does domain experience change the benefits that decision makers derive from the DSS?

It was hypothesized that the Uncertainty DSS would help users to recognize situations in which information was insufficient for making a choice; it would reduce their confidence in choices made in those situations; and it would encourage them to seek clarifying information. It was also hypothesized that designers with less experience would be assisted more by the Uncertainty DSS than those with more experience, because those with less experience are more in need of guidance.

6.2 Experimental design

A 2 x 3 within subjects design was used.

6.2.1 Independent Variables

The independent variables were:

- *Expertise level* (entry-level or intermediate-level designer), and
- *Design Comparison Method* (control system, Certainty DSS, or Uncertainty DSS).

6.2.2 Dependent Variables

The dependent variables were:

- *Perceived information sufficiency,*
- *Effort* to reach a decision*,*
- *Decision confidence*,
- *Plans to seek additional information*, and
- *Preference* between the methods.

Perceived information sufficiency, effort to reach a decision, and decision confidence were measured based on participant's answers to the statements: "I feel I had sufficient information to make an informed decision.", "I feel this method required more effort than should be necessary." "I am not very confident that the design(s) I chose were the best ones." The answers to these questions were marked on 7-point Likert scales from "Strongly Disagree" (scored as 1) to "Strongly Agree" (scored as 7).Perceived effort was measured rather than time because users' perceptions would be more important in influencing their behavior than the actual time spent. Similarly, perceived information sufficiency was measured so that it can be compared against actual information sufficiency (as defined by overlap) in order to assess to what degree the Uncertainty display was successful in changing their understanding.

Plans to seek additional information were assessed by first asking, "Is there a particular aspect of the design(s) on which you would like to know more in order to be confident of your decision? If so, please elaborate below and indicate how strongly you want to know about it." The degree of desire was measured by a 7-point Likert scales ranging from "Strongly Undesired" (scored as 1) to "Strongly Desired" (scored as 7).

Preference between methods was measured by asking participants to "split 100 points among the three decision methods you have just used. The most preferred system should get the most points." The three decision methods were: no DSS, the Certainty DSS, and the Uncertainty DSS.

6.3 Participants

22 participants with engineering design experience were recruited in this study. They were recruited from mechanical engineering and medical device design programs. Of the 22 participants, 12 were entry-level designers (senior undergrads

and one junior). This group was named "entry-level designers" rather than "novice designers" because they had significant design-relevant training in analysis techniques. The other 10 participants were categorized as intermediate-level designers (6 graduate students from the Mechanical Engineering Department and 4 graduate fellows from the Center for Medical Devices at the University of Minnesota. This program is a highly competitive program which recruits mainly highly accomplished practitioners in medical devices. All participants were currently working on design projects which had been underway for at least four weeks.

There was an average age difference of six years between the two groups. Entry-level designers reported on average of 1.5 years (SD = 0.84) design experience working on actual design projects, while intermediate-level designers reported 3.1 years (SD = 1.22) of design experience (See Table 2). This was a statistically significant difference ($F_{(1,21)}$=14.396, p=0.001).

Table 2. Design Experience and Age of Participants

Entry-level Designers			Intermediate-level Designers		
	Design Project Experience (years)	Age (years)		Design Project Experience (years)	Age (years)
Average	1.5	25.8	Average	3.1	31.8
Median	1.8	24.5	Median	3.0	31.0
Minimum	0.0	21.0	Minimum	2.0	27.0
Maximum	2.5	44.0	Maximum	6.0	43.0

It is important to note that the participants may have interpreted the background question "How many years of design experience (working on design projects) do you have?" in the same way. The intermediate-level designers were, on average, six years older than the entry-level designers and had spent more time in engineering jobs, yet they only reported a year and a half more design experience, on average. While this is possibly true it seems unlikely. The author suspect that the more senior participants (e.g. graduate students and fellows) had a higher standard for what they considered to be "design experience" than the undergraduates. In future studies design experience need to be defined more precisely, possibly asking separately about design experience as a student, and as a professional.

6.4 Design Tasks

A different set of alternatives was randomly assigned to each condition, for each participant. An important goal of this work was to understand how decision makers would use the Uncertainty DSS in conditions that were as realistic as possible. In order to do so, it is necessary to use real, as opposed to "controlled," design projects and designers in the experiment. However, using real tasks can create many challenges in the design of a experiment, as has been described in the literature on naturalistic decision making (Klein, 1993). The primary challenge is that real problems and natural settings are harder to control, and may introduce variability into experiments. However, the benefit is that results obtained in this way are more likely to reflect what occurs outside the laboratory. In the following paragraphs, some of the specific issues will be discussed when using real design tasks. In addition, a proposal to achieve a balance between a naturalistic and a controlled experimental design will be introduced.

Design is a complex process and the more artificial the evaluation setting, the more difficult it becomes to assess whether the findings are likely to translate into practical work environments. To reproduce realism in the design decision making tasks, it was important for participants to be knowledgeable about the specific design project, and be seriously vested in the outcome of the decision. Neither of these objectives can be achieved easily if participants were all given the same sets of small and artificial design tasks which they have never seen before.

Instead, of using artificial “laboratory” design tasks, the author recruited participants who were designers currently working on large design projects of several months in duration. The topics of the specific projects ranged from a dryer for plant DNA samples to a conveyer system for packaging beverage bottles. All designers had been working on their design projects for at least four weeks and were knowledgeable in their specific design topic. Several days before the evaluation they were asked to identify three subtasks in their design which were currently under development. For example, if the project was a desk-mounted device to enable a paraplegic person to manipulate objects on a desk, subtasks which the designer might identify included: design of an end effecter, design of a coupling between the end effecter and the device, and the motor system. Participants were asked to e-mail their three subtasks to the experimenters several days before the trials. Subtasks were then randomly assigned to the participant’s three trials (control, certainty DSS and uncertainty DSS).

6.5 Design Comparison Methods

Each participant was asked to compare different sets of design alternatives of their own choosing, using three different methods: a control condition (no DSS), the

Certainty DSS, and the Uncertainty DSS. In some cases the control condition was pencil and paper, and in others, a spreadsheet. The order in which they used the DSSs was systematically varied, so as to counter-balance learning effects.

6.6 Procedure

Each participant completed all tasks in the experiment individually. Participants were randomly assigned to one of the two experiment groups. The difference between groups was the order in which they used the Certainty and Uncertainty DSSs. Both groups used their normal method (control condition) first, then one group used the Certainty DSS followed and the Uncertainty DSS, while the other group did the reverse.

- At the beginning of each experiment, participants were given a brief introduction to the study. Participants signed a consent form which included agreement to audio recording, and answered questions on their demographics and design experience.
- Participants were asked to complete three trials. In each trial they used a different comparison method (control, Certainty DSS or Uncertainty DSS) to compare alternatives from three different design subtasks. Subtasks were assigned randomly to methods, and methods were presented to participants in a systematically varied order. Participants were asked to use each method to identify one alternative from a set which they judged to be most appropriate for further design development. Participants were given no time limit for completing design tasks.

- Participants were given a brief training session on both DSSs (before they used each system). The training sessions took roughly 10 minutes on each system.
- Immediately after using each method, participants were asked to complete a questionnaire measuring the dependent variables: information sufficiency, effort and decision confidence.
- After completion of three trials, participants were asked to report which method they most preferred by dividing 100 points between the three methods.
- Debriefing: at the end of the experimental session participants were asked to use the Uncertainty DSS to reproduce decision tasks used in the Certainty and control conditions. The purpose of this step was to determine the degree of uncertainty (e.g. overlap) that existed between those alternatives. This information was used later in the information sufficiency analysis.
- A week later, the participants were asked to report on their information gathering activities completed for their design project during the week.

6.7 Results

6.7.1 Method Preferences

In all cases, all designers listed the Uncertainty DSS as the most preferred; and no DSS was the least preferred in all cases with only one occasion in which an intermediate-level designer reported that no-DSS was tied with the Certainty DSS for last place. Furthermore, designers not only preferred the Uncertainty DSS, they strongly preferred it over the other approaches, as shown in Figure 7. (Note that since the DSS preference values are not independent measures, ANOVA is not the appropriate analysis to test the difference between systems. However the difference is very dramatic as shown in Figure 8) Participants' comments after the study confirmed these findings. They expressed that the Uncertainty DSS better reflected the uncertain nature of the design tasks. One participant stated, "It [the Uncertainty DSS] adds another dimension of what I can do" and "it is intuitive to draw uncertainty as a range."

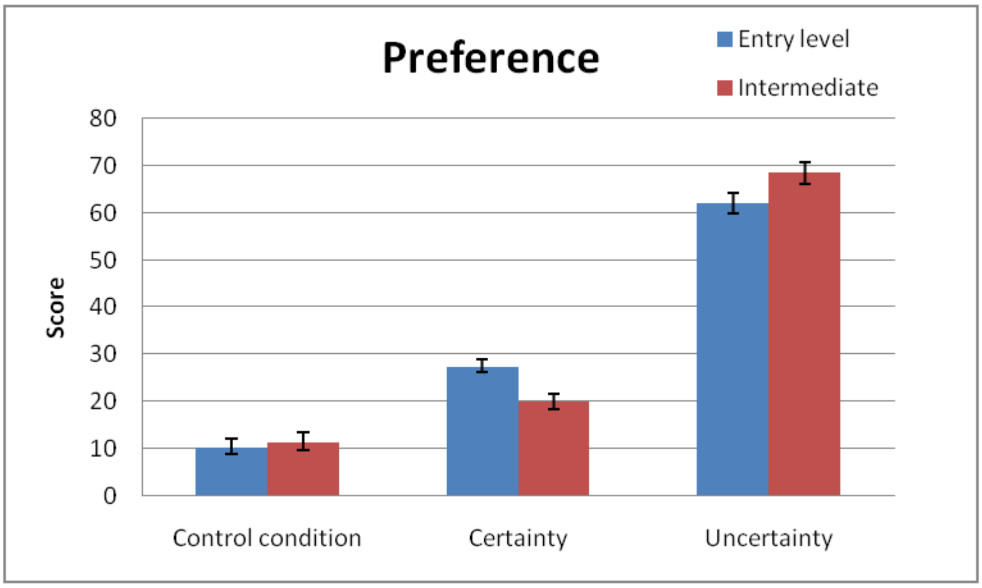


Figure 7.Participants' preferences for the three comparison methods. Participants were asked to divide 100 points between the three approaches. More points indicate greater preference. Error bars stand for standard error.

Note that preferring one approach over another does not necessarily imply that the preferred one will result in better performance. However, preferences are important in that they suggest a willingness to accept the Uncertainty DSS. People are unlikely to benefit from an interface they do not like, and are therefore unwillingly use.

In previous findings reported in Hayes and Akhavi (2008) the less experienced participants preferred a DSS similar to the Certainty DSS over one similar to the Uncertainty DSS, because of its conceptual simplicity and easier data entry relative to the Uncertainty DSS. Since then, the author has put in considerable development effort to make data entry easier for both DSSs. The next section will discuss the effort required for using each of the current DSS approaches.

As the last research question stated, it is valuable to see whether there were differences between the entry-level and intermediate-level designers, in how much they preferred each system. For any given DSS approach, entry and intermediate level designers ratings were given independently, so it was possible to test for significant differences based on experience. There was no difference in entry-level and intermediate-level designers' preferences for no-DSS($F_{(1,20)}$=0.21, p=0.656), a significant difference in their preferences for the Certainty DSS ($F_{(1,20)}$=12.92, p=0.002), and a marginally significant difference in their preference for the Uncertainty DSS ($F_{(1,20)}$=3.82, p=0.065).

6.7.2 Effort

Perceived effort can influence users' willingness to use each approach (No DSS, Certainty DSS, and Uncertainty DSS) outside the laboratory. For example, even if users say prefer the Uncertainty DSS over the other methods in a laboratory setting, they might not actually use it in a work setting with deadlines and time pressures if they perceive it to require more time than other approaches.

In particular, the first design goal with respect to effort was that the Uncertainty DSS should not increase perceived effort over using no DSS; ideally it should decrease effort. Furthermore, it was hypothesized that the Uncertainty DSS may require more effort than the Certainty DSS since users must enter a range in the system for every one number entered in the latter. While it is not desirable that the Uncertainty DSS should require the greater effort than Certainty DSS, it might be a necessary compromise. The second design goal was to minimize the effort required to use the Uncertainty DSS to a level not significantly different from the Certainty DSS.

Figure 8 shows differences in the effort participants reported for using the three different comparison methods.

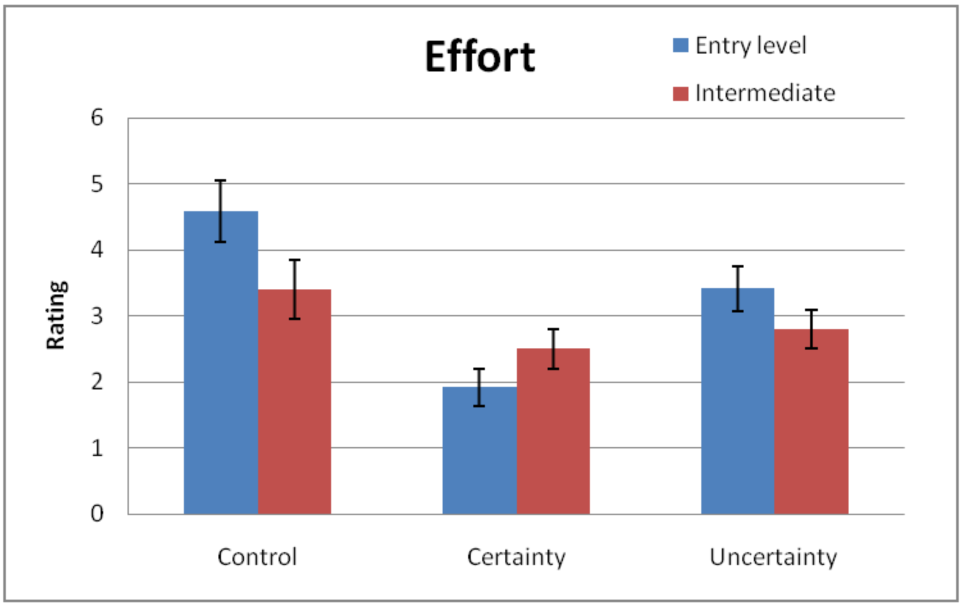


Figure 8. Participants' reported effort for using each method, as reported by designers. The error bars show standard error.

The intermediate-level designers found all DSS approaches to require about the same level of effort, ($F_{(2,18)}$=5.40, p=0.16). The pair wise comparisons were: no DSS vs. Certainty: p=0.24, Uncertainty vs. Certainty: p=0.58, Control vs. Uncertainty:

p=0.33.The entry-level designers reported that the Certainty DSS required significantly less effort than the Uncertainty DSS, or no DSS ($F_{(2, 22)} = 21.020$, $p < 0.001$). The pair wise comparisons were: no DSS vs. Certainty DSS: p=0.001, Uncertainty vs. Certainty: p=0.009, no DSS vs. Uncertainty: p=0.095.

When comparing the entry-level versus intermediate-level designers, the entry-level designers felt it required significantly more effort to use no DSS than did intermediate-level designers: $t_{two\text{-}tail}(20)=1.82$, p=0.08. However there were no significant differences between the two experience groups for the two DSSs: Certainty DSS: $t_{two\text{-}tail}(20)=1.39$, p=0.18; Uncertainty DSS: $t_{two\text{-}tail}(20)=1.39$, p=0.18.

From this it is concluded that the system met the first design goal: the Uncertainty DSS did not increase perceived effort over using no DSS for either group, as mentioned above, this is treated as a good outcome, because it implies that the users do not feel that using the Uncertainty DSS was more work than doing the task by hand or with Certainty DSS. In addition, the second design goal was partially met: only intermediate-level designers experienced no significant difference in effort between the Certainty and Uncertainty DSSs. This is an interesting result given that the entry-level designers preferred the Uncertainty DSS over the Certainty DSS. So entry-level designers found Uncertainty DSS took more effort to use but still like it more. There are several possible explanations for this. First, the Certainty DSS requires half as much data entry as the Uncertainty since the Uncertainty DSS requires entry of a both minimum and maximum for each value in the matrix. But this does not explain why only the entry-level designers found the Certainty to be less effort. Earlier studies of manufacturing planners found that novice machinists spent less time planning because they skipped many steps due to ignorance of all the issues that needed attention (Hayes and Wright, 1989). Similarly, the entry-level designers may have been less inclined to spend effort thinking about uncertainty when they did not have the uncertainty visualizations to remind them. The intermediate-level designers, on the other hand, appeared to be more concerned about uncertainty overall. Finally, the entry-level designers struggled more than intermediate-level designers when working without a DSS. Thus, the effort required would not deter decision makers from using the Uncertainty DSS, and the entry-level users may be positively motivated to adopt decision support system.

6.7.3 Perceived Information Sufficiency

The hypothesis of primary interest in this study was that the Uncertainty DSSs would help participants to become *more aware* of situations in which information was insufficient to clearly identify the best alternative. It would do so by allowing decision makers to visualize overlap in the Total Scores of top alternatives. The author also hypothesized that when participants did not have the Uncertainty DSS available to them (e.g. when using no DSS or the Certainty DSS) they would not distinguish between situations in which there was sufficient information (e.g. no overlap) from those in which there was insufficient information (e.g. overlap).

The first step in the analysis was to apply a two way ANOVA using DSS approach (no DSS, Certainty DSS or Uncertainty DSS) as a within subjects variable, and experience-level (entry-level or intermediate level) as a between subjects variable. The results show that domain experience-level did not have a significant effect on perceived information sufficiency ($F_{(1,20)} = 0.045$, $p = 0.84$), but that the DSS approach did have a significant effect ($F_{(2,40)} = 3.307$, $p = 0.047$). In particular, information sufficiency was lowered (e.g. more ambiguity was perceived) when participants used the Uncertainty DSS.

