HOW MANY OPTIONS? BEHAVIORAL RESPONSES TO TWO VERSUS FIVE ALTERNATIVES PER CHOICE

MARTIN MEISSNER
UNIVERSITY OF SOUTHERN DENMARK/MONASH UNIVERSITY
HARMEN OPPEWAL
MONASH UNIVERSITY
JOEL HUBER
DUKE UNIVERSITY

ABSTRACT

What is an appropriate number of alternatives per choice task? Why are two or five alternatives so rarely used? We characterize the contexts where Choice-Based Conjoint (CBC) on pairs makes sense when projecting to real decisions and use eye-tracking to study how respondents search for information when answering choice tasks with either two or five alternatives.

INTRODUCTION

While there has been substantial work asking how many choice tasks are needed in a CBC study, less attention has been paid to determining the appropriate number of alternatives per choice task. The typical answer to the question of how many alternatives to include revolves around the tradeoff between greater statistical efficiency and increased task difficulty for the respondent. From the perspective of statistical efficiency (as tested using computer simulations) paired comparison choice tasks produce less efficient designs (Bunch, Louviere and Anderson 1996; Louviere and Woodworth 1983). Increasing the number of alternatives in each task provides greater statistical efficiency, because multinomial logit models in fact assume that each final choice is based on a comparison of the selected alternative to all of the available options. However, answering more complex choice questions also makes the choice task more difficult for the respondents and makes it more likely that respondents use simplifying decision heuristics (Bettman, Johnson and Payne 1991; Todd 2007).

As a starting point for our research, we investigated how frequently practitioners who published in the past four Sawtooth Software Proceedings used two, three, four or five alternatives in their CBC studies. We found about 20 studies in each of the four Proceedings. We were surprised by a trend in the last few years which suggests substantial changes in practice. Examining Proceedings in 2010, 65% of the studies mentioned had 5 or more alternatives per choice while 35% had 3 or 4. By 2015 that proportion had reversed, with 10% choosing 5 or more, and almost 90% choosing 3 or 4. Across time, the proportion of studies using two alternatives has consistently stayed under 10%.

The paper by Pinnell and Englert (1997) may be one reason why pairs rarely appear in the recent Sawtooth Software Proceedings. The authors varied the number of alternatives in three experiments and concluded that respondents are capable of accurately answering choice tasks with up to seven options. Compared with two alternatives the authors found that it took about 33% more time for respondents to evaluate four alternatives and about 60% more time to

evaluate seven alternatives. The authors concluded that it is advisable to use more than two options in a choice task because in their studies pairs had lower predictive validity, were less stable and did not save much time relative to choice tasks with more alternatives.

We propose that the performance of pairs should be reevaluated in light of respondents' behavioral responses and that responses to choice tasks depend on how respondents search for information when making their choices. We use eye-tracking to investigate how respondents allocate their attention in CBC choice tasks with two (pairs) or five alternatives (quints). We focus on pairs and quints, because from a practical perspective they reflect the range in the number of alternatives reported in recent Sawtooth Software Proceedings.

We investigate task differences from three different perspectives: First, we examine the ways respondents search for decision-relevant information in pairs and quints, focusing on processing patterns and ways that respondents learn to more efficiently complete their task. Second, we assess to what extent respondents perceive the task as difficult. Third, we compare pairs versus quints in terms of internal consistency and their ability to predict holdout choices from triples.

Behavioral Responses to Choice Complexity

Results from previous studies which have investigated information processing suggest that in choice tasks including only two alternatives respondents use compensatory decision strategies, such as the additive difference strategy (ADD, see e.g., Payne 1976). In line with this research, we expect that in pairs respondents process almost all attribute information that is available (Payne *et al.* 1992) and compare the two alternatives in a step-by-step, attribute-wise and top-to-bottom manner (Russo and Dosher 1983). Using a systematic and complete search process, respondents focus fairly evenly across all attributes, including the less important ones.

In contrast, in more complex choice tasks, respondents have been shown to simplify their choice (Bettman, Johnson and Payne 1991; Ford *et al.* 1989; Payne *et al.* 1992). Compared to a task consisting of pairs, in a task with larger numbers of decision alternatives, respondents can be expected to display greater cognitive effort, but at the same time also show a greater degree of simplification. Respondents can, for example, simplify their search by eliminating alternatives from further consideration based on important attributes. Process tracing studies have shown that respondents often use heuristics, such as the lexicographic rule, to reduce the number of options by excluding those not meeting a minimum level for a particular attribute (Payne *et al.* 1992). We also expect that respondents simplify their search in later tasks because they will learn from the earlier choice tasks which attributes matter most to them (Meissner and Decker 2010; Orquin, Bagger and Mueller Loose 2013).

