

Association for Information Systems AIS Electronic Library (AISeL)

All Sprouts Content

Sprouts

7-1-2008

The Impact of DSS Use and Information Load on Errors and Decision Quality

Michael Williams

Indiana University, miwillia@indiana.edu

Alan R. Dennis

Indiana University, ardennis@indiana.edu

Antonie Stam

University of Missouri, stama@missouri.edu

Jay Aronson

University of Georgia, jaronson@terry.uga.edu

Follow this and additional works at: http://aisel.aisnet.org/sprouts_all

Recommended Citation

Williams, Michael; Dennis, Alan R.; Stam, Antonie; and Aronson, Jay, "The Impact of DSS Use and Information Load on Errors and Decision Quality" (2008). *All Sprouts Content*. 78.

http://aisel.aisnet.org/sprouts_all/78

This material is brought to you by the Sprouts at AIS Electronic Library (AISeL). It has been accepted for inclusion in All Sprouts Content by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

The Impact of DSS Use and Information Load on Errors and Decision Quality

Michael Williams
Indiana University, USA

Alan R. Dennis
Indiana University, USA

Antonie Stam
University of Missouri, USA

Jay Aronson
University of Georgia, USA

Abstract

This paper uses a laboratory experiment to examine the effect of DSS use on the decision maker's error patterns and decision quality. The DSS used in our experiments is the widely used Expert Choice (EC) implementation of the Analytic Hierarchy Process. Perhaps surprisingly, our experiments do not provide general support for the often tacit assumption that the use of a DSS such as EC improves decision quality. Rather, we find that, whereas a DSS can help decision makers develop a better understanding of the essence of a decision problem and can reduce logical errors (especially if the information load is high), it is also susceptible to introducing accidental effects such as mechanical errors. In some cases, as in our study, the accidental errors may outweigh the benefits of using a DSS, leading to lower quality decisions.

Keywords: Decision Support Systems; Multicriteria Decision Making; Analytic Hierarchy Process; Decision Quality

Permanent URL: <http://sprouts.aisnet.org/4-22>

Copyright: [Creative Commons Attribution-Noncommercial-No Derivative Works License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Reference: Williams, M., Dennis, A.R., Stam, A., Aronson, A. (2004). "The Impact of DSS Use and Information Load on Errors and Decision Quality," Indiana University, USA . *Sprouts: Working Papers on Information Systems*, 4(22). <http://sprouts.aisnet.org/4-22>

The Impact of DSS Use and Information Load on Errors and Decision Quality

Michael L. Williams
Indiana University (miwillia@indiana.edu)

Alan R. Dennis
Indiana University (ardennis@indiana.edu)

Antonie Stam
University of Missouri (stama@missouri.edu)*

Jay E. Aronson
University of Georgia (jaronson@terry.uga.edu)

December 17, 2004

*: Corresponding author, mailing address: Department of Management, College of Business, 418 Cornell Hall, University of Missouri, Columbia, MO 65211, USA

The Impact of DSS Use and Information Load on Errors and Decision Quality

Abstract

This paper uses a laboratory experiment to examine the effect of DSS use on the decision maker's error patterns and decision quality. The DSS used in our experiments is the widely used Expert Choice (EC) implementation of the Analytic Hierarchy Process. Perhaps surprisingly, our experiments do not provide general support for the often tacit assumption that the use of a DSS such as EC improves decision quality. Rather, we find that, whereas a DSS can help decision makers develop a better understanding of the essence of a decision problem and can reduce logical errors (especially if the information load is high), it is also susceptible to introducing accidental effects such as mechanical errors. In some cases, as in our study, the accidental errors may outweigh the benefits of using a DSS, leading to lower quality decisions.

Keywords: Decision Support Systems; Multicriteria Decision Making; Analytic Hierarchy Process; Decision Quality

The Impact of DSS Use and Information Load on Errors and Decision Quality

1 Introduction

Decision Support Systems (DSS) are designed to assist a decision maker (DM) with systematically analyzing complex semi-structured decision problems (Gorry and Scott Morton 1971; Power and Meyeraan 1994). By integrating complex mathematical models into user-friendly software a DSS alleviates cognitive limitations that may restrict and bound a DM's ability to make rational decisions (Todd and Benbasat 1999). Presumably, the DSS helps the DM organize his/her thoughts and structure the decision problem, thus facilitating a better understanding of the problem at hand. This, in turn, it is often assumed, should result in improved, better quality decisions.

However, DSS use does not always improve decision quality. One reason for the equivocal performance of DSS is the pervasiveness of human error. Despite the robust algorithms and user-friendly interfaces, DSS are still implemented by humans who are prone to a variety of errors. There is no generally accepted classification of human error, nor is there likely to be one, because no single scheme is likely to satisfy all needs (Reason 1990). In this paper, we adopt the generic error-modeling system (GEMS) (Reason 1990) as the general-purpose framework for understanding errors and their causes. According to GEMS there are two broad categories of errors: *knowledge-based* errors and *skill-based* errors (a more comprehensive description of these error types follows in a later section).

While many studies have paid attention to the mathematical and algorithmic properties of specific DSS tools, there is much less systematic research addressing the effect of DSS use on errors, and few studies have evaluated how errors made during the DSS decision process affect the quality of the final decision. One reason may be that an objective benchmark of decision quality is often not available once the DSS is put to use in practice. Instead, researchers have sometimes resorted to surrogate quality measures based on DM perceptions, focusing on how confident or satisfied the DM is about the solutions suggested by the DSS. Other studies have relied on simulation studies to show the potential decision quality of specific DSS tools, but of course simulations do not take errors made by actual DM in the decision process into account. For example, Lerch and Harter (2001) investigate the effectiveness of several aspects of sequential, real-time decision making tools in a simulation study. The tasks involved assigning

mail to sorting machines. Their findings show that certain kinds of cognitive support degrade performance while simpler approaches based on ‘automatic’ implementation of planned responses to situations were generally considered superior to methods that consider all factors. They found that more complex tools that provided a wealth of information tended to overload the decision maker.

The current study seeks to systematically analyze the types of errors made by DSS users vs. non-DSS users when solving a semi-structured decision task. In this study we seek to identify the reasons why DSS users may not make better quality decisions than non-DSS users. In our study, we conduct an experiment to assess the impact of using a specific DSS tool, the well-regarded and widely used Expert Choice (EC) software implementation of the Analytic Hierarchy Process (AHP) (Saaty 1980), on errors and decision quality, for two levels of information load (low and high).

