International Journal of Information Technology & Decision Making
Vol. 6, No. 4 (2007) 723–750




# INTEGRATING VISUALIZATION AND MULTI-ATTRIBUTE UTILITY THEORY FOR ONLINE PRODUCT SELECTION

CHUREE THEETRANONT* and PETER HADDAWY†

*School of Computer Science and Information Management*
*Asian Institute of Technology, P.O. Box 4, Klong Luang*
*Pathumthani, 12120, Thailand*
**churee@cs.ait.ac.th*
*†haddawy@ait.ac.th*

DONYAPRUETH KRAIRIT

*School of Management, Asian Institute of Technology*
*P.O. Box 4, Klong Luang, Pathumthani, 12120, Thailand*
*donya@ait.ac.th*

Effectively selling products online is a challenging task. Today's product domains often contain a dizzying variety of brands and models with highly complex sets of characteristics. This paper addresses the problem of supporting product search and selection in domains containing large numbers of alternatives with complex sets of features. A number of online shopping websites provide product choice assistance by making direct use of Multi-Attribute Utility Theory (MAUT). While the MAUT approach is appealing due to its solid theoretical foundations, there are several reasons that it does not fit well with people's decision making behavior.

This paper presents an approach designed to better fit with people's natural decision making process. The system is called VMAP for Visualizing Multi-Attribute Preferences. VMAP provides on one screen both a multi-attribute preference tool (MAP-tool) and a product visualization tool (V-tool). The product visualization tool displays the set of available products, with each product displayed as a point in a 3D attribute space. By viewing the product space, users can gain an overview of the range of available products, as well as an understanding of the relationships between their attributes. The MAP-tool integrates expression of preferences and filter conditions, which are then immediately reflected in the V-tool display. In this way, the user can immediately see the consequences of his expressed preferences on the product space.

The VMAP system is evaluated on a number of factors by comparing users' subjective ratings of the system to those of a more traditional MAUT product selection tool. The results show that while VMAP is somewhat more difficult to use than a traditional MAUT product selection tool, it provides better flexibility, provides the ability to more effectively explore the product domain, and produces more confidence in the selected product.

*Keywords*: Visualization; multi-attribute utility theory; choice assistance; online shopping; usability evaluation.

## 1. Introduction

Effectively selling products online is a challenging task. Today's product domains often contain a dizzying variety of brands and models with highly complex sets of characteristics. Electronic products such as laptop computers, digital cameras, and mobile phones are good examples. The task of supporting online shopping is further complicated by the fact that there are many different types of shopping behavior that we may need to support, ranging from hedonic shoppers who shop because they enjoy the process to utilitarian shoppers who shop to find a product that best satisfies a particular need.[1] In this paper, we address the problem of supporting utilitarian shopping in product domains containing large numbers of alternatives with complex sets of features. We can think of a utilitarian shopping as the process of searching for a product that best fits the shopper's preference criteria. The process of choosing an outcome that best satisfies a set of preferences has been extensively studied in the field of Multi-Attribute Utility Theory (MAUT). The traditional MAUT approach to preference elicitation starts by describing each item in the choice set in terms of a number of attributes. The decision maker (the customer in this case) is then asked a series of questions concerning his preferences for various values of the attributes, as well as combinations of values. From the answers to these questions a preference order over the items is constructed and the most preferred item is the recommended. The value of this approach is that it reduces the daunting task of choosing among a large set of complex alternatives to the process of answers a series of relatively simple questions. A number of online shopping websites provide product choice assistance by directly implementing this procedure. Examples include Activebuyersguide[2] and General Motors' vehicle advisor.[3] While this approach is appealing due to its solid theoretical foundations, there are several reasons that it does not fit well with people's decision-making behavior. From the standpoint of shopping behavior, it is rather off putting for a customer to be forced to answer a long series of questions before he is permitted to see the list of available products. It is natural for customers who want to "enter the store" and look over the set of available products before and during the process of selecting the most suitable one. More generally, more recent research in the area of behavioral decision theory[4] indicates that human choice is an adaptive and constructive process rather than one of expressing existing well-defined preferences. This has a number of implications for product choice assistance tools.

- The system should reveal domain knowledge whenever possible.[5] Particularly in domains that are unfamiliar to the user, he may need to learn about the domain before he can construct preferences.[6] For example, suppose that I am in the process of designing a house. Before I can indicate how sensitive I am to price, I must have some general idea of the price range of a house of the size I want to build and what various materials cost. I also need to know things like the fact that wooden beams are more expensive than steel but look better. I may know that I want wooden beams and a stone floor, but to express a tradeoff I need to know the price of each.

- The system should not impose a rigid order for preference elicitation.[5] Users may differ in the part of the problem they wish to focus on first. In the process of constructing their preferences, users may also change their minds about earlier expressed preferences.
- The system should reveal the consequences of the user's partially expressed preferences at all stages. In eliciting preferences, we must choose some interface through which the user will communicate his preferences. Users may not fully understand the consequences of the information they communicate through this interface. Thus it is important to immediately display the consequences of expressed preferences to provide feedback to the user.
- The system should support incremental elicitation of preferences.[5,7] It may not be necessary for the user to provide enough information to identify an optimal alternative. After providing some preference information, the user may be able to easily select an optimal or near-optimal alternative from a list of highly ranked alternatives.

Much research has been done by developing efficient methods for navigating online product catalogs. Some work has concentrated on aiding decision making in the selection of consumer products and services by applying information visualization techniques. For example, FilmFinder[8] supports exploration of available films by using a simple selection tools and providing an overview with zoom and filter and details-on-demand. Other work applying information visualization techniques are InfoZoom,[9] MultiNav,[10] and EZChooser[11] that provide functional product selection tools and comparison tools. However, this technique is based on incremental query construction and does not provide feedback on the utility of product alternatives.[12] Other research applies MAUT for elicitation of preference to help the user search and select products. For example, Product scoring catalog (PSC)[13] provides a soft navigation by embedding preference elicitation mechanism called scoring tree to help the user select the most appropriate product. The scoring tree is used to evaluate and sort products. The high scoring products are more visible than the low scoring products. Other work implementing preference elicitation includes Automated travel assistant (ATA),[14] Apt Decision,[15] Interactive data exploration and analysis (IDEA),[16] and Smart Client.[17]

The integration of visualization and MAUT provides another alternative approach for online product selection. Research from Blythe[12] shows the first implementation of this approach in such a way that preference elicitation is combined with applying the current preference model to the set of alternatives to generate new suggestions. The implemented system is called visual exploration and incremental elicitation (VEIL). VEIL uses a 2D projection to display items with their scores. The use of 2D display makes VEIL's interface easy to understand, but it limits the amount of information that can be displayed. VEIL seems most suited to domains where items can be described in terms of relatively few attributes.

In this paper, we follow this approach and extend the implementation by addressing more issues described above, increasing the capability to comparing more attributes, and providing more flexible tools in one screen so that users can easily grasp an overview of the search process at all times. In addition, we perform extensive evaluation of our approach by measuring user satisfaction after using both systems that implement this approach and MAUT approach. Our system is called VMAP for Visualizing Multi-Attribute Preferences. VMAP provides on one screen both a multi-attribute preference tool (MAP-tool) and a product visualization tool (V-tool). The product visualization tool displays the set of available products, with each product displayed as a point in a 3D space. The V-tool can display three attributes using position and one using shape. Users can select and change the attributes being displayed at any time. By viewing the product space, users can gain an overview of the range of available products, as well as an understanding of the relationships between their attributes. Users can also zoom in on areas of interest. The MAP-tool integrates expression of preferences and filter conditions. Specifying filter conditions immediately eliminates products from the V-tool display. Preferences are displayed in the V-tool using color. In this way, the user can immediately see the consequences of his expressed preferences on the product space. By including the entire preference elicitation interface in one screen, the user can easily specify preferences in any order and change earlier expressed preferences. The V-tool permits the user to select and view the details of any item. So the user can easily select one or more products before he has completely specified his preferences.

We illustrate the concepts in VMAP using the laptop computer domain as a running example. This is a challenging domain because of the large space of products, the large set of attributes, the variety of different types of attributes, and the complex relationships among attributes.

The rest of this paper is organized as follows. Section 2 provides a brief review of the aspects of multi-attribute utility theory used in this paper. Section 3 presents the data and MAUT models used in VMAP. Section 4 presents the details of the V-tool and MAP-tool interfaces. Section 5 discusses the evaluation of the system, including design of the study and analysis of the results. Section 6 discusses related work and Section 7 presents conclusions.