The next step was to determine if perceived information sufficiency was always lowered when using the Uncertainty DSS, or was is associated with overlap situations (e.g. ambiguous decisions?). To analyze these data it was necessary to identify which decision making tasks resulted in insufficient information situations, and which did not. The former situation was referred as "overlap" cases (as in Figure 7a) and the latter one as "no-overlap" cases (as in Figure 7b). Since the participants created the decision tasks (e.g. the sets of alternative designs) the experimenters did not know a priori which design tasks would result in overlap or no-overlap situations. Furthermore, even the experimenters had created the tasks still they would not know a priori because the information brought to bear on a decision is a function of what the individual decision maker knows, not the task. One person may have enough information to make a decision because of what she knows or has researched, while another does not. The "overlap-or-not" information was easy to obtain for conditions in which participants used the Uncertainty DSS by taking a screenshot at the end of the trial since the uncertainty visualization showed which alternatives overlapped. However, when participants used no DSS or the Certainty DSS it required an extra step because participants did not explicitly describe the uncertainty associated with alternatives. Thus, at the end of the experiment they were asked to carry out a debriefing (step 6 in the "Procedure") in which they were asked to use the Uncertainty DSS to recreate the comparisons which they had just completed using no DSS or the Certainty DSS. After the experiment, all trials were separated into "overlap" and "no-overlap" cases after the experiment.

Table 3. Total number of cases with overlap and no-overlap situations.

Situation	**DSS Approach**			**Totals**
	No DSS	Certainty DSS	Uncertainty DSS	
Overlap	12	11	14	37
No-overlap	10	11	8	29
Totals	22	22	22	66
Percentage of Overlap Cases	55%	50%	64%	56%

Table 3 shows the results of sorting all tasks into overlap and no-overlap situations. There were a total of 66 trials (22 participants x 3 conditions). Overlap situations occurred in 37 cases, and no-overlap situations occurred in 29 cases. In order to test whether the DSS approach used initially to compare alternatives (e.g. no DSS, Certainty DSS or Uncertainty DSS) may have biased the participants' later judgments about whether alternatives overlapped, the author applied a test for independence and found no evidence of dependence, $\chi^2(2, N=66)=0.86, p= 0.65$. In other words, the initial problem solving method used to compare alternatives did not appear to have biased participants' later judgments concerning overlap between alternatives. Additionally, there were no significant differences for "overlap-or-not" incidence between the entry-level (overlap case: 61%) and intermediate-level designers (overlap case: 50%), $\chi^2(1, N=66)=0.82, p= 0.37$.

Table 2 was used to separate the information sufficiency data into overlap and no-overlap cases. Figure 9 shows results for entry-level and intermediate-level designers on perceived information sufficiency.

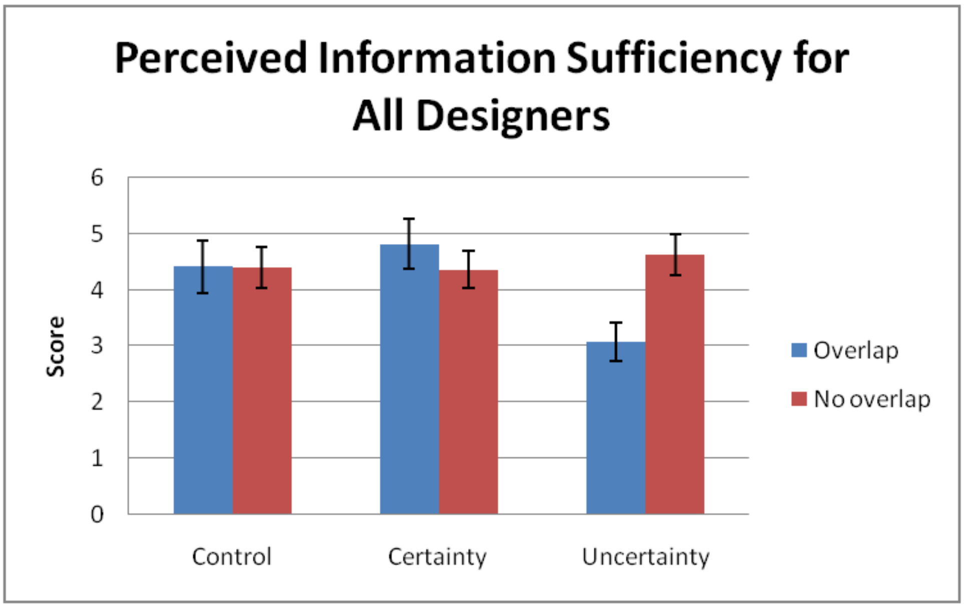


Figure 9.Participant's perceptions of information sufficiency when using each method. The error bars show standard error.

As the graph in Figure 10 suggests, participants perceived reduced information sufficiency only when they used the Uncertainty DSS and it displayed an overlap. It would not have been appropriate to use overlap as a variable in an ANOVA analysis together with other variables since it is not a controlled variable and it is nested with individual designers and the three systems. For example, Mary compared set A designs with no DSS support and found no overlapping for top designs, she compared set B, C designs on Certainty and Uncertainty DSS but with overlapping for top designs in both cases. So for Mary, the experience of using no DSS, Certainty DSS and Uncertainty DSS is no-overlap, overlap and overlap. But this specific pattern may not happen for another designer, say Dave. This is why overlap was not appropriate as a variable run repeated ANOVA together with other variables. However, as a compromise, several two-tailed t-tests compare perceived information sufficiency with and without overlap for each of the DSS approaches. There was a strongly significant difference between the overlap and no-overlap cases only for the Uncertainty DSS: $t_{two\text{-}tail}(20)=-3.07$, $p=0.007$. In contrast, participants using no DSS $t_{two\text{-}tail}(20)=0.03$, $p=0.98$, or the Certainty DSS $t_{two\text{-}tail}(20)=0.82$, $p=0.43$ perceived no differences in information sufficiency, regardless of whether an overlap had occurred or not. Thus, without the Uncertainty DSS's visualizations participants appeared to be unable to distinguish between ambiguous (overlap) and unambiguous (no overlap) decision situations. This effect remains when looking at entry-level and intermediate-level designers separately.

These findings are consistent with the research reported by Cole (1989) and Hoffman et al. (2006) who found that visualization of uncertainty information helps convey the sense of uncertainty. Also, Nadav-Greenberg (2009) found that decision makers were able to recognize the value of uncertainty information, and evaluate the information to produce a more realistic understanding of the situation.

6.7.4 Decision Confidence

For this analysis the data were also sorted into overlap and no-overlap cases, as shown in Figure 11. The main hypothesis was that participants would express less decision confidence when they used the Uncertainty DSS and its visualization showed an overlap between the top alternatives, indicating an ambiguous decision situation. The second hypothesis was that when participants had no visualizations to help them distinguish between ambiguous and unambiguous choices, their decision confidence would be unaffected by the presence or absence of overlap (e.g. ambiguity). A two way ANOVA using DSS approach (no DSS, Certainty DSS or Uncertainty DSS) as a within subjects variable, and experience-level (entry-level or intermediate level) as a between subjects variable showed a significant effect for domain experience-level, and for DSS approach. In particular, the entry-level designers were marginally more confident in their decisions than the more experienced, intermediate-level designers ($F_{(1,20)} = 3.315$, $p = 0.08$). Their greater confidence may have been a reflection of their greater naïveté. Additionally, DSS approach had a significant effect ($F_{(2,40)} = 4.158$, $p = 0.02$)on decision confidence. In particular, use of the Uncertainty DSS reduced decision confidence.

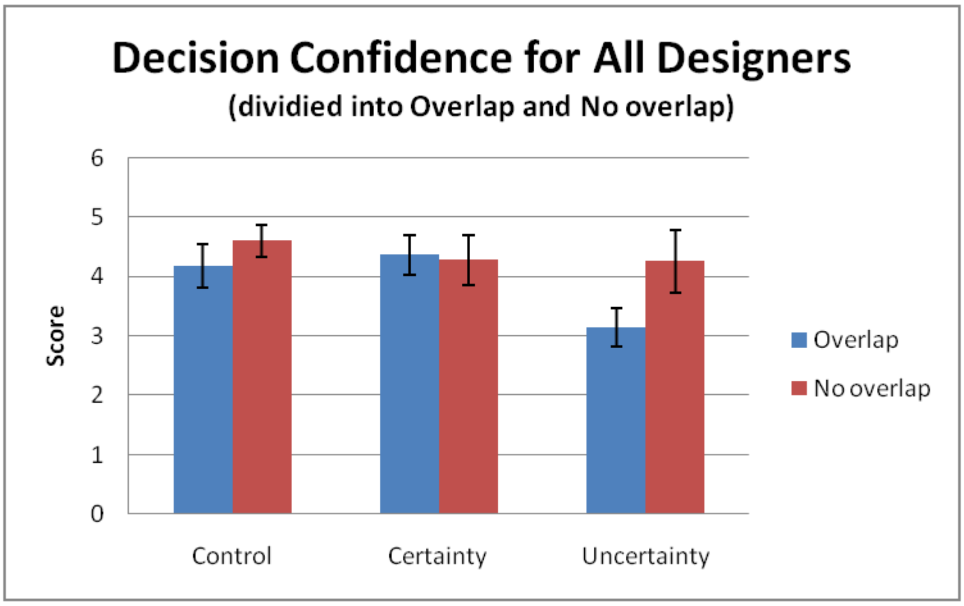


Figure 10. Designers expressed marginally reduced confidence when the Uncertainty

DSS showed a visualization of overlap between top ranked alternatives. Error bars show standard error.

The same as perceived information sufficiency, overlap and no-overlap decision cases were separated. Figure 11 shows the results for the designers after the separation. A two tailed t-test shows that when the Uncertainty DSS's visualization displays an overlap, decision confidence is significantly less than when no-overlap is displayed, $t_{two\text{-}tail}(20)=-1.78$, $p=0.09$. There are no significant differences between the overlap and no-overlap cases when participants used no DSS $t_{two\text{-}tail}(20)=-0.96$, $p=0.35$ or the Certainty DSS $t_{two\text{-}tail}(20)=-0.17$, $p=0.87$. Decision confidence had a similar pattern to perceived information sufficiency.

6.7.5 Plans to seek additional information

While it is important to recognize when one does not have enough information to make an informed decision, it is also important to take action to get more information, and to get the right information. When participants lacked sufficient information to make an informed decision, it is examined whether they:

- Planned to look for the "right" information (i.e. information that would reduce ambiguity between the alternatives), and
- Followed through on these plans a week later.

6.7.5.1 Participants' Plans Immediately After the Experiment

The author wanted to assess whether designers planned to look for the "right" information when they make plans to gather additional information. In this case, the "right" information pertains to the criteria that contribute most to the overall uncertainty (e.g. the width of the Total Score bar). These are the ones that have a high importance weight and large uncertainty in the values associated with them (e.g. many wide bars in the column under the criterion). For example, if manufacturing cost is very important, and the manufacturing cost of the options "three fingered gripper" and "suction device" are largely unknown, then it is a good strategy to investigate the manufacturing costs of those specific alternatives in order to get closer to identifying the most advantageous choice.

So, how good were participants at searching for the "right" information to support the decision? Table 4 below summarizes these results. The discussion following explains the analysis. The results indicated that participants were marginally better at identifying what information to gather when using the Uncertainty DSS than when using the Certainty DSS (z=1.58, p=0.057). There were no significant differences between no DSS and the Certainty DSS (z=0.32, p=0.376), or between no DSS and the Uncertainty DSS (z=1.26, p=0.104).

Table 4. How effective were participants at planning to search for information that would support the design decision?

	Average correlation between: • **expressed interest in a criterion, and** • **contribution to overall uncertainty**		
	Correlation(r)	**Total number of information requests (N)**	**Number of Trials**
No DSS	0.292	65	22
Certainty DSS	0.239	63	22
Uncertainty DSS	0.475	79	22

In Table 4, N is that total number times participants expressed interest in gathering information on a specific criterion in a given trial. Many participants wanted to explore multiple criteria. The results showed the correlation between the strength of participants' expressed interest in finding more information about a specific criterion, and the critionion's actual contribution to overall uncertainty. Expressed interest in finding more information on a criterion was measured on a 7 point Likert scale in a survey given to participants after each trial. A criterion's actual contribution to overall uncertainty was the sum of the widths of all bars in that criterion's column, times its weight. These values were obtained either 1) during trials with the Uncertainty DSS, or 2) for problems solved with no DSS or the Certainty DSS, participants were asked to recreated those problems in the Uncertainty DSS after all of their trials were completed. This was the "Debriefing" step, Step 6 of the procedure).

The correlations in Table 3 were further analyzed by first performing a Fisher z' transformation on the correlations, and then making pair wise comparisons. Results showed that participants were marginally better at identifying what information to gather when using the Uncertainty DSS than when using the Certainty DSS (z=1.58, p=0.057). There were no significant differences between no DSS and the Certainty DSS (z=0.32, p=0.376), or between no DSS and the Uncertainty DSS (z=1.26, p=0.104). Thus, it appeared that the Uncertainty DSS did help participants more than the Uncertainty DSS to identify what additional information to gather to support decision making.

The following describes specifically how the participants used the Uncertainty

DSS's visualizations to identify what additional information would be most helpful in the decision process. The participants looked at the interface to find a heavily weighted criterion with many wide bars in its column. For example, in Figure 2, the first column has the heaviest weight and it also has many wide bars. This criterion is an appropriate candidate for further investigation because it contributes heavily to the uncertainty in the Total Scores of many of the alternatives. By gathering more information on this criterion, one may be able to narrow its range possible values, and the range of the Total Score. If range of the Total Scores can be sufficiently narrowed for the top alternatives, then the overlap between them may be eliminated or significantly reduced, making it possible to identify the best alternative with little or no ambiguity. It may not always be possible to completely eliminate overlap between the top alternatives. However by considerably reducing it, decision makers are expected to place more (well founded) confidence in their choice. While participants were trained to use the interface in this way, not all of them remembered to do this, or understood it. The researchers feel that in future work the interface can be redesigned to make the information more obvious, while requiring less work on the part of the user.

Next the author looked at whether expertise had any impact on information seeking plans specifically for Uncertainty DSS.

Table 5. How does expertise impact the planning to search for information that would support the design decision?

	Correlation between: • **expressed interest in a criterion, and** • **contribution to overall uncertainty**		
	Correlation(r)	**Total number of information requests (N)**	**Number of Trials**
Entry-level	0.522	48	12
Intermediate-level	0.557	31	10

From Table 5, it is not significant (z=0.21, p=0.418) whether intermediate-level designers were better than the entry-level designers at identifying the information that would actually help in their decision. This is different from other studies of domain-experience; people with greater expertise tend to be better at recognizing the relevant features of the problem or environment (Chi, et al 1988; Anderson, 1993). Researchers (Rogers & Fisk, 2001; Rogers et al., 2006) found that novice users may not be as adept as experts at searching for information and selecting the information most relevant to task goals. This is probably due to the Uncertainty DSS effectively improved the entry-level designers' ability of identifying relevant information to support decision making.

6.7.5.2 Participants' Follow-through Actions One Week Later

One week after the experiment, a follow-up survey was sent to the participants to inquire what they did in the week following the experiment. The author found that one week later, the entry-level designers were less likely than intermediate-level designers to have followed through on their expressed plans to seek information; only 42% (5 out of 12) of the entry-level designers followed through on their plans, while 80% (8 out of 10) of the intermediate-level designers followed through. A two-sided hypothesis test on population proportions shows that the proportion of intermediate-level designers who followed through was marginally greater than that for entry-level designers ($\chi^2(1, N=22)=3.316$, $p=0.069$). Much of entry-level designers' information seeking effort was spent seeking non-critical information. Further efforts are needed to identify strategies to help entry-level designers follow through on appropriate information seeking plans.

6.8 Summary and Discussion

6.8.1 Summary

Following are the findings from the evaluation respect to the four research questions:

- Can uncertainty visualizations help designers understand when a lack of information interferes with the ability to make a choice? *Yes. Furthermore, without the uncertainty visualizations, participants did not distinguish between situations in which they had sufficient information, and those in which they did not.*
- When uncertainty visualizations show ambiguous situations, is decision confidence reduced? *Marginally. However, good decisions can often be made even when choices are ambiguous.*
- Can uncertainty visualizations encourage designers to seek clarifying information when appropriate? *Yes. When using the Uncertainty DSS, participants were slightly more accurate in identifying which information*

would eliminate ambiguity (e.g. overlap) between the alternatives. However, there is room for improvement in the interface.

- Does domain experience change the benefits that decision makers derive from the DSS?

No. However, there were differences between the entry-level and intermediate designers. The most important differences were that the more experienced, intermediate-level designers were less confident in their decisions (perhaps because they were more cautious); better at identifying information that would clarify the decision; and better at following through on obtaining that information during the following week. Finally, all designers preferred to use Uncertainty DSS but entry-level and intermediate-level designers were different in perceived effort using different DSS.

6.8.2 Discussion of Experiment one

Uncertainty is present in all situations where decisions are made. Decision makers need tools and approaches that can help them manage uncertainty and its impact on their decision making tasks. This first study demonstrated that the Uncertainty DSS's visualization significantly improved entry and intermediate-level decision makers' ability to recognize situations in which they lacked sufficient information to make an informed decision. It also increased the likelihood that they would seek information that would help clarify the decision. However, while the intermediate-level designers were better than the entry-level designers as identifying what type of information was needed, neither group was very good at this, or carrying out plans to get that information.

