December 2003

Using Effort, Accuracy and Technology Acceptance to Predict Decision Confidence in Online Shopping

Arnold Kamis
Bentley College

Using Effort, Accuracy and Technology Acceptance to Predict Decision Confidence in Online Shopping

Arnold Kamis
Bentley College
akamis@bentley.edu

Abstract

In business-to-consumer (B2C) electronic commerce, the conversion rate of lookers-to-buyers averages 2%, 2 buyers for every 100 lookers. We believe that this rate is due, in part, to decision aids that are not designed to fit the large search space faced by online shoppers. A Decision Aid (DA) is a software tool designed to help decision makers, e.g., online shoppers. We investigate whether some sequences of DAs help shoppers more than other sequences.

We develop a shopping model, which combines Effort, Accuracy and the Technology Acceptance Model (TAM) (Davis 1989) to support the shopper with different electronic store (e-store) designs, which are sequences of two or three DAs. The shopper's goal is consistent with the Effort-Accuracy Model (EAM): maximize accuracy and minimizing effort. Our integrated shopping model shows that TAM extends EAM to better predict Decision Confidence (DC).

We use a controlled experiment on 116 subjects and treatments that are four different e-store designs. We use exploratory second-generation Structural Equation Modeling, namely Partial Least Squares Regression. The analysis helps us determine the best experimental treatment, i.e., e-store design. There are two key findings: 1) some e-store designs minimize effort, maximize accuracy, or maximize DC significantly more than others, and 2) several TAM-related variables are important predictors of DC.

This research could have direct implications for electronic commerce decision aid designers who are trying to increase revenues. The designers could have their decision aids dynamically detect the current task complexity and either recommend or impose a particular decision aid. The dynamic detection could be tailored to the individual's customer profile or real-time behavior.

Keywords: Cognitive effort, choice accuracy, TAM, decision confidence, online shopping, e-commerce, decision aid, product search engine, structural equation modeling, PLS, data envelopment analysis

Introduction

In electronic commerce, buyers benefit from convenient access to content, community, and commerce while sellers benefit from selling to customers anytime and anywhere with low brick-and-mortar and intermediary costs. Despite these benefits for both buyers and sellers, conversion rates have averaged only 2% (i.e., 2 buyers for every 100 lookers) (Gomory et al. 1999; McQuivey et al. 1998). A careful selection and sequencing of Decision Aids (DAs), i.e., a Product Search Engine (PSE) offers the potential to improve conversion rates. For examples of leading PSEs, see ActiveBuyersGuide.com or Epinions.com.

There are consultancy research findings suggesting that product search engines[1] (PSEs) subject online consumers to an inadequately designed shopping process and produce sub-optimal choices, which do not elicit decision confidence:

- Boolean, keyword search does not produce sales. It needs to be replaced by parametric search,[2] recommendation engines, and product configurators[3] (Hagen et al. 1998).
- Shoppers struggle with poor search mechanisms, including poor guidance in formulating a good query, confusing search results, and displays that cannot be manipulated (Hagen et al. 1999).
- Merchants typically fail their customers in a few key areas: page design, site search, and the checkout process (Good 2000).

Decision Aids in Multistage Product Search

The purpose of this section is to investigate and analyze product search on the Web using PSEs that allow users to make decisions with different DAs in multiple stages. The multiple stage model draws from marketing and decision research and provides the basis for our research design. We base our analysis primarily on the "effort-accuracy" framework of Payne et al. (1993). This framework seeks to explain how decision makers tradeoff the effort involved in executing a decision strategy versus the accuracy (quality in a utility maximizing sense) of the final choice.

Objectives of Buyer Search: Effort-Accuracy Framework

We adopt a cognitive cost-benefit perspective known as the effort-accuracy framework (Payne et al. 1993). Effort refers to the cognitive effort required to execute a decision strategy. Accuracy refers to the quality of the final product choice - whether or not a decision strategy "leads to the identification of the best alternative in a choice set" (Payne et al. 1993, p. 72). The best alternative is the one presumed to best satisfy the decision maker's true preferences.

With the effort-accuracy framework, we see tradeoffs and decision strategies at the individual level of analysis. We assume that the individual engages in economical, cost-benefit decision making. Although individuals are free to tradeoff cognitive effort (a cost) for outcome accuracy (a benefit), the literature shows that individuals tend to minimize cognitive effort rather than maximize outcome accuracy *in preferential choice tasks*, in which there are no correct answers (Todd and Benbasat 1992; Kleinmuntz and Schkade 1993; Todd and Benbasat 1993). By contrast, in decision tasks with correct or optimal outcomes, individuals tend to behave differently. For example, in stable environments (e.g., drilling for oil), in narrow decision spaces (trading stock options), or where stakes are high (e.g., maximizing corporate profits), individuals are more prone to maximize outcome accuracy. The emphasis on effort reduction may also be simply because feedback regarding effort tends to more salient than feedback regarding accuracy (Einhorn et al. 1978; Haubl et al. 2000; Kleinmuntz et al. 1993; Payne et al. 1993).

The effort-accuracy framework has been applied fruitfully to research the tradeoff between decision outcome accuracy and cognitive processing effort, given different decision strategies (Payne et al. 1993). The same framework has been tested and validated by a large number of researchers, in psychology (Kleinmuntz et al. 1993; Saad et al. 1996; Tybout 1994), marketing (Bettman et al. 1998; Klein et al. 1989; Widing et al. 1993), and information systems (Benbasat et al. 1996; Fischer et al. 1999; Todd et al. 1992).

The design of a Product Search Engine, i.e., a DA sequence, must take into account the ability of the decision aids to yield products that are highly-valued by the shopper (in a utility maximizing sense) and yet reduce the total amount of cognitive effort involved in the search.

---

[1]PSEs differ from regular search engines such as Yahoo! (**www.yahoo.com**) and AltaVista (**www.altavista.com**) in that PSEs, e.g., shopping.yahoo.com or shopping.altavista.com, focus on product attributes, rather than general information.

[2]A parametric search allows one to specify minimum or maximum values for one or more attributes, while indicating indifference to all other attributes.

[3]Product configurators allow one to configure a product and then adjust its attributes incrementally, e.g., Dell computers (See www.dell.com.).

Decision Aids in Multiple-Stage Search

Few consumers would have the patience to inspect several hundred computers in a single phase, all in an in-depth manner. Therefore, a multiple stage approach to Web shopping could be useful. In the early stages of product search, the shopper would use the available decision aids to narrow the number of products to be considered. In later stages, the shopper would use decision aids to compare products in depth by carefully inspecting and comparing their attributes.