## 2. Multi-Attribute Utility Theory

Multi-attribute utility theory[18] is concerned with the evaluation of the consequences or outcomes of a decision maker's decisions or actions, where outcomes are characterized with sets of features called *attributes*. The attributes will be denoted by $X_1, X_2, \ldots, X_n$, and outcomes by $x = (x_1, \ldots, x_n)$, where $x_i$ designates the value of attribute $X_i$. The outcome space $\Omega$ is just the Cartesian product $X_1 \times X_2 \times \cdots \times X_n$. We will often talk about subsets $Y$ of the set of attributes $X = \{X_1, X_2, \ldots, X_n\}$, and also refer to $Y$ and their complements $Z = X - Y$ as attributes. With respect to such a pair $(Y, Z)$, an outcome $x = (x_1, x_2, \ldots, x_n)$ can be written as $(y, z)$. For example, if $n = 5$ and $Y = \{X_1, X_3\}$, then $y = (x_1, x_3)$ and $z = (x_2, x_4, x_5)$.

When there is no uncertainty involved, in order to make decisions, an agent needs only express preferences among outcomes. If the decision maker's preferences satisfy some commonly accepted rationality constraints, then the preference relation, denoted by $\succeq$, can be captured by an order-preserving, real-valued *value function* $v$.

In order to reduce the complexity of obtaining preferences over items described by large sets of attributes, it is typical to make independence assumptions that permit a high dimensional value function to be decomposed into a combination of lower dimensional functions. Given a preference order $\succeq$ over $\Omega$, an attribute $Y \subset X$ is *preferentially independent* of its complement $Z$, or, for short, $Y$ is PI, if the preference $\succeq$ over outcomes that are fixed in $Z$ at some level does not depend on this level. The attributes $X = \{X_1, X_2, \ldots, X_n\}$ are *mutually preferentially independent* if every subset $Y$ of these attributes is preferentially independent of its complementary set. This property is important because it guarantees the existence of a particularly simple form of value function. If a set of attributes $X_1, \ldots, X_n$ is mutually preferentially independent, then there exists an additive value function

$$v(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} c_i v_i(x_i)$$

that faithfully represents the decision maker's preferences. The $v_i$ are called *sub-value functions* and specify the decision maker's preferences over values of the corresponding attribute $x_i$, while the *tradeoff coefficients* $c_i$ specify the relative importance of each attribute. In this paper, we will assume that preference are mutually preferentially independent and thus will work only with additive value functions.

## 3. Data and MAUT Models

We distinguish two types of attributes that can be used to describe items: interval and nominal.[19] The range of an *interval attribute* is linearly ordered with a distance measure defined on it. The range of a *nominal attribute* is simply a discrete set of values with no ordering and no distance measure. The types of attributes are important because they determine how data is represented in the visual structure and how preferences are represented and elicited. For example, interval attributes can be displayed on an axis and nominal attributes can be represented by differently shaped symbols on a graph.

Laptop computers can be described by a combination of interval and nominal attributes. Examples of interval attributes include processor speed $[500, 2000]$MHz, hard disk capacity $[5, 40]$GB, and amount of installed RAM $[32, 1024]$MB. Examples of nominal attributes include operating system {Linux, Windows XP, Windows 2000, Windows NT} and wireless network capability {yes, no}.

As mentioned previously, our MAUT model assumes an additive form for the value function. We normalize each sub-value function $v_i$ to have a range of values from 1 to 5. The coefficients $c_i$ specify the relative importance of each attribute. The specification of the sub-value functions depends on the kind of attribute. For interval attributes, we assume a linear function between the end points specified by the user. The end points are by default the lowest and highest possible values of the attribute, but can be narrowed if the user specifies filter conditions, e.g., a maximum price of 1500 dollars. We assume a particular directionality of preference for each attribute. So, for example, we assume users prefer less expensive laptops, all else being equal. For nominal attributes the user must specify his preferences for the various values of the attribute. For example, a user might indicate the desirability of different brands as Compaq = 3, Sony = 4, and Toshiba = 5.

We assign initial default constant values for the sub-value functions and tradeoff coefficients so that the system always has a complete specification of the value function and can always provide a ranking of the products. Sub-value functions for nominal attributes are all initialized to a constant value of three as are the tradeoff coefficients. Sub-value functions for the interval attributes are initialized to linear functions over the range of each attribute value.

## 4. The VMAP Interface

A guiding principle in the design of VMAP has been to display information concerning the product space, the elicitation process, and the result of the elicitation process in one screen. In this way, the user can always immediately see the consequences of his expressed constraints and preferences reflected in the item space.

### 4.1. *V-tool interface*

Figure 1 shows a screen shot of the 3D star field[8] type display. We have chosen to use a 3D rather than the usual 2D star field display due to the large number of products in the space. As can be seen from the figure, the display provides an initial overview of the entire product space. Having an overview is important because it helps reduce search time, allows the detection of overall patterns, and aids the user in choosing the next move.[20] The $X, Y$, and $Z$ axes in the figure represent processor speed, installed RAM, and hard disk capacity, respectively. The shapes of the points in the graph represent different laptop brands. The color of each point shows the score according to the currently specified value function. The 3D display allows users to clearly see the relationship between attributes. For example, they can see how screen size and weight are related. It also allows users to see clusters in the product space and to see the ranges of the attribute values. Due to the distortion the 3D perspective causes to the relative positions of products on the axes, we provide the option to view 2D projections along the $XY$, $XZ$, and $YZ$ planes. This is done via the buttons , , and , respectively. The 2D projection facilitates identifying extreme points, like the cheapest laptop. Since objects in a 3D representation of a large product space can often occlude one another, users can also grab and rotate the graph to view it from any angle.

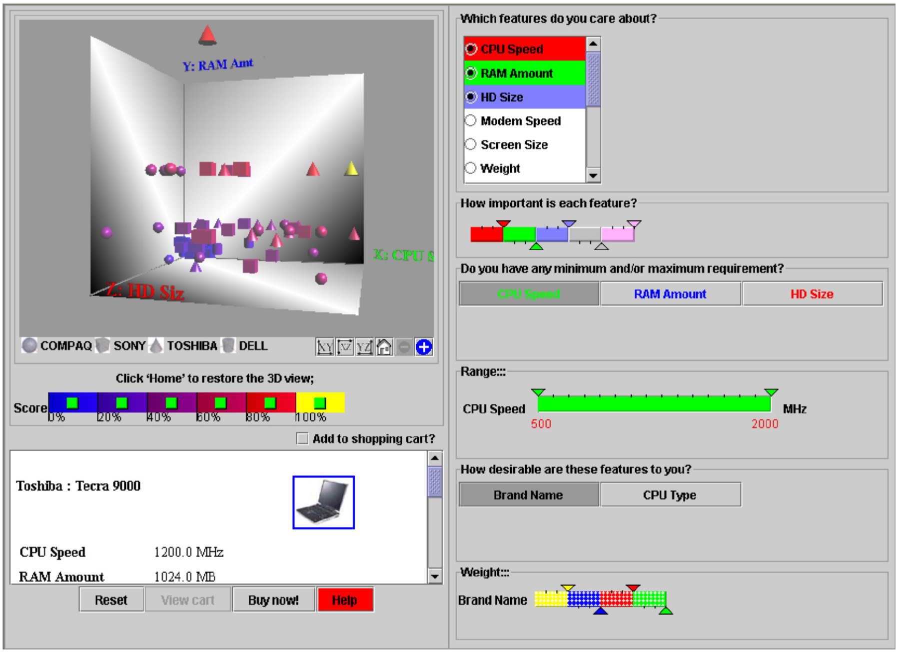


Fig. 1. Screen shot of VMAP applied to the laptop domain (for a VMAP tutorial that includes color screen shots, see www.cs.ait.ac.th/˜churee).

Since users may wish to inspect regions of the product space in various ways, the V-Tool supports several forms of drill-down. First, the user can view the details of any given laptop by right clicking on the point of interest, as shown in Fig. 2. The user can also zoom in or out and pan in a continuous manner by manipulating the mouse. The user can focus on certain attribute values by setting filter conditions. Filter conditions for nominal attributes are set with check boxes and those for interval attributes are set with slider bars. Setting of filter conditions is integrated with preference elicitation in the MAP-Tool. As filter conditions are changed, the items in the display disappear or reappear. Associated with filtering of interval attributes is a display expand/contract feature. Suppose that the total range of prices of laptops is 800 to 3000 dollars and that the user chooses to view only laptops that have a price range between 1200 and 1500 dollars. All those laptops that do not satisfy the filter condition will disappear from the display. If price is one of the axes of the graph, say the $x$ axis, then all the points in the graph will now occupy only a small vertical slice somewhere in the center of the $x$ axis. The user can use the expand button  to set the endpoints of the axis to the endpoints of the filter range, i.e., 1200 to 1500 dollars. In this way, the user can select a subset of the attribute range of interest and then inspect the items in this subset in more detail. The contract button  undoes the effect of the expand button.

