An Experimental Study of Decision Process with Interactive Technology[*]

Anya Samek[a], Inkyoung Hur[b], Sung-Hee Kim[c], and Ji Soo Yi[d]

[a]Center for Economic and Social Research, University of Southern California
[b]Department of Decision Sciences and Information Systems, Florida International University
[c]Department of Computer Science, University of British Columbia
[d]School of Industrial Engineering, Purdue University

September 14, 2015

Abstract

We investigate the effect of different interactive technologies on the decision-making process in an information search laboratory experiment. In our experiment, the participant makes a selection from a list of differently-valued objects with multiple attributes. We compare presenting information in static form to two methods of interactive presentation. In the first, the participant can manually sort objects by attribute, a capability similar to that found in spreadsheet software. In the second, we present an interactive visual tool that (1) automatically sorts all objects by attribute and (2) uses visual cues for comparisons. Manual sorting capability does not cause an improvement in decisions in this context. On the other hand, the visual tool increases the value of the objects selected by the participant and decreases time spent deliberating. We also find that our interactive presentations affect the decision-making process of participants by changing the number of intermediate options considered. Our results highlight the importance of investigating the effect of technology on information search, and suggest that appropriate interactive visual displays may improve search in practice.

JEL Classifications: C91, D83

Keywords: laboratory experiment, information search, decision process, visual tools

1. Introduction

Decision-making problems encountered by individuals on a daily basis range from relatively minor choices such as purchasing products to major life choices such as purchasing a home, selecting a university to attend or choosing a retirement plan. When making these decisions, individuals are often faced with large and complicated choice sets with multiple attributes. For example, when choosing among several universities, the prospective student may need to consider aspects such as the cost of tuition, majors available, quality of campus life and faculty-to-student ratios.

When faced with complex choice sets, many individuals fail to choose the best possible option, often due to incomplete search or bounded rationality (Simon, 1987). Due to cognitive limitations, the full set of available choices (or attributes) cannot always be considered. We propose that different interactive technologies may be utilized to address these problems without limiting search options. We investigate the effect of sorting capability and of interactive visual tools on the decision-making process in an information search experiment. Such technologies are created with the aim of facilitating search for consumers, yet the impact of these tools on decision-making deserves greater study.

We use a laboratory experiment, which has an advantage over collecting data from the field since it allows us to create a choice set in which options can be objectively ranked. In our experiment, each participant is incentivized to select the highest valued object from a matrix of 15 objects (in rows) with multiple attributes (in columns). Each individual object is in a row. The first column contains the object number and the subsequent columns contain numbers representing the object's attributes. The value of each object depends on the sum of the values of each of its 7, 2-digit number attributes. Following Caplin et al. (2011), we use the 'choice process' methodology whereby participants are incentivized on the basis of the final object

selection, but also on selections made during search. That is, participants are incentivized to immediately select an object (initial object), and continue improving on their selection (intermediate object) until they are satisfied with their selection (final object). This implementation allows us to study the decision-making process of participants. Earlier work seeking to understand the decision-making process includes paper-and-pencil 'information boards' (Payne, 1976) and a computerized system called MouseLab, which allowed users to reveal and hide information using a mouse (Payne et al., 1988).[1]

Importantly, different from Caplin et al. (2011) as well as from other related work, we investigate not only the decision process but also how interactive and visual presentation impacts decision-making. In a between-subjects design, we compare presenting information in static form to two different methods of interactive presentation. In the first, the participant is able to manually sort objects by one attribute at a time, a capability that is similar to that found in spreadsheet software. In the second, we evaluate a new interactive visual tool that (1) automatically sorts all objects by attribute and (2) uses visual cues to assist with comparisons (this tool was first introduced in Hur and Yi, 2009).

We find that despite its prevalence in e-commerce and spreadsheet programs, typical sorting capability does not cause a significant improvement in decision-making in this context. On the other hand, with the interactive visual tool, participants were more likely to choose initial and final objects of greater value, yet spent less time deliberating. Our findings have direct practical implications on the design of computerized systems that support consumer search, since we provide new evidence of the impact of such systems on decision-making.

---

[1] MouseLab has been used to study information search in a number of studies, for example Johnson et al. (2002) and Costa-Goma et al. (2001).

In what follows, Section 2 provides background literature, Section 3 describes the experimental design, Section 4 presents our findings, and Section 5 concludes.

2. Background

The information-search process has been an active topic of investigation since Stigler's (1961) seminal paper on search with imperfect information, which showed that individuals often search for alternatives much less than expected and thus pay more than necessary for goods and services. The decision-making problem that we focus on in this paper is multi-attribute object selection, also called search between alternatives or multiple criteria decision-making. A number of papers have gathered empirical evidence on this type of search problem (e.g., Wallenius et al., 2008; Caplin et al. 2011; Savikhin et al., 2011; Besedes et al. 2012a, 2012b; Balcombe et al., 2014; Sanjurjo, 2014; Sonntag, 2015).

The study of the information search process also has practical applications. For example, Schram and Sonnemans (2011) study the choice of health care policies in the laboratory and find that an increased number of alternatives changes the search process and reduces the quality of decisions. In a series of choice tasks designed to mimic the selection of health insurance, prescription drug, or retirement savings plans, Besedes et al. (2012a, b) find that an increase in the number of alternatives decreases the quality of decisions, and this decline is more pronounced for older subjects. Sitzia et al. (2014) find that complexity makes subjects more likely to stick to defaults, resulting in sub-optimal outcomes. In their paper, providing a search engine reduced inertia and translated into more optimal outcomes. While we do not investigate the effect of choice set size in our paper, we are interested in whether new technologies can facilitate search in a setting with a relatively large choice set.

A rational decision-maker should evaluate the outcome of each option and select the optimal one. However, research suggests that the option selected may not necessarily be the highest valued, especially when the choice set is large and search costs exist. For instance, *satisficing*, or stopping search when a reservation value has been met, is one prominent model of behavior that has been suggested (Simon, 1955). Caplin et al. (2011) provide empirical evidence for a reservation-based model of sequential search that supports Simon's model of satisficing. Search behavior is considered "procedurally" rational if the individual's choice is improving over the course of the decision-making process (Manzini and Mariotti, 2007). In the presence of search costs, a directed cognition model, in which decision-makers seek out only most useful information at each decision point, may better match the empirical evidence on behavior (Gabaix et al., 2006).

Depending on the characteristics of the decision task, individuals may utilize any number of different information search strategies, which also play a role in determining the final outcome (Einhorn, 1970, 1971; Tversky, 1972; Payne et al., 1990; Gigerenzer and Goldstein, 1996; Gabaix and Laibson, 2005). If the information set is large, individuals may use heuristics such as selectively processing only subsets of available information or simplifying the processing of particular elements of the problem (Payne et al., 1990). Moreover, consumers may switch between different information acquisition strategies during search (Shi et al., 2013).

In light of the behavioral findings in this literature, the way information is presented may have a significant impact on the decision-making process as well as the ultimate option selected by the decision-maker. Indeed, Caplin et al. (2011) find that changes in the choice environment affect reservation values and the order of search. Besedes et al. (2015) find that even when the choice set is kept constant, a tournament-style choice architecture in which choices are presented

sequentially improves the option selected. Finally, Sonntag (2015) finds that delayed display of information affects the search process but not decision quality.

We propose that the way information is presented may affect the marginal cost of search. For instance, the marginal cost of search may decrease when opportunities to interact with the data are available, such as when using our interactive tools. If reservation value is determined by equating marginal value of search to marginal cost of search, and marginal value is unchanged, this would increase the reservation value and the value of the final object selected. We also propose that our interactive tools encourage the user to adopt different search processes – for instance, while manual sorting capability may encourage search by attribute, the interactive visual tool may encourage sequential search.