The remainder of our paper is organized as follows. In the next section we develop a theoretical perspective for examining errors in DSS use. In particular, we examine the role of decision strategies, the influence of DSS on decision strategy selection, and several types of errors that play a role when using computer-based systems. Additionally, we explain our motivation for using the EC software in our experiment, and explain how EC may induce certain skill-based errors that affect decision quality. This section is followed by a presentation of our methodology and the results from our experiment. The paper concludes with a discussion of the results and suggested avenues of future research in this area.

2 Theory Development

2.1 Decision Making Strategies

We chose to analyze a discrete alternative, multi criteria decision problem. Because such problems tend to be relatively well-structured, DSS tools can play an important role in their analysis (Cats-Baril and Huber 1987; Power and Meyeraan 1994). In this paper, we will refer to this class of problems as discrete alternative multi criteria decision analysis (DAMCDA) problems. The decision task in DAMCDA problems typically revolves around selecting the most preferred (“best”) alternative from a set of candidate alternatives. The selection process involves an evaluation of the alternatives with respect to the set of relevant criteria, for which the DM may use one or more of a variety of decision strategies.

Two commonly used classes of decision strategies are Additive Compensatory (AC), and Elimination by Aspects (EBA) (Biggs et al. 1985; Jarvenpaa 1989). With an EBA strategy the DM narrows the set of alternatives by successively eliminating alternatives that fail to reach a minimum threshold level achievement on each criterion. Therefore, one weakness of EBA strategies is that they use only a portion of the range of information that is available. In contrast, an AC strategy requires the DM to utilize the full range of available information. Using an AC strategy, the DM evaluates all alternatives with respect to all relevant criteria, in the process arriving at relative importance ratings for the criteria and subsequently determining measures of relative preference for each alternative. Thus, while EBA strategies may lead to the premature elimination of certain alternatives based on “poor” achievement on a single criterion, AC strategies should lead to more balanced and accurate results (Todd and Benbasat 1999). It is not surprising that most DAMCDA software, including EC, is based on AC and not on EBA strategies.

2.2 The Influence of DSS on Decision Making Strategies

Many DSS attempt to improve decision making performance by changing the decision strategies used by DMs from “poorer” strategies to “better” strategies (Todd and Benbasat 1994a; 1994b, 1999, 2001). While AC decision strategies are superior to EBA strategies, they also require more cognitive effort to employ. The cost-benefit framework of cognition indicates that decision-makers continually make trade-offs between effort and accuracy in selecting decision strategies (Payne et al. 1993). In this ongoing compromise between effort and accuracy, research indicates that decision-makers generally favor decision strategies that require less effort (Payne et al. 1988; Russo and Doshier, 1983; Todd and Benbasat 1999). Todd and Benbasat (1999) argue that the primary criterion for using a decision strategy is ease of use. In a laboratory setting they demonstrate empirically that the key to inducing a particular strategy is by making it easier to execute than competing alternative strategies.

Numerous software products that implement AC strategies exist in the MCDA and MCDM fields. One popular AC strategy which automates much of the decision process is the EC implementation of the AHP. EC aids the DM in structuring and solving the decision problem, first using pairwise comparisons to assess the relative importance of the criteria, then evaluating the relative attractiveness of the alternatives, again through pairwise comparisons, and finally by

integrating the preference information to determine final preference ratings for each alternative. EC identifies the alternative with the highest final rating as the most preferred (“best”) one.

The AHP philosophy seeks to aid the DM in developing additional insights into and a better understanding of the decision problem at hand, thereby reducing knowledge-based errors. We decided to use EC in our experiment, because of its user friendliness, its popularity and success in practice and the close match between its characteristics and the class of problems that we are interested in (selection of one or more alternatives from a pre-defined set of alternatives)¹. Because the EC software completely automates the superior AC decision strategy, DM should be induced to consider the entire problem space. Thus, we hypothesize:

H₁: DMs who use a DSS to solve a DAMCDA problem will have a better understanding of the relative importance of the criteria than those who do not use such a DSS.

2.3 The Role of Errors in Decision Making with DSS

As mentioned previously, we use GEMS as the general-purpose framework for understanding errors in this paper. In GEMS, Reason (1990) argues that there are two broad categories of errors that stem from different cognitive causes. *Knowledge-based errors* are those that occur when an individual resorts to “attentional processing within the conscious workspace” (Reason 1990, p. 57); that is, knowledge-based errors are made when the decision-maker focuses on the task and consciously thinks and reasons about it. Knowledge-based errors would include errors such as assigning too much or too little relative rating to a specific decision criterion, or incorrectly evaluating an alternative on a criterion. These are the types of errors that DSS are primarily designed to resolve. By developing and using a mathematical model in a DSS, a decision maker can overcome many knowledge-based errors.

A second category of error in GEMS is *skill-based or rule-based errors*. These are errors that occur while human behavior is under the control of largely automatic units within the

¹ Although the AHP has been used successfully in many real-life applications (see, e.g., Golden *et al.* 1989; Saaty 1980), its theoretical foundation has been challenged on a number of grounds (Barzilai and Golani 1994; Dyer 1990; Lootsma 1993; Schoner *et al.* 1992, 1993; Stam and Duarte Silva 1994; Winkler 1990). Despite its widespread use and the scrutiny it has received on the mathematical and algorithmic side, questions related to its effectiveness as a DSS tool have not been evaluated systematically in controlled laboratory experiments. In fact, a particular implementation of the AHP may improve or exacerbate algorithmic issues. While the choice of any single DSS may create a systematic bias in our results, this drawback would not be more or less pronounced based on the software product selected. Thus, caution is encouraged when interpreting the results of our experiments.

cognitive knowledge base (Reason 1990). That is, skill-based errors occur when the decision-maker does not devote his or her complete conscious attention to the task at hand, but instead relies on secondary skills or rules to accomplish the task. Skill-based or rule-based errors would include errors made in touch-typing, in copying values from one list to another, or other activities that frequently do not require a high level of cognitive effort. Regardless of the sophistication and artistry of a DSS, when users do not devote adequate attention to the problem, errors may arise. Many DSS are well-designed in terms of addressing the knowledge-based errors common for a decision problem, but—as we will show below in our experiment—fall short when it comes to preventing unintentionally introduced skill-based errors.