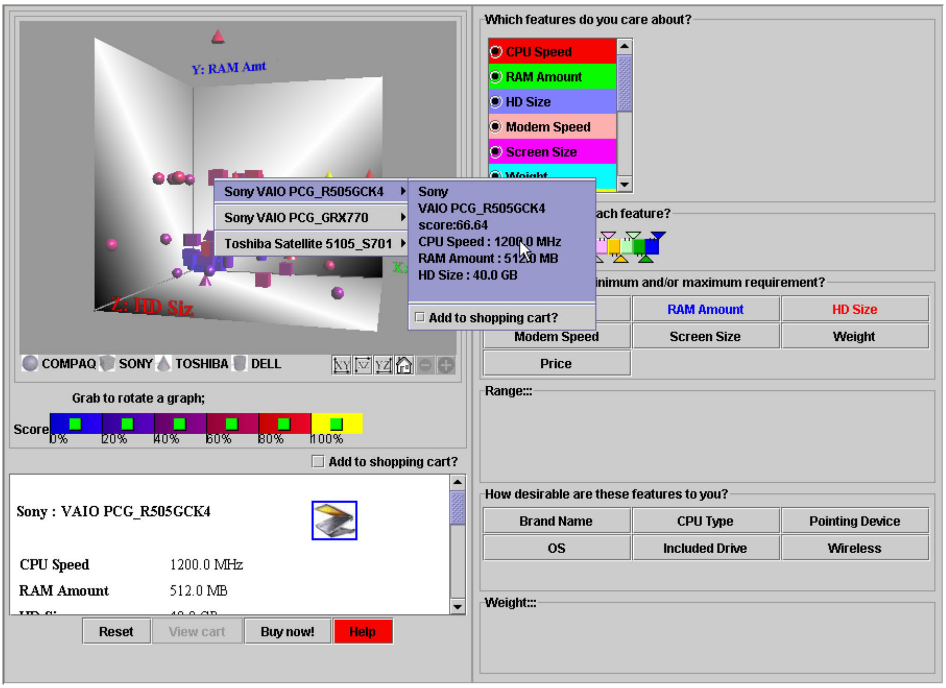


Fig. 2. Product detail viewed from pop-up menu. The menu shows three laptops since multiple products can occupy the same point in attribute space.

## 4.2. *MAP-tool interface*

The MAP-Tool interface integrates the functions of attribute selection, attribute value filtering, and preference elicitation. As shown in the upper right-hand portion of Fig. 1, users can select the attributes that are important to them. This helps to simplify and focus the elicitation process. Next users can specify filter conditions to indicate values of attributes they do not want to consider. For interval attributes the MAP-Tool provides slider bars, as shown in Fig. 3. For nominal attributes check boxes are used, as shown in Fig. 4.

For preference elicitation, we use the decomposition approach with direct rating scale in which users directly assign weights to a set of attributes.[21] We choose this approach because it is a relatively quick way for users to specify their preferences and has been shown to be quite accurate. Reducing the burden of preference elicitation is important in order to retain the engagement of the user. VMAP handles the fact that users may be uncertain about attribute weights by making it easy for the user to change weights and immediately see the effects on the ranking of the product space. An alternative approach is to explicitly model the uncertainty in attribute weights.[22] The technique is also flexible and meshes well with the use of filter conditions and with the visualization — users can easily experiment with

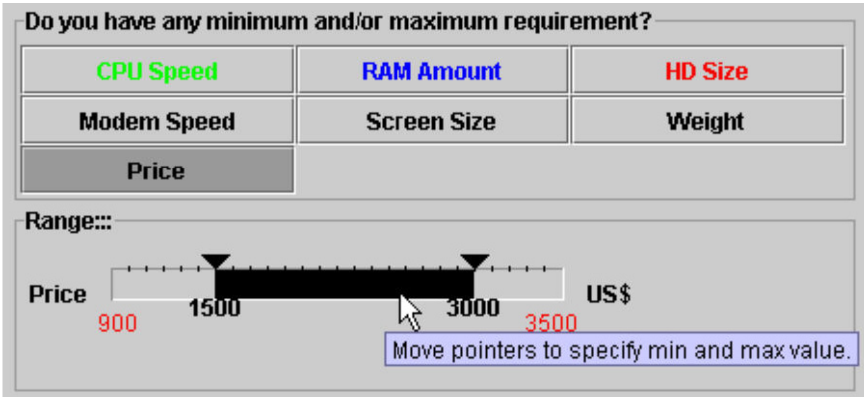


Fig. 3. A filtering slider for interval attributes.

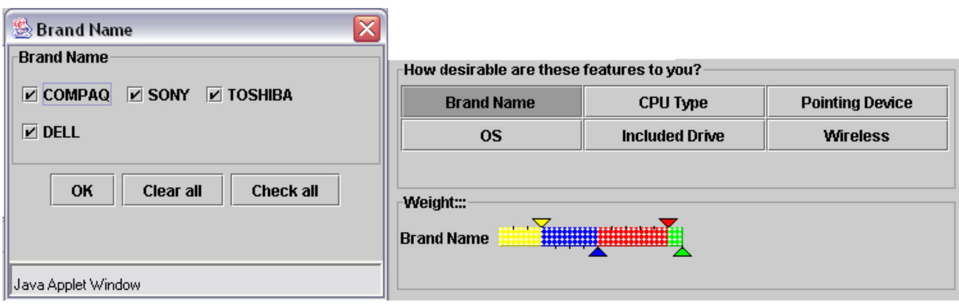


Fig. 4. A filtering check box for nominal attributes (left) and the associated proportional slider bar (bottom right). Selecting the attribute name (top right) brings up the check box and the slider bar.

various weights and see the results reflected in the V-Tool display. Coupling the MAP-Tool and V-Tool would be much more difficult and less natural with an elicitation technique like conjoint analysis that relies on answers to pairwise comparisons of sets of attribute values.

Preference elicitation consists of determining preferences over the values of each attribute and of determining the relative importance of each attribute. For interval attributes we simply assume a linear function over the values of the attribute between the end points set by the filter condition or the extreme values if no filter condition has been set. We assume either an increasing or decreasing function, depending on the attribute. For example, preference is decreasing in price while it is increasing in processor speed. To specify preferences over nominal attributes, as well as to specify relative importance of attributes, we use a common interface technique which we call a *proportional slider bar*. A proportional slider bar contains multiple sections differentiated by color. The position of the separation between each pair of adjacent sections is controlled by a slider arrow. By putting all the information

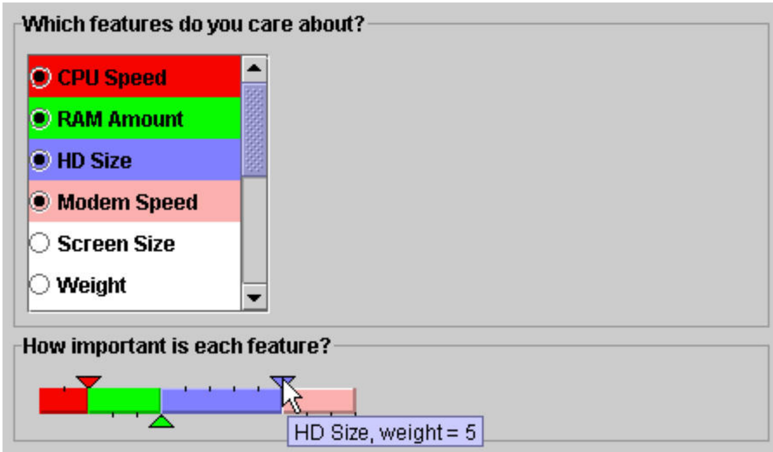


Fig. 5. The proportional slider for specifying the relative importance of attributes.



about attribute or attribute value weights in one slider bar, the user can easily see the values of the weights relative to one another. This technique also uses very little screen space, which is important since fitting the entire VMAP interface into one screen is a key design criterion of the system. The proportional slider for specifying relative importance of attributes is shown in Fig. 5. Each feature that the user specifies he cares about is added to the proportional slider. The colors of the regions on the slider bar are keyed to the colors of the attributes in the list above it. Each subslider represents the relative importance on a scale of 1 to 5.

To specify preferences over nominal attributes, users select the attribute from the list shown in the lower right-hand side of Fig. 1 and shown in detail in Fig. 4. The proportional slider then appears in the area below the list. In this case the user has chosen to specify the weights for the various values of the attribute Brand Name. If the user specified filter conditions that filter out some values of the attribute, these values do not appear in the proportional slider.

The user specified preferences are displayed and continuously updated in the V-Tool 3D display field using a color spectrum. Lowest rated items are shown in blue and highest rated items are shown in red. The top rated item is shown in yellow. A color spectrum bar directly beneath the display field permits the user to hide items with various ratings as shown in Fig. 6. Thus, for example, the user can choose to focus on only the highly rated items, or the items at the extremes, or to display only the top rated item.

## 5. Evaluation

The objective of our evaluation was to determine the benefits, if any, attained from adding visualization to MAUT versus the use of MAUT alone. To do this, we implemented a MAUT-based product selection system for the laptop domain.