Our work is most related to Besedes et al. (2015), who consider how choice architecture (simultaneous or sequential) affects choice overload and the ultimate decision made. Unlike Besedes et al. (2015), in our experiment the entire choice set is presented simultaneously in all treatments, but the way that participants view and interact with the data is varied. An additional contribution of our work relative to Besedes et al. (2015) is that we provide choice process data to document the impact of our treatments on initial, intermediate and final decisions made.

The interactive visual tool that we evaluate is called SimulSort, which makes use of and builds on several techniques in the visualization literature (Hur and Yi, 2009). In this regard, our paper is close to Hur et al. (2013), who investigate the use of SimulSort on accuracy in solving problems. Different from Hur et al. (2013) we focus on use of SimulSort for a decision-making problem. SimulSort, along with a baseline and typical sorting display that we created for the experiment, are described in greater detail in Section 3.

3. Experimental Environment

The experiment was conducted at the Vernon Smith Experimental Economics Laboratory. Participants were recruited by email through ORSEE (Greiner, 2015) from a subject pool of undergraduate students at Purdue University. A total of 120 subjects participated in 6 sessions, with 20 subjects participating in each session. Each subject participated in only one session of this study. In each session, subjects took part in 20, 180-second individual decision-making rounds. The computerized experimental sessions used SimulSort (Hur and Yi, 2009) and two other versions of information displays, specifically created for this experiment, to record subject decisions and the process of information search. Subjects were given their written instructions, shown in Appendix A, at the beginning of the experiment and the experimenter also read the instructions out loud. Subjects took a short quiz on understanding and then proceeded to the experiment. Subjects were not allowed to use calculators or scratch paper.

Each decision-making round of the experiment involved choosing the highest valued object from a matrix of 15 objects (displayed in rows) with 7 attributes (displayed in columns). Each individual object was in a row. The first column contained the object number and the subsequent columns contained numbers representing the object's attributes. Each object's attributes consisted of a set of 2-digit numbers, where the value of each object was calculated as the sum of the value of its attributes, and the value of an attribute depended on the values of other objects' attributes in the same column using the following formula: *value = (number of attribute – minimum number in attribute column) / (maximum number in attribute column – minimum number in attribute column)*. We adopted this formula for value calculation because in practical application, each attribute may have a different range of values that are not easily compared (for example, the efficiency of autos may range from 15 to 60 MPG, but the cost of

7

autos may range from $10,000 to $50,000) and our formula acts to standardize comparisons of attributes. Final object values were equivalent to U.S. dollars, and the distribution of object values in each round roughly resembled a bell-shape as displayed in Figure 1, with values ranging from $0.00 to $7.00. Subjects were given Figure 1 in the instructions and were aware of the possible range of values. We generated a dataset of objects and attributes for each of the decision rounds, and in half of the sessions we used one order of the dataset, while in the remaining sessions we used the reverse order.[2]

[FIGURE 1 – VALUE DISTRIBUTION - ABOUT HERE]

Following Caplin et al. (2011), we used the 'choice process' methodology whereby participants were incentivized on the basis of the final object selection, but also on selections made during search. At the beginning of each round, all objects were deselected. Subjects were instructed to immediately select an object (initial object), and continue improving on their selection (intermediate object) until they were satisfied with their selection (final object). Only one object could be selected at a time. Earnings in each round were calculated as *earnings = (value of final object selected) + 0.2*(value of intermediate object selected at time t)*, where *t* was a number randomly drawn from a uniform distribution of (0 seconds, 180 seconds).[3] In this way, subjects were incentivized to continue improving upon their object selections, revealing their preferred object at all times during the decision-making process.

If no intermediate object was selected at time *t*, the value of the intermediate object was counted as 0. If the final object was selected at time *t* (i.e., the participant had ended the round

---

[2] The inter-correlation coefficient across attributes was set to 0-0.1, which was found to be common in datasets such as nursing home choice selection in an earlier paper evaluating SimulSort (Kim et al., 2012).

[3] Caplin et al. (2011) ran two experiments, one in which the final decision was used to calculate earnings, and one in which the decision at a random time was used to calculate earnings. Both yielded a similar result in the choice process data, except that in the former there was a greater incentive to continue search. Because we were comparing different user interfaces that could be used to make decisions in practice (the outcome of which depends on the final choice), we preferred a payment structure that incentivized both the final and one intermediate choice.

early), the value of the intermediate object was equal to the value of the final object. If, on the other hand, the subject did not have time to select a final object by 180 seconds, the intermediate object selected at 180 seconds was counted as the final object. Subjects were made aware of the earnings calculations, including the procedure when time is up, in the instructions.

Subjects could select an object by clicking on it, on any of its attributes, or on the radio button directly to the left of the object. The object currently under consideration was automatically highlighted in green and its radio button was filled in. Subjects could also move the mouse cursor over alternative objects. When the cursor hovered over a particular object (except the currently-selected object), that object and its attributes were highlighted in yellow for ease of comparison to the currently selected object. Subjects were given up to 180 seconds to complete each round, although a subject who finished in less than 180 seconds could press a button to continue. After submitting the final selection in each round (or after time ran out), subjects moved to the outcome screen. At the end of each round, we displayed to each subject the value of his or her final selection, the time randomly selected to be paid out, the value of his or her intermediate selection, and his or her payment in that round.

We conducted three treatments: Baseline, Typical Sort, and Automated Sort. All treatments retained the matrix of 15 objects (in rows) and 7 attributes (in columns) with the same highlighting features, see Figure 2. Treatments varied only in the way that individuals viewed and interacted with the information. In Baseline, the only interactive capability was to select and mouse-over objects. In Typical Sort, subjects could in addition manually sort objects by one attribute (column) at a time by clicking on the column header (in ascending or descending order). Subjects could engage in this sorting at any time, and the currently selected object would remain highlighted throughout the sorting process (but would be moved to its sorted location on the

screen). In Automated Sort, the SimulSort (Hur and Yi, 2009) interface was used whereby the columns of each attribute were always sorted by value automatically in descending order, and objects could be compared using visual highlighting in each attribute column.

The screenshots in Figure 2 also provide an overview of the Typical Sort and Automated Sort interactive features.[4] Object 8 is selected in all panels. The Typical Sort panel (center) shows that column F is sorted in ascending order (its header is also highlighted in yellow). The Automated Sort panel (bottom) shows the relative value of all of object 8's attributes relative to the other objects. An attribute that is at the top of the screen is higher valued than those below it. Rollover could also be used to highlight the attributes of an alternative object in yellow. In this way, in the Automated Sort treatment subjects had a visual picture of how the object under consideration ranked relative to all other objects in each attribute column.

[FIGURE 2 – SCREENSHOTS OF DISPLAYS – ABOUT HERE]

The purpose of the Typical Sort treatment was to understand how the ability to sort by attribute affects the decision-making process and value of final object selected. While sort features are ubiquitous in many applications, such as Microsoft Excel and in online product lists, the actual impact of sorting on the choice process is less clear. The purpose of the Automated Sort treatment was to test the efficacy of the new SimulSort interaction technique relative to no interaction and to typical sorting capability. In addition, while Typical Sort encouraged a decision-making process that compared objects by one attribute at a time (attribute based search), Automated Sort encouraged a decision-making process that compared two objects at a time to one another (sequential search).