While GEMS is an appropriate general-purpose framework for understanding broad categories of errors, Reason (1990) argues that a unique error classification system must be instantiated to properly analyze errors within specific contexts. Thus, several researchers have developed task context-specific error types in order to define and operationalize distinctly different categories of error. In our study, we use an adaptation of Reason's (1990) generic error-modeling system (GEMS) for measuring and understanding errors and what causes them, and Panko and Halverson's (1996) framework for more precisely classifying errors in spreadsheet-based DSS.

Panko and Halverson (1996) use Reason's (1990) general approach, and categorize human errors in the use of spreadsheet-based DSS into three categories: (1) *logic* errors, (2) *mechanical* errors, and (3) *omission* errors.

Logic errors are one type of *knowledge-based errors*. They are higher-level cognitive-based errors (Rasmussen 1986; Reason 1990) in which the DM makes a mistake while actively, consciously, engaged in the task and not under automated control; such errors may result from a misinterpretation of the problem situation or from distraction and preoccupation (Reason and Mycielska 1982) and become more frequent as task complexity increases (Reason 1990). Logic errors are the specific types of error that DSS are designed to address by integrating complex mathematical models that automate a decision rule for the problem. For example, a logic error occurs when a DM intentionally (but erroneously) enters an incorrect formula into a spreadsheet.

We identified three types of logic error that apply to the EC-based DSS: *criteria* error, *rank inversion* error and *inconsistency* error. *Criteria errors* occur when the DM misjudges the

importance of a criterion – either by interpreting an important criterion as unimportant, or the other way around.

Rank inversion errors (Stam and Duarte Silva 1997; Troutt 1988) are human errors that occur when the DM mistakenly ranks a less attractive alternative higher than a more attractive alternative, when the two are very close in value. Note that in this paper we do *not* address the issue of “rank reversal,” a potential problem with the AHP (and EC) due to its underlying mathematical calculations (Barzilai and Golani 1994; Dyer 1990; Schoner *et al.* 1992, 1993).

Inconsistency errors are a well-known concern with the AHP. Saaty (1980) posits that a moderate amount of inconsistency in the AHP is inherent to the nature of human decision making, as each pairwise comparison judgment will be approximate rather than exact. However, high levels of inconsistency are indicative of problems with the DM’s pairwise preference judgments. An extreme example of inconsistency occurs if preference transitivity fails to hold – for instance, when in the pairwise comparisons alternative A is preferred to B, B is preferred to C and C is preferred to A, all *ceteris paribus* and with respect to the same criterion. Saaty (1980) developed an index to measure the extent of inconsistency in the AHP, higher values indicating more inconsistency. The DM is said to have made an inconsistency error if the inconsistency index exceeds a pre-specified threshold value (.10)².

Because a DSS is designed to structure a DAMCDA problem to reduce logical errors of DM we hypothesize:

H₂: DMs who use a DSS to solve a DAMCDA problem will make fewer logic errors than those who do not use such a DSS.

Mechanical errors are *skill-based errors* that occur in routine activities that are often performed in an automated manner without conscious control (Rasmussen 1986; Reason 1990). In other words, in mechanical error the intended action is correct, but the execution is faulty; they can usually be attributed to a lack of attention at key junctures, interruptions, or perceptual confusion (Reason 1990) and are sometimes called “slips” or “lapses” (Norman 1981; Reason 1984). Mechanical errors include typing errors, pointing errors and other simple mechanical slips

² The AHP literature suggests that an inconsistency index below 0.10 is an acceptable level of consistency for AHP to produce a correct ranking of alternatives (see Saaty, 1982; Saaty and Vargas 1994). In general, an inconsistency index exceeding 0.2 suggests an inversion error (e.g., B is preferred to A is entered into the software, but A is really preferred to B by the DM) or a set of major inconsistent judgments. Values between 0.10 and 0.20 may indicate that a judgment was entered incorrectly, or that the DM really does have a high level of inconsistency.

(Panko and Halverson 1996). For example, a mechanical error occurs when a DM accidentally types a wrong value into a spreadsheet when copying from a printed page.

Previous research indicates that DMs using a DSS may require more time to complete a decision problem than DMs who do not use a DSS (Power and Meyeraan 1994). This is because EC requires the DM to conduct a complete systematic series of pairwise comparisons. Because EC requires more time and potentially more calculations in the form of pairwise comparisons from DMs we hypothesize:

H₃: DMs who use a DSS tool like EC to solve a DAMCDA problem will make more mechanical errors than those who do not use such a DSS.

Omission errors are another type of *knowledge-based errors* – these errors result from the DM leaving crucial elements out of the decision process (Panko and Halverson 1996). Omission errors are more likely when the DM skips from one aspect of the problem to another without considering any aspect in-depth, or focuses on one narrow aspect of the problem to the detriment of others (Doerner 1987). For example, an omission error occurs when a DM accidentally omits a variable from an equation where it is required.

Because EC breaks a potentially complex decision into a series of pairwise comparisons we believe there are fewer opportunities for omission errors. Thus, we hypothesize:

H₄: DMs who use a DSS to solve a DAMCDA problem will make fewer omission errors than those who do not use such a DSS.

2.4 The Influence of DSS on Decision Quality

Just about any decision situation has the potential for logic, mechanical, and omission errors. One justification for using a DSS in decision making is to reduce the occurrence of these kinds of errors. DSS tools like EC employ highly structured decision processes and are designed to address the essence of the decision task. These DSS tools derive much of their value from influencing the decision strategies adopted by the DM (Todd and Benbasat 1994a; 1994b, 1999). DSS attempt to reduce logic and omission errors by leading the user through the decision process in a simple, step-by-step, automated manner that may require relatively little cognitive effort. Again, this points to the fact that ease of use is a primary factor in a DM's selection of a decision strategy (see Todd and Benbasat 1994a, 1994b, 1999, 2001, who demonstrate this empirically). However, at the same time the use of these DSS tools may inadvertently introduce unintended accidental effects, such as increasing the number of mechanical errors. The potential impact of

these kinds of errors in the DSS context is not to be ignored, as research on error in spreadsheet models reveals that 40 to 60 percent of spreadsheets contain errors (Panko and Halverson 1995).