There are many reasons that a designers might want to look for information, one is to learn enough about the top options so to enable one to confidently choose the best of the set. That best option can then be developed further while the others are put aside. However, one might gather more information simply to learn more about one or more of the alternatives whether it helps with the decision or not. However, the latter is an ineffective strategy which may lead to much wasted time. The longer one carries forward multiple alternatives working on all of them, the more time may be wasted on alternatives that are eventually dropped. It is very tempting for designers to "waffle" for a long time, carrying forward multiple alternatives in parallel. However, in order to make progress focus is necessary. For this reason, many design projects set deadlines by which the design team must commit to a specific alternative. As that deadline approaches, time spent on anything other than understanding the specific facets of the alternatives that allow a decision to be made is potentially wasted. If the designer reaches the decision deadline without sufficient information to clearly identify one option as better than the rest, the choice may be suboptimal. Designers must focus their information gathering efforts to support the decision if they are to be effective. This is especially true when designers work in a collaborative environment where group decision making is needed. The time wasted by a group will be easily amplified by the size of the team. Also in a group environment, decision making involves extensive communication, how will Uncertainty impact the communication within a group is an interesting topic. All these facts lead me to design another study with group using of the Uncertainty DSS.

Chapter 7--Changes to DSS Interface and Measurement for Experiment Two

As mentioned in chapter 2, in many professional settings decision making carried out by a group of people is also common. In team situation, shared cognition or shared situation awareness measures the common knowledge base that is built by the team to support joint decision making (Orasanu, 1994; Rouse et al., 1992; Salas, Cooke, & Rosen, 2008; Salas, Sims, &Burke, 2005). Common domain knowledge were found to foster effective team processes (i.e., extent of planning, collaboration, and communication), which in turn increased task performance (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Smith-Jentsch, Cannon-Bowers, Tannenbaum & Salas, 2008). Thus in order to study the team use of Uncertainty DSS, shared situation awareness would be a good factor to measure.

Situation awareness (SA) research is widespread and ongoing in military operations (Endsley et al., 2000b; Matthews et al., 2000), aviation (Kaber et al., 2002; Keller et al., 2004), air traffic control (ATC) (Hauss and Eyferth, 2003; Endsley and Smolensky, 1998), automobile driving (Zheng et al., 2004), and C4i environments (Stanton et al., 2005). Endsley (1995a) defines SA as “The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.” For successful task performance, team members must also have a shared understanding of current task conditions and shared expectations of future events. As teams operate in dynamically changing environments, plans that were appropriate at some point in time may become inadequate. Thus, accurate situation assessment is critical. Team members must monitor their task environment for cues and information that are inconsistent with their original assumptions and may call for a change of action.

The three-level model of SA is by far the most commonly cited and used model of SA. It describes SA through an information processing approach, and proposes that SA is an internally held product comprising three hierarchical levels that is separate to those processes (termed situation assessment) used to achieve it. According to Endsley (1995a), SA consists of level 1 SA (the perception of the elements in the environment), level 2 SA (the comprehension of their meaning) and level 3 SA (the projection of their future status). The model depicts SA as an essential component of human decision-making activity. The achievement and maintenance of SA is influenced by both individual and task factors, such as experience, training, workload and also interface design.

The SAGAT questionnaire is the most popular technique to measure SA, developed to assess pilot SA within the three-level model (Endsley, 1995a). It is the most widely used and validated SA measures available, and variations of SAGAT questionnaire have consistently shown reliability and validity in a number of domains other than military. According to Jones and Kaber (2004) numerous studies have been performed to assess the validity of the SAGAT and the evidence suggests that the method is a valid metric of SA. Endsley (2000) reports that the SAGAT technique has been shown to have a high degree of validity and reliability for measuring SA; a study found SAGAT to have high reliability (test-retest scores of .98, .99, .99 and .92) of mean scores for four fighter pilots participating in two sets of simulation trials. Collier and Folleso (1995) also reported good reliability for SAGAT when measuring nuclear power plant operator SA. In a driving task study Gugerty (1997) reported good reliability for the percentage of cars recalled, recall error and composite recall error. Regarding validity, Endsley et al. (2000a) reported a good level of sensitivity for SAGAT. SAGAT also showed a degree of predictive validity when measuring pilot SA, with SAGAT scores indicative of pilot performance in a combat simulation (Endsley, 1990). The study found that pilots who were able to report on enemy aircraft via SAGAT were three times more likely to later kill that target in the simulation.

7.1 Measurement Changes

In this research, SAGAT approach was adopted to measure SA and developed a questionnaire specifically for engineering designers. It includes 15 questions measuring three levels of SA. Questions 1-5 are around the perception of the facts in the design decision (Perception); questions 6-10 are about the understanding of the design decision (Comprehension); questions 11-15 are about making predictions based on the design decision (Projection). The questions are listed here.

1. *How many designs do you have?*
2. *How many criteria do you have?*
3. *Which criteria is the most important? Write more than one if there is tie.*
4. *Is your first design better than the third one in total score?*
5. *Which design is the best in total score?*

6. What are the top two features needed by your users?

7. What are the pros and cons of each design?

8. Which design is the hardest to manufacture?

9. Which is the most innovative design?

10. Are there several designs where it is not clear which one is the best? (If there are, can you list them?)

11. If you are going to produce your product/system, how are you going to test your designs?

12. What types of information do you need the most right now to continue your design work?

13. Which design worries you the most? How can you improve on it?

14. For which design do you have the most open questions? Which design alternative do you feel like working on the most in the next week?

15. Which design do you feel is least likely to require major changes in the future?

Each individual was asked to respond to these questions after the study, then the whole team work together on these questions to give a group response. The difference between individual and team response is counted as a measurement of shared situation awareness. This approach is widely used by previous researchers to study team situation awareness (Artman 1999, 2000, Fowlkes et al. 2000, Dekker 2000, Rasker et al. 2000, Cooke et al. 2001).

In Experiment One, results showed that control condition (No DSS) is generally dominated by the other conditions. So in Experiment Two, author decided to only keep Certainty and Uncertainty conditions. Furthermore, the author condense the measurements inherited from experiment one. Perceived information sufficiency and decision confidence have very similar results in experiment one, thus only perceived information sufficiency is kept in experiment two. Due to the improvement of Uncertainty DSS, which will be mentioned below, "where to seek additional information" is now more prominent in the new interface than the one in experiment one. Thus the measurement of plan to seek additional information was dropped as well.

7.2 Improvement of Interface

One outcome of experiment one is that when using Uncertainty DSS, there was a correlation between where engineering designers think they need more information and the decision criteria that contribute the most to the uncertainty. However this correlation is not high enough to be satisfied (correlations are in the 0.4-0.5 range). So a more prominent way to show which criteria has the most uncertainty was needed. Thus an improvement was made to the Uncertainty DSS interface to explicitly show the weighted uncertainty of each criterion. Figure 11 shows the new interface. The bar chart at the bottom of the graph shows relatively how much uncertainty each criteria contribute to the overall uncertainty. The uncertainty for each criterion was computed by adding the uncertainty for all the designs and scale it with the weight of the criteria. This added illustration was intended to make it easier for decision makers to browse which criteria contribute more to the overall uncertainty.

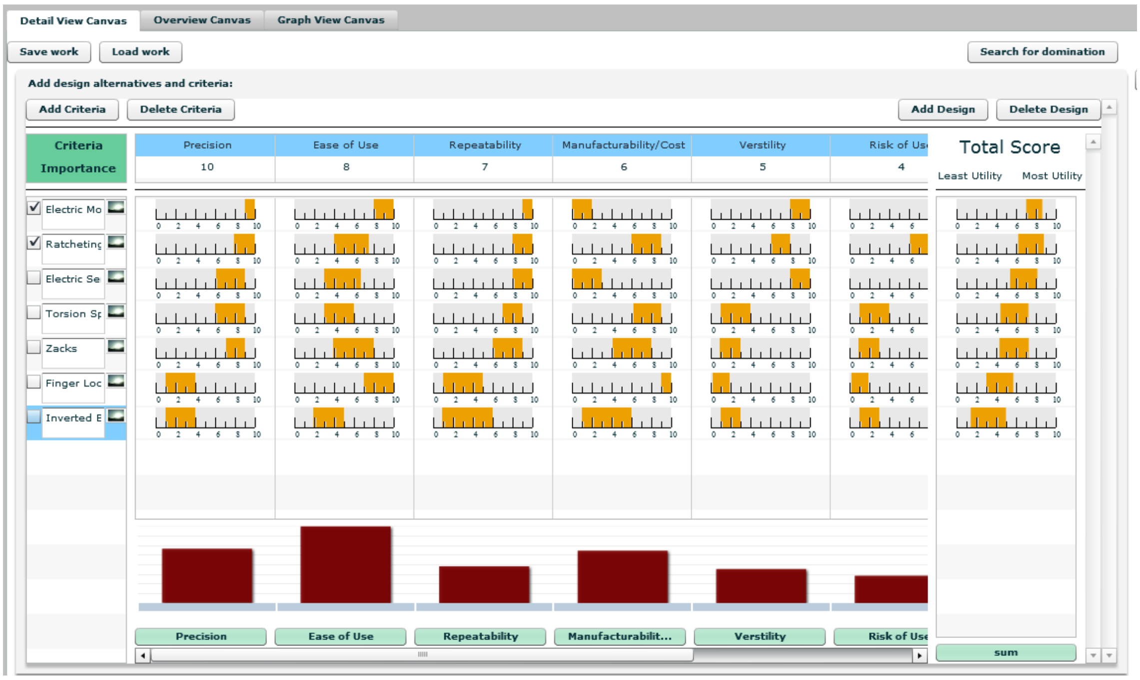


Figure 11. Main interface of Uncertainty DSS used in the second study

Chapter 8--Experiment Two

The author conducted a group study of the design decision making in which 9 teams composed of graduate and undergraduate engineers (33 total participants)were recruited to use two decision making support methods: the Certainty DSS and the Uncertainty DSS. As in experiment 1, the sets of design alternatives were drawn from participants' on-going design projects. Unlike experiment 1, whole team rather than individuals were recruited to participate in the experiment.

8.1 Research questions

In this research, the following questions were investigated:

- Can uncertainty visualizations help a team of designers understand when a lack of information interferes with the ability to make a choice?
- Can uncertainty visualizations improve team members' situation awareness of design alternatives and design situation?
- Can uncertainty visualizations in team settings end up with higher effort demand but is preferred more, just as witnessed in experiment one?

As in the first experiment, it was hypothesized that Uncertainty DSS would help groups of users to recognize situations in which information was insufficient for making an unambiguous choice. It may require more effort to use, but because of the ability to express Uncertainty, Uncertainty DSS was expected to have more preference. It is also hypothesized that groups of designers who used the Uncertainty DSS would have communication more within the team, and exhibit greater situation awareness of design alternatives themselves and future plans than the team using Certainty DSS.

8.2 Experimental design

Unlike experiment one, which used a within subjects design, experiment two used between subjects experiment design. This is mostly due to the longer time it took for a group of designers to compare design alternatives. Experiment two used DSS-type (Certainty or Uncertainty) as the independent variables. The dependent variables were:

- *Perceived information sufficiency*
- *Effort of using DSS*
- *Preference between methods.*
- *Usability*
- *Situation awareness on design alternatives*
- *Group communication.*

Perceived information sufficiency was measured by asking participants the extent to which they agreed with the following four statements "I feel I had sufficient information to make an informed decision." "For some designs, I feel like I still need to learn more about them." "If I am forced to choose a design right now, it is likely that my decision will be suboptimal." "There are things that I do not know about each design alternative." They expressed their agreement or disagreement with these statements on 5 point Likert scales. (Note: in experiment one, 7 point Likert scales were used. However participants expressed difficulty distinguishing subtle differences between the choices on 7 point Likert scales, so in experiment two, the scale was reduced to 5 points.)

The effort of using each DSS was measured using the NASA-TLX inventory. NASA-TLX allows users to perform subjective workload assessments when working with human-machine systems. It is a multi-dimensional rating procedure that assesses overall workload based on a weighted average of six parameters: Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort and Frustration. It has been used to assess cognitive workload in tasks such as aircraft, command, control, and communication (C3); supervisory and process control in manufacturing environments, and many other tasks. (Appendix B).

DSS preference was measured by asking participants whether they agreed with the following three statements "I like the functions provided by the system," "I do not think the system fit well with the design work I am doing," "I think that I would like to use this system frequently." Their agreement was also measured on a 5 point Likert scale.

Usability was measured using the System Usability Scale (SUS) developed by Brooke (1986). System Usability Scale (SUS) is a ten-question scale giving a global view of subjective assessments of usability. It yields a single score on a scale of 0–100 as a composite measure of the overall usability of the system being studied. This metric is described in detail in Appendix B. The SUS has generally been perceived as providing a high-level, subjective view of usability and is often used in carrying out comparisons of usability between systems. The SUS has been widely used in the evaluation of many systems; Bangor, Kortum and Miller (2008) have studied the usage of SUS scale extensively over a ten year period and produced normative data that allow SUS ratings to be compared to other similar systems (e.g. graphical user interface system).

Situation awareness was measured using the Situation Awareness Global Assessment Techniques (SAGAT) (Endsley, 1988, 1995). Customized questions probing the designer's knowledge about design alternatives were developed by the author and categorized into three levels of SA. There were a total of 15 questions used to investigate three levels of SA (SA level 1, Perception: questions 1-5; SA level 2, Comprehension: questions 6-10; SA level 3, Projection: questions 11-15). See Appendix B for more details. The author felt it was important not to interrupt team discussion of decision making. So instead of using a "freeze-probe" (stop participant randomly and ask questions) during design sessions, SA questionnaires were administered after each session with each participant filling in one copy. The "correct" answers to the SA questions were obtained, after the questionnaires were completed, by having a team discussion to gain agreement on one set of answers.

Last but not least, team communication was measured by counting the total amount of utterance which was found by transcribing session video recordings.

8.3 Participants

Nine teams participated in this experiment. Each team had 3 to 4 members. There were a total of 33participants (26 males, 7 females). All participants were students at University of Minnesota enrolled in a graduate level Human Factors class in which both graduate students and senior-level undergraduate were enrolled (14 undergraduates and 19 graduates). The average age was 24.94 years (SD = 4.46), and participants reported an average of 1.77 years (SD = 1.51) of design experience.

Teams were randomly assigned to one of two experiment groups. One experimental group used the Certainty DSS only, while the other used Uncertainty DSS only. Five teams used the Uncertainty DSS (in total 19 participants); and four used the Certainty DSS (in total 14 participants). The statistics describing the two groups can be found in Table 6.

Table 6.Design Experience and Age of Participants

Groups using Uncertainty DSS		
	Design Experience (years)	Age (years)
Average	1.95	25.95
Standard deviation	1.23	5.53
Median	2	23
Minimum	0	22
Maximum	4	39

Groups Using Certainty DSS		
	Design Project Experience (years)	Age (years)
Average	1.54	23.57
Standard deviation	1.83	1.74
Median	1	23
Minimum	0	21
Maximum	7	26

ANOVA tests were used to examine the potential demographic differences between the two groups. There was no statistical difference on age ($F_{(1,32)} = 2.389$, $p = 0.132$)or reported designer experiences ($F_{(1,32)} = 0.595$, $p = 0.446$)between the two experimental groups. The design experience showed that, the participants in experiment two are more similar to entry-level designers in experiment one (averagely 1.58 years of design experience).

8.4 Design Tasks

As in experiment one, instead of giving all participants the same design tasks, each team was asked to compare several options currently under consideration in their own on-going design projects. This was done so that decision behavior in the context of real design tasks could be evaluated. Each team had been working on their design projects for 6 or 7 weeks (the experiment went through the later 6th week till the early 7th week). That gave them sufficient time to develop a reasonably deep knowledge of their specific design topic, and become highly vested in the decision outcome. None of these real world conditions would be possible to recreate using an artificial "laboratory" design task presented to participants for the first time in the experiment.

The topics of the specific projects ranged from redesign of a cable production process to streamlininga small-scale manufacturing line.

8.5 Procedure

The teams were contacted by the author at the beginning of their project timeline about whether they would like to participate in the experiment. The nine teams who agreed to participate were scheduled at weeks 6 and 7 to use one of the two DSS tools to compare design alternatives which the team had developed.

- At the beginning of the experiment, each team was given a brief introduction to the study. Participants signed a consent form which included agreement to video recording, and answered questions on demographics and past design experience.
- Teams were given a brief training session on the DSS they would use (Uncertainty DSS or Certainty DSS). The training sessions took roughly 10 minutes.

- Teams were asked to use the DSS assigned to their team to identify one "best" alternative from a set which they judged to be most appropriate for further design development. The alternatives were created by the teams themselves over the past few weeks as part of their design project. Teams were given no time limit for completing design tasks.
- Immediately after using the DSS, each individual participant was asked to complete the questionnaires measuring all dependent variables, except group communication.
- After completing the questionnaires individually, teams were asked to develop a group consensus on answers to the situation awareness questions.
- At the end of the session, the four teams that used the Certainty DSS were asked to reproduce the decision task they completed on Certainty DSS using Uncertainty DSS. The purpose was to identify the degree of uncertainty (e.g. overlap) that existed between the alternatives. The five teams that were assigned to use the Uncertainty DSS were not asked to do this final task.

8.6 Results

In this study, all nine decision making cases had overlaps between the top alternatives. Thus, it was not possible to compare overlaps against no overlap situations in Experiment Two. Thus the analysis presented here focuses only on ambiguous (e.g. overlap) situations in which decision maker lacked sufficient information to identify the best design alternative.

8.6.1 Perceived Information Sufficiency

As in Experiment One, when the "Total Scores" for top design options overlap, it is a signal that there is not enough information to definitively identify the "best" alternative. Team's perceived information sufficiency when using the different DSSs is shown in Figure 12. A one-way ANOVA test showed that teams using the Uncertainty DSS perceived less information sufficiency than those using the Certainty DSS ($F_{(1,32)} = 58.602$, $p < .001$).This indicates that the Uncertainty DSS did increase team's ability to correctly identify when they lacked sufficient information to make an informed decision. Information was equally lacking for the teams using the Certainty DSS, but they were largely unaware of it. This result for teams is consistent with the results for individuals in Experiment One. Based on the observation during Experiment Two, the uncertainty DSS for teams can be even more effective than for individuals since teams which used Uncertainty DSS spent more time on the discussion of uncertainty and how it could be managed. The group discussion appeared to raise individual team member's awareness of the uncertainty. The discussion potentially led to better situation awareness.