---

[4] The reader can try out the tools at: http://www.cs.ubc.ca/~kim731/econ-demo/demo.html

At the end of the experiment, three rounds were selected for payment using a random draw from a bingo cage and subjects were paid in U.S. dollars for their earnings in those three rounds. Subjects earned about $26 on average, and sessions (including instruction time) lasted approximately 140 minutes.

4. Results

4.1 Summary of Data

In total, 120 subjects participated in the experiment, which gives us 2,400 observations across the 20 rounds. Subjects were allotted 180 seconds in each round to make a decision, and were incentivized to immediately select an object (initial object), and continue improving on their selection (intermediate object) until they were satisfied with their selection (final object). Our main analysis uses Mann-Whitney tests with outcomes averaged for each subject across 20 rounds of the experiment, with regressions provided for additional support.

Since our experimental design limited the time allotted for the decision, we may worry that participants who did not press the "submit" button prior to the end of the round had not completed their decision – that is, the object selected at 180 seconds may be an intermediate object rather than the final object. The final choice was made before time ran out in 83.9%, 79.4%, and 93.6% of cases in Baseline, Typical Sort and Automated Sort, respectively. Significantly more subjects made the final choice before time ran out in Automated Sort relative to Baseline or Typical Sort (tests on the proportion of completed rounds by subjects averaged across all rounds, Mann-Whitney p-value = 0.04 and 0.02, respectively). About half of subjects (52.1%) always made their final choice before time ran out and one fifth (19.0%) almost always made their final choice before time ran out. Later rounds resulted in greater completion rates; for instance, completion rates were 80.3% in rounds 1-5 and 90.5% in rounds 16-20.

Because most subjects made their final decision before time ran out, and because they were aware that the currently-selected object at 180 seconds would count as their final object, in what follows, we report results using all subjects' data, treating the choice made at 180 seconds as the final object, even if in some cases it may have been the intermediate object.[5] However, we provide separate analysis in Appendix B that restricts the sample only to subjects who had pressed the "submit" button prior to the end of the 180-second round. Restricting the sample does not affect our conclusions.

In what follows, sub-section 4.2 explores the outcome of decision-making, 4.3 analyzes the decision-making process, and 4.4 uses mouse tracking and other data to explore potential explanations for our results.

4.2 Decision Outcome Analysis

Our main outcome measure is the value of the final object selected after search concludes, relative to the highest possible value. We find that subjects in the Automated Sort treatment chose final objects that were significantly higher valued as compared to subjects in the Baseline or Typical Sort treatments, but that Typical Sort did not result in significantly different final value relative to Baseline. Table 2 reports summary statistics averaged across rounds and subjects for each treatment. The average value of the final choice in Automated Sort was \$4.82 (s.e.=\$0.03), while average values in Baseline and Typical Sort were \$4.63 (\$0.04) and \$4.58 (\$0.05), respectively. Because each round has a different possible maximum and minimum value, we also use another measure for comparison, which we define as: *efficiency = (Value of final choice – value of minimum rank choice in round)/(value of maximum rank choice in round*

---

[5] Note that all subjects made at least an initial object selection during the experiment.

– *value of minimum rank choice in round)*. Efficiency is also higher in Automated Sort (0.91, *s.e.*=0.01) than Baseline (0.86, *s.e.*=0.01) and Typical Sort (0.84, *s.e.*=0.02). Both value and efficiency were significantly higher in Automated Sort relative to Baseline (Mann-Whitney *p*-values<0.01), as well as in Automated Sort relative to Typical Sort (Mann-Whitney *p*-values<0.01).[6] There were no significant differences in value or efficiency between Baseline and Typical Sort (Mann-Whitney *p*-values>0.10).

An additional measure of decision quality is how often subjects selected the highest possible valued object as their final object – this happened 56% of the time in Automated Sort, 42% of the time in Baseline, and 39% of the time in Typical Sort. Again, subjects in Automated Sort outperformed subjects in Baseline or Typical Sort by selecting the highest valued object significantly more often (Mann-Whitney *p*-values<0.01 for both comparisons). Moreover, participants in Automated Sort rarely choose objects ranked in the lowest 1-5 of the possible objects least often (2% of the time) versus participants in Baseline and Typical Sort, who chose low ranked objects 4-5% of the time.

<center>**[TABLE 2 – SUMMARY STATISTICS – ABOUT HERE]**</center>

In summary, we observe that subjects in the Automated Sort treatment outperform those in the Typical Sort and Baseline treatments by selecting higher valued objects. Table 3 reports regressions that provide additional support for these results. The regressions use subject random effects, cluster errors at the session level and include treatment dummy variables for *Automated Sort*, and *Typical Sort*, a time trend (*Period*) and interaction terms for the time trend with the treatment dummy. We observe positive and significant coefficients for *Automated Sort* in regressions on value in column 1 (0.25, *p*-value<0.01) and in regressions on efficiency in column

---

[6] Note that these statistics are computed by first averaging by participant across all rounds, therefore generating one unique observation per individual.

2 (0.08, *p*-value<0.01). However, we observe a significant *negative* coefficient for *Typical Sort* in regressions on value (-0.11, *p*-value<0.05) and on efficiency (-0.03, *p*-value<0.05). This brings us to the first result:

Result 1: Automated Sort results in higher valued final objects selected relative to Baseline and Typical Sort.

[TABLE 3 -- REGRESSIONS – ABOUT HERE]

We also consider learning over the 20 rounds of the experiment. We see a positive and significant time trend (*Period* is 0.005 for value and 0.002 for efficiency, *p*-values<0.01), suggesting that participants become better at the information search task over time. The interaction term on *Period*Automated Sort* is negative and marginally significant (-0.004, *p*-value<0.10 and -0.001, *p*-value>0.10 for value and efficiency, respectively), while the interaction term on *Period*Sort* is positive and highly significant (0.008 and 0.002 with *p*-values<0.01). The greater improvement in *Typical Sort* across rounds suggests that subjects must learn how to implement this tool, while the lack of improvement in *Automated Sort* across rounds suggests that this is a tool that subjects are able to grasp fairly quickly.

Another interesting measure is the amount of time spent deliberating in each round. As reported in Table 2, despite selecting a higher valued final object, subjects spent significantly less time deliberating in *Automated Sort* (78.64 seconds, s.e.=5.69) relative to *Baseline* (108.03 seconds, s.e.=6.32) and *Typical Sort* (111.49 seconds, s.e.=7.18) (Mann-Whitney *p*-values<0.01). The regressions in Table 3 (column 3) that use time spent as a dependent variable support these results, showing a statistically significant negative coefficient on the *Automated Sort* treatment

dummy (-33.39, *p*-value<0.01).[7] Time spent deliberating also decreased across all treatments over the 20 rounds of the experiment (coefficient on *Period* is negative and significant, -1.98 with *p*-value<0.01), but it decreased less in *Automated Sort* or *Typical Sort* (*Period*Automated Sort* and *Period*Typical Sort* coefficients are positive and significant with *p*-values<0.01, 0.53 and 0.57, respectively). This brings us to the next result:

Result 2: Time spent deliberating is significantly lower in Automated Sort relative to Baseline or Typical Sort.

We also find that while time spent deliberating is positively correlated with the value of the final object selected in Baseline and Typical Sort treatments, the correlation is weakest in the Automated Sort treatment. Average time spent deliberating in each round is positively and significantly correlated with the efficiency of the final decision in Baseline and Typical Sort treatments (Spearman rho = 0.60, p-values<0.01), but the correlation is weaker and only marginally significant in the Automated Sort treatment (Spearman rho = 0.26, *p*-value = 0.10).