Study Design

Upon arrival in the laboratory, participants received a general instruction and were familiarized with the eye-tracker. The main task consisted of a self-guided computer-based conjoint survey. It first introduced vacation packages as the product of interest. Next it asked respondents about their prior purchase experience, future purchase intention as well as purchase familiarity and involvement regarding vacation packages. The following screens then explained the attributes and levels of the vacation packages. Respondents then answered eight choice tasks that were presented on separate screens. The profiles in each task were randomly generated with Sawtooth Software's complete enumeration algorithm (Orme 2013). The statistical strength of

the designs for pairs and for quints was determined based on eight choice tasks, 40 respondents, and using Sawtooth Software's complete enumeration algorithm. The relative D-efficiency ratio of the pairs vs. quints design is 174.4/133.6=1.31. This result means that statistically, 31% more observations are needed for pairs to achieve the same aggregate logit efficiency as quints.

After the sequence of eight choice tasks, respondents were asked about their search goals and perceived task difficulty, their frustration and the perceived similarity of options for the last of their eight choice tasks. Finally, respondents answered two holdout choice tasks consisting of three randomly generated alternatives (triples). The survey ended with socio-demographic questions. The results of these additional survey questions are beyond the scope of the current paper and are not discussed.

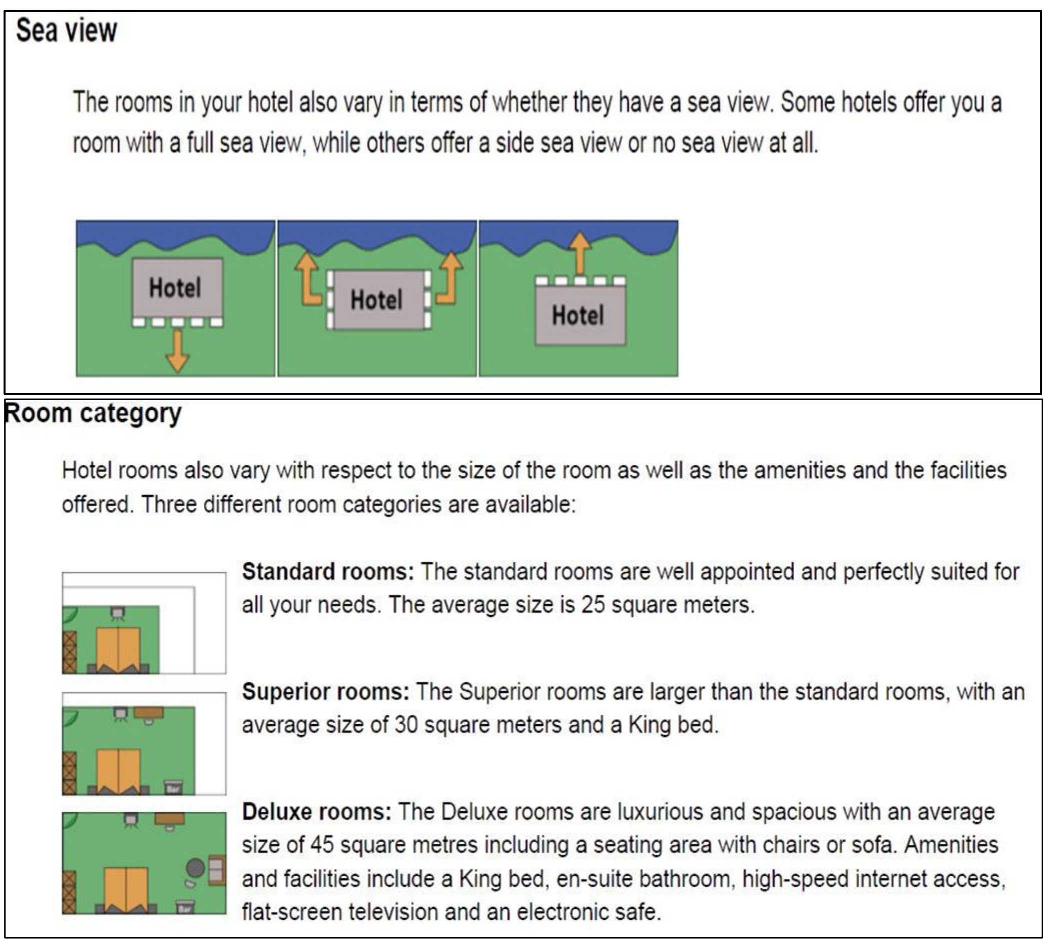
The hypothetical vacation packages were characterized by six attributes, each with three levels. As common in CBC studies, the attributes appeared in a fixed display order. Table 1 describes the attributes and their levels.