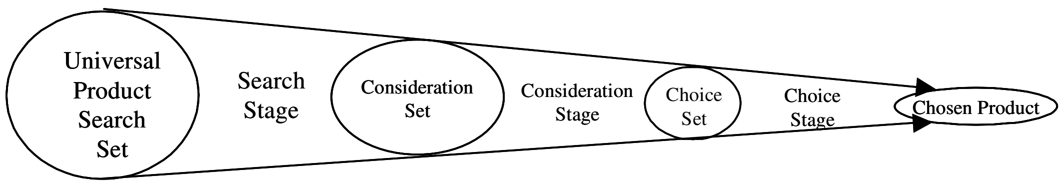


Figure 1. Decision-Making Sets and Stages

Decision strategies have been researched in a single-stage (e.g., EBA) and in multiple stages (e.g., EBA followed by WADD) (Payne, Bettman et al. 1993; Coupey 1994; Levin and Jasper 1995; Levin, Jasper et al. 1998). In addition, the marketing and decision literatures have suggested using a combination of non-compensatory and compensatory model types, depending on the decision stage.

Staged decision making recognizes the reality of shopping on the Web, which is a large search space with low search costs.[4] To search the Web offerings comprehensively, consumers must use one or more DAs. For example, a strategy from the lexicographic family followed by a strategy from the additive utility family. This is typical of how consumers actually shop (Haubl et al. 2000). The elimination strategies, Parametric Search (PS) and Lexicographic (LEX), are generally provided by Web sites to initiate product search, while the compensatory decision strategy of displaying the attributes of products side-by-side in a Comparison Matrix (CM) is generally provided at the end of a consumer's search.

We use this view of multistage product search throughout this study. We point out, however, that not all three stages of decision and choice may be required in many cases of actual product search. Moreover, the product search may often be performed entirely with only two decision aids or even with one decision aid. A major thrust of this research is to discover how these decision aids should be used either singly or in combination to facilitate user search and optimize the effort-accuracy tradeoff mentioned above.

Research Model

In this section, we embed multiple-stage search in a research model that draws from the Effort-Accuracy Model (EAM) (Payne et al. 1993; Todd et al. 1992), as well as key elements of the Technology Acceptance Model (TAM) (Davis 1989). We combine EAM and TAM. We integrate EAM and TAM because we suspect that an integrated model combining objective and subjective data would be more predictive of Decision Confidence than either model alone.

We chose to use the Effort-Accuracy framework, because it has been heavily researched by Payne, Bettman and Johnson, its originators, as well as numerous other researchers (Fennema et al. 1995; Kleinmuntz et al. 1993; Levin et al. 1998; Lohse et al. 1998). Effort and Accuracy are well-established, fundamental variables that individual decision makers weigh and decision technology designers consider (Payne et al. 1996; Todd et al. 1992; Todd et al. 1994). We chose to use TAM, because no matter how effective or efficient a decision aid may be, if it is not accepted, consumers will not use it. TAM has been researched using many kinds of information technology, and has been found to have valid and reliable constructs (Doll et al. 1998; Gefen et al. 1997; Szajna 1996).

---

[4]Contrary to popular belief, everything is not one click away on the Web. Search costs are low, but not zero.

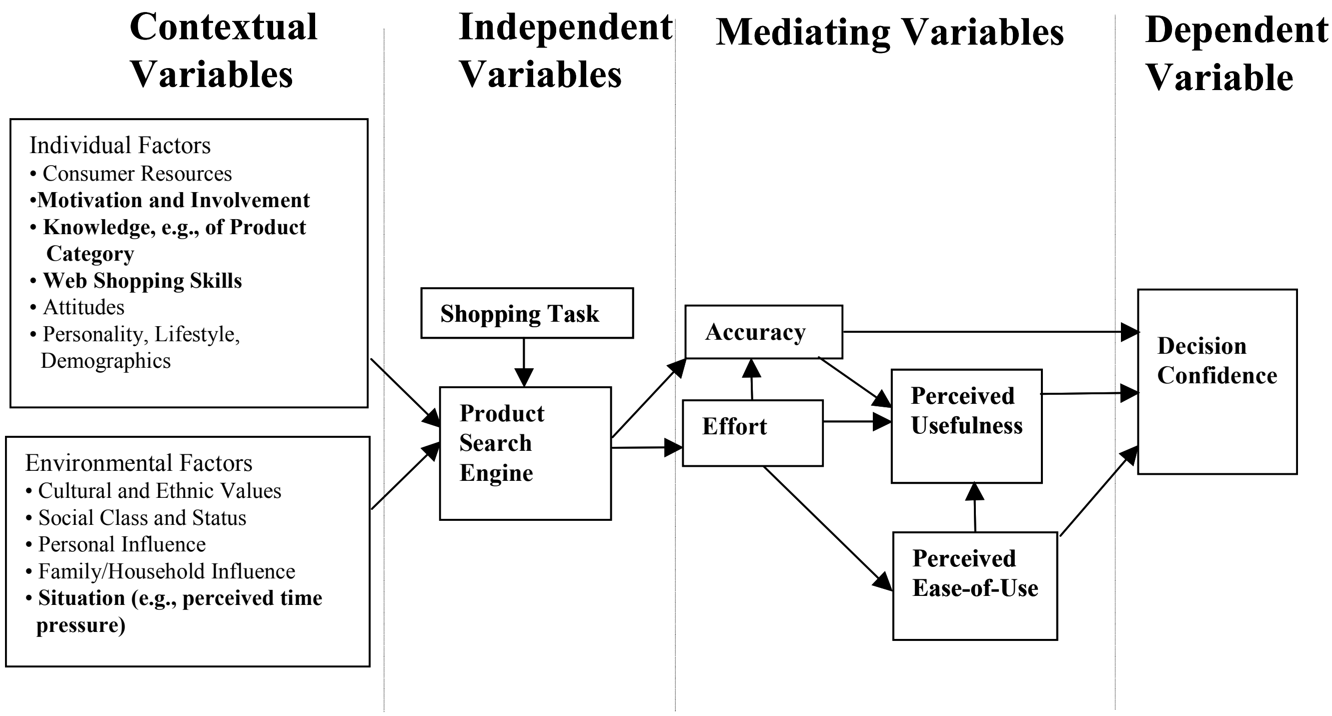


Figure 2. Structural Shopping Model

Figure 2 depicts the four classes of variables in our integrated EAM-TAM shopping model: Contextual, Independent, Mediating and Dependent. On the left side are the two main sources of contextual variables: individual and environmental. We measure some of them and leave the rest for future research. In the middle of the figure, we see the independent variable: decision aid sequences. If we were to open this box, we would see the multiple stage process described in the previous section. On the right side of the figure, we see mediating and dependent variables. The mediating variables follow from the Effort-Accuracy Model and the Technology Acceptance Model, which indicates whether the user accepts the decision-aiding technology. Since Perceived Ease of Use impacts Perceived usefulness (Davis 1989) and PEU/PU parallels Effort/Accuracy, we predict that the direction of impacts are as shown. The items in the figure that are rendered in **boldface** have been manipulated or measured in this research. The non-boldface items represent interesting areas for further research.

Operationalization of Variables

This section operationalizes all the variables. Proceeding from left to right in Figure 2, we examine and define the contextual, independent, mediating and dependent variables used in our study.