4.3 Decision-making Process Analysis

Decision-making process data provide information about whether individuals make 'procedurally rational' decisions by continuing to improve upon their initial decisions during search (Manzini and Mariotti, 2007). The data we use to evaluate whether decisions are 'procedurally rational' are the value of intermediate objects and the number of switches made during search. A switch is defined as making any object selection, such that a participant who selects only an initial object prior to continuing to the next round has switches=1, while a

---

[7] Note that we are missing one observation from the reported time spent, because one subject in *Baseline* experienced a computer crash in period 11. While we have data for his or her final decision from his or her record sheet, we were unable to recover the time data on time spent or on the search process for this subject.

participant who makes a change from the initial object to an intermediate object has switches=2, and so on. A switch is defined as 'efficiency improving' if the current object selected is higher valued than the previously selected object, and not efficiency improving if the current object selected is lower valued than the previously selected object (thus, the first switch is always efficiency improving). We posit that efficiency-improving switches are 'procedurally rational' in the sense of Manzini and Mariotti (2007).

Figure 3 provides histograms of the total number of switches by treatment. Most decision rounds include 2 switches, i.e. subjects make an initial object selection followed by an intermediate object selection, which is also their final selected object. As can be seen in Figure 3, very few decision rounds contain more than 6 switches. Notice also that observations in Automated Sort are almost as likely to contain only 1 switch as 2 switches, and are also more likely than observations in Typical Sort and Baseline to contain only 1 switch. That is, subjects in Automated Sort often select their initial object as their final object.

[FIGURE 3 ABOUT HERE]

We find that subjects' decision-making process is different across treatments. As summarized in Table 2, subjects in Automated Sort made on average of 3.10 (s.e.=0.27) switches, significantly less than subjects in Typical Sort, who made an average of 5.98 (s.e.=1.10) switches (Mann-Whitney p-value<0.01). Subjects in Baseline made an average of 5.48 (s.e.=2.53) switches, which was not significantly different from Automated Sort (Mann-Whitney p-value=0.49). Subjects in Typical Sort also made significantly more switches than subjects in Baseline (Mann-Whitney p-value<0.01). The percentage of switches that were efficiency improving is highest in Automated Sort (86%) and lower in Typical Sort (76%) and Baseline (84%). These percentages represent a statistically significant difference between

16

Typical Sort and Baseline (Mann-Whitney p-value<0.01) and between Typical Sort and Automated Sort (p-value<0.01) but not between Baseline and Automated Sort. In summary, Typical Sort seems to result in the greatest number of switches, yet it has the lowest proportion of efficiency-improving switches. These findings suggest that while 'procedurally rational' decisions are made much of the time, Typical Sort results in the lowest proportion of decisions that are 'procedurally rational.'

Table 5 provides regressions similar to that in Table 3 but using total switches, number of efficiency-improving switches, and proportion of switches that were efficiency improving, in columns 1-3, respectively. We find that the proportion of efficiency-improving switches was significantly higher in Automated Sort (coefficient=0.03 with p-value<0.01) and significantly lower in Typical Sort (coefficient=-0.08, with p-value<0.01) relative to Baseline. This brings us to the next result:

Result 3: Switching behavior is different across treatments; in particular, subjects in Typical Sort make the greatest number of switches but a lower proportion of efficiency-improving switches, while subjects in Automated Sort make the highest proportion of efficiency-improving switches.

We also consider whether search-stopping behavior resembles satisficing in the sense of Simon (1955). A subject searches through the choice set and stops if and only if a reservation level of value (or utility) is achieved. Our results suggest that subjects do engage in satisficing behavior, and this is evidenced by Figure 4, which plots the average value of each intermediate object selected, separately for subjects who engaged in 1, 2, 3, 4 or 5 switches in any round, by treatment. For example, the "1 switch" marker on each graph represents the average value of the

object selected by subjects who made only an initial object selection, while the "2 switches" markers on each graph represent the average value of the initial object and final object selected by subjects who made one initial object selection and one intermediate object selection. As displayed in Figure 4, when subjects select a relatively high valued initial object, they often end their search (see lines with 1 switch); whereas when subjects select a lower valued initial object, they often continue search (see the first data point for lines with 2 or more switches).

[FIGURE 4 – INTERMEDIATE OBJECT SELECTION – ABOUT HERE]

The data also provide us with insight about the point during the decision-making process at which the Automated Sort tool helps subjects. We find that the initial object is selected within 5-7 seconds on average in all treatments.[8] We also find that subjects are able to choose a higher-valued initial object in Automated Sort (77.9% efficiency) as compared to subjects in Baseline or Typical Sort (72.9% and 69.3% efficiency, respectively) (Mann-Whitney p-values<0.01 for both comparisons). There are no significant differences between Baseline or Typical Sort on the value of the initial object (Mann-Whitney p-value>0.10). This brings us to the next result:

Result 4: Automated Sort results in significantly higher-valued initial object selections than either Typical Sort or Baseline.

We follow the methodology of Caplin et al. (2011) in calculating reservation values. We assume that individuals within a given environment have the same distribution of reservation values, and that their reservation value depends on $\bar{v}$, which is constant given the environment, and a stochastic error term ϵ that is by assumption independent and drawn from the standard normal distribution. Subjects decide to stop search if and only if the value of the object, v,

---

[8] The initial switch is selected within 6.7 (s.e.=1.13) seconds in Baseline, 5.38 (s.e.=0.84) seconds in Typical Sort, and 5.99 (s.e.=1.04) seconds in Automated Sort. There are no statistically significant differences between treatments (Mann-Whitney p-values>0.10).

exceeds that of their reservation value $\overline{v} + \epsilon$. Search continues if $v < \overline{v} + \epsilon$. v_k denotes the value of the object at decision node k, and ϵ_k denotes the error. Φ is the cumulative density function of the standard normal distribution. The probability of stopping search is given by $\Phi(v_k - \overline{v})$, while the probability of continuing search is given by $1 - \Phi(v_k - \overline{v}) = \Phi(\overline{v} - v_k)$.

The reservation values $(\overline{v})$ estimated through maximum likelihood are 4.22 (0.602) in Baseline, 4.15 (0.637) in Typical Sort and 4.48 (0.535) in Automated Sort.[9] The reservation values under Automated Sort are significantly higher from those under Typical Sort and under Baseline ($p<0.01$). However, there are no significant differences in reservation values under Baseline and Typical Sort ($p>0.10$). We also calculate percentage of subjects who follow the reservation strategy: they stop search for objects at or above the reservation value and continue search for objects below the reservation value. We find that 76.8% of subjects in Baseline, 72.2% of subjects in Typical Sort, and 72.7% of subjects in Automated Sort follow the reservation strategy. This brings us to the final result:

Result 5: Automated Sort leads to a significantly higher reservation value than Typical Sort or Baseline.