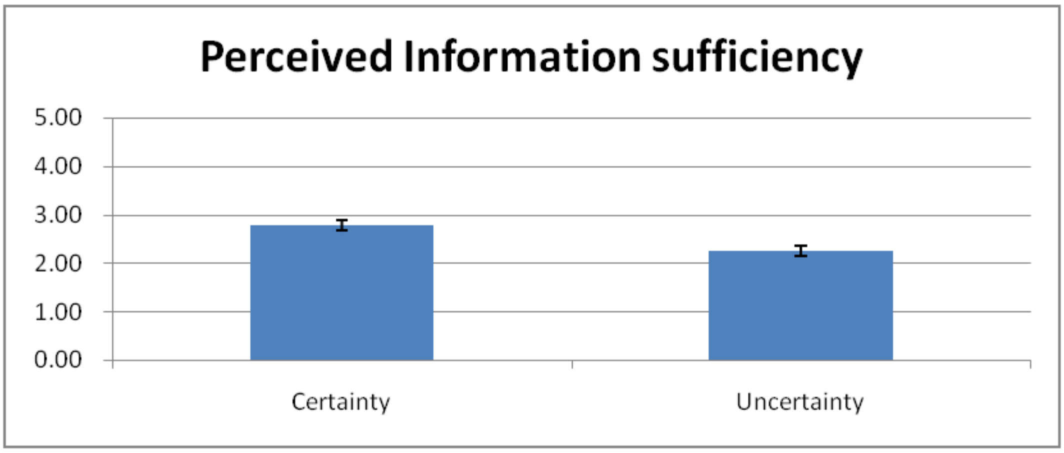


Figure 12. Perceived information sufficiency using Certainty and Uncertainty DSS

(All decision situation in experiment two are the “overlapping” cases). Error bars show standard error.

8.6.2 Workload and DSS Preference

An ANOVA was used to evaluate the effect of DSS-type on workload and DSS preference. Results are summarized in Figure 14 and Table 5. Teams who used the Uncertainty DSS experienced significantly higher workloads than did teams using the Certainty DSS($F_{(1,32)}$ = 10.312, p = 0.003; Cohen *d*> .8).However, teams showed a significant preference for the Uncertainty DSS over the Certainty DSS ($F_{(1,32)}$ = 4.829, p = 0.036; Cohen *d* = .7).

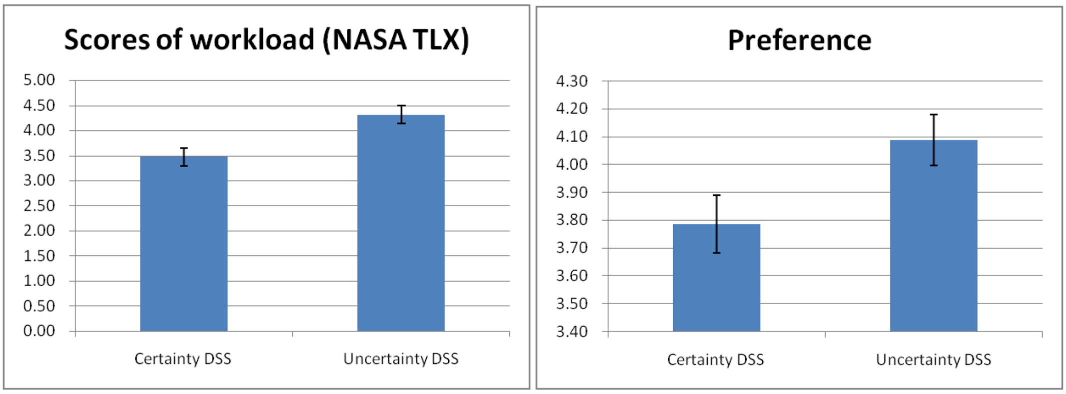


Figure 13.Workload and Preference using Certainty and Uncertainty DSS (All decision situation in experiment two are the “overlapping” cases). Error bars show standard error.

Table 7:Scores of workload (NASA TLX) and preference

	Workload		Preference	
	Mean	SD	Mean	SD
Certainty DSS	3.47	0.67	3.79	0.38
Uncertainty DSS	4.31	0.79	4.09	0.40

The overall workload scores for both DSSs are at the lower range of

NASA-TLX scale's possible score range (0-10). Thus both DSSs are acceptably easy to use.

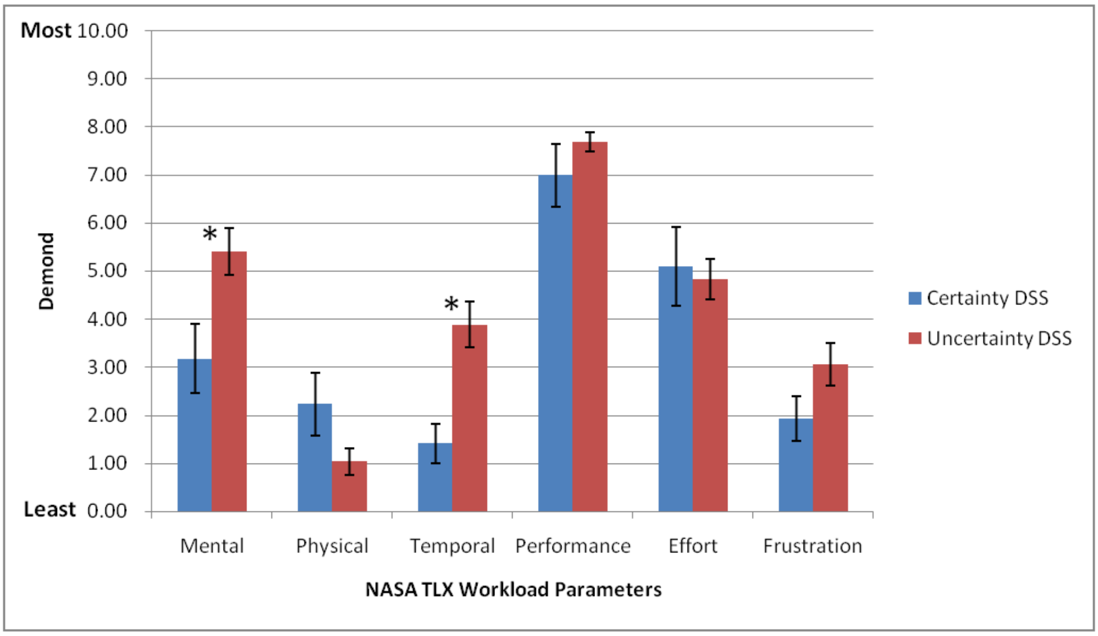


Figure 14.The detailed comparison of NASA TLX scale on different dimensions. Error bars show standard error. *Significantly different in the 0.05 level

Figure 14 shows the component parameters used to produce the aggregate TLX workload score. Several One-way ANOVAs were used to test each dimensions of NASA TLX workload parameters. There were significant differences for the two DSSs in "Mental demand"($F_{(1,32)} = 7.251$, $p = 0.011$) and "Temporal demand"($F_{(1,32)} = 14.430$, $p = 0.001$). In other words, the Uncertainty DSS took more time and mental effort to use than the Certainty DSS, probably because of the additional dimension of uncertainty. The complete summary of the p-values of these parameter are shown in Table 8.

Table 8.One Way ANOVA results of NASA TLX workload parameters

	Mental Demand	Physical Demand	Temporal Demand	Performance Demand	Effort	Frustration
F value	7.251	3.485	14.430	1.264	0.103	2.899
P value	0.011	0.071	0.001	0.270	0.751	0.099

This result makes sense, since using uncertainty DSS requires decision makers to express explicitly the uncertainty which exists in a vague format in their brain. Also the need to develop a group consensus of how much uncertainty should be assigned to each design inspired rich discussion. This process of synthesizing uncertainty judgments may explain the higher mental workload and higher temporal demand associated with the Uncertainty DSS. In contrast, when participants used the Certainty DSS, they did not have the extra job of determining uncertainty. The result of "Physical demand" is counter intuitive, because the experimenter's initial hypothesis was that participants would rate the Uncertainty DSS as requiring more physical effort than the Certainty DSS because the Uncertainty DSS required click-and-drag motion to specify a range of values, versus just moving a slider for the Certainty DSS. However, data showed that participants perceived the opposite trend ($F_{(1,32)} = 3.485$, $p = 0.071$). It is possible that click-and-drag is a natural gesture that users have already been used to based on their software/internet using experience. It is such a natural gesture that users might think it is even less intrusive than the traditional pointer.

It is interesting that users preferred the Uncertainty DSS given that it required greater mental workload. As with individual uses (Experiment One), the experimenter heard comments from team such as, "uncertainty in a DSS is a great compliment of traditional DSS." and "I would like to introduce it to my company once it is commercialized." Designers were willing to adapt and pleased by the idea of integrating an uncertainty dimension into decision making process. Thus, even though the Uncertainty DSS required more effort to use, participants' strong preference for the Uncertainty DSS and comments suggest that decision makers felt the extra work was worth it.

In Experiment One (individuals) and Experiment Two (teams), different methods were used to measure workload (one question answered as Likert scale vs. NASA-TLX inventory of 6 questions) and preference (split 100 points vs. three questions answered on a Likert scale) but the results are similar. In Experiment One, both entry-level and intermediate-level designers had a significant preference for the Uncertainty DSS. Entry-level designers find a significant workload difference between the DSSs (Uncertainty DSS required more workload). The similarity between experiment Two and experiment One was high especially between results of Experiment Two and the results of entry level designers in Experiment One, which is probably because the design experience of participants in Experiment Two (1.77 years) was similar to entry-level designers in Experiment One (1.58 years).

However, the difference between the setup of Experiment One and Experiment Two is significant. There were difference between group and individual decision making, e.g. many verbal exchanges occurred in the team decision making environment, but not in the individual decision making environment. Thus, in the team environment (Experiment Two)the uncertainty visualization appeared to spark much discussion about the uncertainty and the appropriate estimation of it.

Discussion about uncertainty did not happen to any notable extent when teams used the Certainty DSS. So it is possible that the higher workload resulted, when teams used the Uncertainty DSS was at least partly due to the added effort required for team discussions and negotiations of uncertainty values. Unfortunately, due to change of workload measurement, the workload scores of the two experiments were not comparable. However, participants' comments support this explanation: "Using this system forced us to think and talk about uncertainty on each design", "It made us think things through way more thoroughly than what we would have [otherwise]."

8.6.3 Usability

SUS yields a single number representing a composite measure of the overall usability of the system being studied. SUS scores range from 0 to 100.The author followed the procedure described by John Brooke (1996) to calculate the usability scores in the study. The usability scores for Uncertainty and Certainty DSSs are summarized in Table 9.Aone-way ANOVA found no significant differences in usability between the two DSSs. This is good, since the added dimension of uncertainty did not reduce usability ($F_{(1,32)} = 0.545$, $p = 0.466$).

Table 9.Scores of usability (SUS scores)

Usability Scores	Mean	Standard Deviation	Median	Minimum	Maximum
Certainty DSS	82.29	7.87	83.33	62.5	91.67
Uncertainty DSS	86.51	8.98	87.5	70.83	97.92

Bangor, Kortum and Miller (2008) have provided benchmarks correlating usability score with acceptability based on more than 200 studies (composed of more than 2300 individual surveys) that can be used by researchers to see how well a system scores in usability against similar systems that have been studied before. To visually present this benchmark, Bangor, Kortum and Miller put useful scales around the SUS scores as shown in Figure 15. After summarizing the 2300 surveys, they used three different ways (quartile range, acceptability ranges and adjective ratings) to interpret the single SUS score researcher get from usability questionnaire. This helped in interpreting the SUS scores for the two DSSs studied in this thesis work. In Figure 15, the two scores (Uncertainty DSS: 86.51; Certainty DSS: 82.29) are within the "4th quartile", "Acceptable" range, and the Uncertainty DSS scored "Excellent" among the 2300 surveys evaluating different systems.

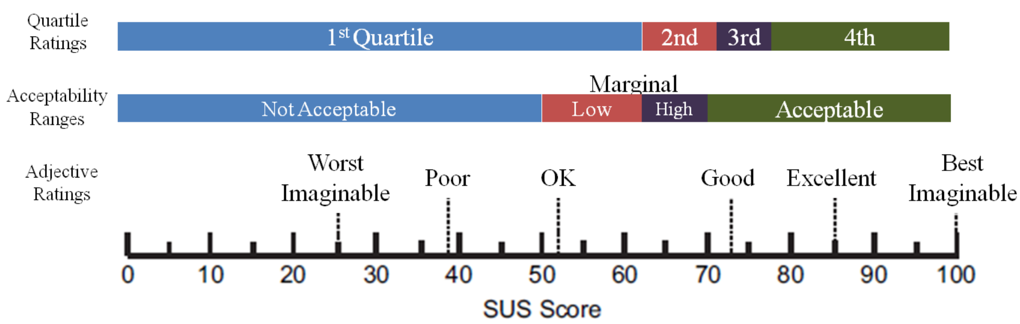


Figure 15.Based on Figure 13 in Bangor, Kortum and Miller (2008)

To separate the surveys in more granular type of UI, Bangor, Kortum and Miller (2008) categorized the 2300 surveys by the type of the user interface. Figure 16 summarizes their analysis.

Interface type	Count	M	SD
Cell	189	66.55	19.84
CPE	219	71.60	21.60
GUI	208	75.24	20.77
IVR	401	73.84	22.15
Web	1180	68.05	21.56
Web/IVR	50	59.45	19.19

Note: An interface type was required to have at least 50 surveys to be included in the analysis. Cell = cell phone equipment; CPE = customer premise equipment (e.g., phones, modems, etc.); GUI = graphical user interface for OS-based computer interfaces; IVR = interactive voice response phone systems, including speech based; Web = Internet-based Web pages and applications.

Figure 16. Based on Table 6 in Bangor, Kortum and Miller (2008)

The DSSs that were created by the author are best categorized as "GUI" type of interface. Regardless of the categorization, it is notable that both DSS scores are better than the benchmark of industry means.

The high usability scores are probably due to the adoption of user centered design from the very beginning of the project. In every major step of the design and development process, the author validated the proposed change in a small scale user testing and/or thorough expert reviews. Low fidelity paper prototypes and mock-up interactive systems were used in these formative evaluations before making any changes to the system's functions or the look-and-feel of the interface. By following a user centered design process, the author made sure to eliminate many sources of confusion and extra effort for the users.

7.7.4 Situation Awareness

Each participant's answers for the situation awareness questionnaire were compared against the team's consensus on the correct answers. Based on how close each participant's answer was to the team consensus (in total 15 questions), the answer to each question from each participant was given a score ranging from "0" to "1" for each subject. "1" means the team consensus and individual answer are the same whereas"0" means totally different. The shared situation awareness when using different DSSs were compared using Overall SA scores as well as SA scores for the three levels. Figure 17 and Table 8show the results:

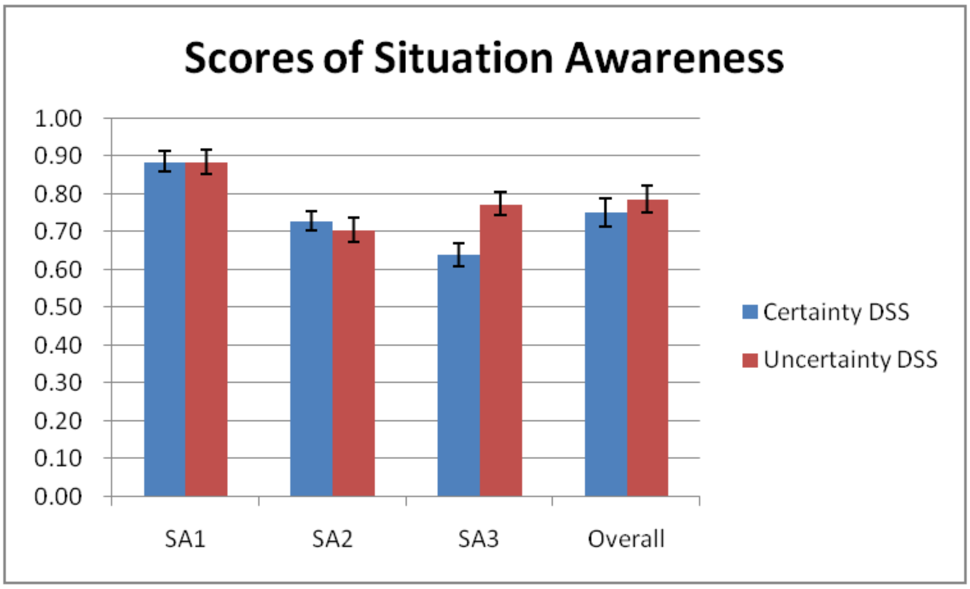


Figure 17. Scores of situation awareness. The higher score, the better awareness. Error bars show standard error.

Table 10. Scores of situation awareness

	SA1 Perception[1]		SA2 Comprehension[2]		SA3 Projection[3]		Overall[4]	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Certainty DSS	0.89	0.10	0.73	0.10	0.64	0.11	0.75	0.14

[1]Be aware of the facts of the design decision (How many design options, criteria, etc.). SA questionnaire (Q: 1-5)
[2]Understand the design decision (Pros and cons of designs, etc.). SA questionnaire (Q:6-10)
[3]Predict what to do after the design decision (How to improve on the design). SA questionnaire (Q: 11-15)
[4]Average values of SA1, SA2, and SA3

Uncertainty DSS	0.88	0.14	0.71	0.14	0.77	0.13	0.79	0.15

In this part of the analysis, a two way ANOVA was conducted on the data with levels of SA as within subject measures and the type of DSS (Certainty vs. Uncertainty) as a between group variable. Results showed that the type of DSS did not impact overall SA ($F_{(1,31)}$=1.582, p=0.218), but the average scores for level of SA differed significantly ($F_{(2,62)}$=25.357, p<0.001). The pair-wise comparisons between the three different SA levels were SA1-SA2: p<0.001, SA2-SA3: p=1.000, SA1-SA3: p<0.001. Furthermore, there was significant interaction effect between type of DSS and level of SA ($F_{(2,62)}$=4.614, p=0.014). From table8it can be seen that when using Certainty DSS the score of mean SA decreased as SA level increased but when using Uncertainty DSS there was not a clear pattern.