Despite the potential introduction of mechanical errors, we expect the use of EC should lead to improved decisional performance for several reasons. First, EC automates the DM's use of the superior AC decision strategy. Todd and Benbasat (1994a, 1994b, 1999) demonstrate empirically that DMs use DSS in a way that minimizes the overall amount of effort they have to expend on the problem. In our context, EC works as a process enabler that automates the superior AC strategy and decreases the cognitive effort required on the part of the DM. Second, EC structures the decision process and requires the DM to engage in a comprehensive and systematic analysis of the decision problem with explicit attention devoted to each criterion and alternative alike. Third, EC provides a logical system for comparing quantitative as well as qualitative measures to rank and rate the alternatives (Forman and Selly 2001; Saaty 1982). Each of these benefits of EC is expected to improve decisional performance of a DM.

Overall, we anticipate the benefits of using a DSS to outweigh the potential costs of DSS use in decision making. Thus, we hypothesize:

H₅: DMs who use a DSS to solve a DAMCDA problem will make better quality decisions than those who do not use such a DSS.

2.5 Information Load

DMs implement DSS to facilitate consistent and correct decisions. As hypothesized above, we anticipate DSS will have beneficial affects on decision quality and will influence the types of errors made by a DM. However, as the nature of a task changes (e.g., in information load, effort required, incentives), the influence of a DSS may also change. In this study, we specifically examine the influence of information load on decision quality and errors of that DMs make when using the EC DSS.

Information load is the amount of input that an information processor receives within a given period of time (Grise and Gallupe 2000; Miller 1960). Information load has been shown to impact decision performance by stimulating the DM to alter his/her decision making strategy (Newell and Simon 1972). Jacoby *et al.* (1974) demonstrate that DMs tend to make worse decisions, yet are more satisfied, more certain and less confused about their decisions, as the information load increases. Grise and Gallupe (2000) note that DMs may use process enablers to facilitate effective decisions when faced with high information load. A process enabler is an

external processor that seeks to increase the DM's processing ability by providing cognitive process support or by reducing the information load, for instance by filtering, organizing or sorting alternatives with respect to selected criteria. With EC, the problem "chunking" decomposes potentially complex problems into smaller, easier to handle cognitive comparisons (Saaty 1982), enabling the DM to narrowly focus on one comparison at the time, instead of multiple simultaneous comparisons. As a result, the relative value of using EC should become more pronounced as the overall information load increases. Consequently, we expect information load to act as a multiplier for all of our hypotheses. Thus, we hypothesize:

H₆: As information load increases, the performance by users of a DSS in comparison with non-DSS users will improve with respect to H_{6a}: the relative understanding of the criteria; H_{6b}: the number of logic errors; H_{6c}: the number of mechanical errors; H_{6d}: the number of omission errors; and H_{6e}: decision quality.

3 Experiment

We used a 2x2 laboratory experiment, varying the use or non-use of the DSS (EC) with two tasks, one with a lower information load and the other with higher information load. Sixty-four undergraduate business students drawn from a core business course participated in the study. Their average age was 21.6 years and 52% were female. Sixteen subjects were randomly assigned to each of the four treatments.

3.1 The Task and Procedures

We chose to use a modified version of the task used by Dennis (1996) to select one student from a set of four candidates for admission to the university. This task is appropriate for the use of the AHP and EC; in fact, Saaty and Vargas (1994) apply the AHP to a similar setting, the selection of students for an academic program. The original task (Dennis 1996) was revised with the help of the university Admissions Office and validated by three different admissions officers – the director of admissions and two associate directors – to ensure that one alternative (*i.e.*, one particular student) was clearly better qualified than the others for admission. Therefore, we knew the correct decision in advance, providing us with a benchmark against which to measure each subject's decision performance.³

³ Research in multi criteria decision making is typified by the fact that no absolutely correct decision exists beyond that which is most satisfactory to the decision maker. That is, correctness of the decision typically is a subjective matter rather than an objective one. However, for the purpose of our study we designed the

Upon arrival at the experimental lab, each subject was given the task and asked to make a decision. No time limits were imposed on completion of the decision task. Consistent with previous research (*e.g.*, Power and Meyeraan 1994), subjects using the DSS took longer to complete the task, an average of 35 minutes versus 7 minutes. Once the task was completed, each subject filled out a short questionnaire, asking them to record their decision and allocate the relative preference points to the criteria, after which they were debriefed and released.

3.2 Treatments

As mentioned above, the DSS treatment was implemented by subjects either using or not using EC. As part of a DSS course prior to and separate from the experimental treatment, all subjects (regardless of whether they subsequently used EC in the experiment itself) had a one-hour in-class training session on decision making, the AHP and EC. During this classroom training session all subjects had to complete an assignment similar to the experimental task using EC.

Upon arriving at the experiment setting, subjects in the non-DSS treatment were given pencil and paper and were asked to make a decision without instructions on how to do so. Subjects in the DSS treatment were stationed at a PC with EC running and asked to make a decision using EC.

The range of the two information load conditions was selected with two constraints in mind. First, in conformity with Saaty's suggestion that "elements should be clustered into homogeneous groups of five to nine so they can be meaningfully compared to elements in the next higher level" (Saaty 1982, p.36) we limited our range from five to nine. Second, in order to adequately test our hypothesis that a DSS would improve a DM's ability to identify the importance of each attribute, we chose a minimum of five unimportant or distracting attributes. Thus, we ultimately limited our range of attributes from seven to nine. The lower information load version of the task provided information on seven criteria for each of the four alternatives.

decision problem such that there exists a single "correct" alternative that any decision maker should select, regardless of the relative preferences for the relevant criteria. The correct alternative almost dominates the other alternatives – it is the top alternative for all but one of the relevant criteria, and a close second on the criterion for which it is not the best. The fact that the three different experts (admissions officers) agreed upon the one "correct" alternative provides a legitimate benchmark and validating the single correct alternative. Furthermore, a sensitivity analysis revealed that the correct alternative was selected (ranked first) for any criterion weight distribution in which the relevant criteria received at least 80% of the total weight, as long as the pairwise comparisons of the alternatives with respect to the relevant criteria were reasonable. Hence, in the experiment we conclude that some kind of error occurred when a decision maker did not select the correct alternative.