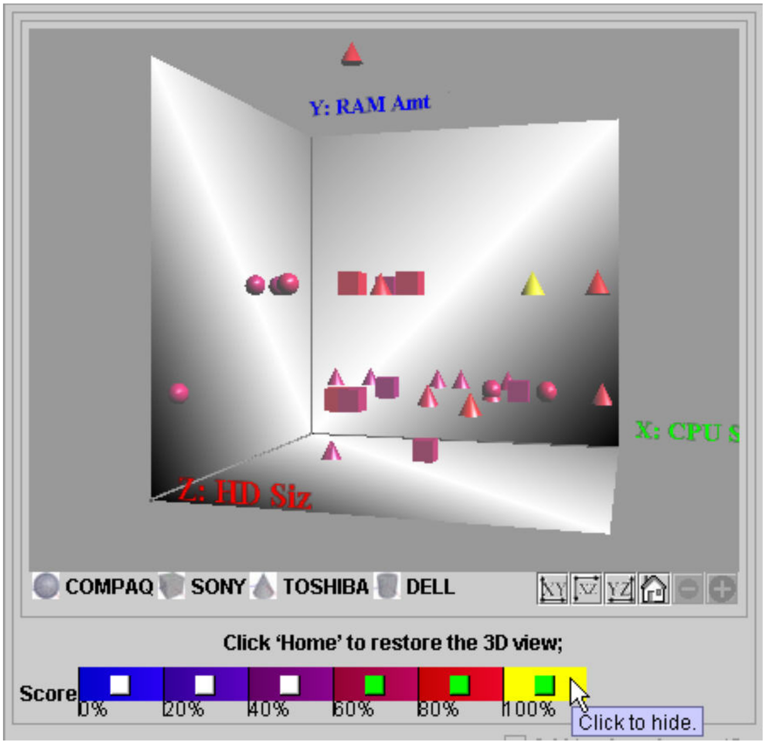


Fig. 6. A color spectrum bar represents the product score.

The system was designed following some of the more popular commercial online product selection tools. We implemented the MAUT system ourselves rather than using an existing system so that we could run the system locally and avoid any variability in response time due to Internet and server traffic. We had a group of participants who carry out a number of tasks on the MAUT-based product selection system and on our VMAP-based system. We then asked the participants to rate each system according to six categories: ease of use, speed of use, system interactability, system comprehensibility, user confidence, and overall satisfaction with the system. The questionnaire used was designed and then tested for construct validity and reliability on a pilot group of subjects. After reliability and validity were determined to be sufficient, a full-scale evaluation was conducted. Since the pilot study did not indicate any needed changes, the pilot and full-scale evaluations were conducted in the same manner.

### 5.1. *The experiment*

#### 5.1.1. *Comparison system*

The construction of the comparison system was based on a survey of existing online shopping websites that use MAUT. Most sites allow the user to indicate which attributes are of interest and then specify both preferences and filter conditions. Almost all sites use decomposition with direct rating scale to elicit preferences. Most sites display results in a tabular form on a page separate from the one used to specify preferences. To change preferences, the user must go back to the preference page and make the changes.

Accordingly, our comparison system first provides a query screen to specify filter conditions and preferences (Fig. 7). The user is prompted to specify filter conditions and/or preferences only for those attributes that he indicates he is interested in. The system then displays a ranked list of alternatives in a tabular form on a result page (Fig. 8). The user can go back to the query page and make changes to the specifications at any time. The user can select some items from the result page and see the details of the items displayed side-by-side in a tabular format (Fig. 9).



### 5.1.2. *Tasks*

In the introduction we listed four requirements for an effective product choice assistance tool. The tasks we selected for the participants were particularly designed to test these aspects of the two systems being compared. Each criterion was evaluated

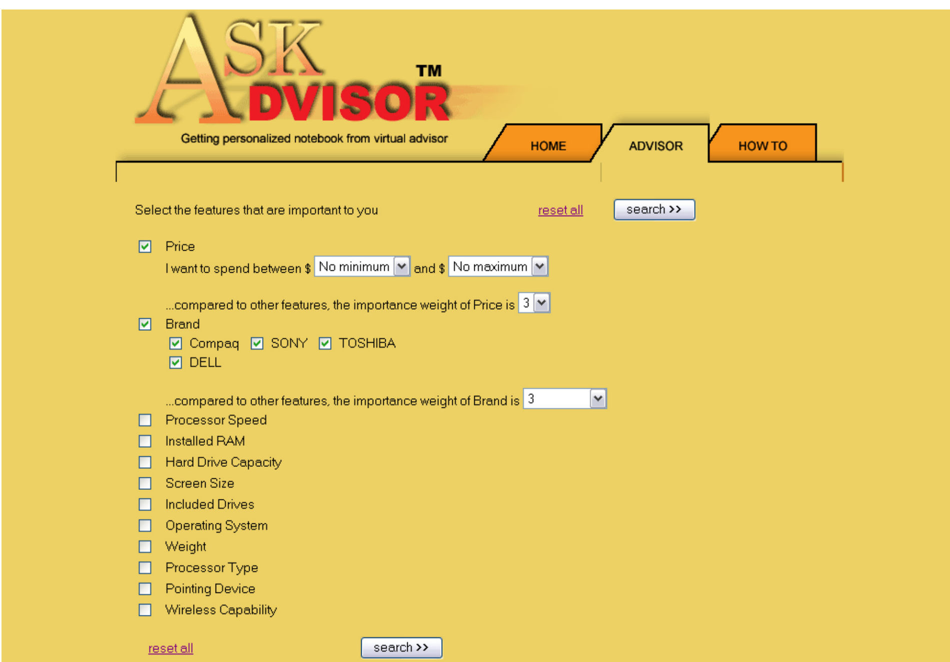


Fig. 7. The search screen of the MAUT system.


| Compare | Add to Shopping Cart | Score Percentage | Product Image | Brand and Model | Price | Processor Type | Processor Speed | Installed RAM | HD Capacity | Screen Size | Included Drive | OS | Weight | Pointing Device |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 100.00% | | Sony VAIO PCG_R505EZK1 | 900 US$ | Pentium III | 750 MHz | 64 MB | 10 GB | 12.1 in. | CD ROM (24) | Windows 2000 | 7.3 lb | Touchpad |
| | | 100.00% | | Toshiba Satellite 1415_S105 | 900 US$ | Celeron | 840 MHz | 256 MB | 20 GB | 15 in. | CD RW (24) | Windows XP | 7.6 lb | Touchpad |
| | | 98.12% | | Compaq Presario 906US | 949 US$ | AMD K6_24 | 1333 MHz | 256 MB | 20 GB | 14.1 in. | CD RW (24) | Windows XP | 6.5 lb | Touchpad |
| | | 98.12% | | Compaq Presario 2100US | 949 US$ | Celeron | 1600 MHz | 256 MB | 30 GB | 14.1 in. | CD RW (24) | Windows XP | 6.7 lb | Touchpad |
| | | 98.12% | | Compaq Presario 2100US | 949 US$ | Celeron | 650 MHz | 64 MB | 10 GB | 14.1 in. | CD ROM (24) | Windows 98 2Ed | 6.75 lb | Touchpad |

Fig. 8. The result list of the MAUT system.


| | | | | | |
|---|---|---|---|---|---|
| Score | 100% | 100% | 98.12% | 98.12% | 98.12% |
| Brand Name | Sony | Toshiba | Compaq | Compaq | Compaq |
| Model | VAIO PCG R505EZK1 | Satellite 1415 S105 | Presario 906US | Presario 2100US | Presario 2100US |
| Price | 900 US$ | 900 US$ | 949 US$ | 949 US$ | 949 US$ |
| CPU Type | Pentium III | Celeron | AMD K6 24 | Celeron | Celeron |
| CPU Speed | 750 MHz | 840 MHz | 1333 MHz | 1600 MHz | 650 MHz |
| RAM Amount | 64 MB | 256 MB | 256 MB | 256 MB | 64 MB |
| HD Size | 10 GB | 20 GB | 20 GB | 30 GB | 10 GB |
| Modem Speed | 56 bps | 56 bps | 56 bps | 56 bps | 56 bps |
| Screen Size | 12.1 in. | 15 in. | 14.1 in. | 14.1 in. | 14.1 in. |
| Weight | 7.3 lb | 7.6 lb | 6.5 lb | 6.7 lb | 6.75 lb |
| Pointing Device | Touchpad | Touchpad | Touchpad | Touchpad | Touchpad |
| Battery Type | LiIon | LiIon | LiIon | LiIon | LiIon |
| OS | Windows 2000 | Windows XP | Windows XP | Windows XP | Windows 98 2Ed |
| Warranty | 1 year | 1 year | 1 year | 1 year | 1 year |
| Speaker | y | y | y | y | y |
| Microphone | y | y | y | y | y |
| On Site Service | n | n | n | n | n |
| Wireless Capability | y | n | n | n | n |
| Drive Setting | external | internal | internal | internal | internal |

Fig. 9. The comparison screen of the MAUT system.

in a number of ways:

- The system should reveal domain knowledge whenever possible: This was evaluated by asking the user about attribute ranges and about relationships between attributes.
- The system should not impose a rigid order for preference elicitation: This was evaluated by asking the user to change previously specified preferences.
- The system should reveal the consequences of the user's partially expressed preferences at all stages: This was evaluated by asking the user to observe the consequences of changes in his specified preferences.
- The system should support incremental elicitation of preferences: This was evaluated by asking the user questions about laptops in the top 40% of the ranking and indirectly by determining how quickly and easily he could find an appropriate laptop given some very general guidelines.