Contextual Variables (Covariates)

From Figure 2, we know there are a number of potentially confounding factors that can impact our primary variables of interest. We can measure only a few of them without creating a data collection burden; we therefore chose to check the following statistically:

- *Task Involvement* (TI): the individual's internal state of arousal regarding the Task. (See Appendix B, questionnaire items 3-6.) TI is an important motivational concept (Beatty and Smith, 1987; Andrews, Durvasula et al., 1990; Mishra, Umesh et al., 1993). If a consumer is uninvolved, he is unlikely to absorb information fully or search efficiently.

- *Product Category Knowledge* (PCK): the individual's general knowledge of the product category, e.g., computers, printers or cars. (See Appendix B, questionnaire items 7-11.) Product search increases when prior knowledge is general to the product category, rather than specific to products (Punj et al. 1983). Search decreases as prior knowledge of and experience with specific products increases.

- *Perceived Time Pressure* (PTP): the individual's perception that time pressure was a significant influence. (See Appendix B, questionnaire items 12-15) Across multiple product categories, a positive association has been found between PTP and search effort (Beatty et al. 1987), up to a limit (Payne et al. 1993; Payne et al. 1996). According to Jarvenpaa and Todd (1997), their subjects indicated that PTP is the primary factor driving their desire for convenience, and convenience is one of the main reasons they shop online.

- *Perceived Web Shopping Skills* (PWSS): the individual's perception that his web shopping skills are strong (See Appendix B, questionnaire items 16-18). We expect that greater PWSS will increase the perceived usefulness and perceived ease-of-use of web shopping tools.

Independent Variables (Treatments)

The structure of the Decision-Aided Sequence, consists of Attribute Weight Elicitation, followed by a DA for the Search Stage, followed by a DA for the Consideration Stage, followed by a DA for the Choice Stage. The DA in the first two stages is Parametric Search or LEX. The DA in the third stage is Comparison Matrix. **Error! Reference source not found.** shows the independent variables as our four web store treatments:

Table 1. Independent Variables as Web Store Treatments

Treatment	Attribute Weight Elicitation	Search Stage	Consideration Stage	Choice Stage
T1	WADD	PS	LEX	CM
T2	WADD	PS	PS	CM
T3	WADD	LEX	LEX	CM
T4	WADD	LEX	PS	CM

Dependent Variables

In this section, we define a number of objective constructs pertaining to EAM and TAM variables. We operationalize Accuracy in two ways and Effort in two ways. For operationalizing accuracy, both Data Envelopment Analysis (DEA) and Subjective Utility are useful. This combination of accuracy measures is a methodological contribution.

Dominance (δ)

DEA, a linear optimization technique, is applied to the universal set of products to determine each product's δ, an objective measure of input- or output-inefficiency. Delta (δ) ranges from 0 (no inefficiency) to 9 (maximum inefficiency). Technically, the DEA model is of the standard model type, VRS Surface, Base-Orientation, with Standard Evaluation. For standard treatments of DEA, see Banker et al. (1984) and Charnes et al. (1978) and for more recent developments: Cooper et al. (1999) Halme et al. (1999), Joro et al. (1998). In a computer shopping task, for example, DEA minimizes one input (price) and maximizes four outputs (memory, disk, speed, monitor size). It also allows variable returns to scale, which makes sense; subjects may not think of each attribute in linearly increasing or decreasing fashion. This is the most flexible model, accommodating both price minimizers and feature maximizers, and any combination of the two.

Utility (U)

Utility (U) is based on the WADD score, according to the user's elicited attribute weights and standardized scales for each attribute. This is the linear, weighted additive (WADD) Utility score, computed by weighting each subject's attribute weights by standardized attribute levels (lowest = 1, middle = 2, highest = 3), and summing for features, subtracting for price. In the computer shopping task, for example, the five attribute weights are memory, disk space, chip speed, monitor size, and price. The score is computed as follows:

Score = mem weight*memLevel + disk weight*diskLevel + speed weight*speedLevel + mon weight*monLevel - price weight*priceLevel

Each set of scores is ranked from highest to lowest. Utility (U) is the rank of each subject's choice, ranging from 1 to 241 (for the car or printer shopping task) or 242 (for the computer shopping task).

Effort (t)

Effort is measured by a proxy: elapsed time spent in seconds from the beginning of the task, after reading directions, to the end of the task, before completing a survey. We also measure it as the time spent strictly within the decision stages.

Perceived Ease of Use and Perceived Usefulness

Perceived ease-of-use (PEU) and Perceived usefulness (PU) are the standard subjective measures of the cognitive effort expended in the use of a decision aid. Standard questionnaires are available to measure these items (Davis 1989).

- *Perceived usefulness* (PU): the degree to which an individual believes that the Web store enhanced the accomplishment of the shopping task, on a seven point Likert scale. (See Appendix B, questionnaire items 22-26.)

- *Perceived ease-of-use* (PEU): the degree to which an individual believes that the Web store was easy to use, on a seven point Likert scale. (See Appendix B, questionnaire items 27-31.)

To the subjects, a web store is implemented as a decision aid sequence. We measure PU and PEU after each store is used by an individual subject. Each subject uses one store in a practice task and two different stores in tasks that count toward a nominal reward. After both real stores are visited and used, the subjects rank them in terms of PU and PEU. (See Experimental Procedure, steps 9-10). The diagram below summarizes the objective and subjective measures that should correlate.

Accuracy Maximization And Effort Minimization	≈	Perceived Usefulness and Ease of Use

We thus obtain objective and subjective assessments of the value of decision aid sequences.

Decision Confidence

We are ultimately interested in Decision Confidence (DC), because without it, an e-store visitor is unlikely to purchase. DC is superior to Accuracy because it reflects ones assessment of the item in question as well as the process used to find it. DC is also an attractive construct for the seller to manipulate, because it may be possible to increase DC without increasing Accuracy, thereby preserving profits.

- *Decision Confidence* (DC): how satisfied the subject is with the final product, on a seven point Likert scale. The subject is asked how satisfied he is with his choice and how likely he would be to recommend purchase of it, on a seven point Likert scale. (See Appendix B, questionnaire items 19-21.)

Predictions

Our predictions hinge on the fact that the first DA determines the products that are seen in the second stage. If it is PS, effort reduction will likely be the driving factor, because PS is good for eliminating many products quickly. If it is LEX, accuracy maximization will likely be the driving factor, because it supports one attribute at a time. We make analogous predictions for TAM variables:

Treatment 1: PS-LEX-CM

- P1: For treatment 1, Effort will have a greater impact on Accuracy, PU and PEU than Accuracy will have on PU and DC.
- P2: For treatment 1, PEU will have a greater impact on PU and DC than PU will have on DC.

Treatment 2: PS-PS-CM

- P3: For treatment 2, Effort will have a greater impact on Accuracy, PU and PEU than Accuracy will have on PU and DC.
- P4: For treatment 2, PEU will have a greater impact on PU and DC than PU will have on DC.

Treatment 3: LEX-LEX-CM

- P5: For treatment 3, Accuracy will have a greater impact on PU and DC than Effort will have on Accuracy, PU and PEU.
- P6: For treatment 3, PU will have a greater impact on DC than PEU will have on PU and DC.