To better understand the interaction effect, a few one-way ANOVA tests were used, which included the test of overall SA. Overall, there was no difference ($F_{(1,31)}$=1.582, p=0.218) in situation awareness when using the different DSSs. There was also no significant difference on level 1($F_{(1,31)}$=0.001, p=0.973)or level 2($F_{(1,31)}$=0.287, p=0.596)situation awareness. The only significant result is on level 3 situation awareness: projection. Teams that used the Uncertainty DSS had an improved ($F_{(1,31)}$=9.650, p=0.004) ability to foresee what they need to do based on the knowledge they had.

This result is very interesting, and resonates with what was found in Experiment One (individual study). In Experiment One, the Uncertainty DSS helped decision makers identify which design parameters, if researched further, were most likely to reduce uncertainty. It is reasonable to argue then that using the Uncertainty DSS in a team environment would foster conversations on these uncertainty related topics, and thus augment each individual's situation awareness. Because of this prediction, situation awareness of using Uncertainty DSS was expected to be better than Certainty DSS. However this prediction is only proved for level 3 situation awareness a.k.a. the ability to project from what is known to the future. The author looked at the group discussion transcripts, and found that groups using Uncertainty DSS spent a good portion of their discussion on uncertainty including estimation, comparison and how to reduce the uncertainty[5]. For groups who used the Certainty DSS, the discussion around uncertainty was largely missing. It is possible this back and forth team discussion of uncertainty and how to manage it fostered the shared cognition between group members. But it is not clear there was only difference on level 3 situation awareness based on team communication. One possible explanation is that some questions for level 1 and level 2 situation awareness were not complex enough to identify differences between the two DSSs. For example the questionnaire asked“ how many designs do you have?” and “how many criteria do you have?” as two questions for first level situation awareness: perception. Almost all of the participants answered these two questions 100% correctly regardless of which DSS

[5]Sample quotes for uncertainty discussions:
Group A
Participants 1: if we are pretty sure the tool is going to cost a lot of money to make, these would have low uncertainty. but we don't know really how much the tool is going to cost. So potential there is going to be a range…
Participants 2: no I do know the cost. (P2 assigned a very small uncertainty in the system)
Group B
Participants 1: I think there are definitely some risks there with the Design One. There is higher risk that the user may misuse it or may miss the fill in spot (Technical term) or whatever.
Participants 2: I think the higher end (Uncertainty) is good; maybe you want it to be wider.
Participants 3: Let's adjust it to be 2-8 (the range of the uncertainty bar)

was used. This kind of "perception" questions weren't affected by what type of system was used. So it is likely that Uncertainty DSS benefit engineering designers when they deal with hard, projection problems better. Another possibility would be, unlike the Army or Air Force SA questionnaires, this study is the first one to measure situation awareness in design decision making using SAGAT-like questionnaires. It is likely the power of the questionnaire needs to be improved. With a better crafted set of questions, it maybe possible that Uncertainty DSS can have evidence impacting level 1 and 2 of situation awareness.

8.6.4 Team communication

There are many ways to measure team communication. One can look at the frequency of information exchange, or the number of comments with "good quality". A qualitative way is to code the whole team conversation and then try to generalize the theme over the communication. But given the time constraints, the author chose to include only the total number of utterances per participant per meeting and leave the other possibilities for the future studies.

Team communication was measured by the average number of utterance made by each participants on a team. Each participant's number of utterance was identified and taken average within teams who used Certainty DSS and teams who used Uncertainty DSS. Unfortunately two videos were not able to be recorded due to recording machine failure. The author was only able to go back and transcribe the meetings of the other seven teams (4 teams used the Uncertainty DSS, 3 teams used the Certainty DSS). Figure 18 illustrates the difference between the teams who used the Uncertainty DSS, and the teams who used the Certainty DSS.

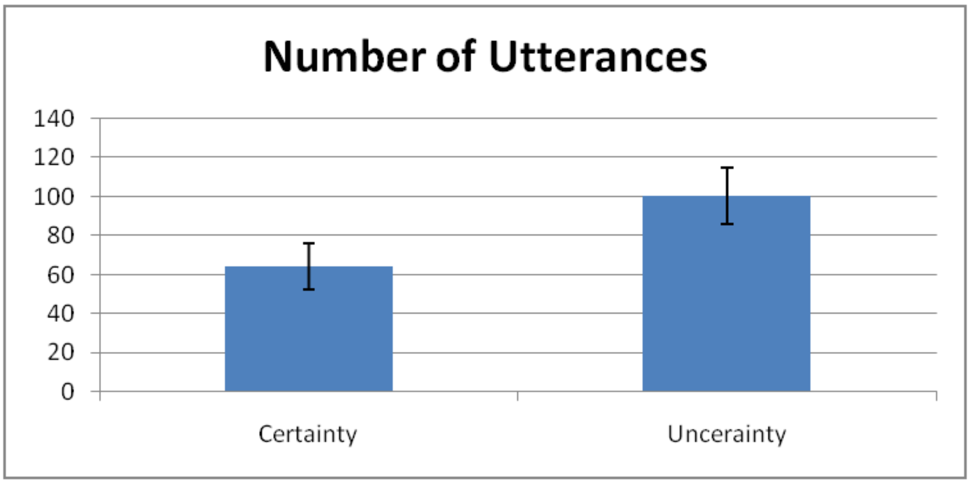


Figure 18.The difference of the number of utterances spoken by team members who used the Uncertainty DSS and the Certainty DSSs. Error bars show standard error.

There were marginally more utterances spoken in the teams who used the Uncertainty DSS ($F_{(1,31)}$=3.257, p=0.085). This result suggests that the Uncertainty DSS encouraged or inspired more communication. Based on the experimenter's experiences observing the teams, there were two areas where Uncertainty DSS triggered discussion. First, when they enter uncertainty values into the DSS, they felt the need to develop a consensus of how much uncertainty each design particularly had. This process of discussion and conflict resolution increased the amount of knowledge shared within the team. Second, when top designs alternative overlapped, teams typically revisited the decision matrix they built with the DSS and looked for actions which they could take to reduce uncertainty. One typical action is to look at the criteria's uncertainty graph to find out which criteria contribute the most to the uncertainty. The process of revisiting the decision matrix and identifying possible uncertainty reducing actions produced shared knowledge of what activities to focus on next. This is likely a reason that teams who used Uncertainty DSS had better situation awareness than the Certainty DSS team.

8.7 Summary and Discussion

8.7.1 Summary

Following are the findings for Experiment Two respecting to the three research questions:

- Can uncertainty visualizations help a team of designers understand when a lack of information interferes with the ability to make a choice?

  Yes. The Uncertainty DSS help design teams realize the uncertainty with their design alternatives and design decision.

- Can uncertainty visualizations improve team members' situation awareness of design alternatives?

  Yes, but an improvement was only observed for level 3 situation awareness, a.k.a. the ability to make prediction of further plans. This is a compelling result since when making decisions it is valuable to understand the caveat and be able to understand the risks and consequences and how to react to them. The Uncertainty DSS can improve the shared awareness of the future plans for a team.

- Will uncertainty visualization in team settings result in with higher workload demand but more preference?

  Yes. The same pattern was observed in team settings as was observed for individual decision makers. In addition, usability benchmarking was conducted for both DSSs and found no significant difference. The *benchmarks indicated that the uncertainty system is exceptional in terms of usability.*

8.7.2 Discussion

Much decision making occurs in a collaborative team environment. When working in a team more time is spent on group communication but the team's collective understanding can augment an individual's limitations in information seeking, and knowledge comprehension. (Cooke, Salas, Cannon-Bowers, & Stout, 2000;Mohammed & Dumville, 2001; Wright & Endsley, 2008).

The results of this experiment demonstrated that a DSS which allows teams to express and visualize uncertainty in alternatives, changes the way which teams thought about their designs, and consequently their discussions of those designs. When teams used the Certainty DSS, the topic of uncertainty rarely entered into their discussions. However, when teams used the Uncertainty DSS, there was much more discussion. Moreover, there was uncertainty related discussion in the design, and how that uncertainty should be managed or reduced. As a result, the rich discussion generated higher demand of workload but resulted in higher preference towards Uncertainty DSS. Furthermore, the Uncertainty DSS helped teams to more accurately perceive when they lack sufficient information to definitively identify the "best" alternative. Without the uncertainty visualization, teams were less aware of that they lacked sufficient information. This is important for helping teams to make better informed decisions, and to better manage their information search efforts. The results on shared situation awareness within team support that uncertainty visualization fostered discussion among team members, resulting in decision makers' understanding of future plans to reduce uncertainty. The extra communication on time and effort exposed on the Uncertainty DSS is arguably worth the effort and the impact is positive. The result of workload and preference on DSSs are even more convincing after usability benchmarking for both DSSs on a scale generated by over 200 systems. Both the systems scored highly on the scale and there were no significant difference between them, which ruled out the possibility of poorly designed system or design with bias.

Chapter 9--Research Contribution and Future Work

9.1 Development and Evaluation of Uncertainty DSS

Uncertainty is present in all situations where decisions are made. Decision makers need tools and approaches that can help them manage uncertainty and its impact on their decision making tasks. In this dissertation the author developed an uncertainty visualization to help decision makers identify when they need more information, and implemented it in a decision support system called the Uncertainty DSS. The tool adds an uncertainty dimension to the traditional multi criteria decision making process. Decision makers can formulate their understanding about the design options similar as single point multi criteria decision matrix, but the beauty of Uncertainty DSS is that it takes ranges as potential scores and uses that to compute uncertainty that decision makers have on the design. On one hand this relaxed the constraint that decision makers have to formulate a single score for each evaluation criteria, which is often very hard to do based on the author's observation. On the other hand, the uncertainty information is very valuable and can be used to create graphs to guide decision makers focusing their information seeking effort on things that will be most effective. Unlike other researchers (LeBoeuf & Shafir, 2005; Stanovich, 1999) who argued in their studies that uncertainty information is hard for decision makers to understand. The results showed engineering designers can easily pick up the concept of uncertainty and use the decision support system without much supervision. This is partly because the good usability design of the tool. Surprisingly even though there was extra workload to use the Uncertainty DSS, it was preferred in both studies. The

Uncertainty DSS created by the author is understandable, usable, and useful.

In this research, it is critical to the researchers to test the Uncertainty DSS in both individual as well as team environment because the majority amount of design scenarios fall into these two cases. It is attractive to validate the effective of Uncertainty DSS for both cases especially since individual and team decision making can be vastly different. When making decisions by one's own, it is tempting for him/her to believe that there is enough information on designs while in team settings, the consensus requires discussion and communication. Also for information seeking, teams can have allocation and cooperation, while individuals need to take care of everything no matter if he/she is familiar with the specific subject. Given these differences, the most compelling results of this research are that in both team and individual scenarios, Uncertainty DSS significantly reduced the unrealistic perception of information sufficiency and provided decision makers guidance to search for clarifying information to make the best design prominent. Engineering designs' information seeking plan was correlated with how much uncertainty the subject matter had in the first experiment. There was a similar result when looking at situation awareness in the second experiment, teams had better projection to future plans (level 3 SA) using Uncertainty DSS.

9.2 User centered approach

Another hallmark of this research is that the idea of the system is coming from use cases "in the wild" and the evaluation of the uncertainty DSS in both studies was through actual designers and design problems. This work sets a good example of the use of Naturalistic Decision Making/Cogitative Engineering methods to motivate design and use of empirical evaluation studies to establish that the proposed design does infect influence decision-making. Design is a complex process and the more artificial the evaluation setting, the more difficult it becomes to assess whether the findings are likely to translate into practical work environments. By using real design tasks with real designers, the realism in the design decision making tasks was preserved. This is a well adoption of cognitive engineering and decision making approach and the make the results more convincing. For both the studies, designers had been working on their design projects for a few weeks and were knowledgeable in their specific design topic. This step is critical because it was important for participants to be knowledgeable about the specific design project, and be seriously vested in the outcome of the decision. When looking at previous researches on uncertainty, there is a collection of psychological literature demonstrating that people violate the axioms of utility theory, assess probabilities/uncertainties incorrectly, fail to incorporate prior probabilities in judgments, and respond to probabilities nonlinearly (Baron, 2000; Cole, 1989; Kahneman & Tversky, 1973, 1979; Leboeuf & Shafir, 2005; Stanovich, 1999).However, many of these studies are based on artificial tasks, such as one-time choices between hypothetical gambles, that have no real consequences and for which the participants receive no feedback about the outcome. It has been argued that this approach underestimates people's ability to reason with uncertainty because the situation is unfamiliar and because there is little motivation for participants to put forth their best effort (Beach, 1996; Cosmides & Tooby, 1996;

Hertwig & Gigerenzer, 1999; McKenzie, 2004).The research in this dissertation justify that uncertainty information in decision making process is understandable, and useful.

9.3 Generalization of the results

The Uncertainty DSS was developed based on the needs and use cases of engineering designers in mind and from this research it was found effective for both individual and team decision makings. Furthermore the applicable domains of the tool are not limited to engineering design. It has potentially broad applicability to any decision making process that involves making comparisons of multiple options along multiple dimensions where there is uncertainty as to the value of options on different dimensions. Based on the observation from the two studies, the tool would be most effective to be used to support complex projects where decision makers are making decisions while developing understanding about the subject matter. Also the use of the tool should start from the early stage of decision making (or project), the earlier the better and the longer the better. That being said, the tool was not intended for one time use when the final decision is needed, rather it is better off for repeated use where decision makers have complex problem and can constantly learn from the tool what information to search for, go out and find information, refine their knowledge and change the scores in the tool, then repeat the process. It is expected that the general trend of uncertainty that decision makers put in the tool is being reduced, but there may be situations when information seeking raises more questions than it answers. So if looking at the degree of uncertainty overtime, it is expected to be a zigzag line going downwards.

9.4 Future work

In Experiment One (individual decision makers), a week after the experiment, neither entry-level nor intermediate-level designers were found to be very consistent at carrying out their plans to gather more information on specific, critical criteria. Further work is needed to develop tools and visualizations that can help them be more effective in these tasks. This step, if successful, will lead to tools that can help decision makers better focus their information search activities, and hence manage their time more effectively, and make better informed and timely decisions.

A longitudinal study on how individuals and teams use the Uncertainty DSS overtime would give researchers more insights into how the uncertainty DSS would be used in practice and how to make it more effective. For example, it would be useful to see how the degree of perceived uncertainty evolves over the lifetime of a design project, and as the result of various information seeking activities. Once decision makers develop more trust of the Uncertainty DSS, will there be a behavioral change (i.e. a better correlation) between what information decision makers plan to seek and what they actually do.

Bibliography

Anderson, J. R. (1993) The Development of Expertise, in Readings in Knowledge Acquisition and Learning, Morgan Kaufman Publishers, San Francisco, CA USA.

Artman, H., 1999, Situation awareness and co-operation within and between hierarchical units in dynamicdecision making, Ergonomics, 42, 1404–1417.

ARTMAN, H., 2000, Team situation assessment and information distribution, Ergonomics, 43, 1111–1128.

Aughernbaugh&Paredis, 2006, The value of using imprecise probabilities in engineering design, journal of mechanical design, 128(4), 969-979

Ball, L.J., Evans, J.St.B.T., Dennis, I., 1994. Cognitive processes in engineering design: a longitudinal study. Ergonomics 37 (11), 1753–1786.

Bangor, Aaron, Kortum, Philip T. and Miller, James T. (2008)."An Empirical Evaluation of the System Usability Scale". International Journal of Human-Computer Interaction 24 (6): 574 - 594.

Baron, J. (2000). Thinking and deciding (3rd Ed.). New York: Cambridge University Press. Bellman, R. E and L. A. Zadeh, (1970), "Decision making in a fuzzy environment," Management Sciences 17 B141-B164.

Beach, L.R. (1996). Decision making in the workplace: A unified perspective. Mahwah, NJ: Erlbaum.

Belkin, N.J. (1980). Anomalous states of knowledge as a basis for information retrieval. Canadian Journal of Information Science, 5, 133–143.

Borgers, R., Mullen, P.D., Mertens, R., Rijken, M., Eussen, G., Plagge, I., Visser, A.P., &Blijham, G.H. (1993).The information-seeking behavior of cancer outpatients—A description of the situation. Patient Education and Counseling, 22(1), 35–46.

Bradley, S.R. &Agogino, A.M., 1994, An Intelligent Real Time Design Methodology for Catalog Selection, Transaction of the ASME, Journal of Mechanical Design, 116, pp. 980-988.

Brooke, J. (1996). "SUS: a "quick and dirty" usability scale". In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland. Usability Evaluation in Industry. London: Taylor and Francis.

Buchanan, R., 1992. Wicked problems in design thinking. Design Issues 8 (2), 5–12.

Cannon-Bowers, J. A., & Salas, E. (2001).Reflections on shared cognition. Journal of Organizational Behavior,22, 195–202.

Carlisle E, Spranca M, KanouseDE.Empirical Studies of Decision Aids for Consumers. Santa Monica:

The RAND Corporation. June 2003.

Cellier, J., Eyrolle, H., & Marine, C. (1997).Expertise in dynamic environments. Ergonomics, 40, 28–50.

Charness, N., &Tuffiash, M. (2008).The role of expertise research and human factors in capturing, explaining,and producing superior performance. Human Factors, 50, 427–432.

Chi, M., Feltovich, P., & Glaser, R. (1981).Categorization and representation of physics problems by experts andnovices. Cognitive Science, 5, 121–152.

Chi, M., Glaser, R., & Farr, M. (1988).The nature of expertise. Mahwah, NJ: Erlbaum.