Tasks were designed with varying levels of difficulty. We were careful to design tasks in such a way that all participants who passed the screening could complete them. To make the comparison fair, the set of tasks for the VMAP and MAUT systems differed only in the parameter values used. The following are the tasks that participants performed with the VMAP system.

T1: Suppose you are interested in the features — CPU speed, Installed RAM, Hard disk capacity, Price, Screen size, Weight, and Brand name. Specify importance weight of 1 for CPU speed, 2 for Screen size, and 5 for Installed RAM, Hard disk capacity, and Weight.

(a) What is the top-ranked laptop?
(b) How does changing the importance weight of 3 for Price to 5 effect the top-ranked laptop?
(c) What is the range of screen sizes?

T2: Now specify that you are interested in a laptop with at least a 14-inch screen and a weight of not more than 6 lbs.

(a) What is the price range of the candidate laptops?
(b) What is the weight of the lightest candidate laptop?
(c) Are there any laptops with a 14-inch screen and a weight of less than 5 lbs?
(d) What is the model of the cheapest laptop with a score of at least 60%?

Specify that you are not interested in Sony laptops.

(e) What is the brand of the laptop with the largest screen size?

Change the importance weight of Screen size to the highest value (5).

(f) What is the top-ranked laptop before and after the change?
(g) What is the price range of laptops ranked above 60% before and after the change?

T3: Suppose you have been given a budget of $2500 to purchase a laptop and you would like to buy a high-performance machine. Select the features that you care about and specify their relative importance. Also specify any restrictions on attribute ranges that you have. Continue until you have found a laptop that satisfies your requirements.

(a) Which laptop did you choose?

#### 5.1.3. *Procedure*

The study was conducted in the computer laboratory at the Information Technology Service Center, Chiang Mai University. Participants were screened to ensure that they satisfied the minimum requirements to use the VMAP system. The requirements included the ability to see colors, some experience in using an Internet browser, and an understanding of the attributes used to describe laptops. One hundred subjects were recruited, but only 80 subjects met the three requirements and were selected to participate. The evaluation had a within-subject design, with each subject using both systems to perform all three tasks. In order to compensate for learning effects,[23] half of the subjects first used the MAUT system and the other half first used the VMAP system.

At the beginning of each session, participants were briefly introduced to the purpose of the study and given overview of the tasks to perform. Participants were then provided with online tutorials on the use of the VMAP and MAUT systems and given 45 minutes to go through the tutorials. They were then given 15 minutes to practice using each system before beginning their tasks. We opted for the use of an online tutorial rather than human instruction because it guaranteed that each participant received exactly the same instruction in the use of the systems.

Next, participants were introduced to the evaluation tasks. Two different but similar sets of task questions were used to avoid repetition. Participants wrote their answers to each task performed in order on the answer sheet. After finishing the tasks for both the VMAP and MAUT systems, participants completed the questionnaire summarizing their experience. Participants rated each system in six categories: ease of use, speed of use, system interactability, system comprehensibility, user confidence, and overall satisfaction with the system. The questionnaire contained 30 questions, with 2 to 7 questions in each category. The questionnaire is included in Appendix A. Each participant took about 3 hours to complete the study and was paid 100 baht (approx. $2.50) for participating.

### 5.2. *Questionnaire construction*

A number of well-known questionnaires exist for general evaluation of software usability. They include CSUQ,[24] PSSUQ,[25] PUEU,[26] SUMI,[27] USE,[28] and SUS.[29] We used questions from these questionnaires for those criteria that our evaluation had in common with them. For other criteria, we developed our own questions.

5.2.1. *Procedure*

We selected items for our questionnaire that would evaluate whether the VMAP system achieved its stated design objectives. These included: reducing the time and effort for finding the best laptop; providing feedback to the users concerning the effects of their specific requirements and preferences; helping users understand the relationships between product attributes; giving users confidence in their selection of a product from among the available alternatives; and achieving a high overall user satisfaction. For those factors that had been included in previously published questionnaires, we used the corresponding questionnaire items. Thus we selected item numbers 1, 5, 11, 25, 26, 27, 30 from CSSUQ, item numbers 2, 6, 8, 9 from PSSUQ, item number 3 from PUEU, item number 4 from USE, item number 21 from SUS, and item number 29 from SUMI. The remaining 15 items were newly constructed. In the end, we had a pool of 30 items in our questionnaire in categories according to the above objectives. (See Appendix A for the content of questionnaire items).

We chose to use a 7-point Likert scale with strongly disagree (−3), neither agree nor disagree (0), and strongly agree (3) because it has been shown to strike a good balance between reliability and discriminative burden on the respondent.[24,30] Since all participants were Thai, the questionnaire was translated into Thai to minimize any misunderstanding of the questions. We ran a pilot study to test the reliability and validity of the questionnaire. The questionnaire was tested on 42 participants following the procedure described in Sec. 5.1.3.

5.2.2. *Factor analysis and construct validity test*

Factor analysis is used to detect structure in the relationships among questions and to determine the number of factors. We performed factor analysis by first producing a scree plot, which showed a point of discontinuity at six components, indicating the questionnaire contains six factors.

The varimax-rotated factor pattern discloses the structure of the subscales. An item is assigned to a factor if the value of factor loading on that item is high. If it is not highly loaded on any factor, we have to consider which factor the item should be assigned to based on the content of the item. Of the 30 items, six items (7, 12, 14, 24, 27, and 19) were not highly loaded on any factor and were thus assigned to factors based on their content. (See Appendix A for the contents of questionnaire items). We then perform a reliability test to validate the assignments of items to factors (described in the next section). The result of the factor analysis is shown in Table 1. The definitions of the factors are shown in Table 2.

5.2.3. *Reliability test*

The most common way to estimate the reliability of summative scales is with Cronbach's coefficient alpha.[30] Coefficient alpha[32] is a measure of the internal

Table 1. Result from factor analysis showing each factor with items.

| Factors | Item No. |
|---|---|
| Speed of use | 4, 5, 6, 7, 10, 16, 28 |
| Ease of use | 1, 2, 3, 20, 21 |
| User confidence | 9, 19, 23, 24, 27 |
| System comprehensibility | 12, 13, 14, 17, 18 |
| System interactability | 11, 15 |
| Satisfaction with the system | 8, 22, 25, 26, 29, 30 |

Table 2. Definitions of the factors.

| Factors | Definitions |
|---|---|
| Ease of use | The degree to which a person believes that using a particular system would be free of effort[26] |
| Speed of use | The degree to which a person is satisfied with the time it takes to use a particular system |
| System interactability | The degree to which the system is easy to interact with |
| System comprehensibility | The degree to which the information the system provides is easy to comprehend |
| User confidence | The degree to which a user is confident in the results presented by the system |
| Satisfaction with the system | The degree to which the system is pleasant to use, so that users are subjectively satisfied when using it[31] |

Table 3. Results from reliability analysis with Cronbach alpha.

| Factors | Cronbach Alpha |
|---|---|
| Speed of use | 0.9134 |
| Ease of use | 0.8525 |
| User confidence | 0.8180 |
| System comprehensibility | 0.8141 |
| System interactability | 0.7497 |
| Satisfaction with the system | 0.8833 |
| Overall | 0.9486 |

consistency of a scale and ranges from 0 (no reliability) to 1 (perfect reliability). A high alpha indicates that the items share the same objective and thereby are measuring the same underlying property. The literature on software evaluation generally considers a value above 0.70 as indicating acceptable reliability.[33] The overall reliability of the questionnaire is determined by calculating the Chronbach alpha over all items in the questionnaire, while the reliability of each factor is determined by calculating the Chronbach alpha over only the items in that factor. Table 3 shows the reliability values for the six factor subscales and the overall summative scale. Since the values are all sufficiently high, it was not necessary to delete any items or factors.

### 5.3. *Statistical analysis*

The hypothesis we wished to test was whether subjects would score VMAP higher than the MAUT system in each of the six categories. In addition, we analyzed the subject demographic data to determine if there was any correlation between subject background and the scoring patterns. Since the experimental procedure in our pilot study was identical to that of the full-scale evaluation, we were able to combine subjects from the pilot study (42) with subjects from the full-scale evaluation (80), giving a total of 122 subjects for response analysis.