Table 1. Vacation Package Attributes and Levels

Attributes	Attribute levels		
	1	2	3
Food quality	good	very good	excellent
Customers recommending	50%	70%	90%
Distance to CBD	3km	2km	1km
Sea view	no sea view	side sea view	full sea view
Price per person	$899	$799	$699
Room category	standard	superior	deluxe

Separate screens helped respondents become familiar with each of the attributes and their levels before respondents answered the choice tasks. Figure 1 gives an example for the attributes "sea view" and "room category." We agree with Eggers, Hauser and Selove (2016, this Proceedings) that craft is important in designing a preference measurement study and used images and ceteris paribus instructions to enhance precision and accuracy. We did not use training videos and did not incentive-align our respondents but do not believe our results comparing pairs vs. quints would change if we included those changes.

Figure 1. Example Instructions Explaining the Attribute Levels of the Attributes "Sea View" (Top) and "Room Category" (Bottom)



The eye-tracking study was carried out at Monash University (Australia). A total of 39 respondents finished the questionnaire with pairs and 38 respondents finished the questionnaire with quints.

Information Processing and Eye-Tracking

Eye-tracking is one of the most reliable approaches for observing humans' attentional processes.[1] Modern eye-tracking systems use video images of the eyes to determine the so-called "point of regard." It reveals that human perception is primarily based on two states of the eyes: fixations and saccades. A fixation is defined as a state where the eyes are relatively stable and "rest on" a certain stimulus. A rapid movement of the eyes between two consecutive fixations is called a saccade. Typically a fixation is between 100 and 500 milliseconds (ms) long with an average of about 250 ms. The fixation duration largely depends on the viewed stimuli and their characteristics (Rayner 1998). A saccade typically lasts between 30 and 50 ms. Studies have shown that humans cannot acquire information during a saccade (Rayner 1998) because the brain

[1] The interested reader is referred to Holmqvist *et al.* (2011) for a more comprehensive introduction to eye-tracking and for a discussion of adequate measures.

blocks visual processing during eye movements in a way that neither the motion of the eye nor the gap in visual perception is noticeable to the individual.

A Tobii T120 recorded the eye movements in our study. This system has an accuracy of 0.4° of visual angle and a sampling rate of 120 Hz. The infrared sensors are built into a 17" thin-film transistor (TFT) monitor with a resolution of 1280 x 1024 pixels. A standard 9-point calibration routine was used to calibrate participants' eye movements (Tobii Software 2016). When placing the respondents in front of the eye-tracker, we made sure that the distance indicator provided by the Tobii software displayed a value between 50 and 80 cm (ideally 60 cm) as recommended by the Tobii handbook. Respondent answers were given solely with a computer mouse.

The areas of interest were defined as cells in the display matrix; they were all of equal size, non-overlapping, and the number ranged between 12 (2 alternatives * 6 attributes) and 30 (5 alternatives * 6 attributes) cells. Fixations were defined as continuous gazes within each area of interest. We used the standard Tobii fixation filter to determine fixations (Tobii Software 2016).

Moreover, it is important to emphasize that we used only simple text labels to describe the features (see Figures 2 and 3). It is therefore unlikely that the feature stimuli differed regarding their saliency, which could have produced differences with respect to the number of fixations to features (effects of bottom-up attention). All respondents reported to have normal or corrected to normal vision. In order to simplify the analyses, we only used the data of the right eye. However, the results do not differ if we use the data for the left eye or the average of the left and the right eye as provided by the Tobii software.