Treatment 4: LEX-PS-CM

- P7: For treatment 4, Accuracy will have a greater impact on PU and DC than Effort will have on Accuracy, PU and PEU.
- P8: For treatment 4, PU will have a greater impact on DC than PEU will have on PU and DC.

We test these predictions by examining the statistically significant weights in Structural Equation Models.

Methods

A Web store was built for testing subjects who shopped for a variety of products, printer, car and computer, with a variety of PSEs. The shopping task was to obtain the best printer/car/computer for the best price, i.e., the best “bang-for-the-buck,” given a task scenario of corporate procurement. The subject’s reward was based on the accuracy (optimality) of the product selected and the effort (time) spent in the process. The PSEs were sequences of decision aids that implement different decision strategies. Specifically, they were Parametric Search (PS) and Lexicographic (LEX), the former supporting a quick, low effort strategy, the latter supporting a gradual, high optimality strategy. The order of tasks was random, as was the assignment of treatments.

We implemented the decision aids as realistically and professionally as possible, in a university computer lab. In our pilot tests, we verified that the shopping task required neither too much, nor too little time. We therefore expected subjects to complete all treatments without growing bored or frustrated. The Subjects (Ss) were 116 college students participating for a nominal incentive. These students were relatively homogeneous with regard to age and level of education, thus eliminating systematic sources of variation.

We populated the product database with a variety of printers for the practice task, cars and computers for the real tasks. The method of generating products from their attribute levels was the same regardless of the product. For each product attribute, we used current market data for high, medium and low levels on each of five attributes, including price. We used every combination to generate the database. This method of product construction ensured that our product database was preference indifferent; i.e., we did not skew the database toward any combination of attribute values. In addition, we removed the dominant product, the one with High levels on the first four attributes and a Low level for the price, to prevent the task from being too easy. We also removed the 5 next most dominant products. Thus each product database contained 11 objectively dominant products. We verified that the task was reasonable with two pilot tests. This method of product database construction was similar to that used by Haubl et al. (2000), carefully constructing the database to contain only a few objectively dominant products. The task given to each subject, regardless of PSE or product database, was to identify their most preferred product, the one that gave them the most “bang-for-the-buck”.

The screen design of the experimental treatments was web-based. To proceed from start to finish, subjects simply clicked on hyperlinks and entered values in forms. The page design minimized screen clutter and distractions by having no advertisements. Subjects were allowed to progress from one stage to the next, whenever they wish to do so. Survey instruments were presented after each task and required to proceed.

We conducted two pilot tests, a Procederal Pilot and a Statistical Power Pilot. The purpose of the Procederal Pilot was to ensure that the Web store simulation was bug-free and the instructions were clear. The purpose of the Statistical Power Pilot was to ensure that we would obtain statistically significant results in the final study.

Findings

For each treatment, we construct a Structural Equation Model (SEM) with Partial Least Squares (PLS) Regression. PLS maximizes the R-squared in the overall model suggested by the path representation of all the Regression Analyses. (See Appendix A for a brief introduction to PLS and the notation used in PLS-Graph Version 3, build 279) The following SEMs show only the latent variables, not the indicator variables. All indicator variables, which were checked for statistical, convergent and divergent validity, are omitted for clarity.

For Treatment 1 (PS-LEX-CM), the SEM is the following:

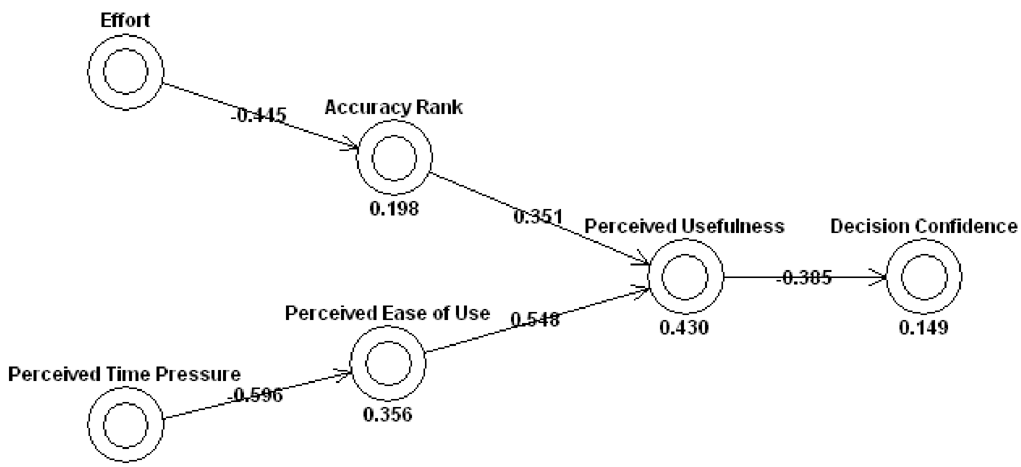


Figure 3. PS-LEX-CM (treatment 1)

Decision Confidence was impacted only by Perceived usefulness, which accounted for 14.9% of variance. Accuracy Rank and Perceived Ease of Use account for 43% of the variance in Perceived usefulness. Note that Accuracy is a rank variable; the smaller, the better. Predictions P1 and P2 are supported.

For Treatment 2 (PS-PS-CM), the final Structural Equation Model is the following:

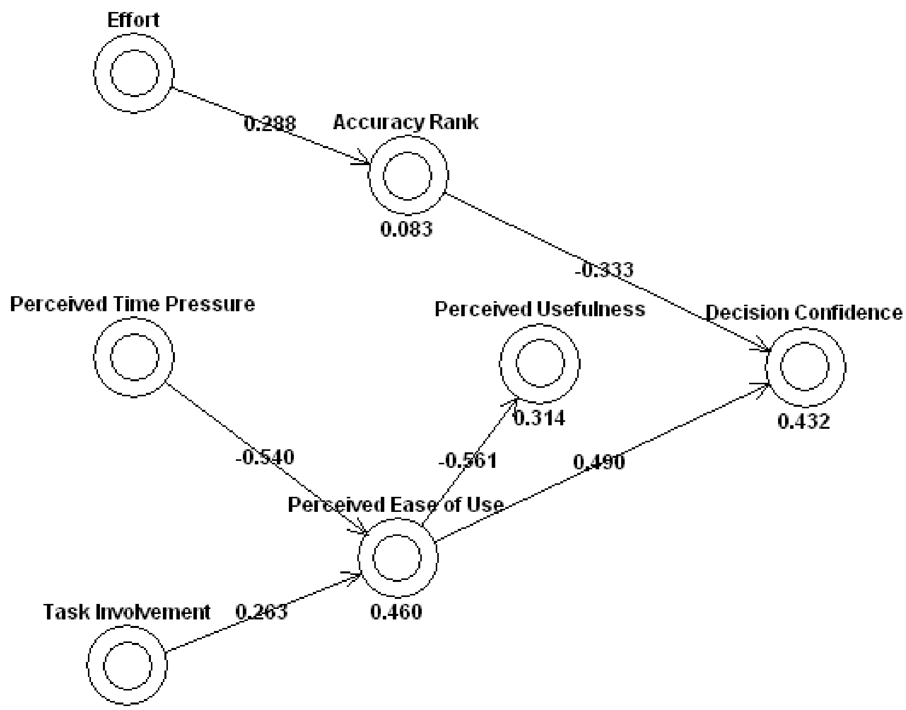


Figure 4. PS-PS-CM (treatment 2)

Decision Confidence was impacted by both Accuracy and Perceived Ease of Use, explaining 43.2% of variance. Interestingly, Perceived usefulness was not an influence. Little Accuracy was explained by effort (8.3%). Prediction P4 is supported but P3 is not.