4.4 Examination of Mechanisms

We find that Automated Sort causes subjects to select final objects that are higher valued and improves the decision-making process by shortening time spent and increasing the number of switches that are considered 'procedurally rational.' In this section, we consider why this

---

[9] As in Caplin et al. (2011) we use the log likelihood function $\ln \mathcal{L} = \sum_{k=1}^{K}[x_k \ln(\Phi(v_k - \overline{v})) + (1 - x_k) \ln(\Phi(\overline{v} - v_k))]$ to find the value of $\overline{v}$ for each subject that maximizes $\ln \mathcal{L}$ and report the average of these values in the text with standard errors in parentheses. Estimated reservation efficiency is 0.721 (0.193) in Baseline, 0.693 (0.207) in Typical Sort and 0.803 (0.175) in Automated Sort.

might be the case. As noted by Payne (1990), subjects may selectively process only subsets of available information. We propose that the different interaction capabilities in Baseline, Typical Sort, and Automated Sort may encourage subjects to choose different subsets of information to process, which may then lead to the differences in final outcomes that we observe. In particular, since our English-speaking subjects are used to reading from left to right and top to bottom, subjects in the Baseline treatment may use the heuristic of processing information about objects at the top of the screen. Typical Sort encourages sorting by attribute, which may cause subjects to consider some attributes more than others, and one possibility is to evaluate objects based on the value of attributes on the left side of the screen. On the other hand, Automated Sort may encourage subjects to consider a larger sub-set of data since it is easier to calculate value by looking at the location of the object's attributes.

Indeed, in a complementary paper to this one, Kim et al. (2012) used eye-tracking to investigate another group of subjects' browsing behavior in the same decision task, with Typical Sort and SimulSort (i.e., our Automated Sort treatment). Kim et al. (2012) found long periods of eye fixation in top-left areas of the screen in Typical Sort, while eye fixation in SimulSort was of shorter duration. These results suggest that Typical Sort may have caused subjects to overweight the importance of maximally valued objects along only a few attributes on the left of the screen, while SimulSort encouraged patterns of quick browsing.

We also collected data on how long subjects hovered their mouse over each cell. Figure 5 provides scatter plots weighted by total seconds spent hovering on each cell, by treatment. While subjects in Baseline hovered over most cells about equal amounts of the time, subjects in Typical Sort focused attention on sorting (as evidenced by the large amount of time spent in row 0) while subjects in Automated Sort focused on selecting different objects to view the automatic sorting

and highlighting (as evidenced by the large amount of time spent on column 0). Moreover, subjects in Typical Sort spent the most time sorting column 1 relative to columns 2-7, in line with the eye tracking results of Kim et al. (2012).

As an additional test to investigate whether subjects use a heuristic that is based on processing sub-sets of information based on location, we create variables that categorize final objects selected based on the ranking of some or all of its attributes, and conduct regressions reported in Table 6. The dependent variables are *Average Rank*, which indicates the average rank of each objects' attributes (Column 1), *Left Rank,* which indicates the average rank of each objects' 4 left-most attributes on the screen (Column 2), *Right Rank,* which indicates the average rank of each objects' 4 right-most attributes on the screen (Column 3), *Object Number*, which indicates the number of the object, where objects are numbered 1-15 from top to bottom on the screen (Column 4), and *Top Object*, which is a 0/1 measure of whether the final object selected is within the top 7 of objects on the screen (Column 5). We find that Automated Sort results in final object selections that are on average higher ranked (-0.514, p-value<0.01) and are also higher ranked on the left side of the screen (-0.511, p-value<0.05). We find suggestive evidence that Automated Sort, relative to Baseline, results in less likelihood of choosing a final object that is at the top of the screen (-0.0105) but the result is not statistically significant. We predicted that Typical Sort may result in more final objects selected that have high ranked left attributes, but we do not find that this is the case.

In summary, while evidence from eye tracking and mouse movement data seem to suggest support for our conjectures, the data from final object selections is less clear. More work is needed to explore the mechanisms underlying the impact of different interactive visual tools.

5. Discussion & Conclusion

We experimentally investigated the decision-making process in an information search task in which the individual chooses an object from a list of objects with multiple attributes. We studied different ways of presenting information by establishing the ability to interactively sort objects by attribute (Typical Sort treatment) and by introducing the SimulSort tool (Automated Sort treatment). Our results confirmed our prediction that the type of decision support tool used influences the process of search and the value of the final object selected. In particular, we found that Automated Sort resulted in higher valued initial and final objects selected relative to Baseline and Typical Sort, while time spent deliberating in Automated Sort was significantly lower. We also found that subjects in Automated Sort make the highest proportion of efficiency-improving switches and therefore exhibit more 'procedurally rational' behavior relative to Typical Sort and Baseline. Finally, we observed that Automated Sort led to significantly higher reservation values than Typical Sort or Baseline. In summary, it seems that the design of decision support tools not only affects the final decision, but also affects how users approach the decision-making problem.

Our results suggest that interactive displays may be a new way to augment and improve decision-making in similar tasks in practice. While sorting capabilities are ubiquitous in many environments in practice such as online shopping sites or word processing/spreadsheet computer programs, our results suggest that they may not be particularly effective for certain classes of tasks. In fact, users of Typical Sort spent more time and did not arrive at better final decisions relative to no tool used. For this reason, new visual analytics tools may be needed to support the decision-making process, and SimulSort (used in the Automated Sort treatment) is a step in this direction. Importantly, the way that individuals are able to interact with data actually influences

22

the process of search, which we have documented here. This is the first study to carefully consider the result of several interactive technologies on the search process. It joins a small but growing literature on the use of interactive or visual tools to improve decision-making (Rudolph et al., 2009; Kaufman et al., 2013).

Our environment placed objects into rows and attributes in columns, which is similar to the design of spreadsheet software and therefore may be more intuitive for subjects to understand. Several online decision-making tools include such an approach, including, for example, Medicare.gov's 'hospital comparison' tool[10] and Anthem's BlueCross BlueShield 'find a doctor' tool.[11] It is feasible that implementing a tool like SimulSort on these platforms would benefit consumers in practice. However, another approach is placing objects into columns and attributes in rows. Examples include online shopping sites with the capability to compare objects along several dimensions, such as on Apple.com.[12] More work is needed to learn whether the SimulSort tool would be equally effective with objects in columns and attributes in rows.

One possible reason for the greater effectiveness of the Automated Sort treatment is that it allowed users to consider each object's value as a whole, whereas in the Typical Sort treatment users could compare only by attribute. Our mouse movement data provides suggestive evidence for this possibility. Considering all attributes equally is important in this environment where each attribute is equally valued. One could design an environment in which Typical Sort might be more appropriate; for instance, an environment with differently valued columns may make Typical Sort more valuable. In fact, this may be one reason why sorting is ubiquitous in online search and shopping sites: sorting by the most valuable attribute may be a much more rational

---

[10] See https://www.medicare.gov/hospitalcompare/results.html#dist=25&loc=33331&lat=26.0621374&lng=-80.3769999
[11] See https://www.anthem.com/health-insurance/provider-directory/searchcriteria?qs=*5YTNo/dAJqDa8088PcRLjA==&brand=abcbs
[12] http://www.apple.com/mac/compare/

strategy in environments where one attribute (such as the price) plays the largest role in adding value to the object. We hope to encourage future work on the use of interactive visual tools in economic decision-making across many different environments to shed light on these questions.

References

Balcombe, K., Fraser, I., Mcsorley, E., 2014. Visual attention and attribute attendance in multi-attribute choice experiments. Journal of Applied Economics. doi:10.1002/jae.2383

Besedeš, T., Deck, C., Sarangi, S., Shor, M., 2012a. Age Effects and Heuristics in Decision Making. Review of Economics and Statistics. 2, 580-595. doi:10.1162/REST_a_00174

Besedeš, T., Deck, C., Sarangi, S., Shor, M., 2012b. Decision-making strategies and performance among seniors. Journal of Economic Behavior and Organization. 81, 524–533. doi:10.1016/j.jebo.2011.07.016

Besedeš, T., Deck, C., Sarangi, S., Shor, M., 2015. Reducing Choice Overload without Reducing Choice. Review of Economics and Statistics. In press.