We used the Yamane's simplified formula,[34] a well-known formula used in many usability studies, to statistically measure the sample size needed for our study. We first determined a large number of users that our evaluation would affect. We assumed 20,000 users as a finite number of population size. The sample size for $\pm 10\%$ precision levels with 95% confidence level is 100. When we increased the population size to 25,000; 50,000; 100,000 or more than 100,000; the sample size was constantly 100. We needed 100 users for a number of sample size in our study. Therefore, 122 users were sufficient to evaluate the statistical results.

#### 5.3.1. *Results comparing VMAP and MAUT*

The results were analyzed for each category of data: speed of use, ease of use, user confidence, system comprehensibility, system interactability, and satisfaction with the system. Paired samples $t$-test was used for hypothesis testing of user satisfaction scores for VMAP and MAUT systems. We accepted results at 99% confidence interval ($p = 0.01$) with significance (sig.(2-tailed)$/2 < \alpha$; $\alpha = p$) and $t > 0$.

Figure 10 shows the mean scores of VMAP and the MAUT system over the six factors. The hypothesis that users rated the VMAP system higher than the MAUT system was confirmed in all categories except ease of use at $p = 0.01$, as shown in Table 4. Users scored VMAP highest for comprehensibility and interactability, two of the primary design criteria. Users scored the MAUT system highest for ease of use, although the scores of the MAUT system and VMAP are almost identical in that category, with a difference of only 0.14. Over the remaining categories, the differences between the mean scores ranged from 0.71 for satisfaction with the system to 1.37 for system comprehensibility. In addition, while the overall rating combining all categories was high for both systems (VMAP = 5.75 and MAUT = 5.0), VMAP was rated higher, with the difference between the overall ratings statistically significant at $p = 0.01$.

We interviewed some of the subjects in order to determine the reasons that they rated one system higher than the other. A summary of the answers for each category follows:

(i) Speed of use: Subjects felt that they could rapidly modify their searches and obtain rapid feedback from the 3D graph of the VMAP system.

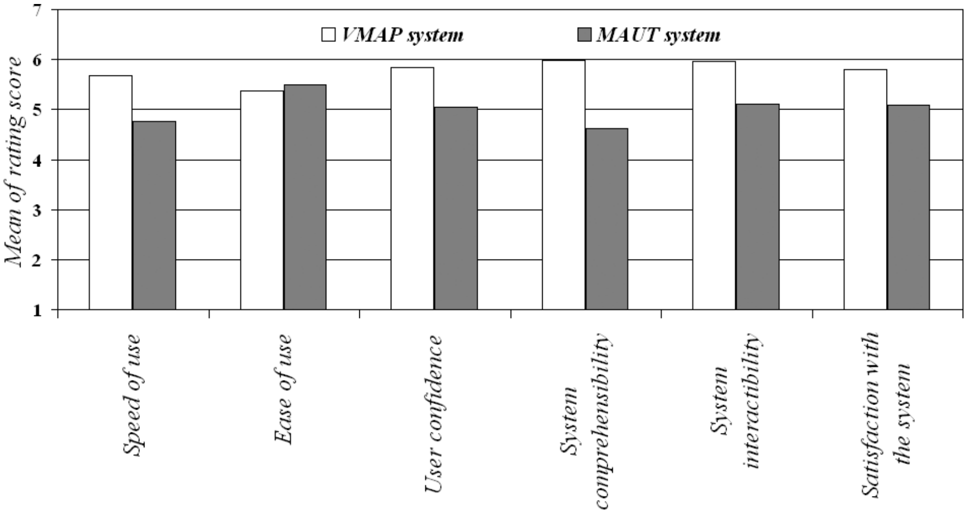


Fig. 10. Average user subjective satisfaction in a full-scale evaluation.

Table 4. Statistic results from paired samples $t$-test at $p = 0.01$.

| Factors | Mean Difference (VMAP − MAUT) | $t$ | df | Sig.(2-tailed) |
|---|---|---|---|---|
| Speed of use | 0.9145 | 4.348 | 121 | 0.000 |
| Ease of use | −0.1361 | −0.853 | 121 | 0.396 |
| User confidence | 0.7984 | 5.185 | 121 | 0.000 |
| System comprehensibility | 1.3672 | 7.465 | 121 | 0.000 |
| System interactability | 0.8566 | 4.463 | 121 | 0.000 |
| Satisfaction with the system | 0.7090 | 4.336 | 121 | 0.000 |
| Overall | 0.7505 | 4.709 | 121 | 0.000 |

(ii) Ease of use: Subjects said that they were more familiar and had more experience with the interface of MAUT system than the interface of VMAP system. The VMAP interface required some learning and practice. Users believed that if they became more familiar with the VMAP interface it would be a better system for selecting products.

(iii) User confidence: Participants said that the color spectrum bar representing score percentages was very useful. It helped them easily cluster groups of highly recommended laptops and kept them focused on the group of most interesting laptops. They were also able to see the top-ranked laptop among other alternatives. These factors increased their confidence in the results.

(iv) System comprehensibility: Participants said they could easily see the relationships between attributes (e.g., CPU speed vs. Price) directly from the 3D comparison graph. The VMAP system supported comparison of groups of laptops and provided a simple comparison on multiple attributes. In contrast, users found it difficult to see the relationships between attributes and to compare groups of laptops with the tabular format of the MAUT system.

(v) System interactability: Subjects liked having the entire interface contained in one screen and being able to see the results of their actions immediately reflected in the V-Tool display. Users felt that such tight coupling supported rapid exploration and helped them to recover easily and quickly when they made errors in specifying their preferences. This was in contrast to the MAUT system in which results were displayed on a separate page and users had to toggle between pages to change settings and view the results.

(vi) Satisfaction with the system: Subjects described the VMAP interface as more interesting and attractive than that of the MAUT system.



### 5.3.2. *Demographic analysis*

Demographic information collected from the participants included sex (male, female), age (below $25, 25$ to 40), education level (high school, Bachelor degree, Master degree), specialty of education (science and technology, social science), and occupation (student, government employee, business owner). The demographic data shows that 80 participants were male and 42 participants were female. Ninety-one participants were below 25 years old, and 31 were between 25–40 years old. Most participants had a Bachelor degree (65). Forty-one participants had Masters degrees and 16 studied only through high school. Most of them had studied science and technology (112) and ten participants had studied social sciences. There were 99 students, 12 government employees, and 11 business owners. We analyzed the dependence of the rating in each of the six categories on these demographic factors.

Independent samples $t$-test and ANOVA were used for analysis. We accepted results at 99% confidence interval ($p = 0.01$) with significance (sig.(2-tailed)/2 $< \alpha$; $\alpha = p$) and $t > 0$. We found the following statistically significant results. In the four categories ease of use, user confidence, system interactability, and satisfaction with the system the rating did not depend on any of the demographic factors. The factor speed of use was found to depend on occupation, with the biggest difference between the ratings given by students and by government employees. The mean student rating for VMAP on this factor was 5.59 and for MAUT was 5.02, while the mean government employee rating for VMAP was somewhat higher at 5.80 and for MAUT was only 3.21. The last factor, comprehensibility, was found to depend on the level of education, with the biggest difference between the ratings given by participants with a Bachelor degree versus a Master degree. The mean rating given to VMAP by participants with a Bachelor degree was 5.80 and for MAUT was 5.24. Participants with a Master degree had a much larger gap between the ratings given to VMAP and MAUT on this factor, with the mean rating for VMAP at 6.31 and for MAUT at only 3.86.

Combining data from the six categories into an overall satisfaction rating, analysis shows that the rating does not depend on any demographic factors. So we can conclude that overall the ratings of the participants depend more on the VMAP and MAUT systems themselves than on the demographic factors.

## 6. Related Work

Much research has been done on effective methods for navigating online product catalogs. Most of this research has concentrated on applications that use various information visualization techniques to support search and navigation. The main focus of the work is on aiding decision making in the selection of consumer products and services.

Ahlberg and Shneiderman developed FilmFinder,[8] which helps users explore a film database in a 2D space. FilmFinder supports exploration of available items, providing an overview with zoom and filter and details-on-demand by using a simple selection tool. Several important features of this work have been incorporated into the design of VMAP. First, query components (e.g., sliders and buttons) are used as filters to reduce the number of items in the result set. Second, a two-dimensional scatterplot called a Starfield display, supports selection and zooming for large item sets. Third, tight coupling means that every output is also a candidate for input, resulting in efficient use of screen space and focusing users' attention on a single location for collecting information.

Callahan and Koenemann present a table-based approach as demonstrated in a system called InfoZoom.[9] InfoZoom displays tables with attributes as rows and objects as columns. The column is expandable depending on the number of products in the result set. This approach can suffer from information cluttering when the number of products and product attributes becomes large.