For Treatment 3 (LEX-LEX-CM), the final SEM is the following:

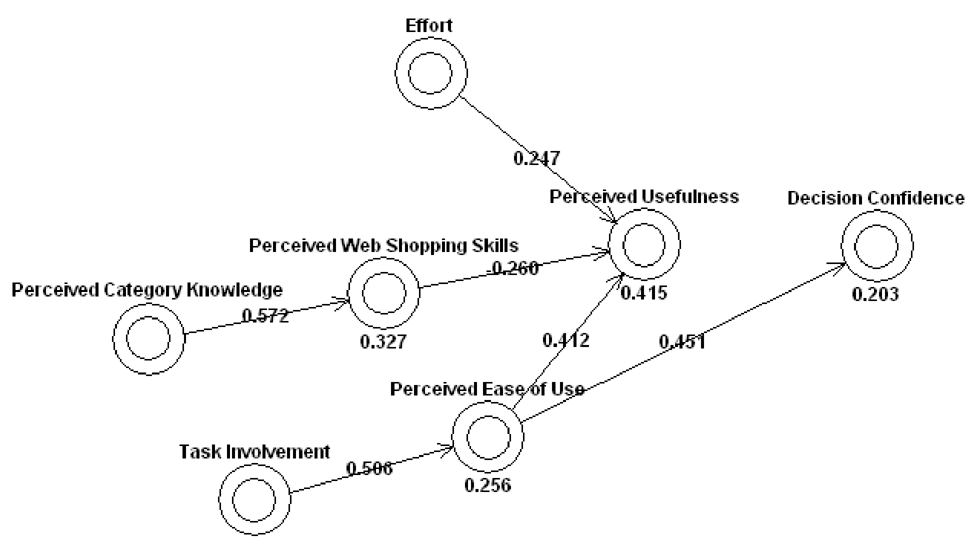


Figure 5. LEX-LEX-CM (treatment 3)

Decision Confidence was impacted by only Perceived Ease of Use, explaining 20.3% of variance. Interestingly, neither Accuracy nor Perceived usefulness is an influence. It appears that subjective variables have far better explanatory power than do the objective variables. Neither P5 nor P6 is supported.

For Treatment 4 (LEX-PS-CM), the final SEM is the following:

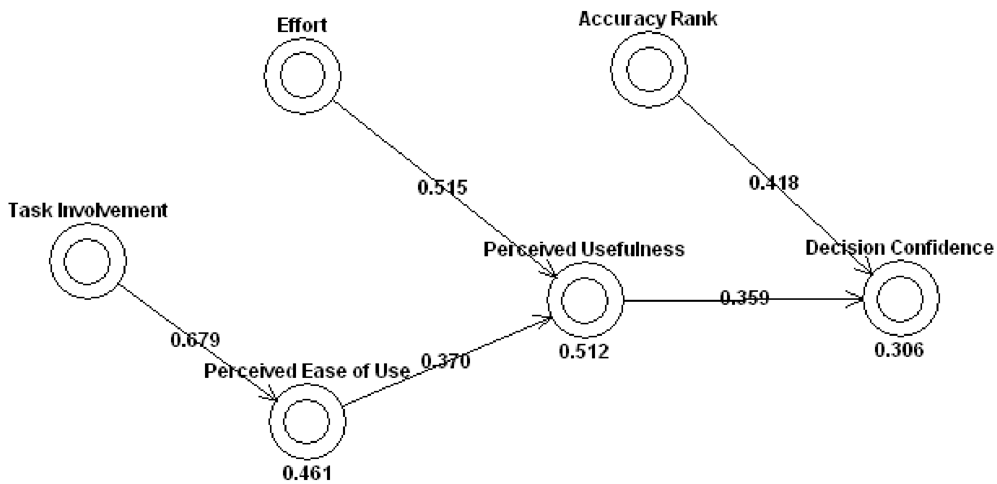


Figure 6. LEX-PS-CM (treatment 4)

Decision Confidence is impacted by both Accuracy Rank and Perceived usefulness, explaining 30.6% of variance. Interestingly, Effort influences Perceived usefulness rather than Accuracy. Together with Perceived Ease of Use, Effort accounts for 51.2% of the variance in Perceived usefulness. Both P7 and P8 are supported.

Table 2. Summary of Prediction Results

Prediction 1	Supported
Prediction 2	Supported
Prediction 3	Not Supported
Prediction 4	Supported
Prediction 5	Not Supported
Prediction 6	Not Supported
Prediction 7	Supported
Prediction 8	Supported

The Structural Equation Models (SEMs) are quite different for all four treatments. This suggests that subjects are highly sensitive to the particular decision aids, as well as their sequencing.

Summary of Results for the EAM-TAM Model

We obtained some insights from our Structural Equation Modeling, which we summarize here. The insights provide ideas for future research and managerial implications regarding both objective and subjective variables.

Overall, treatment does matter and the effort-accuracy model of decision making (Levin et al. 1998; Payne et al. 1996; Todd et al. 1992) is supported here. We also show that Accuracy can be a significant positive or negative influence on Decision Confidence (Kasper 1996).

We have made a possible contribution to the literature of Technology Acceptance in that we have found antecedents to PEU. The greater the Task Involvement, the greater the PEU. The greater the Perceived Time Pressure, the lesser the PEU.

Effort (time) spent by the customer, not just response time by the system, is a key mediating variable between the decision strategy offered and the accuracy of decisions made. Sellers should try to find out how much time the user wishes to spend, whether minimum, maximum, or moderate. This would help determine the decision aids to offer.

The variables in the TAM sub-model had a substantial impact on Decision Confidence, explaining more variance than that explained by Effort and Accuracy. This means that sellers do need to devote resources to perceptions and attitudes in addition to objective product attributes and concrete consumer behaviors.

Discussion of Findings and Contributions

We made several contributions in this research. We tested a number of Structural Equation Models to capture both direct and indirect impacts of experimental treatments on Decision Confidence. Effort, Accuracy, Perceived usefulness, and Perceived ease-of-use are all key variables leading to Decision Confidence.

Does decreased Effort or increased Accuracy or Technology Acceptance have a positive impact on Decision Confidence (DC)? Yes, Effort has an indirect impact on DC. Accuracy, Perceived-Perceived ease-of-use and Perceived usefulness have direct or indirect impacts on DC.