Caplin, A., Dean, M., Martin, D., 2011. Search and satisficing. American Economic Review. 101, 2899–2922. doi:10.1257/aer.101.7.2899

Costa-Goma, M., Crawford, V., Broseta, B., 2001. Cognition and Behavior in Normal-Form Games: An Experimental Study. Econometrica. 69.5, 1193-1235.

Einhorn, H.J., 1970. The use of nonlinear, noncompensatory models in decision making. Psychological Bulletin. 73, 221–230. doi:10.1037/h0028695

Einhorn, H.J., 1971. Use of nonlinear, noncompensatory models as a function of task and amount of information. Organizational Behavior and Human Performance. 3, 221-230. doi:10.1016/0030-5073(71)90002-X

Gabaix, X., Laibson, D., 2005. Bounded Rationality and Directed Cognition. Harvard University.

Gabaix, X., Laibson, D., Moloche, G., Weinberg, S., 2006. Costly information acquisition: Experimental analysis of a boundedly rational model. American Economic Review. 96, 1043–1068. doi:10.1257/aer.96.4.1043

Gigerenzer, G., Goldstein, D.G., 1996. Reasoning the fast and frugal way: models of bounded rationality. Psychological Review. 103, 650–669. doi:10.1037/0033-295X.103.4.650

Greiner, Ben (2015). Subject pool recruitment procedures organizing experiments with ORSEE," Journal of the Economic Science Association 1:1 114-125.

Hur, I., Yi, J.S., 2009. SimulSort: Multivariate data exploration through an enhanced sorting technique, in: Jacko, J. (Ed.). Human-Computer Interaction. Novel Interaction Methods and Techniques. pp. 684–693. Springer Berlin Heidelberg. doi:10.1007/978-3-642-02577-8_75

Hur, Inkyoung, Sung-Hee Kim, Anya Samak, and Ji Soo Yi. "A Comparative Study of Three Sorting Techniques in Performing Cognitive Tasks on a Tabular Representation." *International Journal of Human-Computer Interaction* 29, no. 6 (2013): 379-390.

Johnson, E., Camerer, C., Sen, S., Rymonk, T., 2002. Detecting failures of backward induction: monitoring information search insequential bargaining. Journal of Economic Theory.

Kaufmann, C., Weber, M., Haisley, E., 2013. The Role of Experience Sampling and Graphical Displays on One's Investment Risk Appetite. Management Science 59, 323–340. doi:10.1287/mnsc.1120.1607

Kim, S.H., Dong, Z., Xian, H., Upatising, B., Yi, J.S., 2012. Does an eye tracker tell the truth about visualizations?: Findings while investigating visualizations for decision making. IEEE Trans. Vis. Comput. Graph. 18, 2421–2430. doi:10.1109/TVCG.2012.215

Manzini, P., Mariotti, M., 2007. Sequentially rationalizable choice. American Economic Review. 97, 1824–1839. doi:10.1257/aer.97.5.1824

Payne, J.W., 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. Organizational Behavior and Human Performance. 16, 366-387. doi:10.1016/0030-5073(76)90022-2

Payne, J.W., Bettman, J.R., Johnson, E.J., 1988. Adaptive strategy selection in decision making. Journal of Experimental Psychology. Learning, Memory, and Cognition. 14, 534–552. doi:10.1037/0278-7393.14.3.534

Payne, J.W., Bettman, J.R., Johnson, E.J., 1990. The Adaptive Decision Maker: Effort and Accuracy in Choice, in: Insights in Decision Making: A Tribute to Hillel J. Einhorn. pp. 129–153.

Rudolph, S., Savikhin, A., Ebert, D.S., 2009. FinVis: Applied visual analytics for personal financial planning, in: VAST 09 - IEEE Symposium on Visual Analytics Science and Technology, Proceedings. pp. 195–202. doi:10.1109/VAST.2009.5333920

Sanjurjo, A., 2015. Search with Multiple Attributes: Theory and Empirics. Social Science Research Network. http://ssrn.com/abstract=2460129 Last accessed: July 14, 2015. doi:10.2139/ssrn.2460129

Savikhin, A., Fisher, B., Ebert, D.S., 2011. An Experimental Study of Financial Portfolio Selection with Visual Analytics for Decision Support, in: 2011 44th Hawaii International Conference on System Sciences. pp. 1–10. doi:10.1109/HICSS.2011.54

Schram, A., Sonnemans, J., 2011. How individuals choose health insurance: An experimental analysis. European Economic Review. 55, 799–819. doi:10.1016/j.euroecorev.2011.01.001

Shi, S.W., Wedel, M., Pieters, F.G.M.R., 2013. Decision Making : A Model-Based Exploration Using Eye-Tracking Data. Management Science. 59, 1009–1026. doi:10.1287/mnsc.1120.1625

Simon, H., 1955. A Behavioural Model of Rational Choice. The Quarterly Journal of Economics. 69, 99–118. doi:10.2307/1884852

Simon, H., 1987. Bounded Rationality. in: Eatwell, J., et al. (Eds.). The New Palgrave: A Dictionary of Economics. New York: Macmillan Press Limited, 15-18.

Sitzia, Stefania, Jiwei Zheng, and Daniel John Zizzo. "Inattentive consumers in markets for services." *Theory and Decision* (2014): 1-26.

Sonntag, A., 2015. Search costs and adaptive consumers: Short time delays do not affect choice quality. Journal of Economic Behavior and Organization. 113, 64-79.

Stigler, G.J., 1961. The Economics of Information. The Journal of Political Economy. 69, 213–225.

Tversky, A., 1972. Elimination by aspects: A theory of choice. Psychological Review. 79, 281-299. doi:10.1037/h0032955

Wallenius, J., Dyer, J.S., Fishburn, P.C., Steuer, R.E., Zionts, S., Deb, K., 2008. Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead. Management Science. 54, 1336–1349. doi:10.1287/mnsc.1070.0838

Instructions for SimulSort; Typical Sort and Baseline are similar.

Note that while we use the term 'objects' throughout the paper, we used 'items' to describe the objects in the instructions.

Displays are available at: http://www.cs.ubc.ca/~kim731/econ-demo/demo.html

INSTRUCTIONS

This is an experiment in the economics of decision making. Various research agencies have provided funds for this research. The instructions are simple. If you follow them closely and make appropriate decisions, you can earn an appreciable amount of money which will be paid to you at the end. It is very important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. In addition, in this experiment you will not be allowed to use calculators or scratch paper to make calculations. We expect and appreciate your cooperation.

In this experiment you will participate in a task for 20 rounds. At the end of the experiment **3 out of 20** rounds will be randomly selected for payment. After you have completed all rounds three tokens will be randomly drawn out of a bingo cage containing tokens numbered from **1 to 20**. The token numbers determine which three rounds are going to be paid. You do not know which rounds will be paid when you make your choices. Hence, you should pay attention to the choice you make in every round.

Your Decision

In each round, you will be shown a table of 15 items from which you will be asked to make a selection. Each item has seven (7) numbers shown in seven (7) different columns, which represent different attributes of the item. Your earnings will depend on the **value** of the item you choose in each round.

This table is different from a typical table found in spreadsheet software, in which only one column can be sorted at any given moment. In this table, every column is sorted at the same time. Thus, in order to find a corresponding attribute of each item, you should move your mouse over an item or a cell in the table, and corresponding cells will be highlighted. Since all columns are sorted, cells near the top of the screen will have a higher number, and cells near the bottom of the screen will have a lower number within that column.