Lanning *et al.* proposed MultiNav[10] which implements a parallel coordinate approach. MultiNav provides a visualization interface for multidimensional data exploration and browsing tasks. The primary interface for exploration is a sliding rod. Users can browse for items that are near a given item in any of the dimensions that matter to them and uncover hidden relationships in the data set. Due to the linear data representation, comparing several products can be difficult. For example, if the user wants to compare memory size in two price ranges for digital cameras, he or she must specify the first price range by moving a sliding rod to the specified position and determine the memory size. The user must then repeat the task for the second price range, and then compare the memory sizes. The difficulty of comparison increases with the number of attributes and alternatives.

Wittenburg *et al.* introduced a multi-dimensional visualization and interaction technique based on parallel bargrams and dynamic querying approaches. Their EZChooser system uses bargrams for displaying information. A bargram is derived from tipping over the columns (bars) of a histogram and laying them end-to-end, ignoring any null bins, and showing the relative count through the relative width of the bars.[11] EZChooser provides a good tool for visual data exploration, but does not adequately support the task of comparison when a large number of attributes and alternatives are present.

Other research addresses flexible elicitation of preference, or combining visualization with preference elicitation, or both.

Linden *et al.*[14] propose an iterative approach to preference elicitation, which they implement in the Automated Travel Assistant (ATA) system, that helps users to select airline flights. The ATA initially elicits minimal preference information from the user and then supplements this with default information to construct a complete value function. The system assumes that preferences are additive independent. After obtaining initial preference information, the ATA presents the user with a set of five candidate solutions, chosen from among the top-ranked flights. In selecting this set, the system uses two primary heuristics to aid the user in exploring and understanding the space of alternatives. The first is to display extrema, which in the flight domain means displaying the cheapest trip and the best non-stop trip. The other heuristic is to display solutions that differ sufficiently from one another. The difference between two solutions is simply defined as the sum of the differences between their attribute values. After being presented with the list of candidate flights, the user is free to adjust the sub-value functions and the weights in the additive value function. The ideas of incrementally eliciting preferences and of helping the user to explore and understand the space of alternatives are common to ATA and VMAP, although the particular ways in which they are implemented is different. ATA iteratively presents small sets of solutions, while VMAP provides a constant overview of the product space.

Stolze[13] presents a product scoring catalog (PSC) using a soft navigation. When using soft navigation, the user expresses preferences of various strengths for product features, and the stated preferences are used to evaluate and sort products. The product alternatives are displayed in a way that higher-scoring products are more visible than lower-scoring products. PSC consists of three main parts. First, the product list shows the offered products with price and overall scoring. The user can view and compare product alternatives in the product list. Second, the scoring tree describes how products are evaluated. Scoring tree can be one of four different types: mix rule(M), scoring rule(R), scoring function(F), and product features that are referenced by the rules. Scoring rules and scoring functions describe how to score products with respect to a base preference, and mix rules describe the relative importance of base preferences. In addition, scoring tree displays product features and their values for the product selected in the product list. Third, the editing panel is used to indicate relative importance values of product features. PSC and VMAP have a number of similarities but VMAP adds the ability to view the product space at a high level and then zoom in on portions of the space to examine them in more detail.

Shearin and Lieberman[15] present a system called Apt Decision for helping users to choose apartments. The design of Apt Decision emphasizes flexibility of the preference elicitation process. The system starts with a default preference model and then allows the user to provide a small number of criteria in the initial interaction. The system then displays a ranked list of apartments, with each apartment listing containing far more features than in the initial user query. The user can then select any of these features as being desirable or undesirable. In this way, the user can incrementally discover features of importance and incrementally specify his preferences. Features are added to the preference model by dragging and dropping them into slots. The slots with features are constantly displayed on the same screen as the apartment ranking so that the user always has an overview of the current preference model. In addition to dragging and dropping features, the user can specify preferences by choosing between presented pairs of apartments. Differences between the pair are then translated into modifications to the preference model. The user can at any time remove features from the preference model or change the weights of features. While Apt Decision provides a flexible interface for construction of preferences, it does not support the user in learning about the relationships among attributes.

Carenini and Moore[16] propose combining interactive data exploration and analysis (IDEA) with a multimedia explanation for product search and selection. IDEA systems provide environments that include visualization and interaction to enable users to explore a set of data or items. They permit users to control the level of detail, eliminate part of the information, and create new information by transforming the original information. Carenini and Moore argue that IDEAs are an appropriate decision making framework for computer- and graphics-literature users, particularly when the decision is the selection of an object out of a large set of possible alternatives and the selection is performed with respect to the object values over a set of attributes. In addition to the issue of helping the user find an optimal alternative, Carenini and Moore are interested in issues of decision confidence and satisfaction, and of flexible use of information to make a decision. They point out that even users who are comfortable using IDEA systems may need assistance when they encounter situations they do not expect or understand. For this they propose integrating a multimedia explanation facility. They illustrate their ideas with a system for helping users choose houses. The system provides a 2D visualization of attributes such as location, price, number of rooms, and lot size. It is capable of generating explanations to questions such as "Why are these houses much cheaper than those houses?"

Motivated by the belief that the preference elicitation process should be an integral part of the search process, Pu and Faltings[17] advocate an example-based approach to product search and selection. Their SmartClient system supports example-based search along with tradeoff analysis. A user starts the search by specifying a query including any number of preferences. Based on this initial preference model, the search engine finds and displays a set of matching results. The user then either accepts a result, or takes a near solution and posts critiques to that result. Critiques are small revisions to the current preference values. A critique may involve asking for a cheaper product or a better quality product. In addition, SmartClient permits the user to modify the relative importance of a preference, thus enabling the user to perform tradeoff analysis while searching for products. The preference model used can be either a value function or a set of soft constraints.[35] Pu and Kumar[5] evaluate their example critiquing approach against a standard ranked list search. The ranked list approach they compare to permit the user to rank a set of products along any single dimension at a time, such as price. They compare SmartClient with this ranking approach in the domain of apartment search, using two databases of 50 apartments each and a set of 22 subjects. They conclude that SmartClient performs only marginally better than the ranked list search for simple tradeoff tasks. But as the complexity of the tradeoff tasks increases, the performance of the ranked list degrades significantly in terms of average completion time and error rate, and SmartClient's performance becomes much stronger. Interestingly, their results also show that more subjects expressed a higher confidence level with answers found using the ranked list search than using SmartClient. They observe that this is due primarily to the fact that users feel they can see everything with the ranked list search, while the SmartClient system is hiding information. This result provides support for the VMAP design which constantly provides the user with an overview of the available alternatives.

The approach closest to that of VMAP is embodied in Blythe's VEIL (Visual Exploration and Incremental Elicitation) system.[12] Like VMAP, VEIL integrates MAUT with visualization to support product search and selection. The primary motivation behind VEIL is to provide an effective interface that implements incremental preference elicitation, in which preference elicitation is interleaved with applying the current preference model to the set of alternatives to generate new suggestions.[7] Blythe illustrates VEIL in the domain of flight selection. VEIL assumes a linear additive value function and permits the user to directly specify the weights of the function or to specify preferences among pairs of alternative items. Pairwise preferences are then translated into linear constraints on the value function and linear programming is used to infer a value function consistent with the constraints. The exact interface for specifying preferences is not described. In VMAP, a major design effort has been to integrate the V-Tool and MAP-Tool components in a single screen in such a way that the user has an easily graspable overview of the search process at all times. The visualization component of VEIL uses a 2D projection to display items. The attributes of the two axes are fixed. Other attributes are depicted with shape, width, height, and a textual label. The scoring according to the value function is depicted with shading, as well as the score shown as a number. The use of a 2D display makes VEIL's interface easier to understand than VMAP's, particularly for users who are not used to graphical displays, but it also limits the amount of information that can be displayed. Due to the less flexible nature of its visualization interface in comparison to that of VMAP, VEIL seems most suited to domains where items can be described in terms of relatively few attributes.

Morningstar's stock-trading guide[36] supports users in searching for investment funds. One tool on the site combines preference elicitation with a simple form of visualization. Users can specify criteria and weights and see the ranked list of funds dynamically updated. The funds are displayed in order, with their scores shown as a bar graph. A second tool, the investment radar, provides a visualization of funds according to a fixed set of attributes. The site does not combine the visualization of the ranking with the visualization of the attribute space as is done in VMAP.

## 7. Conclusions

In this paper, we have presented a novel approach to integration visualization and MAUT to support product search and selection. Our VMAP approach directly addresses the following important design criteria, motivated by behavioral aspects of online shopping and decision making:

- The system should reveal domain knowledge whenever possible.
- The system should not impose a rigid order for preference elicitation.
- The system should reveal the consequences of the user's partially expressed preferences at all stages.
- The system should support incremental elicitation of preferences.

VMAP does this by providing and integrating on one screen both a multi-attribute preference tool (MAP-tool) and a product visualization tool (V-tool). The V-tool displays the set of available products, with each product displayed as a point in a 3D attribute space. By viewing the product space, users can gain an overview of the range of available products, as well as an understanding of the relationships between their attributes. The MAP-tool integrates expression of preferences and filter conditions, which are then immediately reflected in the V-tool display. In this way, the user can immediately see the consequences of his expressed preferences on the product space. By including the entire preference elicitation interface in one screen, the user can easily specify preferences in any order and change earlier expressed preferences.