DA sequences found in PSEs do matter. Redesigning them can have significant impacts on decision confidence, through both Effort-Accuracy and Technology Acceptance variables. Our modeling supports several specific insights for web site designers and managers.

Decision Aids Matter

The four treatments, which combined PS and LEX in four different sequences, showed significant differences between two-stage treatments, three-stage treatments, and overall, comparing all treatments. We have shown that effort is the first consideration. Treatments impact effort directly, but accuracy only indirectly, through effort. The treatments, in effect, help subjects to decide how much effort to spend. Given the effort level constraint, subjects then maximize accuracy, This confirms prior research (Benbasat et al. 1996; Davis 1989; Todd et al. 1992).

Implications for Web Stores

PS-PS-CM is the most common Product Search Engine found on the Web today. It is a parametric search, often designated as appropriate for “power users,” who can mentally juggle several attributes and values simultaneously. It is also borne out in this research as having the strongest model for predicting Decision Confidence (variance explained = 43.2%). The second strongest model for predicting Decision Confidence is LEX-PS-CM (variance explained = 30.6%). This model can be characterized as “Learn and then be a Power User”. This is because LEX is identical to PS, except that LEX shows one attribute at a time, whereas PS shows all attributes simultaneously. These two treatments have in common that PS immediately precedes CM (average variance explained = 36.9%). These treatments instill confidence in the decision maker. The other two treatments have LEX immediately preceding CM, with an average variance explained of only 17.6%. These treatments seem to diminish confidence in the decision maker, perhaps because the learning of LEX has not solidified into the knowledge required of PS.

If web store managers and designers were to reconceptualize their PSEs as sequences of DAs, they may very well increase the decision confidence of their customers. Ideally, according to this research, the decision aids would detect the learning rate or knowledge level of the customer. By doing so, a particular DA could be recommended or imposed for each stage of decision making. The dynamic detection could be also tailored to the individual's customer profile, if available and reliable. Sales would likely increase, thus increasing the store's conversion rate of lookers to buyers.

Current Limitations and Future Research

There are several limitations in the current research that pertain to assumptions we had to make to make our problem tractable.

Conceptual Limitations

We assumed that consumers are rational, information-seeking, and utility maximizing, and that their utility weights are stable and can be elicited. We assumed that online consumers are driven primarily by cost-benefit tradeoffs, including the cognitive cost of search. For consumers who do not engage in active search, e.g., consumers who simply consult references, e.g., C|NET or Consumer Reports or reference groups, e.g., family or friends, for recommendations, our model does not apply. It might apply to the friends or family, the opinion leaders, who do engage in active search to make optimal choices. In other words, to study the opinion-followers, we could measure or manipulate additional contextual variables in our conceptual model.

Methodological Limitations

We assumed that undergraduate students were an adequate proxy for online consumers and that behavioral data combined with survey data would be sufficient. We assumed that the experimental method together with a realistic shopping situation would elicit authentic subject participation. We assumed that a random assignment of subjects to treatments equalized the capabilities and other characteristics of subjects. Since no subjects abandoned their shopping tasks or complained about lack of realism, we infer that using a true experiment was a good choice.

We verified statistically that our constructs had good reliability and validity (see appendix C). In the future, however, we may wish to measure effort in multiple ways, not only as time spent.

We assumed that behavioral data, from web logs and final choices, combined with survey data would be sufficient. Consequently, we did not record a verbal protocol of subjects' thoughts during the experimental tasks. Despite the possible reactivity of verbal protocol, some researchers have found that such a running commentary can be useful for glimpsing decision makers' verbalized processes as they happen (Payne 1976; Schkade et al. 1994; Todd et al. 1992).

Our databases were fixed in size, 241 cars, 242 computers, or 241 printers. This size represents a small web store. A larger database of products, an order of magnitude larger, would be more realistic. In other words, we started small and intend to scale our experiments in the future. Note that having product sets of this size is much more realistic for this task than those used commonly in the literature, e.g., choosing from among 5 jobs (Keller et al. 1987), 10 or 30 apartments (Todd et al. 1993), or 54 backpacking tents (Haubl et al. 2000).

Our experimental subjects were undergraduate students. We chose this population for convenience and because it is rather homogeneous in age, knowledge, skills, etc. For a more representative sample of web shoppers and greater statistical power, we should use a broader sampling frame and a larger sample.

Future Research

Knowing our current limitations, both conceptual and methodological, gives us ideas for what to investigate next. Besides addressing the specific assumptions we have made, we could also

- Sample from a population of purchasing managers. Are they more confident or accurate with their work budget or their own, personal money?
- Determine whether the user's attribute utility weights change during the decision making process? How do they converge? Can they be manipulated?
- Explore the tradeoff between relinquishing control to autonomous agents and retaining control with PSEs.

Finally, with actual purchase/non-purchase field data, we will extend the integrated EAM-TAM model beyond decision confidence, to model and predict purchase likelihood, which is of great interest to online sellers.

References

Banker, R.D., Charnes, A., and Cooper, W.W. "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management Science* (30:9) 1984, pp 1078-1092.

Beatty, S.E., and Smith, S.M. "External Search Effort: An Investigation Across Several Product Categories," *Journal of Consumer Research* (14:1 (Jun)) 1987, pp 83-95.

Benbasat, I., and Todd, P. "The effects of decision support and task contingencies on model formulation: A cognitive perspective," *Decision Support Systems* (17:4 (Aug)) 1996, pp 241-252.

Bettman, J.R., Luce, M.F., and Payne, J.W. "Constructive consumer choice processes," *Journal of Consumer Research* (25:3) 1998, pp 187-217.

Charnes, A., Cooper, W.W., and Rhodes, E. "Measuring the efficiency of decision making units," *European Journal of Operational Research* (2:6) 1978, pp 429-444.

Cooper, W.W., Park, K.S., and Yu, G. "IDEA and AR-IDEA: Models for dealing with imprecise data in DEA," *Management Science* (45:4 Apr) 1999, pp 597-607.

Davis, F.D. "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly* (13:3) 1989, pp 319-340.

Doll, W.J., Hendrickson, A., and Deng, X.D. "Using Davis's perceived usefulness and ease-of-use instruments for decision making: A confirmatory and multigroup invariance analysis," *Decision Sciences* (29:4) 1998, pp 839-869.

Einhorn, H., and Hogarth, R. "Confidence in judgment: persistence of the illusion of validity," *Psychological Review* (85) 1978, pp 395-416.

Fennema, M.G., and Kleinmuntz, D.N. "Anticipations of Effort and Accuracy in Multiattribute Choice," *Organizational Behavior and Human Decision Processes* (63:1 (Jul)) 1995, pp 21-32.

Fischer, G.W., Carmon, Z., Ariely, D., and Zauberman, G. "Goal-based construction of preferences: Task goals and the prominence effect," *Management Science* (45:8 (Aug)) 1999, pp 1057-1075.