Of these criteria, two (SAT and GPA) were important, while the other five (applicant's gender, planned major, level of commitment to major, level of commitment to the university, and author of recommendation letter) were distracters that are required on the university's admission form, but play no role in the admission decision. The higher information load version of the task provided information on nine criteria for each of the four alternatives. The same two criteria as above were important, and the remaining seven (the above five plus parents' education and extracurricular activities) were distracters.

3.3 Measures

Two measures of interest were collected from each subject in the study: the relative importance of the criteria and decision quality. A third measure, type of error, was calculated based on the results collected from each subject.

Each subject in the DSS treatments reported the relative importance of the criteria, calculated automatically by EC, on a post-session questionnaire. The extent to which a subject correctly identified the two important criteria (SAT and GPA) was measured by the sum of the EC relative importance ratings of these criteria, expressed as a percentage (*i.e.*, out of 100 points). Subjects in the non-DSS treatments were asked to allocate a total of 100 points across the individual criteria on the post-session questionnaire. The extent to which the two important criteria were identified was again determined by summing the scores of the two important criteria (SAT and GPA). One subject failed to report this measure.

Decision quality for each subject was measured by a 0-1 binary variable, which equaled one if the subject selected the correct alternative, and zero otherwise. Thus our mean measure of decision quality is the proportion of subjects in each treatment making the correct decision.

The third measure, error type, was calculated from the responses we collected from the subjects. When subjects failed to select the correct alternative, their error was coded as a logic error (criteria, rank inversion or inconsistency error), omission error, or as a mechanical error. We used the proportion of subjects in each treatment making each type of error as our measure of error commission.

We identified three types of logic errors: criteria errors, rank inversion errors, and inconsistency errors. Criteria errors, *i.e.*, the error of misidentifying important and/or unimportant criteria, were defined to have occurred when a subject assigned only 80% or less of the criteria

points to the two important criteria (either based on the EC report or the post-session questionnaire). We picked the cut-off of 80% because it corresponds roughly to a B- grade. As this cut-off may be viewed as somewhat arbitrary; we conducted a sensitivity analysis to check the impact of various different cut-offs between 65% and 95%. All cut-offs in this range yielded very similar results - *e.g.*, lowering the cut-off to 65% (a D- grade) would have changed three criteria errors into mechanical errors and raising it to 95% (an A grade) would have changed three mechanical errors into criteria errors.

Rank inversion was defined to have occurred when the subject (1) chose an incorrect alternative but (2) successfully identified the important criteria (with at least 80% of the criteria points assigned to the two important criteria), and (3) the differences in scores between the correct alternative and the chosen one was at most 2 percentage points.

Following Saaty's (1982) guidelines, we defined an inconsistency error to have occurred when an EC user had an inconsistency index of at least 0.10. While the EC report includes the inconsistency index, no comparable value can be imputed for non-EC users and for these subjects this variable is not included. Although this admittedly introduces a systematic bias in our coding of errors, we believe that the alternatives of ignoring the inconsistency error data provided by EC altogether, or categorizing inconsistency error as a mechanical error, are both theoretically inappropriate. Our approach will tend to inflate the number of errors committed by non-EC users that are categorized as mechanical errors, because some of the errors we would categorize as mechanical would actually be unrecognized inconsistency errors. Thus, our error coding process will overstate mechanical errors for non-EC users in a way that is counter to our hypothesis that DSS users will make more mechanical errors; confirming hypothesis H₃ will therefore be more difficult than it should.

Omission errors were defined to have occurred when important criteria were omitted from the criteria analysis. That is, when a subject chose the wrong alternative and either left an important criteria blank or assigned a zero value to it.. They were coded as a 0-1 binary variable which equaled one if omission occurred and zero otherwise.

Any incorrect decision not categorized as a criteria, rank inversion, inconsistency or omission error, was coded as a mechanical error.

4 Results

The analysis involved a 2 x 2 ANOVA with EC use and information load as the two factors. Table 1 shows the means and standard deviations of the dependent measures. Since the dependent measure for whether a subject made the correct decision was a 0-1 indicator variable, the mean score can be interpreted as the proportion of subjects who selected the correct decision. Examining the error patterns of the EC users in detail, we found that approximately 35% made errors serious enough to affect their performance. We view this as high, but it is comparable to error rates found in spreadsheet research (Panko 1998; Panko and Sprague 1998). There was about an even split between mechanical and logic errors made by EC users. The high rate of mechanical errors is consistent with findings by Panko and Sprague (1998) in spreadsheet research (65% mechanical errors). About 18% of all errors made by EC users were caused by a poor understanding of the criteria, 14% by rank inversion, while 14% were inconsistency errors – the remaining 54% being mechanical errors.

 Tables 1, 2 and 3 About Here

The results of the ANOVA analysis in Table 2 show a strong main effect of EC use in identifying the important criteria ($F_{1,59}=14.409, p<.001$), but no effects due to information load or the interaction term. Therefore, H_1 is supported, but H_{6a} is not. There is also a significant main effect due to information load on logic errors ($F_{1,60}=4.35, p=.041$) and an interaction effect ($F_{1,60}=4.35, p=.041$). EC users committed significantly more mechanical errors ($F_{1,60}=5.60, p=.021$). Therefore, H_3 and H_{6b} are supported, but H_2 is not. No main effects are found for information load or EC use on decision quality, but interestingly the interaction term is significant ($F_{1,60}=4.66, p<.035$), so that H_5 is not supported, but H_{6e} is. A single subject in the non-DSS, high information load condition committed an error of omission refuting H_4 and H_{6d} .

These results reveal an interesting pattern. As expected, EC users show a better understanding of the relative importance of the criteria than non-EC users (H_1 supported). However, EC users did not fare as well at selecting the correct alternative. In the low information load condition only 63% correctly identified the alternative, and only 69% did so in the high information load condition. In contrast, non-EC users far outperformed EC users in the low information load condition (87% correct), but their performance dropped significantly ($p=.008$) in the high information load condition (44% correct). Therefore, while the use of EC reduced logic errors for the high information load task, it also increased the number of mechanical errors for

both the low and high load tasks. Simply stated, EC successfully attacked the essence of the problem, thwarting both logic and omission errors. The importance of this cannot be minimized, as omission errors can be particularly difficult to detect (Panko 1999). The EC was also more helpful, in relative terms, in the higher load task. However, on the negative side the EC also induced users to make more mechanical, or accidental errors. Our results indicate that, by introducing new patterns of action and decision processes that take longer to perform and have more steps, DSS may introduce more opportunities for errors despite integrating robust mathematical models and automating a superior AC decision strategy. We believe that these findings suggest that *unintended consequences* of a DSS may potentially overwhelm the benefits intended by the DSS creator.