Clyde W. Holsapple and Andrew B. Whinston, 1996. Decision Support Systems: A Knowledge-based Approach. West Group; 10th edition

Cole, W.G. (1989). Understanding Bayesian Reasoning via Graphical Displays. Proceedings of CHI'89 Human Factors in Computing Systems. Austin, Texas, ACM Press, 381-386.

Collier, S.G., Folleso, K., 1995. SACRI: a measure of situation awareness for nuclear power control rooms. In: Garland, D.J., Endsley, M.R. (Eds.), Experimental Analysis and Measurement of Situation Awareness. Embry-Riddle University Press, Daytona Beach, FL, pp. 115–122.

Connolly, T. (1988).Hedge-clipping, tree-felling, and the management of ambiguity. In M. B. McCaskey,L. R. Pondy, & H. Thomas (Eds.),Managing the challenge of ambiguity and change (pp. 37–50). New York:Wiley.

Cooper, R. G, &Kleinschmidt, E. J. (1994). Determinants of timeliness in product development. Journal of Product Innovation Management, 11, 381-396

Cosmides, L., and Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature of judgment under uncertainty. Cognition, 58, 1-78

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000).Measuring team knowledge.Human Factors,42, 151–173.

Cooke, N.J., STOUT, R. and SALAS, E., 2001, A knowledge elicitation approach to the measurement of teamsituation awareness. In New Trends in Cooperative Activities: Understanding System Dynamics in ComplexEnvironments, M. McNeese, E. Salas and M.R. Endsley (Eds.), pp. 114–139 (Santa Monica, CA: HumanFactors and Ergonomics Society).

Cushing, S. (1994). Fatal words: Communication clashes and aircraft crashes. Chicago: University of Chicago Press.

Davies, S.P., 1991. Characterizing the program design activity: neither strictly top-down nor globally opportunistic. Behaviour& Information Technology 10 (3), 173–190.

Dekker, S.W.A., 2000, Crew situation awareness in high-tech settings: tactics for research into an ill-definedphenomenon. Transportation Human Factors, 2, 49–62.

Dreyfus, S. E. (1981). Formal models vs. human situational understanding: Inherent limitations on the modeling of business expertise (ORC 81-3). Berkeley: Operations Research Center, University of California.

Ehrlenspiel, K., 1999. Practicians—how are they designing? And why? In: Proceedings of International Conference on Engineering Design (ICED '99), Munich.

Endsley, M.R., 1995a. Towards a theory of situation awareness in dynamic systems. Hum. Factors 37, 32–64.

Endsley, M.R., 1989. Final Report: Situation Awareness in an Advanced Strategic Mission (NOR DOC 89-32). Northrop Corporation, Hawthorne, CA.

Endsley, M.R., 1990. Predictive utility of an objective measure of situation awareness. In: Proceedings of the Human Factors Society 34th Annual Meeting, Human Factors Society, Santa Monica, CA, pp. 41–45.

Endsley, M.R., 2000. Theoretical underpinnings of situation awareness: a critical review. In: Endsley, M.R., Garland, D.J. (Eds.), Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates, London.

Endsley, M.R., Sollenberger, R., Stein, E., 2000a. Situation awareness: a comparison of measures. In: Proceedings of the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium. SA Technologies, Inc., Savannah, GA.

Endsley, M.R., Holder, C.D., Leibricht, B.C., Garland, D.C., Wampler, R.L., Matthews, M.D., 2000b.Modelling and Measuring Situation Awareness in the Infantry Operational Environment (1753). Army Research Institute, Alexandria, VA.

Endsley, M.R., Smolensky, M., 1998. Situation awareness in air traffic control: the picture. In: Smolensky, M., Stein, E. (Eds.), Human Factors in Air Traffic Control. Academic Press, New York, pp. 115–154.

Entin, E. E., &Serfaty, D. (1999).Adaptive team coordination. Human Factors, 41, 312–325.

Erev, I., Bornstein, G. and Wallsten, T. W. (1993).The negative effect of probability assessment on decision quality. Organizational Behavior and Human Decision Processes, 55, 78–94.

Fischer, U., &Orasanu, J. (2000). Error-challenging strategies: Their role in preventing and correcting errors.Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th AnnualMeeting of the Human Factors and Ergonomics Society (pp. 1.30–1.33). Santa Monica, CA: Human Factorsand Ergonomics Society.

Fischer, U., Orasanu, J., &Montalvo,M. (1993). Efficient decision strategies on the flight deck. In R. S. Jensen &D. Neumeister (Eds.), Proceedings of the 7th International Symposium on Aviation Psychology (pp. 238–243).Columbus: Ohio State University.

Flin, R., O'Connor, P., &Crichton,M. (2008). Safety at the sharp end: A guide to non-technical skills. Aldershot,UK: Ashgate.

Fowlkes, J.E., SALAS, E., BAKER, D.P., CANNON-BOWERS, J.A. and STOUT, R.J., 2000, The utility of event-basedknowledge elicitation. Human Factors, 42, 24–35.

Gillan, C. A. (2003). Analysis of multicrew decision making from a cognitive perspective.Proceedings of the12th International Symposium on Aviation Psychology (pp. 427–432). Dayton, OH:Wright State University.

Gugerty, L.J., 1997. Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. J. Exp. Psychol.: Appl. 3, 42–66.

Hauss, Y., Eyferth, K., 2003. Securing future ATM-concepts' safety by measuring situation awareness in ATC. Aerospace Sci Technol. 7, 417–427.

Hayes, C. C. and Akhavi, F. 2008 Cost Effective Decision Aids for Complex Tasks, Journal of Usability Studies, 2(4) pp. 152-172

Hayes, J.R. 1981, The Complete Problem Solver (pp. 145-160). Philadelphia, PA: The Franklin Institute Press

Hertwig, R., and Gigerenzer, G. (1999). The "conjunction fallacy" revisited: how intelligent inferences look like reasoning errors. Journal of Behavioral Decision Making, 12, 275-305.

Hertzum, M. and Pejtersen, A.M. 2000. The information-seeking practices of engineers: searching for documents as well as for people. Information Processing and Management, 36(5): 761-778.

Heuer, R. J.m Jr. (1999). Psychology of intelligence analysis. Washington, DC: Center for the Study of Intelligence, Central Intelligence Agency.

Hoffman, J.R., Wilkes, M.S., Day, F.C., Bell, D.S., and Higa, J.K. (2006). The roulette wheel: An aid to informed decision making. PLoS Medicine, 3, 743-748

Hogg, D.N., Folleso, K., Strand-Volden, F., Torralba, B., 1995.Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. Ergonomics 38 (11), 2394–2413.

Johannesen, L. (2008).Maintaining common ground: An analysis of cooperative communication in the operating room. In C. P. Nemeth (Ed.), Improving healthcare team communication: Building on lessons from aviation and aerospace (pp. 179–203). Aldershot, UK: Ashgate.

Jones, D.G., Kaber, D.B., 2004. Situation awareness measurement and the situation awareness global assessment technique. In: Stanton, N., Hedge, A., Hendrick, H., Brookhuis, K., Salas, E. (Eds.), Handbook of Human Factors and Ergonomics Methods. CRC Press, Boca Raton, USA, pp. 42-1–42-8.

Kaber, D.B., Endsley, M.R., Wright, M.C., Warren, H., 2002.The effects of levels of automation on performance, situation awareness, and workload in an advanced commercial aircraft flight simulation. Final Rep.: NASA Langley Research Center Grant #NAG-1-01002). NASA Langley Research Center, Hampton, VA.

Kahneman, D., and Tversky, A. (1973).On the psychology of prediction. Psychological Review, 80, 237-251.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Keller, J., Lebiere, C., Shay, R., Latorella, K., 2004. Cockpit system situational awareness modelling tool. In: Proceedings of the Fifth Human Performance, Situation Awareness and Automation Conference, Daytona Beach, FL, March 22–25, 2004.

Kemper E. Lewis, Wei Chen, Linda C. Schmidt, 2006, Decision making in engineering design.page 3

Kessler, E., &Chakrabarti, A. K. (1996). Innovation speed: A contextual model of context, antecedents, and outcomes. Academy of Management Review, 21, 1143- 1191.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986).Rapid decision making on the fire ground.In proceedings of the Human Factors Society 30th Annual Meeting (pp. 576-580). Santa Monica, CA: Human Factors Society.

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

Klein, G. A., Orasanu J., Calderwood R., Zsambok C. E., 1993, Decision Making in Action. Norwood, New Jersey, USA: Ablexpublishingcorporation

Klein, G. A. (1993a). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein,J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods(pp. 138–147). Norwood, NJ: Ablex Publishing.

Klein, G. A. (1993b). Naturalistic decision making: Implications for design (CSERIAC SOAR Rep. 93-1).WrightPatterson Air Force Base, OH: CSERIAC.

Krifka, M.,Martens, S., & Schwarz, F. (2004). Linguistic factors.In R. Dietrich & T. M. Childress (Eds.), Groupinteraction in high risk environments (pp. 75–85).Aldershot, UK: Ashgate.

Kuhlthau, C.C. (1993a). Seeking meaning: A process approach to library and information services.

Norwood, NJ: Ablex.) linking of affective states to information-seeking stages. Kuhlthau's (1993a Kuhlthau, C.C. (1993a). Seeking meaning: A process approach to library and information services. Norwood, NJ: Ablex.

Kuhlthau, C.C. (1993b). A principle of uncertainty for information seeking. Journal of Documentation, 49(4), 339–355.

Larson, Adam D. and Caroline Hayes, 2005."An Assessment of Weasel: A Decision Support System to Assist in Military Planning" Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. pp. 287-291.

Law, W. S. 1996, Evaluating imprecision in engineering design. Ph.D. Dissertation, California Institute of Technology, Pasadena, California.

LeBoeuf, R.A., and Shafir, E.B. (2005).Decision making.In J. Holyoak and R.G. Morrison (Eds.), The Cambridge handbook of thinking and reasoning (pp. 243-266). New York: Cambridge University Press.

Leonard, M., Graham, S., &Bonacum, D. (2004). The human factor: The critical importance of effective teamworkand communication in providing safe care. Quality and Safety in Healthcare, 13(Suppl. 1), 85–90.

Lipshitz, R. (1993). Converging themes in the study of decision-making in realistic settings. In G. A. Klein,J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods(pp. 103–137). Norwood, NJ: Ablex.

McKenzie, C.R.M. (2004). Judgment and decision-making.In K. Lamberts and R.L. Goldstone (Eds.), Handbook of cognition (pp. 321-340). London: sage.

MacMillan, J., Entin, E. E., &Serfaty, D. (2004). Communication overhead: The hidden cost of team cognition.In E. Salas & S. M. Fiori (Eds.), Team cognition (pp. 61–82).Washington, DC: American Psychological Association.

Marek J. Druzdzel, Roger R. Flynn. 2002. Decision Support Systems in Encyclopedia of Library and Information Science, Second Edition by Allen Kent. New York: Marcel Dekker, Inc.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000).The influence of sharedmental models on team process and performance. Journal of Applied Psychology, 85, 273–283.

Matthews, M.D., Pleban, R.J., Endsley, M.R., Strater, L.D., 2000. Measures of infantry situation awareness for a virtual MOUT environment. Proceedings of the Human Performance, Situation Awareness and Automation: User Centred Design for the New Millennium Conference, October 2000.

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., et al. (2009).Surgical team behaviorand patient outcomes. American Journal of Surgery, 197, 678–685.

Mignerey, J.T., Rubin, R.B., &Gorden, W.I. (1995).Organizational entry-An investigation of newcomer communication behavior and uncertainty. Communication Research, 22(1) 54–85.

Mohammed, S., &Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expandingtheory and measurement across disciplinary boundaries. Journal of Organizational Behavior, 22, 89–106.

Mosier, K. L.& Fischer, Ute M., (2010) Judgment and Decision Making by Individualsand Teams: Issues, Models, and Applications. Reviews of Human Factors and Ergonomics, pp198-256

Morineau, T., Morandi, X., LeMoëllic, N., Diabira, S., Riffaud, L., Haegelen, C., et al. (2009).Decision makingduring preoperative surgical planning. Human Factors, 51, 67–77.

Nadav-Greenberg, L., Joslyn,S. L., Taing, M. U.. 2007. The Effect of Uncertainty Visualizations on Decision Making in Weather Forecasting. Proceedings of the Eighth International NDM Conference (Eds. K. Mosier & U. Fischer), Pacific Grove, CA, June 2007

Nadav-Greenberg, L., Joslyn, S.L. (2009). Uncertainty Forecasts Improve Decision Making Among Nonexperts. Journal of Cognitive Engineering and Decision Making, 3(3), 209-227.

National Transportation Safety Board. (1994). Safety study: A review of flightcrew-involved, major accidents ofU.S. air carriers, 1978 through 1990 (NTSB/SS-94/01). Washington, DC: National Technical InformationService.

Nemeth (Ed.), Improving healthcare team communication: Building on lessons from aviation and aerospace(pp. 97–114). Aldershot, UK: Ashgate.

Newell, A., 1973. Artificial Intelligence and the concept of mind. In: Schank, R.C., Colby, K.M. (Eds.), Computer Models of Thought and Language. Freeman, San Francisco, CA.

O'Connor AM, Llewellyn-Thomas HA, Barry Flood A. Modifying Unwarranted Variations in Health Care: Shared Decision Making Using Patient Decision Aids. Health Affairs: Web Exclusive. October 7, 2004; 63-72

Omodei, M. M., Wearing, A. J., McLennan, J., Elliott, G. C., & Clancy, J. M. (2004). More is better? Problems of self-regulation in naturalistic decision making settings. In B. Brehmer, F. Lipshitz& H. Montgomery (Eds.), How professionals make decisions. Mahwah, NJ: Lawrence Erlbaum Associates.

Orasanu, J. (1990). Shared mental models and crew decision making (Tech. Rep. 46). Princeton, NJ: PrincetonUniversity, Cognitive Science Laboratory.

Orasanu, J. (1994). Shared problem models and flight crew performance. In N. Johnston, N. McDonald, &R. Fuller (Eds.), Aviation psychology in practice (pp. 255–285). Aldershot, UK:

Ashgate.

Orasanu, J.,& Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood,& C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 3–20). Norwood, NJ: Ablex.

Orasanu, J., & Fischer, U. (1992).Distributed cognition in the cockpit: Linguistic control of shared problem solving. Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society (pp. 189–194). Hillsdale, NJ: Erlbaum.

Orasanu, J., Fischer, U., & Davison, J. (1997).Cross-cultural barriers to effective communication in aviation.In C. S. Granrose& S. Oskamp (Eds.), Cross-cultural work groups (pp. 134–160). Thousand Oaks, CA: Sage.

Orasanu, J., Fischer, U.,McDonnell, L. K., Davison, J.,Haars, K. E.,Villeda, E., et al. (1998).How do flight crewsdetect and prevent errors? Findings from a flight simulation study. Proceedings of the Human Factors andErgonomics Society 42nd Annual Meeting (pp. 191–195). Santa Monica, CA: Human Factors and ErgonomicsSociety

Oskamp, S. (1965).Overconfidence in case study judgments. Journal of Consulting Psychology, 29, 261-265

Patel, V. L., &Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A.Ericsson& J. Smith (Eds.), Toward a general theory of expertise: Prospects and limits (pp. 93–125). Cambridge,UK: Cambridge University Press.

Perry, S. J.,& Wears, R. L. 2011. Large scale coordination of work: Coping with complex chaos within healthcare.In K. Mosier & U. Fischer (Eds.), Informed by knowledge: Expert performance in complex situations.New York: Taylor & Francis.

Rasker, P.C., POST, W.M. and SCHRAAGEN, J.M.C., 2000, Effects of two types of intra-team feedback ondeveloping a shared mental model in command and control teams, Ergonomics, 43, 1167–1189.

Reader, T., Flin, R., &Cuthbertson, B. (2008). Factors affecting team communication in the intensive care unit(ICU). In C. P. Nemeth (Ed.), Improving healthcare team communication: Building on lessons from aviationand aerospace (pp. 117–133). Aldershot, UK: Ashgate.

Rittel, H.W.J., Webber, M.M., 1973. Dilemmas in a general theory of planning. Policy Sciences 4 (2), 155–169.

Rouse,W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complexsystems. IEEE Transactions on Systems, Man, and Cybernetics, 22, 1296–1308.

Saaty, T. L., 1980 The Analytical Hierarchy Process. New York, NY, USA :MacGrawHill

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries anddevelopments. Human Factors, 50, 540–547.

Salas, E., and Klein, G. 2001 Linking expertise and naturalistic decision making. Mahwah, NJ: Lawrence Erlbaum Associates.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a big five in teamwork? Small Group Research, 36, 555–599.

Schraagen, J. M. Militello, L. G. Ormerod, T. Lipshitz, R. 2008, Naturalistic Decision Making and Macrocognition, Ashgate Publishing Limited.

Sexton, J. B., &Helmreich, R. L. (2000).Analyzing cockpit communication.The links between language, performance,error and workload. Human Performance in Extreme Environments, 5, 63–68.

Simon, H.A., 1973. The structure of ill-structured problems. Artificial Intelligence 4, 181–201.

Simon, H. A., &Chase,W. G. (1973). Skill in chess. American Scientist, 61, 394–403.

Smallman, H. S. &Hegarty, M. (2007).Expertise, spatial ability and intuition in the use of complex displays. Paper accepted for presentation at The 51st Annual Conference of the Human Factors and Ergonomics Society, Baltimore, MD, October 1st-5th.

Smith-Jentsch, K. A., Cannon-Bowers, J. A., Tannenbaum, S. I., & Salas, E. (2008).Guided team self-correction:Impacts on team mental models, processes, and effectiveness. Small Group Research, 39, 303–327.

Smith, Philip J., C. Elaine McCoy, and Charles Layton, 1997. "Brittleness in the Design of Cooperative Problem-Solving Systems: The Effects on User Performance." IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans. Vol. 27, No. 3, pp. 360-371.

Stanovich, K.E. (1999). Who is rational? Studies of individual differences in reasoning. Mahwah, NJ: Erlbaum.

Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Young, M.S., Linsell, M., Dymott, R., Green, D., 2005. Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. Ergonomics, submitted.

Stempfle, J., Badke-Schaub, P., 2002. Thinking in design teams—an analysis of team communication. Design Studies 23, 473–496.

Stout, R. J., Cannon-Bowers, J. A., Salas, E., &Milanovich, D. M. (1999). Planning, shared mental models, andcoordinated performance: An empirical link is established. Human Factors, 41, 61–71.

Triantaphyllou, E.; Kovalerchuk, B.; Mann, L. and Knapp, G. M. (1995). Determining the Most Important Criteria in Maintenance Decision Making, Journal of Quality In Maintenance Engineering, 3(1)

Triantaphyllou, E. (1997). A Sensitivity Analysis Approach for Some Deterministic Multi-Criteria Decision Methods, Decision Sciences, 28(1):151-194.

Tschan, F., Semmer, N. K., Gurtner, A., Bizzari, L., Spychiger, M., Breuer, M., et al. (2009). Explicit reasoning,confirmation bias, and illusory transactive memory: A simulation study of group medical decision making.Small Group Research, 40, 271–300.

Waag,W., &Bell,H. (1997). Situation assessment and decision-making in skilled fighter pilots. In C. E. Zsambok& G. Klein (Eds.), Naturalistic decision making (pp. 247–354). Mahwah, NJ: Erlbaum.

Weick, K. E. (1995). Sensemaking in organizations. Thousand Oaks, CA: Sage.

Wright, M. C., &Endsley, M. R. (2008). Building shared situation awareness in healthcare settings. In C. P.

Xiao, Y., Milgram, P., & Doyle, D. J. (1997a). Capturing and modeling planning expertise in anesthesiology.In C. Zsambok& G. Klein (Eds.), Naturalistic decision making (pp. 197–215). Mahwah, NJ: Erlbaum.

Xiao,Y.,Milgram, P., &Doyle,D. J. (1997b). Planning behavior and its functional role in interactions with complexsystems. IEEE Transactions on Systems,Man, and Cybernetics: Part A, Systems and Humans, 27, 313–324.

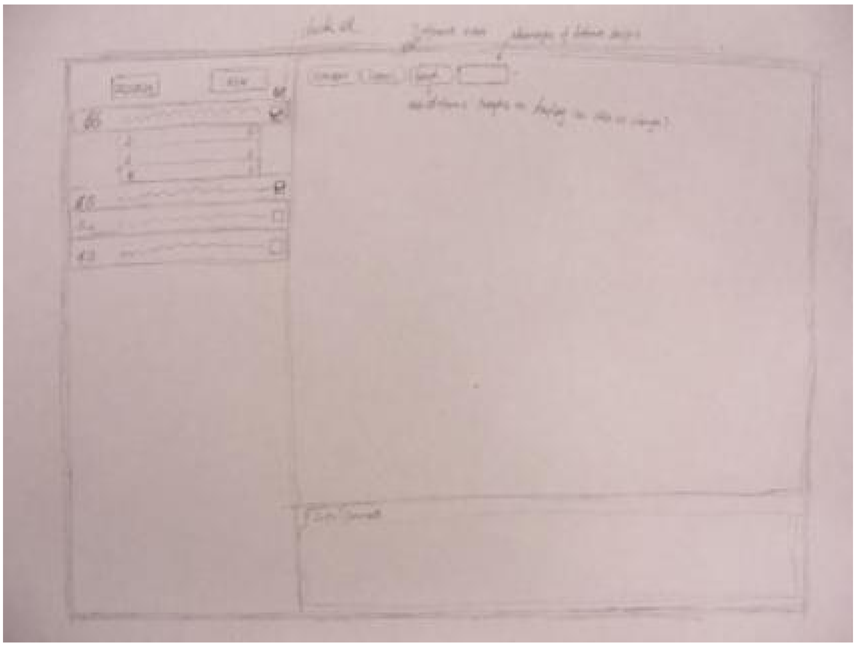
Zheng, X.S., McConkie, G.W., Tai, Y., 2004. Dynamic monitoring of traffic flow: the driver's situation awareness. In: Proceedings of the Fifth Human Performance, Situation Awareness and Automation Conference, Daytona Beach, FL, March 22–25, 2004.

Zsambok C. E., Klein G., 1997, Naturalistic Decision Making, Psychology Press, 1997

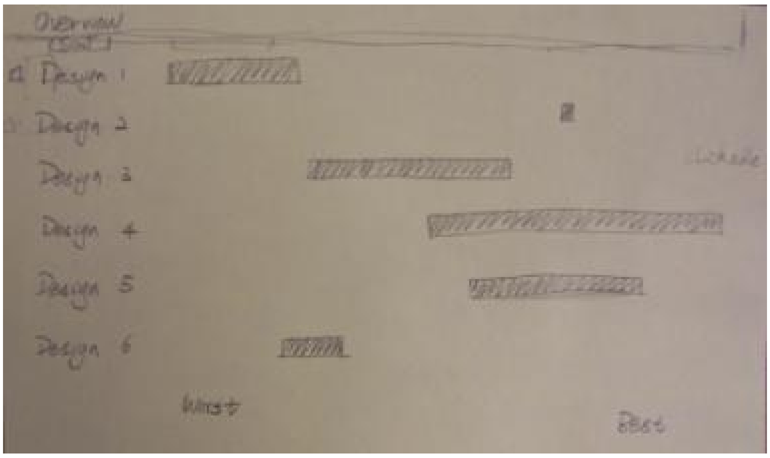
Appendices

A. Paper prototype

Main View:

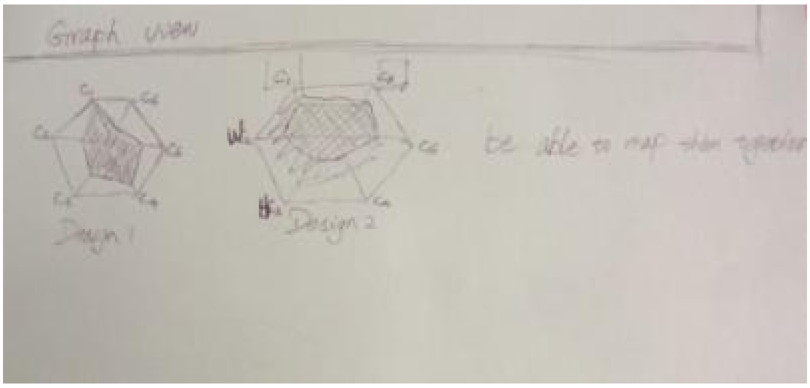


Overview View:

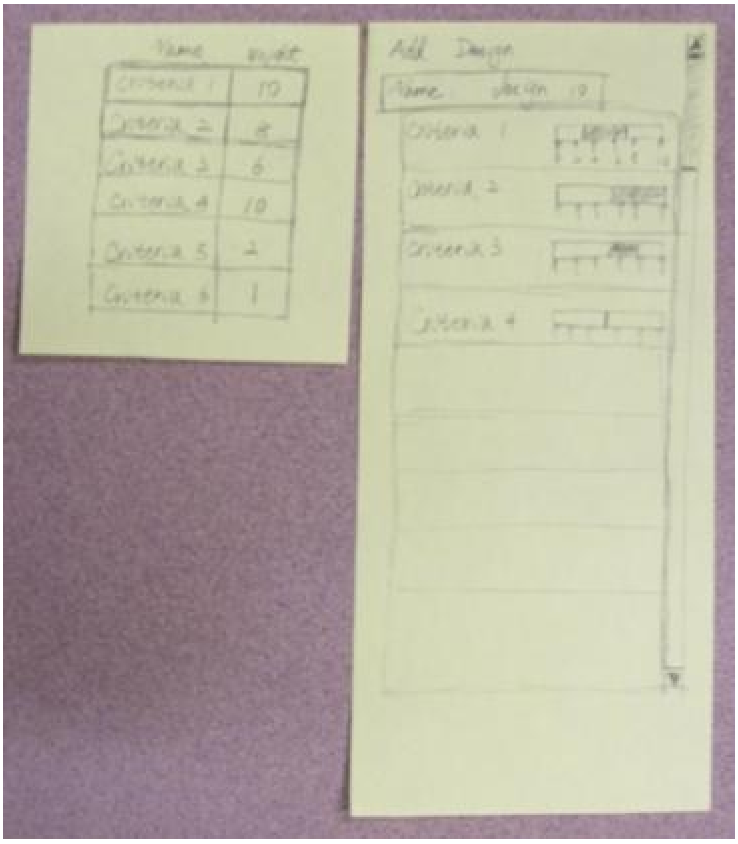


Detail View:

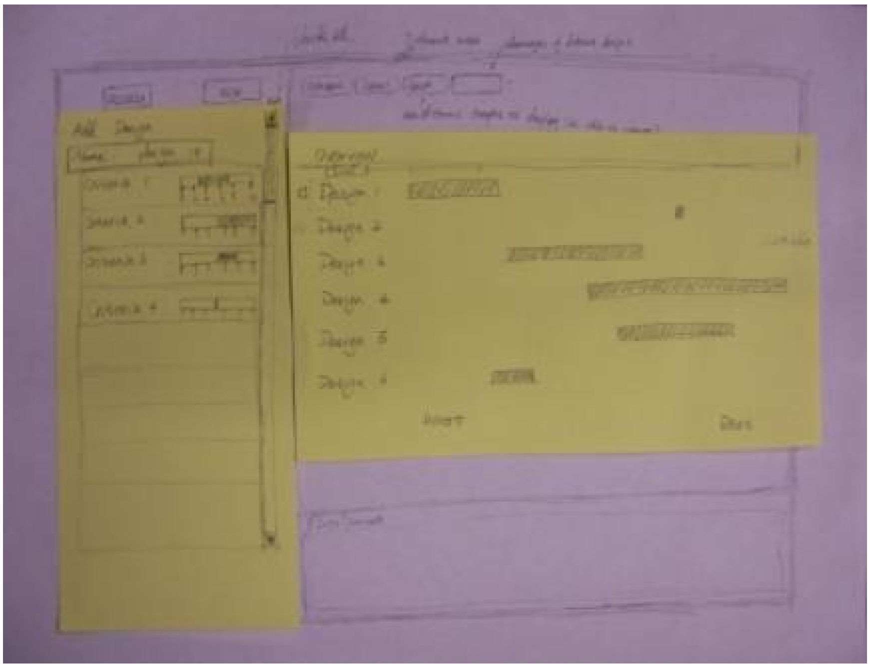
Graph View:



Adding Criteria and Designs:



Window Integration:



B The questionnaire for experiment one

Background information before the experiment

We appreciate your participation in this study. Before starting the experiment, we would like you to fill out this anonymous questionnaire. All information will be kept confidential. Please ask an experimenter if you have any questions. Thank you!

Background information

Age:

Sex : M F

What year student are you?	Sophomore Junior Senior Graduate student
How many years of design experience do you have?	___________________________________
Have you ever used any type of decision support tool? Please specify:	Y N

____

Questionnaire after the experiment

This set of questions is asking about your experience **using your own method** (no decision support tool) to make design decisions. Please indicate how much you agree or disagree with the following statement. (Mark an option that applies)

- It was hard to identify the best design options.

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "I feel I had sufficient information to make an informed decision."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "Going through this process required me to think carefully and thoroughly."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "I feel this method required more effort than should be necessary."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "I am not very confident that the design(s) I chose were the best ones."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- If there is a particular aspect of the design(s) on which you would like to know more in order to be confident of your decision? If so, please elaborate below and indicate how strongly you want to know about it. Give it a score from 1-7 with 1 being strongly undesired and 7 being strongly desired.

This set of questions is asking about your experience using **decision support tool without uncertainty** to make design decisions.

- It was hard to identify the best design options.

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "I feel I had sufficient information to make an informed decision."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "Going through this process required me to think carefully and thoroughly."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "I feel this method required more effort than should be necessary."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- "I am not very confident that the design(s) I chose were the best ones."

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- If there is a particular aspect of the design(s) on which you would like to know more in order to be confident of your decision? If so, please elaborate below and indicate how strongly you want to know about it. Give it a score from 1-7 with 1 being strongly undesired and 7 being strongly desired.

This set of questions is asking about your experience using **decision support tool with uncertainty** to make design decisions

- It was hard to identify the best design options.

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- “I feel I had sufficient information to make an informed decision.”

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- “Going through this process required me to think carefully and thoroughly.”

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- “I feel this method required more effort than should be necessary.”

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- “I am not very confident that the design(s) I chose were the best ones.”

Strongly Disagree	*Somewhat Disagree*	*Slightly Disagree*	*Neutral*	*Slightly Agree*	*Somewhat Agree*	*Strongly Agree*
3	*2*	*1*	*0*	*1*	*2*	*3*

- If there is a particular aspect of the design(s) on which you would like to know more in order to be confident of your decision? If so, please elaborate below and indicate how strongly you want to know about it. Give it a score from 1-7 with 1 being strongly undesired and 7 being strongly desired.

Please split 100 points among the three decision making methods you have just used. The most preferred tool should get more points. The difference in points should reflect the difference in preference.

No decision support tool: ______________
Tool without uncertainty: ______________
Tool with uncertainty: ______________

The questionnaire for experiment two

Background information before the experiment

We appreciate your participation in this study. Before starting the experiment, we would like you to fill out this anonymous questionnaire. All information will be kept confidential. Please ask an experimenter if you have any questions. Thank you!

Background information

Age:			
Sex :	M	F	
What year student are you?	Sophomore	Junior	Senior Graduate student
How many years of design experience do you have?	______________________________		
Have you ever used any type of decision support tool? Please specify:	Y N		

____

Questionnaire after the experiment

Definition of Task Demand Factor

Mental demand

How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Physical demand

How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal demand

How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Frustration level

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

NASA-TLX Mental Workload Rating Scale

Please place an "X" along each scale at the point that best indicates your experience with the display configuration.

The following evaluation is based on your experience using the decision support system. Please think of the demand of using the system itself when you answer the following questions.

Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc)? Was the mission easy or demanding, simple or complex, exacting or forgiving?

Low High

Physical Demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the mission easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low High

Temporal Demand: How much time pressure did you feel due to the rate or pace at which the mission occurred? Was the pace slow and leisurely or rapid and frantic?

Low High

Performance: How successful do you think you were in accomplishing the goals of the mission? How satisfied were you with your performance in accomplishing these goals?

Low High

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low High

Frustration: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during your mission?

Low High

The following evaluation is based on your experience making decisions as a group. Please think of the group decision making process of this meeting when you answer the following questions.

<u>Mental Demand</u>: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc)? Was the mission easy or demanding, simple or complex, exacting or forgiving?

Low High

<u>Physical Demand</u>: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the mission easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low High

<u>Temporal Demand</u>: How much time pressure did you feel due to the rate or pace at which the mission occurred? Was the pace slow and leisurely or rapid and frantic?

Low High

<u>Performance</u>: How successful do you think you were in accomplishing the goals of the mission? How satisfied were you with your performance in accomplishing these goals?

Low High

<u>Effort</u>: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low High

<u>Frustration</u>: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during your mission?

Low High

Measure of situation awareness

1. How many designs do you have?

2. How many criteria do you have?

3. Which criteria is the most important? Write more than one if there is tie.

4. Is your first design better than the third one in total score?

5. Which design is the best in total score

6. What are the top two features needed by your users?

7. What are the pros and cons of each design?

8. Which design is the hardest to manufacture?

9. Which is the most innovative design?

10. Are there several designs where it is not clear which one is the best? (If there are, can you list them?)

11. If you are going to produce your product/system, how are you going to test your designs?

12. What types of information do you need the most right now to continue your design work?

13. Which design worries you the most? How can you improve on it?

14. For which design do you have the most open questions?
    Which design alternative do you feel like working on the most in the next week?

---

15. Which design do you feel is least likely to require major changes in the future?

What is more helpful to support your decision making?

What do you think of this decision support system?

Perceived information sufficiency

Circle one answer for each statement according to your agreement of the statement.

- "Going through this process required me to think carefully and thoroughly."

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

- "I feel I had sufficient information to make an informed decision."

Strongly	*Disagree*	*Neutral*	*Agree*	*Strongly*

Disagree				*Agree*
-2	*-1*	*0*	*1*	*2*

- "For some designs, I feel like I still need to learn more about them."

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

- "If I am forced to choose a design right now, it is likely that my decision will be sub optimal."

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

- "There are things that I don't know about each design alternative."

