8-16-1996

Decision Support System Differences: Do They Really Matter?

Christian Wagner
Information Systems Department, City University of Hong Kong, iscw@cityu.edu.hk

Zoonky Lee
School of Business Administration, University of Southern California, zoonkyl@rcf.usc.edu

David Kang
School of Business Administration, University of Southern California, wkang@rcf.usc.edu

Decision Support System Differences: Do They Really Matter?

Christian Wagner
Information Systems Department
City University of Hong Kong
Kowloon, Hong Kong
ISCW@CITYU.EDU.HK
Zoonky Lee
School of Business Administration
University of Southern California
Los Angeles, CA 90089-1421
ZOONKYL@RCF.USC.EDU
David Kang
School of Business Administration
University of Southern California
Los Angeles, CA 90089-1421
WKANG@RCF.USC.EDU

Background

After two decades of research into the effectiveness of decision support systems and their components, we still have only few prescriptions for good system design. Are graphs better than tables? Which database interface or underlying data model permits faster and more accurate queries? How important are the "right" models in the model base for enhanced user performance? Frequently, half of the research studies state that one treatment, (e.g., graphs, entity-relationship model), is better than the alternative treatment, while the other half claims the opposite

Our approach to this problem is to suggest that results of previous research are so undecided, not because of lack of researcher care, design errors, or the like (as suggested elsewhere), but rather because of the adaptability of the people using decision support systems and their components. Following the argument of the *Power Law of Practice* (Snoddy, 1926), we believe that with enough practice, people are so adaptable, that their performance will only marginally depend on the DSS features.

Research on Design Features: Half in Favor, Half Against

To remind the reader of the controversial findings in DSS research, a brief (and non-exhaustive) review of studies of all three main DSS design elements is in order: the user interface, the database, and the model base component.

User Interface

DSS user interface studies have largely focused on the issue of superiority of graphs versus tables. Key findings include:

- DeSanctis (1984): 12 studies support superiority of tables, 7 support the superiority of graphs, 10 indicate no difference.
- Remus (1984, 1987): Tables better in low complexity environment, graphs in high complexity environment.

- Dickson et al. (1986): Graphical displays more difficult to use. No difference in decision making performance.
- Benbasat et al. (1986): No difference in decision making performance.

Based on these results, no strong statements can be made regarding the superiority of either design alternative. In other words, were these results due to experimenter sloppiness, lack of appropriate tasks, lack of good statistical procedures, or similar shortcomings, as suggested by some (e.g., Remus, 1987)?

Database

The database discussion, as it is relevant to our observations on DSS research, revolves around the query language (also an interface issue of sorts), and the underlying data model. Repeatedly, researchers have discussed, and tried to verify, the superiority of one type of query language over another, or one data model over another. Listed below are several representative research efforts, the first three discussing query languages, the last two, data models.

- Turner et al., 1984: No significant differences bwteen natural and linear keyword language (LKL).
- Jarke et al., 1985: LKL more successful (query completion), more bugs with natural language.
- Suh and Jenkins, 1992: Natural language better in correctness and time to write queries.
- Jih et al., 1989: No difference in query accuracy for relational versus entity-relationship (ER) model.
- Chan et al., 1993: ER queries completed with higher accuracy than relational.

Again, the results are ambivalent. Some studies identify the superiority of one type of language/model, while others find just the opposite.

Model Base

Research on the impact of the model base on decision support has encountered some confusing evidence with respect to the value of models. Results of several studies illustrate this phenomenon.

- Sharda et al. (1988): Confirms the value of models, particularly when used repeatedly.
- Elam and Mead (1990): Very similar models which are part of the same software are both more and less effective than control treatment.
- Mackay and Elam (1992): Lack of DSS expertise inhibited use of functional knowledge in decision making.
- MacCrimmon and Wagner (1994): Treatment (model base) produces little performance improvements for group at large. Significant performance improvement for good subjects.

Essence of Feature Comparisons

Although a short summary cannot do justice to the differentiations made in, and the special nuances observed by different studies, the essential message is that neither design feature is unequivocally better than its alternative. Yet at the same time, studies such as Sharda et al. (1988) and Mackay and Elam (1992) point toward the need of considering expertise and repeat performance in the evaluation of decision support system models, a thesis in line with our concerns for the power law.

Explaining Performance Differences with DSS: The Power Law of Practice

The essence of our argument explaining performance differences may be simply stated as follows. ***People are so adaptive in using technology, that with sufficient practice, they master even poor tools and produce good results with them.*** Three elements are important here. First, people can produce similar levels of performance with different tools, even if one tool is "objectively" inferior to the other one.

Second, performance with the same tool can very much differ for a group of users. Third, each individual may show a wide performance range, depending on the level of practice. We base our argument on findings of research on motor tasks and simple problem solving tasks described elsewhere (e.g., Snoddy, 1926, Kolers, 1975, Laird et al., 1986) and on our research on practice based performance for DSS model use and user interface use, summarized in the next section.

Power Law of Practice

Observations congruent with a power law of practice go back many decades, when they were encountered for motor tasks. Later, similar findings were made for recognition tasks, as well as for simple problem solving. Altogether they point out that with practice (task repetition), the time required to perform a task diminishes exponentially and that performance variance also diminishes with increased practice. Thus, an experienced person will complete a task in a small fraction (even as low as 1/10) it takes a less experienced one.