Determining Value

Each column has different ranges of numbers. For example, a column may have numbers between 10 and 60, and another column may have numbers between 50 and 90. The range of numbers determines the **US dollar value** of each attribute.

Your earnings will depend on the **total US dollar value** of the item that you select. The **total value** of each item is the sum of the **values** of all of its attributes. The **value** of a particular attribute varies depending on which column the attribute is located on. The **value** of an attribute can range from $0.00 to $1.00. For example, the same attribute of 50 can mean pretty high **value**

($0.80) on a column with numbers between 10 to 60, but the same 50 can mean the lowest **value** ($0.00) on another column with numbers between 50 and 90, and mean the highest **value** ($1.00) on another column with numbers between 20 and 50. The **total value** of an item is calculated by adding up the **values** of its attributes. Since an item consists of seven (7) attributes, the **total value** can range from $0.00 to $7.00.

Examples

Each item has 7 attributes, and each item has exactly 1 number per attribute column. Here is an example of an item:

Economics Experiment

Question 1 out of 20 : Current Selection

item 08

Item Number	A	B	C	D	E	F	G
item 15	88	76	73	85	78	97	41
item 14	87	73	70	80	75	95	41
item 13	86	73	68	80	73	95	40
item 12	86	73	67	80	71	90	40
item 11	85	72	64	79	71	84	39
item 10	85	69	64	76	71	84	39
item 09	85	68	63	70	49	79	38
item 08	84	65	51	65	41	64	38
item 07	84	64	51	65	38	64	37
item 06	83	61	48	63	33	62	37
item 05	80	52	48	53	33	51	34
item 04	79	46	44	47	32	48	34
item 03	78	43	38	46	32	46	33
item 02	78	43	38	40	23	44	32
item 01	78	41	37	29	16	44	32

To help you understand how to find the **value** of each attribute of item 8, look at column G. The number of attribute G for item 8 is 33. The lowest number in column G is 32, while the highest number in column G is 41. There are many ways to determine the **value** of an item, and all of them will help you determine approximate value. Which way you use is up to you.

1. **Approximate the value:** Think about the range of numbers and think about where 33 lies in that range. 33 is closer to the low value than to the high value, so the approximate **value** is $0.10 to $0.20.

2. **Approximate the value using the visual:** Values closer to 0 will be sorted near the bottom of the screen, while values closer to 1 will be sorted near the top of the screen, so you can approximate the value by checking where the value shows up on the screen. 33 is near the bottom of the screen, suggesting an approximate **value** of $0.10 to $0.20.

3. **Calculate the value:**

Value = ([number of attribute]-[minimum number in column])/([maximum number in column]/[minimum number in column]).

Solving for the value, we find the value to be (33-32)/(41-32) = $0.11

As another example, take a look at column C. The number of attribute C for item 8 is 64. The lowest number in column C is 37. The highest number in column C is 73.

1. Approximating, we note that 64 is in between 37 and 73, but closer to 73. So we believe the **value** of the attribute in column C may be about $0.65-0.75.

2. Approximating by looking at the screen, we find that 64 is near the top of the screen, suggesting that the **value** of the attribute in column C may be about $0.65 to $0.75.

3. Calculating the value, we find that **value** = (64-37)/(73-37) = $0.75

Notice that each of the ways help you find the approximate value. That is, the ways in which you approximate value help you guess a value that is very close to the calculated value.

Remember that to find the **total value** of each item you have to **consider the values** of all of its 7 attributes. You can do this by either adding values explicitly if you have calculated them, or by approximating the total value using the methods described above.

<u>What are the values of the available items?</u>

The value of each item is independent of the value of other items, and the value of each item is determined by drawing a random integer number. The probability distribution of values will be as follows:

Probability Distribution
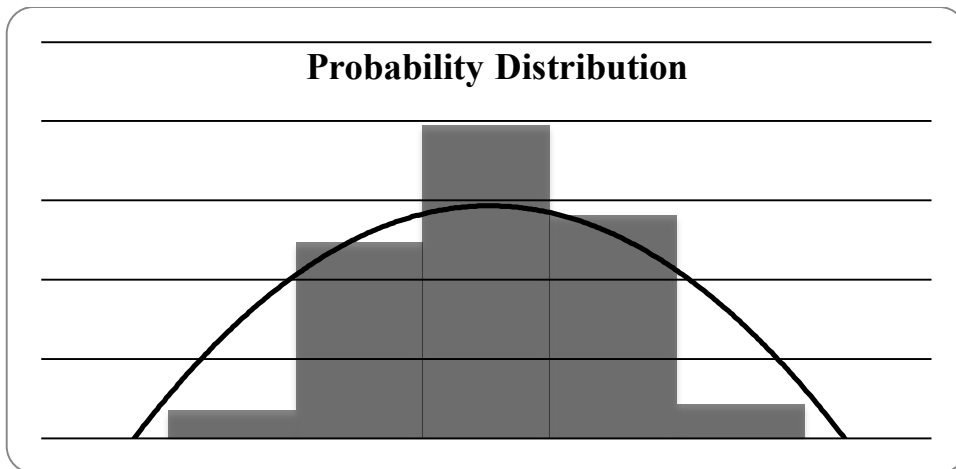
For example, there is around a 30% chance of an item having a total value of between $3.00 and $4.00.

<u>Decision-making</u>

In this part we will guide you through the process of making item selections. Please look at the overhead projector screen.

To begin the period, click on "Start" in the center of the screen.

(Important: Do **NOT** click "Done" (in the bottom right corner) until you are satisfied with all your decisions in that period. You will NOT be able to come back.)

When a round begins, all 15 items will be presented to you on the computer screen. Each round will last up to 3 minutes. Note that at any time, you can end the round by clicking on the "Done" button, but there is no penalty for using the entire 3 minutes. Clicking the "Done" button is the same as letting the remaining time run out with your current selection still selected. If you click the "Done" button, you cannot make any more changes in that round.

When a round begins, none of the items will be selected. You can select an item by clicking on the radio button to the left of the item, the name of the item in the first column or by clicking anywhere on the item's attributes. When an item is selected, all of its attributes will be highlighted in Green. You are free to change which item is selected at any time and as many times as you like.

When the item is selected, you can see the currently selected item at the top of your screen.

You can highlight an item without selecting it by moving the mouse cursor over that option. The item will remain highlighted in Yellow as long as the cursor stays anywhere on top of it.

After 3 minutes or whenever you click the "Done" button, the first round will come to an end. At that point, a time will be picked at random from between 0 and 180 seconds (as described below) and the item that you chose at that time will be recorded. Your earnings in each round will depend in part on **the total value of the final item chosen**, and in part on **the total value of the item selected at that point in time**. Therefore, **if at any point you prefer a different item to the one you have selected, you should change your selection** since this will increase the chance of getting higher earnings.

Dollar earnings =
   (Total value of final item chosen) + 0.2 * (Total value of item at randomly selected time)

The time that the selected item is recorded as your choice is determined by drawing a random time between 0:00 and 3:00 (or 180 seconds). Any randomly chosen time is equally likely. For example, there is an equally likely chance that 1 second will be drawn, or 65 seconds, or 170 seconds, and so on. If no item is selected at the random time selected, your (Value of Choice at Randomly Selected Time) will be zero.

After the round has ended, you will be told which item was recorded as your choice and which item was your final choice. You will be told your earnings in that round. Record the results in your record sheet before you click "continue" to move on to the next period. You will also be told the total value of the item with the greatest value in each round.

The second round and all following rounds will proceed exactly like the first round. Please ONLY use your record sheet for recording outcomes, and not for making calculations.