Results

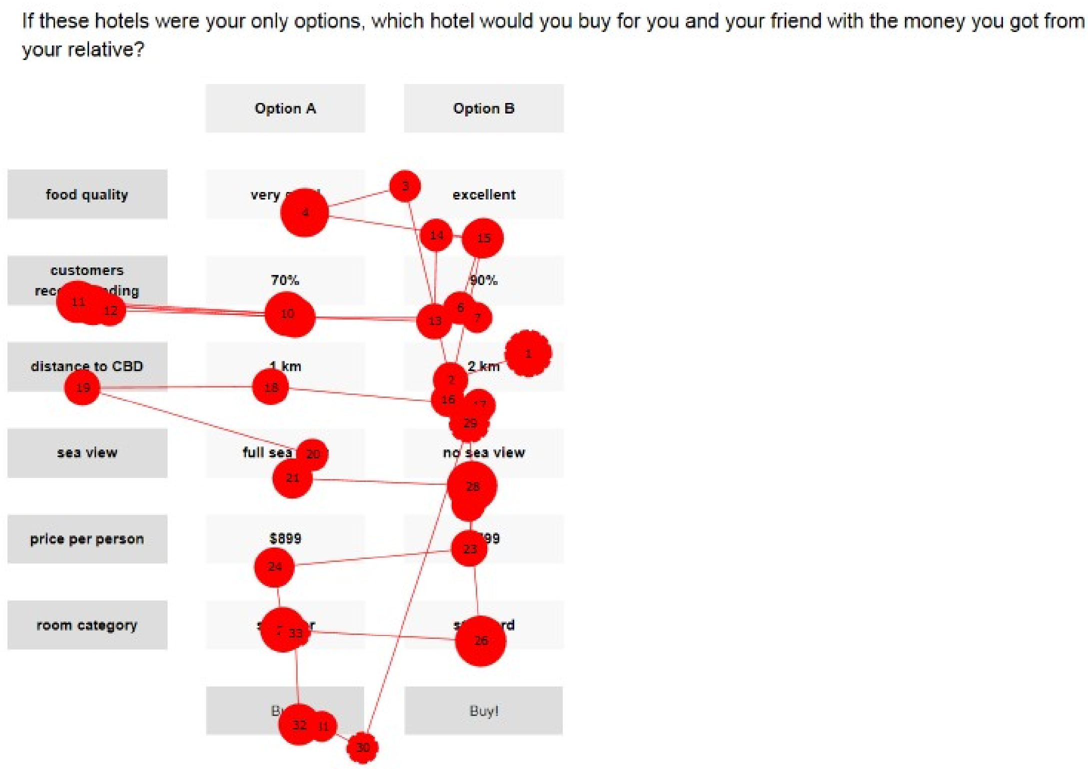
Observing What Respondents Do

Before analyzing respondents' search patterns using known measures of information search, we visually inspected the scanpaths of the choice tasks for every respondent. The inspection of the videos showed that the search patterns often matched our expectations as outlined above in the section "Behavioral Responses to Choice Complexity." Two example paths of fixations that are easily interpreted are depicted in Figure 2 for a pairs task and in Figure 3 for a quints task. We encourage the reader to watch the corresponding videos which are available on YouTube in fast (https://youtu.be/wmpy7O-dZFY) and slow (https://youtu.be/9YZBVyI9TZM) motion for pairs and in fast (https://youtu.be/qXZYIz8eEdc) and in slow (https://youtu.be/he9SjPYVP8Q) motion for quints.

The search process in Figure 2 for pairs follows a typical additive difference model. After two initial fixations to the center of the screen, the respondent starts the search by looking at the top attribute “food quality.” The search continues by comparing the two options with respect to each attribute, moving from the top to the bottom of the screen. In this example all attribute levels are fixated at least once. After having fixated the last attribute, the respondent checks three attribute levels of option B before then choosing option A. The respondent possibly is reassured by the undesired aspects of option B before making the final decision. It is also interesting to see that this respondent does not look at the question text, in this case because she has seen other choice tasks before. She also looks at the description of the attributes only two times, i.e., when comparing the alternatives with respect to “customers recommending” and “distance to CBD.”

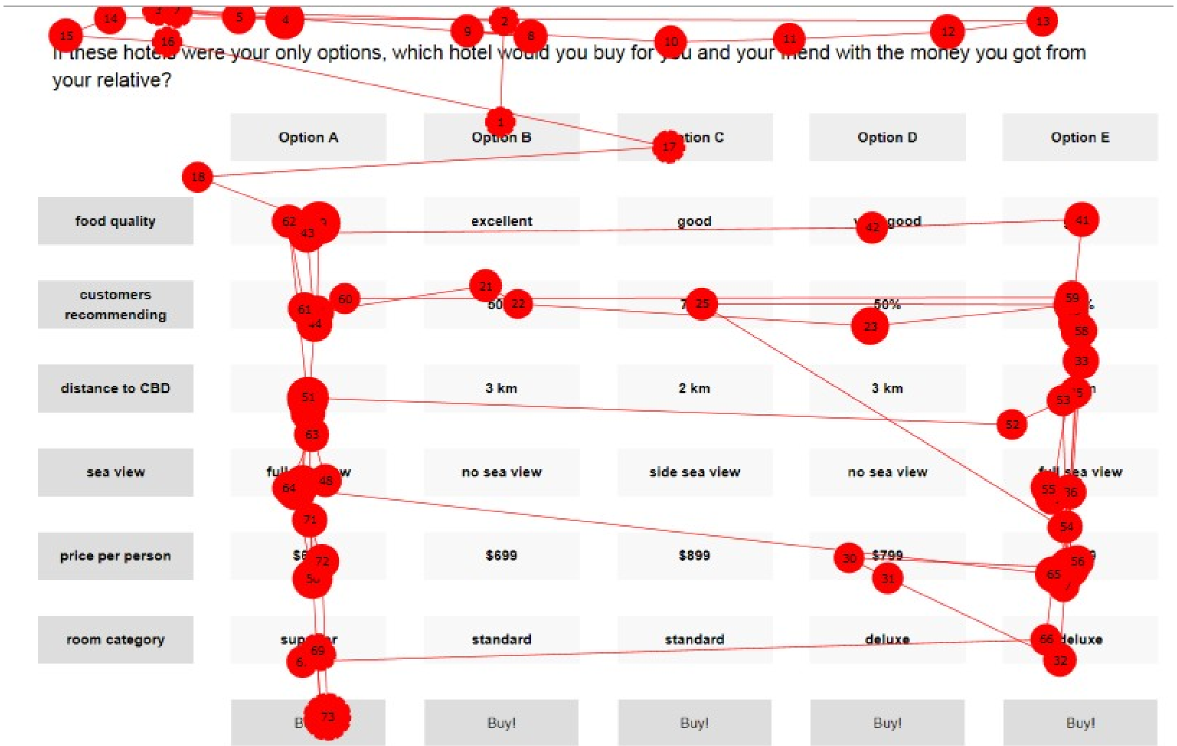
We speculate that she is looking at the attribute labels because the attribute levels are not self-explanatory, as it is not clear what “70%” means without looking at the attribute label.