Gefen, D., and Straub, D.W. "Gender differences in the perception and use of E-mail: An extension to the technology acceptance model," *MIS Quarterly* (21:4 Dec) 1997, pp 389-400.

Gefen, D., Straub, D.W., and Boudreau, M.-C. "Structural Equation Modeling and Regression: Guidelines for Research Practice," *Communications of the Association for Information Systems* (4:7) 2000, pp 1-77.

Gomory, S., Hoch, R., Lee, J., Podlaseck, M., and Schonberg, E. "E-Commerce Intelligence: Measuring, Analyzing, and Reporting on Merchandising Effectiveness of Online Stores," IBM T.J. Watson Research Center, Yorktown Heights, NY, pp. 1-11.

Good, C. "The Dotcom Survival Guide," Creative Good, New York, pp. 1-101.

Hagen, P., Weisman, D.E., Manning, H., and Souza, R.K. "Guided Search for eCommerce," Forrester Research, Inc., Cambridge, MA.

Hagen, P.R., Delhagen, K., Ardito, C., and Aber, A. "On-line Selling Tools," Forrester Research, Inc., Cambridge, MA.

Halme, M., Joro, T., Korhonen, P., Salo, S., and Wallenius, J. "A value efficiency approach to incorporating preference information in data envelopment analysis," *Management Science* (45:1 January) 1999, pp 103-115.

Haubl, G., and Trifts, V. "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science* (19:1 Winter) 2000, pp 4-21.

Joro, T., Korhonen, P., and Wallenius, J. "Structural comparison of data envelopment analysis and multiple objective linear programming," *Management Science* (44:7 Jul) 1998, pp 962-970.

Kasper, G.M. "A theory of decision support system design for user calibration," *Information Systems Research* (7:2 Jun) 1996, pp 215-232.

Keller, K.L., and Staelin, R. "Effects of quality and quantity of information on decision effectiveness," *Journal of Consumer Research* (14:2 Sep) 1987, pp 200-213.

Klein, N.M., and Yadav, M.S. "Context Effects on Effort and Accuracy in Choice - An Inquiry into Adaptive Decision-Making," *Journal of Consumer Research* (15:4 (Mar)) 1989, pp 411-421.

Kleinmuntz, D.N., and Schkade, D.A. "Information displays and decision processes," *Psychological Science* (4:4 (Jul)) 1993, pp 221-227.

Koufaris, M. "System Design and Consumer Behavior in Electronic Commerce," in: *Information Systems*, New York University, New York, 2000.

Laroche, M., Kim, C., and Zhou, L. "Brand familiarity and confidence as determinants of purchase intention: An empirical test in a multiple brand context," *Journal of Business Research* (37:2) 1995, pp 115-121.
Levin, I.P., Jasper, J.D., and Forbes, W.S. "Choosing versus rejecting options at different stages of decision making," *Journal of Behavioral Decision Making* (11) 1998, pp 193-210.
Lohse, G.L., and Johnson, E.J. "What You See is What You Choose: The Effect of Rearranging Product Information on Consumer Choice."
McQuivey, J.L., Delhagen, K., Levin, K., and Kadison, M.L. "Retail's Growth Spiral," Forrester Research, Inc., Cambridge, MA, pp. 1-16.
Mishra, S., Umesh, U.N., and Donald E. Stem, J. "Antecedents of the attraction effect: an information-processing approach," *Journal of Marketing Research* (30:August) 1993, pp 331-349.
Payne, J.W. "Task complexity and contingent processing in decision-making - an information search and protocol analysis," *Organizational Behavior and Human Decision Processes* (16:2 (August)) 1976, pp 366-387.
Payne, J.W., Bettman, J.R., and Johnson, E.J. *The Adaptive Decision Maker* Cambridge University Press, Cambridge, 1993, p. 330.
Payne, J.W., Bettman, J.R., and Luce, M.F. "When time is money: Decision behavior under opportunity-cost time pressure," *Organizational Behavior and Human Decision Processes* (66:2 (May)) 1996, pp 131-152.
Punj, G.N., and Staelin, R. "A model of consumer information search behavior for new automobiles," *Journal of Consumer Research* (9:4 (March)) 1983, pp 366-380.
Saad, G., and Russo, J.E. "Stopping criteria in sequential choice," *Organizational Behavior and Human Decision Processes* (67:3 (Sep)) 1996, pp 258-270.
Schkade, D.A., and Kleinmuntz, D.N. "Information Displays and Choice Processes - Differential-Effects of Organization, Form, and Sequence," *Organizational Behavior and Human Decision Processes* (57:3 (Mar)) 1994, pp 319-337.
Sujan, M. "Consumer knowledge: effects on evaluation strategies mediating consumer judgments," *Journal of Consumer Research* (12:June) 1985, pp 31-46.
Szajna, B. "Empirical evaluation of the revised technology acceptance model," *Management Science* (42:1 Jan) 1996, pp 85-92.
Todd, P., and Benbasat, I. "The use of information in decision making: An experimental investigation of the impact of computer-based decision aids," *MIS Quarterly* (16:3 (Sep)) 1992, pp 373-393.
Todd, P., and Benbasat, I. "An experimental investigation of the relationship between decision makers, decision aids and decision making effort," *INFOR* (31:2 (May)) 1993, pp 80-100.
Todd, P., and Benbasat, I. "The influence of decision aids on choice strategies: An experimental analysis of the role of cognitive effort," *Organizational Behavior and Human Decision Processes* (60:1 (Oct)) 1994, pp 36-74.
Tybout, A. "Consumer psychology," *Annual Review of Psychology* (45) 1994, pp 131-169.
Widing, R.E.I., and Talarzyk, W.W. "Electronic information systems for consumers: An evaluation," *Journal of Marketing Research* (30:2 May) 1993, pp 125-141.

Appendix A. Introduction to Structural Equation Modeling Using PLS-Graph v. 3, Build 279

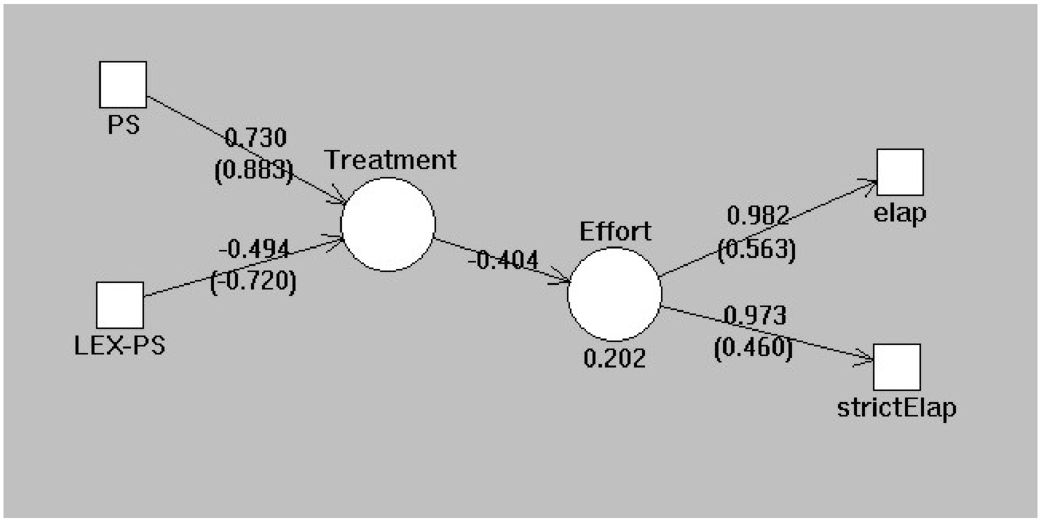
Structural Equation Modeling (SEM) combines a Structural Model, which shows the influence between Latent Variables (underlying factors), and a Measurement Model, which shows the indicators for each latent variable. Statistically, SEM combines Multiple Regression and Exploratory Factor Analysis. The goal is similar to that of Multiple Regression: maximize variance explained while ensuring that all linkages are statistically significant.