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

Preference

I like the functions provided by the system

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I don't think the system fit well into the design work I am doing

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I think that I would like to use this system frequently

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

System Usability Scale

I found the system unnecessarily complex

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I thought the system was easy to use

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I think that I would need the support of a technical person to be able to use this system

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I found the various functions in this system were well integrated

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I thought there was too much inconsistency in this system

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I would imagine that most people would learn to use this system very quickly

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I found the system very cumbersome to use

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I felt very confident using the system

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

I needed to learn a lot of things before I could get going with this system

Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*
-2	*-1*	*0*	*1*	*2*

C The analysis of experiment one

Preference and effort using different DSSs

One way ANOVA for preference of using Control DSS (entry vs. intermediate engineering designers)

ANOVA

VAR00002

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.402	1	6.402	.205	.656
Within Groups	625.417	20	31.271		
Total	631.818	21			

One way ANOVA for preference of using Certainty DSS (entry vs. intermediate engineering designers)

ANOVA

Certainty DSS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	306.818	1	306.818	12.919	.002
Within Groups	475.000	20	23.750		
Total	781.818	21			

One way ANOVA for preference of using Uncertainty DSS (entry vs. intermediate engineering designers)

ANOVA

VAR00004

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	224.583	1	224.583	3.821	.065
Within Groups	1175.417	20	58.771		
Total	1400.000	21			

One way ANOVA for entry-level designers' effort of using three different systems

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Type of DSS	Sphericity Assumed	42.889	2	21.444	21.020	.000
	Greenhouse-Geisser	42.889	1.792	23.931	21.020	.000
	Huynh-Feldt	42.889	2.000	21.444	21.020	.000
	Lower-bound	42.889	1.000	42.889	21.020	.001
Error(type of DSS)	Sphericity Assumed	22.444	22	1.020		
	Greenhouse-Geisser	22.444	19.714	1.139		
	Huynh-Feldt	22.444	22.000	1.020		
	Lower-bound	22.444	11.000	2.040		

Pairwise Comparisons

Measure:MEASURE_1

(I) type of DSS	(J) Type of DSS	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
Control	Certainty	2.667*	.355	.000	1.665	3.669
	Uncertainty	1.167	.474	.095	-.170	2.504
Certainty	Control	-2.667*	.355	.000	-3.669	-1.665
	Uncertainty	-1.500*	.399	.009	-2.625	-.375
Uncertainty	Control	-1.167	.474	.095	-2.504	.170
	Certainty	1.500*	.399	.009	.375	2.625

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

One way ANOVA for intermediate-level designers' effort of using three different systems

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Type of DSS	Sphericity Assumed	4.200	2	2.100	5.395	.156
	Greenhouse-Geisser	4.200	1.255	3.346	5.395	.185
	Huynh-Feldt	4.200	1.362	3.083	5.395	.180
	Lower-bound	4.200	1.000	4.200	5.395	.199
Error(Type of DSS)	Sphericity Assumed	11.133	18	.619		
	Greenhouse-Geisser	11.133	11.297	.986		
	Huynh-Feldt	11.133	12.262	.908		
	Lower-bound	11.133	9.000	1.237		

Pairwise Comparisons

Measure:MEASURE_1

(I) type of DSS	(J) Type of DSS	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Control	Certainty	.900	.458	.243	-.444	2.244
	Uncertainty	.600	.340	.334	-.397	1.597
Certainty	Control	-.900	.458	.243	-2.244	.444
	Uncertainty	-.300	.213	.580	-.926	.326

Uncertainty	Control	-.600	.340	.334	-1.597	.397
	Certainty	.300	.213	.580	-.326	.926

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

T-test of effort between entry-level and intermediate-level designers with No DSS support.

Group Statistics

	Expertise	N	Mean	Std. Deviation	Std. Error Mean
Effort	Entry	12	4.5833	1.62135	.46804
	Intermediate	10	3.4000	1.42984	.45216

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Expertise	Equal variances assumed	.129	.723	1.797	20	.087	1.18333	.65859	-.19045	2.55712
	Equal variances not assumed			1.818	19.914	.084	1.18333	.65078	-.17454	2.54121

T-test of effort between entry-level and intermediate-level designers with Certainty DSS.

Group Statistics

	Expertise	N	Mean	Std. Deviation	Std. Error Mean
Effort	Entry	12	1.9167	.99620	.28758
	Intermediate	10	2.5000	.97183	.30732

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Expertise	Equal variances assumed	.024	.879	-1.383	20	.182	-.58333	.42188	-1.46337	.29670
	Equal variances not assumed			-1.386	19.457	.181	-.58333	.42089	-1.46286	.29620

T-test of effort between entry-level and intermediate-level designers with Uncertainty DSS.

Group Statistics

	Expertise	N	Mean	Std. Deviation	Std. Error Mean
Effort	Entry	12	3.4167	1.16450	.33616
	Intermediate	10	2.8000	.91894	.29059

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper

Expertise	Equal variances assumed	1.642	.215	1.357	20	.190	.61667	.45432	-.33102	1.56435
	Equal variances not assumed			1.388	19.960	.181	.61667	.44435	-.31036	1.54369

Perceived information sufficiency

Two way ANOVA test result of perceived information sufficiency with DSS approach (no DSS, Certainty DSS or Uncertainty DSS) as a within subjects variable, and experience-level (entry-level or intermediate level) as a between subjects variable.

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Type_of_DSS	Sphericity Assumed	10.755	2	5.377	3.307	.047
	Greenhouse-Geisser	10.755	1.798	5.981	3.307	.053
	Huynh-Feldt	10.755	2.000	5.377	3.307	.047
	Lower-bound	10.755	1.000	10.755	3.307	.084
Type_of_DSS * Category_experience	Sphericity Assumed	1.664	2	.832	.512	.603
	Greenhouse-Geisser	1.664	1.798	.925	.512	.585
	Huynh-Feldt	1.664	2.000	.832	.512	.603
	Lower-bound	1.664	1.000	1.664	.512	.483
Error(Type_of_DSS)	Sphericity Assumed	65.033	40	1.626		
	Greenhouse-Geisser	65.033	35.963	1.808		
	Huynh-Feldt	65.033	40.000	1.626		
	Lower-bound	65.033	20.000	3.252		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1159.205	1	1159.205	455.334	.000
Category_experience	.114	1	.114	.045	.835
Error	50.917	20	2.546		

Chi square test of independence of overlapping with type of DSS

System * Overlap Crosstabulation

Count

		Overlap		
		1.00	2.00	Total
System	Control	12	10	22
	Certainty DSS	11	11	22
	Uncertainty DSS	14	8	22
Total		37	29	66

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.861[a]	2	.650
Likelihood Ratio	.867	2	.648
Linear-by-Linear Association	.363	1	.547
N of Valid Cases	66		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.67.

Chi square test of independence of overlapping with expertise

Expert * Overlap Crosstabulation

Count

		Overlap		
		1.00	2.00	Total
Expert	Entry level	22	14	36

Intermediate level	15	15	30
Total	37	29	66

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.820[a]	1	.365		
Continuity Correction[b]	.431	1	.511		
Likelihood Ratio	.821	1	.365		
Fisher's Exact Test				.457	.256
Linear-by-Linear Association	.808	1	.369		
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.18.

b. Computed only for a 2x2 table

T-test of perceived information sufficiency between overlap and no-overlap cases when using Certainty DSS.

Group Statistics

	"Overlap"	N	Mean	Std. Deviation	Std. Error Mean
Perceived Information Sufficiency	Overlap	11	4.8182	1.47093	.44350
	No-overlap	11	4.3636	1.12006	.33771

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
"Overlap" Equal variances assumed	3.367	.081	.815	20	.424	.45455	.55744	-.70826	1.61735

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
"Overlap"	Equal variances assumed	3.367	.081	.815	20	.424	.45455	.55744	-.70826	1.61735
	Equal variances not assumed			.815	18.679	.425	.45455	.55744	-.71356	1.62265

T-test of perceived information sufficiency between overlap and no-overlap cases when using Uncertainty DSS.

Group Statistics

	"Overlap"	N	Mean	Std. Deviation	Std. Error Mean
Perceived Information Sufficiency	Overlap	14	3.0714	1.26881	.33910
	No-overlap	8	4.6250	1.06066	.37500

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper

"Overlap"	Equal variances assumed	.055	.816	-2.921	20	.008	-1.55357	.53188	-2.66305	-.44410
	Equal variances not assumed			-3.073	17.006	.007	-1.55357	.50559	-2.62024	-.48691

T-test of perceived information sufficiency between overlap and no-overlap cases when there is no DSS.

Group Statistics

	"Overlap"	N	Mean	Std. Deviation	Std. Error Mean
Perceived Information Sufficiency	Overlap	12	4.4167	1.62135	.46804
	No-overlap	10	4.4000	1.17379	.37118

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
"Overlap"	Equal variances assumed	3.069	.095	.027	20	.979	.01667	.61542	-1.26707	1.30040
	Equal variances not assumed			.028	19.675	.978	.01667	.59736	-1.23073	1.26406

Decision confidence

Two way ANOVA test result of decision confidence with DSS approach (no DSS, Certainty DSS or Uncertainty DSS) as a within subjects variable, and experience-level (entry-level or intermediate level) as a between subjects variable.

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Type_of_DSS	Sphericity Assumed	9.165	2	4.582	4.158	.023
	Greenhouse-Geisser	9.165	1.749	5.240	4.158	.028
	Huynh-Feldt	9.165	1.999	4.585	4.158	.023
	Lower-bound	9.165	1.000	9.165	4.158	.055
Type_of_DSS * Category_experience	Sphericity Assumed	.619	2	.310	.281	.757
	Greenhouse-Geisser	.619	1.749	.354	.281	.727
	Huynh-Feldt	.619	1.999	.310	.281	.756
	Lower-bound	.619	1.000	.619	.281	.602
Error(Type_of_DSS)	Sphericity Assumed	44.078	40	1.102		
	Greenhouse-Geisser	44.078	34.981	1.260		
	Huynh-Feldt	44.078	39.976	1.103		
	Lower-bound	44.078	20.000	2.204		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1070.675	1	1070.675	457.011	.000
Category_experience	7.766	1	7.766	3.315	.084
Error	46.856	20	2.343		

T-test of decision confidence between overlap and no-overlap cases when there is no DSS.

Group Statistics

	"Overlap"	N	Mean	Std. Deviation	Std. Error Mean
Decision confidence	Overlap	12	4.1667	1.26730	.36584
	No-overlap	10	4.6000	.84327	.26667

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
"Overlap" Equal variances assumed	2.534	.127	-.923	20	.367	-.43333	.46969	-1.41309	.54643
Equal variances not assumed			-.957	19.177	.350	-.43333	.45271	-1.38028	.51361

T-test of decision confidence between overlap and no-overlap cases when using Uncertainty DSS.

Group Statistics

	"Overlap"	N	Mean	Std. Deviation	Std. Error Mean
Decision confidence	Overlap	14	3.1429	1.23146	.32912
	No-overlap	8	4.2500	1.48805	.52610

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
"Overlap" Equal variances assumed	1.516	.232	-1.883	20	.074	-1.10714	.58809	-2.33389	.11960
Equal variances not assumed			-1.784	12.519	.099	-1.10714	.62057	-2.45306	.23877

T-test of decision confidence between overlap and no-overlap cases when using Certainty DSS.

Group Statistics

	"Overlap"	N	Mean	Std. Deviation	Std. Error Mean
Decision confidence	Overlap	11	4.3636	1.12006	.33771
	No-overlap	11	4.2727	1.42063	.42834

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Equal variances assumed	.187	.670	.167	20	.869	.09091	.54545	-1.04689	1.22871
“Overlap” Equal variances not assumed			.167	18.967	.869	.09091	.54545	-1.05087	1.23269

Plans to seek additional information

Test for Difference in Magnitude between Two Independent Correlations

Often, it is necessary to determine whether one correlation is significantly different from another. To peform this test, one has to ascertain a Fisher z' transformation of the correlation. Each of the two correlations needs to be transformed as
$z' = 1/2\ [\ \ln(1 + r) - \ln(1 - r)]$
From these transformations, z1' and z2' are obtained. ln is the natural logarithm. That is the logarithm to the base e. Then it is necessary to compute the standard error for the Fisher's z transformation.

$$std\ err_{z'} = \frac{1}{\sqrt{n-3}}$$

Often the difference is computed between different sized random samples. The differrence between the two transformed correlations is divided by the standard error to yield a normal curve deviate.

$$z = \frac{z_1' - z_2'}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

If this is greater than 1.96, then the difference between the correlations is significant at the .05 level.

Results of our study

Table 3: How effective were participants at planning to search for information that would support the design decision?

	Average correlation between: · **expressed interest in a criterion, and** · **contribution to overall uncertainty**	
	r	**N**
Control Condition	0.292	65
Certainty DSS	0.239	63
Uncertainty DSS	0.475	79

$Z_{\text{control condition}}$=1/2 [ln(1 + 0.292) - ln (1 -0.292)]=0.301
$Z_{\text{certainty dss}}$=1/2 [ln(1 + 0.239) - ln (1 -0. 239)]=0.243
$Z_{\text{uncertainty dss}}$=1/2 [ln(1 + 0.475) - ln (1 -0. 475)]=0.517

$$z = \frac{z_1' - z_2'}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

The comparison between control condition and certainty dss:
Z=(0.301-0.243)/sqrt(1/(65-3)+1/(63-3))=0.315
P=0.376
Statistics (z=0.315, p=0.376) – No significant difference
The comparison between control condition and Uncertainty dss:
Z=(0. 517-0. 301)/sqrt(1/(79-3)+1/(65-3))=1.261
P=0.104
Statistics (z=1.261, p=0.104) – marginal difference
The comparison between Certainty DSs and Uncertainty dss:
Z=(0. 517-0.243)/sqrt(1/(79-3)+1/(63-3))=1.580
P=0.057
Statistics (z=1.580, p=0.057) – significant difference

Participants' follow through actions one week later

Count of numbers of people

		Expertise		
		Entry-level	Intermediate-level	Total
Carried thought expressed plan?	Yes	5	8	13
	No	7	2	9
Total		12	10	22

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.316[a]	1	.069		
Continuity Correction[b]	1.920	1	.166		
Likelihood Ratio	3.458	1	.063		
Fisher's Exact Test				.099	.082
Linear-by-Linear Association	3.165	1	.075		
N of Valid Cases	22				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 4.09.

b. Computed only for a 2x2 table

The $\chi^2(1)=0.069$ demonstrates a close-to-significant effect that intermediate-level designers had a higher chance (80% (8 out of 10)) to followed through the expressed plan than entry-level designers (only 42% (5 out of 12))

The analysis of experiment two

Perceived information sufficiency

One-way ANOVA on perceived information sufficiency between groups who use "certainty" DSS and "uncertainty" DSS

ANOVA

Information_sufficiency

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	8.995	1	8.995	58.602	.000
Within Groups	4.758	31	.153		
Total	13.754	32			

One-way ANOVA on perceived information sufficiency when using "certainty" DSS between experiment one and experiment two

ANOVA

Perceived information sufficiency

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.193	1	5.193	4.592	.043
Within Groups	26.011	23	1.131		
Total	31.205	24			

Workload and DSS Preference

One-way ANOVA on workload between groups who use “certainty” DSS and “uncertainty” DSS

ANOVA

Workload

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.682	1	5.682	10.312	.003
Within Groups	17.082	31	.551		
Total	22.764	32			

One-way ANOVA on sub-dimensions within workload measurement between groups who use “certainty” DSS and “uncertainty” DSS

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Workload_Mental	Between Groups	39.966	1	39.966	7.251	.011
	Within Groups	170.873	31	5.512		
	Total	210.839	32			
Workload_Physical	Between Groups	11.549	1	11.549	3.485	.071
	Within Groups	102.730	31	3.314		
	Total	114.279	32			
Workload_Time	Between Groups	49.459	1	49.459	14.430	.001
	Within Groups	106.255	31	3.428		
	Total	155.713	32			
Workload_Performance	Between Groups	3.794	1	3.794	1.264	.270
	Within Groups	93.081	31	3.003		
	Total	96.875	32			
Workload_Effort	Between Groups	.604	1	.604	.103	.751
	Within Groups	182.077	31	5.873		
	Total	182.681	32			

Workload_Frustration	Between Groups	10.185	1	10.185	2.899	.099
	Within Groups	108.916	31	3.513		
	Total	119.101	32			

One-way ANOVA on preference between groups who use “certainty” DSS and “uncertainty” DSS

ANOVA

Preference

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.741	1	.741	4.829	.036
Within Groups	4.758	31	.153		
Total	5.499	32			

Usability

One-way ANOVA on usability between groups who use "certainty" DSS and "uncertainty" DSS

ANOVA

Usability

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.047	1	.047	.545	.466
Within Groups	2.648	31	.085		
Total	2.694	32			

Situation Awareness

One-way ANOVA on overall situation awareness between groups who use "certainty" DSS and "uncertainty" DSS

ANOVA

Situation awareness

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.011	1	.011	1.582	.218
Within Groups	.211	31	.007		
Total	.221	32			

Two-way ANOVA on three levels of situation awareness and the use of different DSS

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
SA_Levels	Sphericity Assumed	.647	2	.324	25.357	.000
	Greenhouse-Geisser	.647	1.860	.348	25.357	.000
	Huynh-Feldt	.647	2.000	.324	25.357	.000
	Lower-bound	.647	1.000	.647	25.357	.000
SA_Levels * C_U	Sphericity Assumed	.118	2	.059	4.614	.014
	Greenhouse-Geisser	.118	1.860	.063	4.614	.016
	Huynh-Feldt	.118	2.000	.059	4.614	.014
	Lower-bound	.118	1.000	.118	4.614	.040
Error(SA_Levels)	Sphericity Assumed	.791	62	.013		
	Greenhouse-Geisser	.791	57.662	.014		
	Huynh-Feldt	.791	62.000	.013		
	Lower-bound	.791	31.000	.026		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	57.268	1	57.268	2808.265	.000
C_U	.032	1	.032	1.582	.218
Error	.632	31	.020		

Estimates

Measure:MEASURE_1

SA_Levels	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Level 1	.885	.022	.840	.930
Level 2	.717	.022	.673	.761
Level 3	.706	.022	.662	.751

Pairwise Comparisons

Measure:MEASURE_1

(I) SA_Levels	(J) SA_Levels	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
Level 1	Level 2	.168*	.031	.000	.090	.246
	Level 3	.178*	.024	.000	.117	.239
Level 2	Level 1	-.168*	.031	.000	-.246	-.090
	Level 3	.010	.029	1.000	-.063	.084
Level 3	Level 1	-.178*	.024	.000	-.239	-.117
	Level 2	-.010	.029	1.000	-.084	.063

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

One-way ANOVA on three levels of situation awareness between groups who use "certainty" DSS and "uncertainty" DSS

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
SA1	Between Groups	.000	1	.000	.001	.973
	Within Groups	.482	31	.016		
	Total	.482	32			
SA2	Between Groups	.004	1	.004	.287	.596
	Within Groups	.473	31	.015		
	Total	.477	32			
SA3	Between Groups	.146	1	.146	9.650	.004
	Within Groups	.468	31	.015		
	Total	.613	32			

Group communication

One-way ANOVA on the number of utterance between groups who use “certainty” DSS and “uncertainty” DSS

ANOVA

VAR00002

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7212.969	1	7212.969	3.257	.085
Within Groups	46502.857	21	2214.422		
Total	53715.826	22			