Figure 2. An Example Path of Fixations for Pairs



The example search process for quints shown in Figure 3 is quite different. In this case, the respondent starts the task by reading the question text. Next, the respondent looks at the attribute “customers recommending” and compares all five alternatives with respect to that attribute. Only options A and E have a customer recommending rating of 90% in this choice task. The customer rating seems to be most important for the respondent and that is why she probably starts the search process by looking at that particular attribute. After identifying options A and E as promising the respondent’s search process changes substantially. The respondent evaluates option E in detail by looking at all the attribute levels of that alternative. The respondent then jumps to option A, which is also evaluated in detail. In what follows, we can see that the respondent focuses only on these two options, A and E, by going back and forth between them. Options B, C and D are only fixated incidentally, perhaps because they are in the way. Many of the levels for the attributes “food quality,” “distance to CBD,” “sea view” and “room category” are not fixated at all for these three options. In all, the search process can be best described as a staged simplification process. In the first step the respondent uses one attribute to identify promising alternatives, in a second step the remaining alternatives are evaluated holistically and compared to one another. We can also see that the search process in the second step includes more transitions within alternatives compared to the search process for pairs.

Figure 3. An Example Path of Fixations for Quints



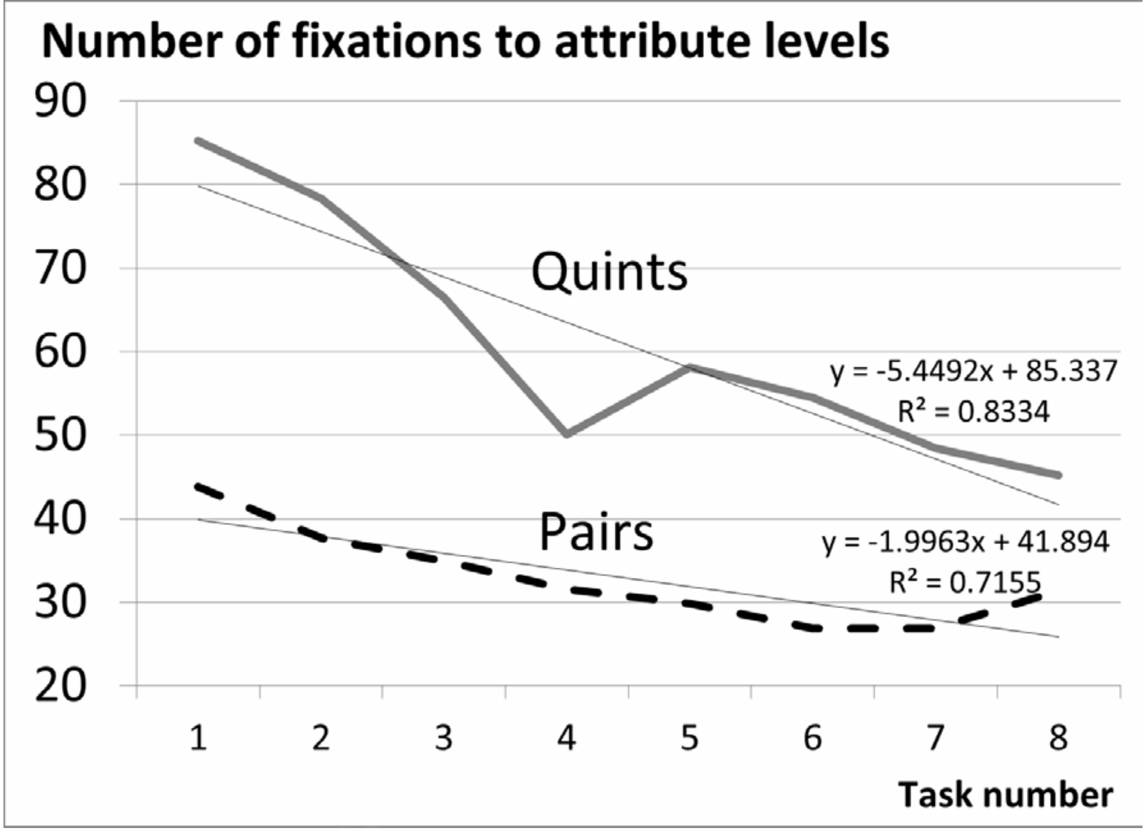
These examples are chosen because they cleanly illustrate the use of expected decision rules. In fact we found that the processes followed many patterns that changed within as well as across respondents. Next we show how quints and pairs differed in using five general measures of information processing: the number of fixations, percent of information accessed, frequency of within vs. between attribute transitions, top-down vs. bottom-up order processing, and average duration of each fixation.