REMEMBER:

Selecting an item:
- You can select an item by clicking anywhere on the item
- Only one item can be selected at a time
- Initially, no item will be selected
- You are free to change your selected item to any other item at any time, whether or not you have picked that item previously
- You can change the selected item as many times as you would like
- After clicking the "Done" button, the round will end, and it will be as if the selected item remained selected for the remainder of the round.

How your item choice is recorded
- You should select an item as soon as you know that it is better than your currently selected item
- After each round, a time between 0 and 180 seconds will be picked at random
- The item that was selected at that time will be recorded

- Your earnings will be determined in part by the item recorded at the random time, and in part by your final item choice
- At the end of each round, you will be told which item was recorded at the random time, your final choice, and your earnings in that round

How you are paid
- At the end of the experiment, we will pick **three** rounds at random, and you will be paid your earnings in those rounds.
- At the beginning of each round, no item is selected. If no item is selected at the random time selected, you will receive 0. Thus, you should select an item as quickly as possible at the beginning of the round.

STATION NUMBER: _____

RECORD SHEET

Round	Your final choice	Total value of final choice	Time chosen for random selection	Choice at randomly selected time	Total value of choice at selected time	$$ Earnings this Round (Final+0.2*random)
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						

This section presents the same analysis as the main results in the paper, but restricting to observations where the final decision was made prior to the end of the 180-second round. For Table A1, we replicated Table 2 in the paper but dropped all observations in which participants did not complete the round, and then averaged across all rounds for each subject. This resulted in a total of 2,055 observations, with 40 individual subjects in Baseline, 38 in Typical Sort, and 40 in Automated Sort.

Supporting results 1 and 2 in the main paper, we observe higher-valued objects in Automated Sort relative to Baseline and Typical Sort. Mann-Whitney tests are statistically significant for Automated Sort versus Baseline and Typical Sort for both total value and efficiency (all p-values<0.01). No significant differences are observed between Baseline and Typical Sort (Mann-Whitney p-value on efficiency is 0.92 and on total value is 0.85). Supporting result 2, we observe significantly less time spent in *Automated Sort* relative to *Baseline* and *Typical Sort* (Mann-Whitney p-values<0.01 for both comparisons).

Table A1: Summary Statistics: Restrict to rounds that finished within 180 seconds

	Baseline	Typical Sort	Automated Sort
Summary Statistics			
Total value	$4.61 ($0.04)	$4.58 ($0.06)	$4.82 ($0.03)
Efficiency	0.85 (0.01)	0.85 (0.02)	0.92 (0.01)
Rate Highest Value Selected	0.42 (0.02)	0.39 (0.02)	0.56 (0.02)
Time Spent (seconds)	98.33 (5.36)	102.39 (6.54)	74.00 (5.03)
Switching Behavior			
Average switches each round	2.98 (0.22)	6.79 (1.10)	3.10 (0.49)
Efficiency improving switches	2.29 (0.11)	4.08 (1.02)	2.38 (0.14)
Sorting Behavior			
Average sort each round		23.82 (3.19)	

Note: This data is averaged across all 20 periods. Standard errors in parentheses. Efficiency is defined as (value of final choice – value of minimum ranked choice in dataset) / (value of maximum rank choice in dataset – value of minimum rank choice in dataset). Switches are defined as making any intermediate

object selection (1 switch represents making an initial object selection). Efficiency improving switches are defined as making a switch that increases the efficiency of the object that is selected.

Table A2 provides regressions mimicking those in Table 3 in main text. We observe similar patterns of behavior, with positive and statistically significant coefficients for *Automated Sort* in columns 1 and 2, in support of Result 1. Even after excluding anyone who did not finish on time, the coefficient for *Automated Sort* in column 3 in the regression on time spent is still a negative and statistically significant, in support of Result 2.

Table A2: Final Decision Regressions: Restrict to rounds that finished within 180 seconds

	(1) Final Object Value	(2) Final Object Efficiency	(5) Time Spent
Automated Sort	0.276***	0.0942***	-29.77***
=1 for A.S. treatment, 0 o.w.	(0.0297)	(0.0115)	(5.722)
Typical Sort	-0.112***	-0.0292**	2.254
=1 for T.S. treatment, 0 o.w.	(0.0338)	(0.0116)	(8.614)
Period	0.00469***	0.00185***	-1.528***
Period trend, 1-20	(0.000271)	(0.000133)	(0.160)
Period*Automated Sort	-0.00429*	-0.00176	0.608**
Period-A.S. interaction term	(0.00236)	(0.00127)	(0.310)
Period*Typical Sort	0.00923***	0.00243***	0.0806
Period-T.S. interaction term	(0.00208)	(0.000900)	(0.175)
Constant	4.540***	0.827***	113.6***
	(0.0285)	(0.0113)	(4.267)
Observations	2,055	2,055	2,054
Subjects	118	118	118

Note: This table reports random effects regressions with session clustering. One observation is missing in specification 3 because data on time spent is not available for one subject in period 11 in the Baseline treatment. Observations in which participants did not press the "continue" button prior to the 180-second ending time of the period are also dropped. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Next, we turn to switching behavior. As displayed in Table A1, total switches are significantly higher in Typical Sort relative to Baseline or Automated Sort (Mann-Whitney *p*-values<0.01) but not different between Automated Sort and Baseline. By virtue of having the

greatest number of total switches, Typical Sort also has a significantly higher number of efficiency-improving switches relative to Baseline or Automated Sort (p-values<0.01). However, as in the main data, the percentage of efficiency-improving switches relative to total switches is highest in Automated Sort (86.1%) followed by Baseline (84.6%) and Typical Sort (75%). The proportion of efficiency-improving switches is significantly lower in Typical Sort relative to Baseline or Automated Sort (p-values<0.01). Table A3 conducts the same regressions as Table 5 in the main text, using the restricted sample.

Table A3: Decision-making Regressions: Restrict to rounds that finished within 180 seconds

	(1) Total Switches	(2) Efficiency-Improving Switches	(3) Proportion Efficiency Improving
Automated Sort	-0.159	-0.0447	0.0333***
=1 for A.S. treatment, 0 o.w.	(0.341)	(0.263)	(0.0107)
Typical Sort	5.144	2.477	-0.0895***
=1 for T.S. treatment, 0 o.w.	(3.273)	(1.549)	(0.0247)
Period	-0.0436***	-0.0163*	0.00224***
Period trend, 1-20	(0.0162)	(0.00836)	(0.000498)
Period*Automated Sort	0.0244	0.0116	-0.00149*
Period-A.S. interaction term	(0.0221)	(0.0136)	(0.000764)
Period*Typical Sort	-0.127	-0.0654	0.000827
Period-T.S. interaction term	(0.118)	(0.0609)	(0.000651)
Constant	3.470***	2.479***	0.820***
	(0.140)	(0.0445)	(0.00842)
Observations	2,054	2,054	2,054
Subjects	118	118	118

*Note: This table reports random effects regressions with session clustering. One observation is missing in each specification because data on time spent is not available for one subject in period 11 in the Baseline treatment. Observations in which participants did not press the "continue" button prior to the 180-second ending time of the period are also dropped. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

Results on initial switching are similar when dropping the observations that did not complete search, with initial objects valued at 77.9% in Automated Sort, 72.9% in Baseline and

69.3% in Typical Sort. Automated Sort results in statistically significantly higher valued initial objects selected relative to Baseline (p-value<0.05) and Typical Sort (p-value<0.01). Different from the results in the main paper, we also observe a marginally significant difference between Baseline and Typical Sort (p-value<0.10).