5 Discussion

In this study, we set out to empirically test the ability of a DSS, specifically the popular EC implementation of the AHP, to effectively analyze discrete alternative MCDA problems. In the literature, much attention has been paid to the theoretical and mathematical soundness of various different DSS algorithms and software products. However, due to the nature of DSS applications, it is very hard to measure how effective the use of DSS tools is in practice. Questions regarding actual decision quality when using a DSS are difficult to answer, because once a DSS has been used it is hard to track what would have happened had the DSS not been used. In other words, in practice it is difficult to establish a reliable benchmark to measure improvement in decision quality that can be directly attributed to the use of the DSS.

As a result, some researchers have resorted to surrogate measures of improved decision quality, such as user confidence and satisfaction. In our experiment, we deliberately designed the decision task such that the “correct” decision was known. We realize that no experimental design is perfect, and one can argue that aspects of our experiment are subjective, but the correct decision in our experiment was validated unambiguously by three different experts. Therefore, we have built a threshold into our experiment by which decision quality can be measured directly.

Our experiment shows that, whereas a DM may have developed a better understanding of the decision problem and may feel confident and satisfied with the DSS, actual decision quality may not improve at all; in fact it may deteriorate due to human error when using the DSS – although the relative decision quality when using a DSS appears to improve as the information load is higher. The increased confidence and satisfaction of the DM with the decisions suggested

by the DSS may result from the fact that DSS tend to enhance the DM's understanding of the decision problem at hand – this was one of the major findings in our study. However, our findings indicate that the common assumption that decision quality is improved by DSS use is tenuous at best. Further research should address the reasons why more complex tools that provide a wealth of information tend to overload the decision maker (Lerch and Harter 2001), a finding which is only partially supported by the findings in our study.

5.1 Implications for research and practice

The implications of our findings open a new opportunity for operations researchers. We have shown that, in addition to the extensive research on complex mathematical models and algorithms designed to prevent knowledge-based errors, DSS researchers and software designers, need to pay careful attention to anticipating and preventing the unintended opportunities for skill-based errors introduced by DSS use. To be clear, when using a mathematically sound and rigorously tested DSS like EC users are apparently vulnerable to making skill-based errors to an extent that may reduce or negate the overall benefit of the DSS tool. The impact of skill-based errors can be substantial and may “crowd out” the positive effects the DSS has by improving the DM's understanding of the decision problem itself. DSS software designers and operations researchers need to redouble their efforts to develop systems that minimize the pitfalls of lapses in cognitive attention by users, and therefore the occurrence of mechanical errors and other “slips” and “lapses.” Particular attention is called for in designing those steps in the decision process that are routine, repetitive and simplified.

One method of addressing the skills-based errors potentially introduced by DSS use is with “model inspections.” Software designers often implement code inspections as a way of reducing errors in software engineering (Panko 1999; Panko and Sprague 1998). In order to reduce the uncorrected error rate in the use of DSS, we believe that formal “model inspections” should be implemented in which DMs review and validate the models they have built and the rankings or ratings determined in the decision process. We note that EC already has several such “checks,” for instance the sensitivity analysis feature, which graphically illustrates the sensitivity of the current rankings, and the option that allows the DM to review the contribution of each pairwise comparison to the current value of the inconsistency index.

In our experiment, we found an interaction effect of DSS use and information load, suggesting that DSS are more helpful, in terms of actual decision quality, as the information load

(and therefore the complexity of the decision problem) increases. This implies that DM may need to exercise discretion about when a decision problem is well-suited for DSS use. Decision problems with a lower information load may be better handled without introducing the opportunity for skill-based errors potentially associated with DSS. More complex decision problems, on the other hand, should benefit from DSS use by providing the DM with a clearer understanding of the decision problem. This issue needs further exploration in future research.

5.2 Limitations

Clearly, our experiment is only a first step in the process of trying to develop a better understanding of the issues raised, and should be followed up in various different ways. Questions remain as to the generalizability of the results in our experiment to other DMs, decision domains and DSS software tools. This study suffers from the weakness inherent in laboratory research. Our experiment is based on one instantiation of a DSS – the EC implementation of the AHP. Additional experiments involving other DSS software products that address DAMCDA in different ways would be useful, as would different types of decision tasks, other than DAMCDA problems. The subjects were students with only a few hours of experience in using EC and may have been unaccustomed to performing this type of task using EC.

Skill-based errors such as the mechanical errors in our study are most likely to occur when decision makers are performing routine well-understood tasks and are less likely to occur in the performance of tasks new to them (Reason, 1990). Thus it is likely that our subjects, working with a tool on a task for which they had not yet developed well-established routines, would be *less* likely to commit skill-based mechanical errors than DMs who routinely use DSS. The rate of mechanical errors made in organizational settings in which the use of DSS is routine may actually be much higher than it was in our study. The routine nature of many DSS should allow reduction of manual steps, thus reducing the opportunity for mechanical errors.

5.3 Future Research

Perhaps the definition, categorization and measurement of the errors can be further refined in future experiments. This can help improve our understanding of the causes of accidental errors in DSS, and can help address whether the accidental errors result from a faulty software design or the nature of human decision making itself. If we can isolate and better understand the causes of accidental errors, then we may be able to better design DSS to reduce accidental errors. This, in

turn, should improve decision quality, and therefore facilitate the design of DSS that more fully realize both dimensions of their promise--a deeper understanding of the problem at hand as well as achieving better quality decisions.

Error patterns similar to those found in our study have also been observed in research on the use of spreadsheet models. The body of research on spreadsheet errors may hold information that is relevant to and can contribute to our understanding of how we can develop effective DSS.

The role of information load in DSS use is an interesting one as well. The widely held assumption that the use of a DSS will be more fruitful as the information load is higher was confirmed by our experiment, but task complexity and information load can take on many different forms and characteristics. Additional research analyzing the effect of various forms of information load and task complexity on DSS use should help us understand these issues better.