Information Processing

First, we investigated the number of fixations. Changing the number of alternatives had a substantial impact on the number of fixations required for each choice. As shown in Figure 4, respondents in choice tasks comprising five alternatives expend about twice as many fixations (M=60.82, SD=39.85) to make a decision than those encountering sets with two alternatives (M=32.91, SD=18.40). This difference is highly significant. Further, as respondents become more experienced with the task they expend fewer fixations. Respondents in the quints condition adapt faster. The number of fixations dropped about 27% from the first to the last choice task for pairs. For the quints the drop is about 43%. This result replicates Pinnell and Englert's (1997) finding that respondents accelerate processing more in choice tasks with seven than in choice tasks with two alternatives. The result suggests that respondents largely change the way they process the information in quints by simplifying more in later choice tasks. For pairs, there is minimal simplification in later choice tasks. The observed reduction in the number of fixations is also in line with findings by Meissner, Musalem and Huber (2016) who used eye-tracking to show that respondents become more efficient; that is, they need fewer fixations and become more consistent as they progress in a decision sequence of multi-attribute choice tasks. Stüttgen,

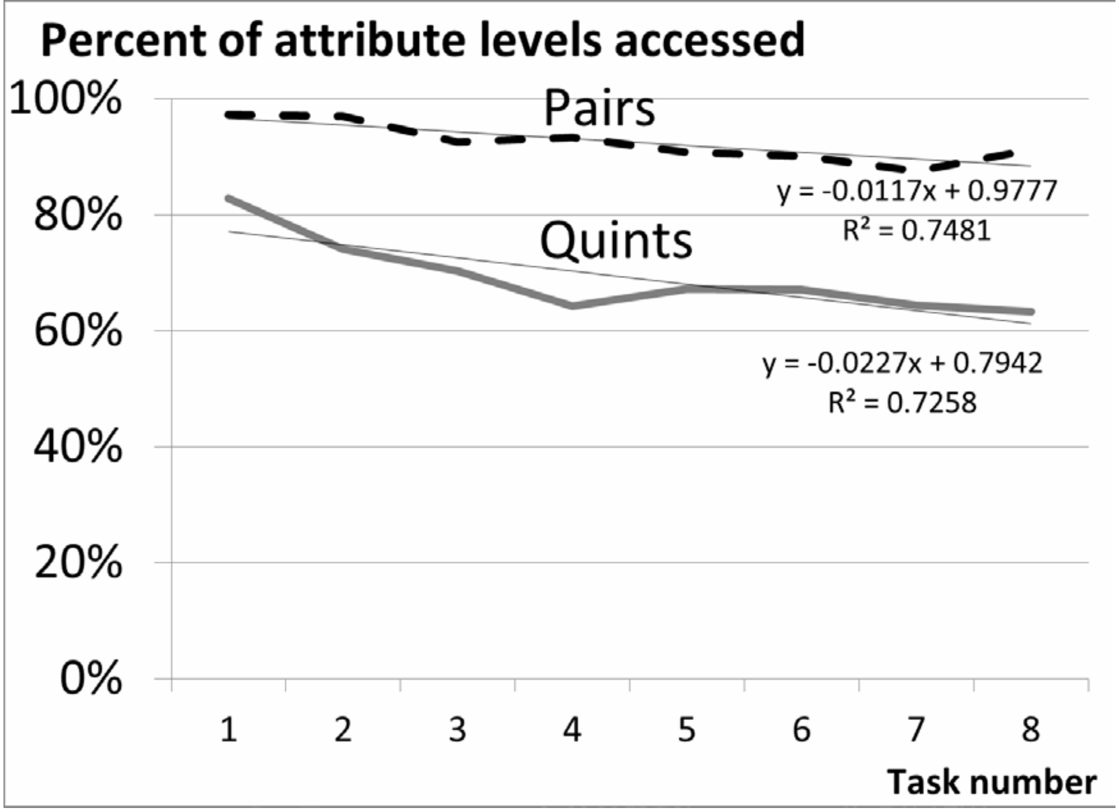
Boatwright and Monroe (2012) found a similar decrease in the number of fixations when testing choice from simulated product shelves.