PLS can handle both reflective and formative latent variables (LVs). Reflective indicators should be highly intercorrelated. Such indicators should be measuring the same exact LV. Formative LVs, on the other hand are formed by multiple indicators, each representing only a component of the LV. Consequently, the indicators are related to each other, but they do not have to be highly inter-correlated.

Traditional SEM, using LISREL, Amos, or EQS, handles only reflective LVs. Traditional SEM cannot handle formative LVs. Consequently, we chose PLS-Graph. For a good review of SEM, including the appropriateness of each technique, see "PLS and Regression: Guidelines for Research Practice," (Gefen et al. 2000).

PLS allows one to examine several sources of statistical validity: Convergent and Discriminant validity of latent variables and the t-statistic/p-value of each indicator variable (loading and weight) as well as the t-statistic/p-value of each path in the graph.

How to interpret a PLS graph



In this excerpt from a PLS graph, we have two LVs, Treatment and Effort. Treatment is a formative LV, with two indicators, PS and LEX-PS, binary variables representing experimental treatments. Effort is a reflective LV, with two indicators, elap and strictElap. Elap is elapsed time over the entire task. StrictElap is elapsed time strictly within stages, which excludes inter-stage time. The notation for reflective and formative LVs is different.

For a formative LV, the number on top is the weight, analogous to a Regression beta coefficient. The number on the bottom, in parentheses, is the loading, analogous to a factor analysis loading. The weight for PS is 0.730, whereas the weight for LEX-PS is smaller and of opposite sign: -0.494. Their loadings are more similar: 0.883 and –0.720. (Sign is not important for loadings.)

For a reflective LV, the notation is the reverse. The loading is on top and the weight is on the bottom. Elap and StrictElap load very highly, 0.982 and 0.973 respectively, and their weights are also similar.

The path from Treatment to Effort has a weight itself, -0.404, which shows a strong negative relationship. To interpret the weight, backup to a treatment indicator. PS has a positive weight, so using the PS treatment strongly decreases Effort. LEX-PS has a negative weight, so using the LEX-PS treatment strongly increases Effort. (Negative * Negative = Positive) Finally, the number under a reflective LV shows the amount of variance explained, analogous to an R^2 in Regression; 20.2% of the variance in Effort is accounted for by the treatment variable.

Every weight and loading, whether for an LV indicator or for a path, has a t-statistic (and p-value). These are what we checked painstakingly, to arrive at the final models, in which all indicators and paths were statistically significant. We tried also to make sure the weights were statistically significant, which would represent a magnitude effect, not merely statistical significance.

Appendix B. Questionnaire Instrument

Task Involvement (Mishra et al. 1993)

3	I found the shopping tasks enjoyable.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Not at all … Very Much
4	I found the shopping tasks interesting.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Not at all … Very Much
5	I found the shopping tasks exciting.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Not at all … Very Much
6	I found the shopping tasks stimulating.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Not at all … Very Much

Product Category Knowledge, adapted from Sujan (1985)

7	How comfortable do you feel using computers?	○ ○ ○ ○ ○ ○ ○ Very Comfortable … Neutral … Very Uncomfortable
8	How satisfied are you with your current computer skills	○ ○ ○ ○ ○ ○ ○ Very Comfortable … Neutral … Very Uncomfortable
9	How knowledgeable are you regarding computers?	○ ○ ○ ○ ○ ○ ○ Very Comfortable … Neutral … Very Uncomfortable
10	For general computer tasks, how often do you become frustrated?	○ ○ ○ ○ ○ ○ ○ Very Comfortable … Neutral … Very Uncomfortable
11	I know _____ about computers, compared to the average computer user?	○ ○ ○ ○ ○ ○ ○ Very Comfortable … Neutral … Very Uncomfortable

Perceived Time Pressure, adapted from Payne et al. (1996)

12	I had enough time to perform the shopping tasks.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
13	I had enough time to perform the shopping tasks effectively.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
14	When performing the shopping tasks, I felt hurried.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
15	I could have made a better choice, if I had had more time.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree

Perceived Web Shopping Skills, adapted from Koufaris (2000)

16	I am skilled at using the Web for shopping.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
17	I know how to find what I want on the Web.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
18	I know more about Web shopping than most users.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree

Decision Confidence, adapted from Laroche et al. (1995)

19	What is your overall reaction to this computer?	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Do not like at all … Like very much
20	How likely is it that you would recommend this computer to the committee for purchase?	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Extremely unlikely … Extremely likely

Perceived usefulness, adapted from Davis (1989)

22	Using the decision tools in this web store enabled me to make a decision quickly.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
23	Using the decision tools in this web store improved my decision making.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
24	Using the decision tools in this web store improved my decision making efficiency.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
25	Using the decision tools in this web store improved my decision making results.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
26	In a real shopping situation, the decision tools in this web store would be useful.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree

Perceived Ease of Use, adapted from Davis (1989)

27	I found it easy to get the decision tools to do what I want it to do.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
28	My interaction with the decision tools was clear and understandable.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
29	I found the decision tools to be flexible to interact with.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
30	It would be easy for me to become skillful at using the decision tools.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree
31	I found the decision tools easy to use.	○ ○ ○ ○ ○ ○ ○ 1 2 3 4 5 6 7 Strongly agree … Strongly Disagree

Appendix C. Reliability and Validity of Survey-Based Constructs

All the survey-based variables have good reliability and validity, as shown in **Error! Reference source not found.**.

Table C1. Reliability and Validity of Survey-Based Constructs

Construct	Final Items	Cronbach Alpha (Reliability)	Average Factor Loading (Validity)
Task Involvement	3-6	0.894	0.822
Product Category Knowledge	7, 8, 9, 11	0.768	0.733
Perceived Time Pressure	12, 14, 15	0.675	0.726
Perceived Web Shopping Skills	16, 18	0.869	0.879
Perceived usefulness	23, 24, 26	0.777	0.791
Perceived ease-of-use	27-29	0.730	0.723
Decision Confidence	19, 20	0.710	0.890