Finally, one aspect not directly studied was time. DMs who used the EC DSS took about five times longer to perform the task than did those working without the DSS. The number of steps in the AC-based strategy implemented by the EC DSS required more time than did the presumably simpler decision strategies adopted by those working without the DSS, which is typical for DSS use (*e.g.*, Power and Meyeraan 1994). This increased time may have contributed to the increased number of mechanical errors. It may also have implications for the adoption of DSS, because users often prefer to expend less effort and the increased time and effort required to use DSS may dissuade some users from adopting them.

5.4 Conclusion

Much past research has focused on the mathematical and algorithmic properties of specific DSS models and tools. We believe that this research is important and useful in advancing our understanding of DSS. However, we believe that operations researchers may also find it fruitful to focus on what happens when DSS move into the hands of the decision makers who actually use them in practice. In our study, use of the DSS reduced the amount of logic errors made by decision makers as information load increased. However, DSS use also introduced unexpected mechanical errors; about 15% of the decision makers in our study selected an incorrect alternative due to mechanical errors.

We believe that understanding errors in the use of DSS is a very important issue, an issue that, to date, has not received much attention in DSS theory, DSS software development or DSS application in practice. By understanding how DSS can be intentionally designed to reduce knowledge-based errors and skill-based errors, we have the potential to significantly improve decision making performance in organizations.

Acknowledgements

The authors would like to acknowledge the helpful comments of two anonymous reviewers. Their comments led to a substantial strengthening of the original manuscript.

References

- Barr, S. H. and R. Sharda (1997), "Effectiveness of Decision Support Systems: Development or Reliance Effect?" *Decision Support Systems*, **21** (2), 133-146.
- Barzilai, J. and B. Golani (1994), "AHP Rank Reversal, Normalization and Aggregation Rules," *INFOR*, **3**, 57-63.
- Biggs, S. F., J. Bedard, B. G. Gaber and T. J. Linsmeier (1985), "The Effects of Task Size and Similarity on the Decision Behavior of Bank Loan Officers," *Management Science*, **31**, 970-987.
- Brooks, F. P. (1987), "No Silver Bullets," *Unix Review*, 39-48.
- Cats-Baril, W. L. and G. P. Huber (1987), "Decision Support Systems for Ill-Structured Problems: An Empirical Study," *Decision Sciences*, **18** (3), 683-697.
- Dennis, A. R. (1996), "Information Exchange and Use in Group Decision Making: You Can Lead a Group to Information, But You Can't Make it Think," *MIS Quarterly*, **20** (4), 433-457.
- Doerner, D., "On the Difficulties People Have in Dealing with Complexity," in *New Technology and Human Errors*, J. Rasmussen, K. Dunan and J. Leplat (Eds), Wiley, London, 1987.
- Dyer, J. S. (1990), "Remarks on the Analytic Hierarchy Process," *Management Science*, **36** (3), 249-258.
- Forman, E. H. and M. A. Selly, *Decision and Objectives*, World Scientific Publishing Co., Singapore, 2001.
- Golden, B. L., E. A. Wasil and D. E. Levy, *The Analytic Hierarchy Process: Applications and Studies*, Springer-Verlag, New York, 1989.
- Gorry, G. M. and M. S. Scott Morton (1971), "A Framework for Management Information Systems," *Sloan Management Review*, **13** (1), 55-71.
- Grise, M. and R. B. Gallupe (2000), "Information Overload: Addressing the Productivity Paradox in Face-to-Face Electronic Meetings," *Journal of Management Information Systems*, **16** (3), 157-185.
- Jarvenpaa, S. (1989), "The Effect of Task and Graphical Format Congruence on Information Processing Strategies and Decision Making Performance," *Management Science*, **35**, 285-303.
- Lerch, F. J. and D. E. Harter (2001). "Cognitive Support for Real-Time Dynamic Decision Making," *Information Systems Research*, **12** (1), 63-82.
- Lootsma, F. A. (1993), "Scale Sensitivity in the Multiplicative AHP and SMART," *Journal of Multi-Criteria Decision Analysis*, **2**, 87-110.
- Miller, J. C. (1960), "Information Input Overload and Psychopathology," *Journal of Psychiatry*, February, 696-704.
- Newell, A. and H. A. Simon, *Human Problem Solving*, Prentice Hall, Englewood Cliffs, NJ, 1972.
- Norman, D. A. (1981), "Categorization of Action Slips," *Psychological Review*, **88** (1), 1-15.
- Panko, R. R. (1998), "What we Know About Spreadsheet Errors," *Journal of End User Computing*, **10** (2), 15-21.
- Panko, R. R. (1999), "Applying Code Inspection to Spreadsheet Testing," *Journal of Management Information Systems*, **16** (2), 159-176.

- Panko, R. R. and R. H. Halverson, Jr. (1995), "Patterns of Errors in Spreadsheet Development I: Quantitative Errors," *Working Paper*, University of Hawaii, 1-33.
- Panko, R. R. and R. H. Halverson, Jr., "Spreadsheets on Trial: A Survey of Research on Spreadsheet Risks," in *Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences*, Maui, Hawaii, 1996, 326-335.
- Panko, R. R. and R. H. Sprague (1998), "'Hitting the Wall': Errors in Developing and Code Inspecting a 'Simple' Spreadsheet Model," *Decision Support Systems*, **22** (4), 337-353.
- Power, D. J. and S. L. Meyeraan (1994), "Impacts of Problem Structure and Computerized Decision Aids on Decision Attitudes and Behaviors," *Information & Management*, **26**, 281-294.
- Rasmussen, J., *Information Processing and Human-Machine Interaction*, North Holland, Amsterdam, 1986.
- Reason, J. T., "Lapses of Attention," in *Varieties of Attention*, R. Davies (Ed.), Academic Press, New York, 1984.
- Reason, J. T., *Human Error*, Cambridge University Press, New York, 1990.
- Saaty, T. L., *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980.
- Saaty, T. L., *Decision Making for Leaders*, Lifetime Learning Publications, Belmont, CA, 1982.
- Saaty, T. L. and L. G. Vargas, *Decision Making in Economic, Political, Social and Technological Environments*, RWS Publications, Pittsburgh, PA, 1994.
- Schoner, B., W. C. Wedley and E. U. Choo (1992), "A Rejoinder to Forman on AHP, With Emphasis on the Requirements of Composite Ratio Scales," *Decision Sciences*, **23**, 509-517.
- Schoner, B., W. C. Wedley and E. U. Choo (1993), "A Unified Approach to AHP with Linking Pins," *European Journal of Operational Research*, **64**, 384-392.
- Stam, A. and A. P. Duarte Silva (1997), "Stochastic Judgments in the AHP: The Measurement of Rank Reversal Probabilities," *Decision Sciences*, **28** (3), 655-688.
- Todd, P. and I. Benbasat (1994a), "The Influence of DSS on Choice Strategies Under Conditions of High Cognitive Load," *IEEE Transactions on Systems, Man and Cybernetics*, **24** (4), 537-547.
- Todd, P. and I. Benbasat (1994b), "The Influence of DSS on Choice Strategies: An Experimental Analysis of the Role of Cognitive Effort," *Organizational Behavior and Human Decision Processes*, **60**, 36-74.
- Todd, P. and I. Benbasat (1999), "Evaluating the Impact of DSS, Cognitive Effort, and Incentives on Strategy Selection," *Information Systems Research*, **10** (4), 356-374.
- Todd, P. and I. Benbasat (2001), "Inducing Compensatory Information Processing through Decision Aids that Facilitate Effort Reduction: An Experimental Assessment," *Journal of Behavioral Decision Making*, **13** (1), 91-106.
- Troutt, M. D. (1988), "Rank Reversal and the Dependence of Priorities on the Underlying MAV Function," *Omega*, **16** (4), 365-367.
- Winkler, R. L. (1990), "Decision Modeling and Rational Choice: AHP and Utility Theory," *Management Science*, **36** (3), 247-248.