Figure 4. Number of Fixations on Attribute Levels



Second, in order to assess the degree of simplification we investigate how many attribute levels the respondents fixated on at least once. In line with our expectations, respondents in the pairs condition accessed 92% of the information available, compared with 69% for quints. As can be seen from Figure 5, pairs access a greater proportion of information and are less likely to reduce that coverage with task experience. Our finding is in line with Yang, Toubia and De Jong (2015) who investigated a sequence of 20 choice tasks including four alternatives. Yang *et al.* found that respondents across all tasks looked at about 70% of the available information and found a similar downward trend with respect to the percent of attribute levels fixated in later tasks.

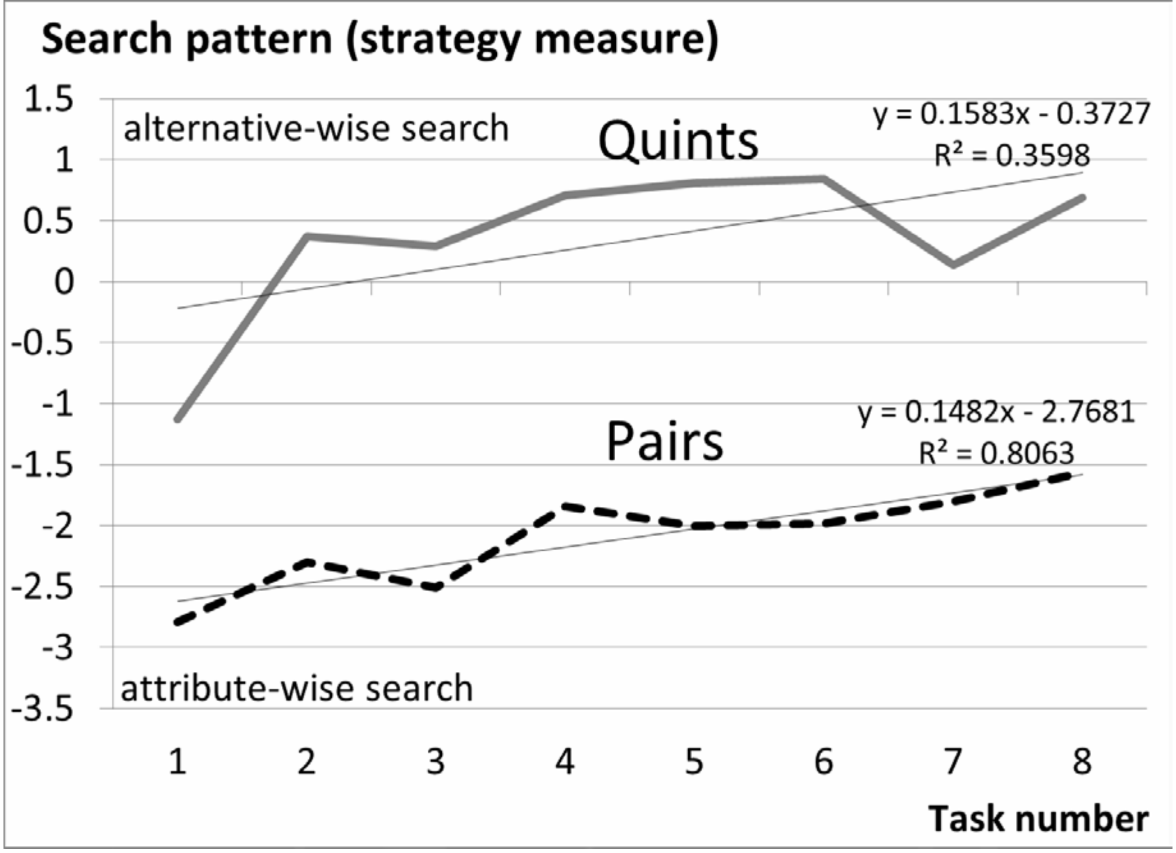
Figure 5. Percent of Attribute Levels Accessed by Task Number



Third, we compare the search (or saccade) pattern used in pairs and quints. A frequently used measure to describe the search pattern is the strategy measure (Böckenholt and Hynan 1994) which quantifies the extent to which information is searched attribute-wise, i.e., comparing alternatives within attributes, or alternative-wise, comparing attributes within alternatives. Because the strategy measure takes into account that the probability of attribute-wise and alternative-wise transitions changes for different numbers of alternatives, the strategy measure is the preferred index for assessing how respondents process task-relevant information (Schulte-Mecklenbeck, Kühberger and Ranyard 2011). A negative value of the strategy measure indicates attribute-wise processing whereas positive values indicate alternative-wise processing.