Given the strong impact practice can have on performance, much of the advantage created by a "better" DSS feature may be more than compensated for by practice. Furthermore, subject groups which perform a task with either treatment or control, may be sampled from a larger pool of individuals with different levels of prior applicable practice, thus resulting in large variance in both the treatment and in the control groups, variances which easily mask differences between treatment and control.

Empirical Evidence

To confirm our belief in the applicability of the Power Law of Practice, we have begun experiments on repeat user performance with both model and interface features.

Study of Performance with the DSS Model Base

The first experiment assessed the decision making performance of 15 subjects in a 2-person, zero-sum game. Every subject competed against a computer opponent (which always played the same fixed strategy). Subjects participated in six, each session consisting of three games. Eight subjects were supported from beginning to end by a decision support system with about 20 models, seven subjects could use the DSS only during their last session (last three games).

All subjects in this experiment learnt from experience. Winning percentages went up significantly ($p = .05$, T-test), from about 55% for the first session (players win slightly more than every second game) to about 75% (players win 3 out of four games) in the last session. No significant performance difference between DSS and no-DSS groups was observed. Those who used the DSS all six sessions became "DSS Experts", however, reducing their time to complete the task significantly (19.0 minutes for DSS novices versus 12.2 minutes for DSS experts). Repeated DSS use also resulted in less planning errors in this strategy game ($p = .001$, T-test). We also observed high variances in the performance of individuals within each group.

Study of Performance with Two Different User Interfaces

In another study, we compared the performances of 32 users with two different interfaces, related to the completion of transactions. Subjects had to complete transactions similar to those one carries out at an automatic teller machine (ATM). This is not *per se* a decision support system oriented task, but it requires planning, and, based on our choice of interfaces also introduced uncertainty. Uncertainty was introduced because subjects had to complete "banking transactions" either using a completely icon-based interface, or a foreign language (German) interface.

Half of the subjects began by using the icon interface for 15 minutes, followed by 15 minutes with the German interface. The other half of the subjects used the interfaces in opposite order. None of the participants knew German.

The icon interface was the "superior" interface, at least during the first round. Transaction completion speeds were significantly higher than for the German text interface (p = .049). In the second round, however (icon users switched to German, German users to icons), the differences became much smaller and insignificant (T-test, p = .17). We also observed high variance in the German interface user group.

Furthermore, we monitored an increase in transaction completion speeds during each of the 15 minute rounds and found speed improvement to be significantly greater for icons in the first round (p = .038), but not in the second round (p = .12). In each round, we saw transaction completion times drop by a factor of about 4 from first to last transaction (e.g., from about 40s to about 10s, icon ATM, first round).

Interpretation and Implications

In both experiments, we saw clear signs of learning, congruent with the Power Law of Practice. Repeat use had a strong effect on performance, and high variance in both treatment and control groups.

Our findings highlight the need to much more than before consider subject experience when conducting experiments. If novice subjects are studied, results should not be generalized beyond that group. If "average" users are studied and an average is created through random assignment to treatment and control groups, potential mismatches may result. Mixing high and low performers within a group can result in unacceptably high variance, spurious effects created by "outliers," and inconsistent results.

References

Chan, H.C., Wei, K.K., and Siau, K.L. User-Database interface: the effect of abstraction levels on query performance. *MIS Quarterly*, 17(4), 1993, 441-464.

Elam, J. and M. Mead "Can software influence creativity?," *Information Systems Research,* 1(1), March 1990, 1-22.

Jarke, M., Turner, J.A., Stohr, E.A., Vassiliou, Y., White, N.H., and Michelsen, K. "A field evaluation of natural language and data retrieval," *IEEE Transactions on Software Engineering*, 11(1), 1985, 97-113.

Jih., K. Bradbard, D., Snyder, C., and Thompson, N. "The effects of relational and entity-relationship data models on query performance of end users," *International Journal of Man-Machine Studies*, 31(3), 1989, 257-267.

Kolers, P.A. "Memorial consequences of automatized encoding," *Journal of Experimental Psychology: Human Learning and Memory,* 1(6), 1975, 689-701.

Laird, J., Rosenbloom, P., and Newell, A. *Universal Subgoaling and Chunking*, Kluwer Academic Publishers, Boston, 1986

MacCrimmon, K.R. and Wagner, C. "Stimulating ideas through creativity software," *Management Science*, 40(11), November 1994, 1514-1532.

Mackay, J.M. and Elam, J.J. "A comparative study of how experts and novices use a decision aid to solve problems in complex knowledge domains," *Information Systems Research*, 3(2), 1992, pp. 150-172.

Remus, W. E. "An empirical investigation of the impact of graphical and tabular data presentations on decision making," *Management Science*, 30(5), 1984, 533-542.

Remus, W. E. "A study of graphical and tabular displays and their with environmental complexity," *Management Science,* 33(9), September 1987, 1200-1205.

Sharda, R., Barr, S.H., and McDonnell, J.C. "Decision support system effectiveness: a review and empirical test," *Management Science*, 34(2), February 1988, 139-159.

Snoddy, G. S. "*Learning and Stability," Journal of Applied Psychology*, 10, 1926, 1-36

Suh, K.S. and Jenkins, M. "A comparison of linear keyword language and restricted natural language database interfaces for novice users." *Information Systems Research*, 3(3), 1992, 252-272.

Turner, J. Jarke, M., Stohr, E.A., Vassiliou, Y., and White, N.H. "Using restricted natural language for data retrieval: a plan for field evaluation," *In Human Factors and Interactive Computer Systems*, Vassiliou, Y. (Ed.), Ablex, Norwood, 1984, 163-190.