Dependent Measures	No DSS		DSS	
	Lower Load	Higher Load	Lower Load	Higher Load
Relative Importance of Criteria	66.66 (11.50)	65.00 (15.17)	80.06 (11.51)	77.49 (15.55)
Logic Errors	0.125 (0.342)	0.563 (0.512)	0.188 (0.403)	0.188 (0.403)
Omission Errors	0.000 (0.000)	.0714 (.267)	0.000 (0.000)	0.000 (0.000)
Mechanical Errors	0 (0)	0 (0)	0.188 (0.403)	0.125 (0.342)
Decision Quality	0.875 (0.342)	0.438 (0.512)	0.625 (0.500)	0.688 (0.479)

Table 1: Means (and Standard Deviations) of Dependent Measures

<u>Dependent Measures</u>	<u>Information Load</u>	<u>DSS Use</u>	<u>Information Load x DSS Use</u>
Relative Importance of Criteria	0.39 (0.537)	14.41 (0.001)***	0.02 (0.893)
Logic Errors	4.35 (0.041)*	2.22 (0.142)	4.35 (0.041)*
Omission Errors	1.19 (.280)	1.19 (.280)	1.19 (.280)
Mechanical Errors	0.22 (0.638)	5.60 (0.021)*	0.22 (0.638)
Decision Quality	2.62 (0.111)	0.00 (1.000)	4.66 (0.035)*

* : $p < .05$; **: $p < .01$; ***: $p < .001$

Table 2: *F*- and *p*-Values for Statistical Tests on Dependent Measures

Dependent Measures	No DSS		DSS		Total
	Lower Load	Higher Load	Lower Load	Higher Load	
Logic Errors	2	8	3	4	17
Criteria	2	8	2	2	14
Inconsistency	0	0	1	1	2
Rank Inversion	0	0	0	1	1
Omission Errors	0	1	0	0	1
Mechanical Errors	0	0	3	1	4

Table 3: Count of Errors by Type

Editors:

Michel Avital, University of Amsterdam
Kevin Crowston, Syracuse University

Advisory Board:

Kalle Lyytinen, Case Western Reserve University
Roger Clarke, Australian National University
Sue Conger, University of Dallas
Marco De Marco, Università Cattolica di Milano
Guy Fitzgerald, Brunel University
Rudy Hirschheim, Louisiana State University
Blake Ives, University of Houston
Sirkka Jarvenpaa, University of Texas at Austin
John King, University of Michigan
Rik Maes, University of Amsterdam
Dan Robey, Georgia State University
Frantz Rowe, University of Nantes
Detmar Straub, Georgia State University
Richard T. Watson, University of Georgia
Ron Weber, Monash University
Kwok Kee Wei, City University of Hong Kong

Sponsors:

Association for Information Systems (AIS)
AIM
itAIS
Addis Ababa University, Ethiopia
American University, USA
Case Western Reserve University, USA
City University of Hong Kong, China
Copenhagen Business School, Denmark
Hanken School of Economics, Finland
Helsinki School of Economics, Finland
Indiana University, USA
Katholieke Universiteit Leuven, Belgium
Lancaster University, UK
Leeds Metropolitan University, UK
National University of Ireland Galway, Ireland
New York University, USA
Pennsylvania State University, USA
Pepperdine University, USA
Syracuse University, USA
University of Amsterdam, Netherlands
University of Dallas, USA
University of Georgia, USA
University of Groningen, Netherlands
University of Limerick, Ireland
University of Oslo, Norway
University of San Francisco, USA
University of Washington, USA
Victoria University of Wellington, New Zealand
Viktoria Institute, Sweden

Editorial Board:

Margunn Aanestad, University of Oslo
Steven Alter, University of San Francisco
Egon Berghout, University of Groningen
Bo-Christer Bjork, Hanken School of Economics
Tony Bryant, Leeds Metropolitan University
Erran Carmel, American University
Kieran Conboy, National U. of Ireland Galway
Jan Damsgaard, Copenhagen Business School
Robert Davison, City University of Hong Kong
Guido Dedene, Katholieke Universiteit Leuven
Alan Dennis, Indiana University
Brian Fitzgerald, University of Limerick
Ole Hanseth, University of Oslo
Ola Henfridsson, Viktoria Institute
Sid Huff, Victoria University of Wellington
Ard Huizing, University of Amsterdam
Lucas Introna, Lancaster University
Panos Ipeirotis, New York University
Robert Mason, University of Washington
John Mooney, Pepperdine University
Steve Sawyer, Pennsylvania State University
Virpi Tuunainen, Helsinki School of Economics
Francesco Virili, Università degli Studi di Cassino

Managing Editor:

Bas Smit, University of Amsterdam

Office:

Sprouts
University of Amsterdam
Roetersstraat 11, Room E 2.74
1018 WB Amsterdam, Netherlands
Email: admin@sprouts.aisnet.org