Comparison of VMAP with a traditional MAUT-based product selection tool over a group of 122 users shows that although VMAP is somewhat more difficult to use than a traditional MAUT product selection tool, it provides better flexibility, provides the ability to more effectively explore the product domain, and produces more confidence in the selected product. VMAP thus presents an important step in achieving more natural decision support for online shopping.

## Appendix A. The Questionnaire for Evaluation of VMAP and the MAUT System

Please indicate your satisfaction of the design of the system you visited with the following statements. Circle a number on the scale such as 1 indicates strongly disagree and 7 strongly agree.

1. Overall, I am satisfied with how easy it is to use this system.
2. It was simple to use this system.
3. I would find it easy to get this system to do what I want it to do.
4. It required the fewest steps possible to accomplish what I want to do with it.
5. I was able to find laptops quickly using this system.
6. I was able to complete the tasks and scenarios quickly using this system.
7. It was fast enough to find laptops that suit my preference.
8. I believe I could become productive quickly using this system.
9. I was able to efficiently complete the tasks and scenarios using this system.
10. It saved my time when I used this system.
11. Whenever I made a mistake using the system, I could recover easily and quickly.
12. This system was very helpful in assisting me in identifying the selected laptop from all the available alternatives.
13. I was able to easily observe the effects of my specified preferences.
14. It was easy to recognize my previously specified preferences while I was observing the effects of these preferences.
15. I found the loop of the exploration was tight.
16. This system required minimal steps to get feedback.
17. I was able to easily observe and learn about the relationships between laptop features (e.g., CPU Speed vs. Price).
18. It was easy to understand the differences of a set of compared laptops.
19. I was able to quickly perceive a list of recommending laptops from all the available alternatives.
20. The system information was presented in a clear and understandable way.
21. I felt very confident using this system.
22. I felt very confident that I had selected the most suitable laptop from all the available alternatives.
23. I felt very confident in the process of the system (e.g., searching).
24. I felt very confident that this system would recommend laptops based on my preferences.
25. The interface of this system was pleasant.
26. I liked using the interface of this system.
27. This system has all the functions and capabilities I expected it to have.
28. The interface of this system helps me simplify tasks.
29. I would recommend this system to my colleagues for finding their new laptops.
30. Overall, I am satisfied with this system.

## References

1. B. J. Babin, W. R. Darden and M. Griffin, Work and/or fun: Measuring hedonic and utilitarian shopping value, *Journal of Consumer Research* **20** (1994) 644–656.
2. Activebuyersguide, Online available at http://www.activebuyersguide.com, Last accessed April 2005.

3. General Motors' vehicle advisor, Online available at http://www.gm.com, Last accessed April 2005.
4. J. W. Payne, J. R. Bettman and E. J. Johnson, *The Adaptive Decision Maker* (Cambridge University Press, 1993).
5. P. Pu and P. Kumar, Evaluating example-based search tools, in *ACM Conference on Electronic Commerce (EC'04)* (New York, May 2004), pp. 208–217.
6. G. Carenini and D. Poole, Constructed preferences and value-focused thinking: Implications for AI research on preference elicitation, in *AAAI-02 Workshop on Preferences in AI and CP: Symbolic Approaches* (Edmonton, Canada, 2002).
7. V. Ha and P. Haddawy, Problem-focused incremental elicitation of multi-attribute utility models, in *Proc. 13th Conf. Uncertainty in Artificial Intelligence* (August 1997), pp. 215–222.
8. C. Ahlberg and B. Shneiderman, Visual information seeking: Tight coupling of dynamic query filters with starfield displays, in *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems* (New York, 1994), pp. 313–317.
9. E. Callahan and J. Koenemann, A comparative usability evaluation of user interfaces for online product catalogs, in *Proc. 2nd ACM Conf. Electronic Commerce* (Minnesota, USA, October 2000), pp. 197–206.
10. T. Lanning, K. Wittenberg, M. Heinrichs, C. Fyock and G. Li, Multidimensional information visualization through sliding rods, in *Proceedings of Working Conference on Advanced Visual Interfaces* (Palermo, Italy, 2000), pp. 173–180.
11. K. Wittenburg, T. Lanning, M. Heinrichs and M. Stanton, Parallel bargrams for consumer-based information exploration and choice, in *Proc. 14th Annual ACM Symposium on User Interface Software and Technology* (Florida, November 2001), pp. 51–60.
12. J. Blythe, Visual exploration and incremental utility elicitation, in *Proc. 18th National Conf. Artificial Intelligence* (Alberta, Canada, 2002), pp. 526–532.
13. M. Stolze, Soft navigation in electronic product catalogs, *International Journal of Digital Libraries*, **3**(1) (2000) 60–66.
14. G. Linden, S. Hanks and N. Lesh, Interactive assessment of user preference models: The Automated Travel Assistant, in *User Modeling: Proceedings of the Sixth International Conference, UM97*, eds. A. Jameson, C. Paris and C. Tasso (Springer Wien New York, Vienna, New York, 1997), pp. 67–78.
15. S. Shearin and H. Lieberman, Intelligent profiling by example, in *Proc. Int. Conf. Intelligent User Interfaces (IUI'01)*, (New Mexico, January 2001), pp. 145–151.
16. G. Carenini and J. D. Moore, Multimedia explanations in idea decision support systems, in *Working Notes of AAAI Spring Symposium on Interactive and Mixed-Initiative Decision Theoretic Systems* (California, 1998), pp. 16–22.
17. P. Pu and B. Faltings, Enriching buyers' experiences: The smartclient approach, in *Proc. SIGCHI Conf. Human Factors in Computing Systems* (The Hague, The Netherlands, April 2000), pp. 289–296.
18. R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* (Cambrige University Press, Cambridge, England, 1993).
19. S. S. Stevens, On the theory of scales and measurement, *Science* **103** (1946) 667–680.
20. S. K. Card, J. D. Mackinlay and B. Shneiderman, Introduction to Section 3.3, in *Readings in Information Visualization Using Vision to Think*, eds. S. K. Card, J. D. Mackinlay, and B. Shneiderman (Morgan Kaufmann Publishers, San Francisco, CA, 1999), Chap. 3, p. 285.
21. G. W. Fischer, Experimental applications of multi-attribute utility models, in *Utility, Probability and Human Decision Making*, eds. D. Wendt and C. A. J. Vlek (D. Reidel Publishing Company, Dordrecht, Holland, 1975), pp. 7–46.

22. Z. Xu and Q. Da, Projection method for uncertain multi-attribute decision making with preference information on alternatives, *International Journal of Information Technology and Decision Making* **3**(3) (2004) 429–434.
23. A. J. Dix, J. E. Finlay, G. D. Abowd and R. Beale, *Human-Computer Interaction*, England: Prentice Hall Europe, 2nd ed. (1998).
24. J. R. Lewis, IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use, *International Journal of Human-Computer Interaction* **7**(1) (1995) 57–78.
25. J. R. Lewis, Psychometric evaluation of the PSSUQ using data from five years of usability studies, *International Journal of Human-Computer Interaction* **14** (2002) 463–488.
26. F. D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS Quarterly* **13**(3) (1989) 319–340.
27. J. Kirakowski, The use of questionnaire methods for usability assessment, 1994. Online available at http://www.ucc.ie/hfrg/questionnaires/sumi/index.html, Last accessed April 2005.
28. A. M. Lund, Measuring usability with the USE questionnaire, *Usability Interface* **8**(2) (2001).
29. J. Brooke, SUS: A "quick and dirty" usability scale, in *Usability Evaluation in Industry*, eds. P. W. Jordan, B. Thomas, B. A. Weerdmeester and I. L. McClelland (Taylor and Francis, London, UK, 1996), pp. 189–194.
30. J. C. Nunnally, *Psychometric Theory* (McGraw-Hill, New York, 1978).
31. J. Nielsen, *Usability Engineering* (AP Professional, Boston, MA, 1993).
32. L. J. Cronbach, Test validation, in *Educational Measurement*, ed. R. L. Thorndike (American Council on Education, Washington, D.C., 2nd ed., 1971).
33. T. K. Landauer, Research methods in human-computer interaction, in *Handbook of Human-Computer Interaction*, ed. M. Helander (Elsevier Science, New York, 1988).
34. T. Yamane, *Statistics, An Introductory Analysis* (Harper and Row, New York, 2 ed., 1967).
35. M. Torrens, B. Faltings and P. Pu, Smartclients: Constraint satisfaction as a paradigm for scaleable intelligent information systems, *International Journal of Constraints* **7**(1) (2002) 49–69.
36. Morningstar's stock-trading guide, Online available at http://www.morningstar.com, Last accessed April 2005.