With respect to the search patterns we find that respondents conducted more within-attribute processing for pairs. This result therefore is in line with previous work (Russo and Dosher 1983) showing that decision makers process the information primarily attribute-wise on pairs. For quints respondents used a mixture of attribute-wise and alternative-wise processing, but in both conditions greater task experience resulted in greater alternative-wise processing. That result for both quints and pairs is consistent with Meissner and Decker (2010) who observe a progression to within-alternative transitions. In our pairs condition, however, the shift towards alternative-wise processing is small. Across all tasks respondents process the information attribute-wise which suggests that respondents continued to emphasize an additive difference strategy throughout their eight choices.

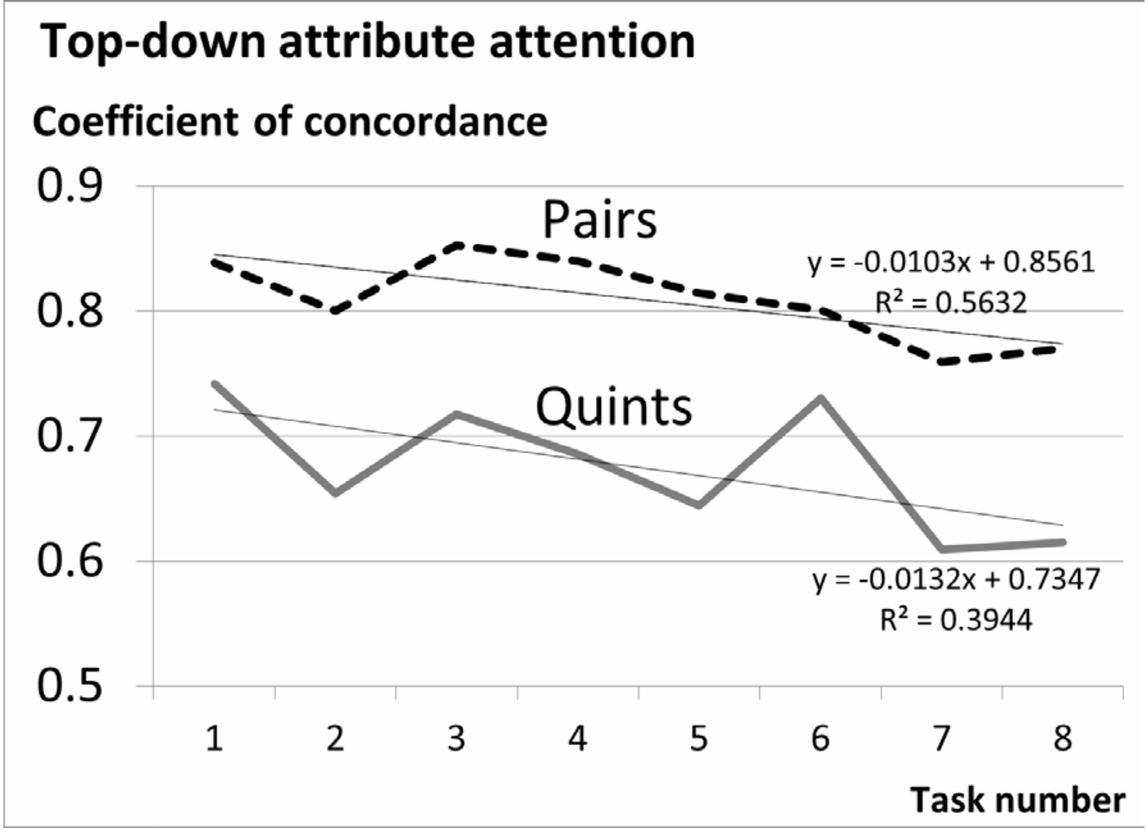
Figure 6. Search Pattern (Strategy Measure)



Fourth, we test whether respondents searched the information following a systematic pattern from the top to the bottom of the screen. Because respondents are found to simplify more in quints, they should also be less likely to search information from the top to the bottom of the screen. We therefore defined the following measure to quantify systematic top-down search: We rank all attribute levels with respect to when they were first fixated in a task. The average rank of all levels belonging to an attribute indicates how early the attribute was considered in the search process. We then compare the average ranking of the attributes with a top-down ranking and calculated the coefficient of concordance between the two sets of ranks. A value close to 1 will indicate top-down processing whereas a value close to zero will indicate that attributes are not considered in a schematic way from top to bottom.

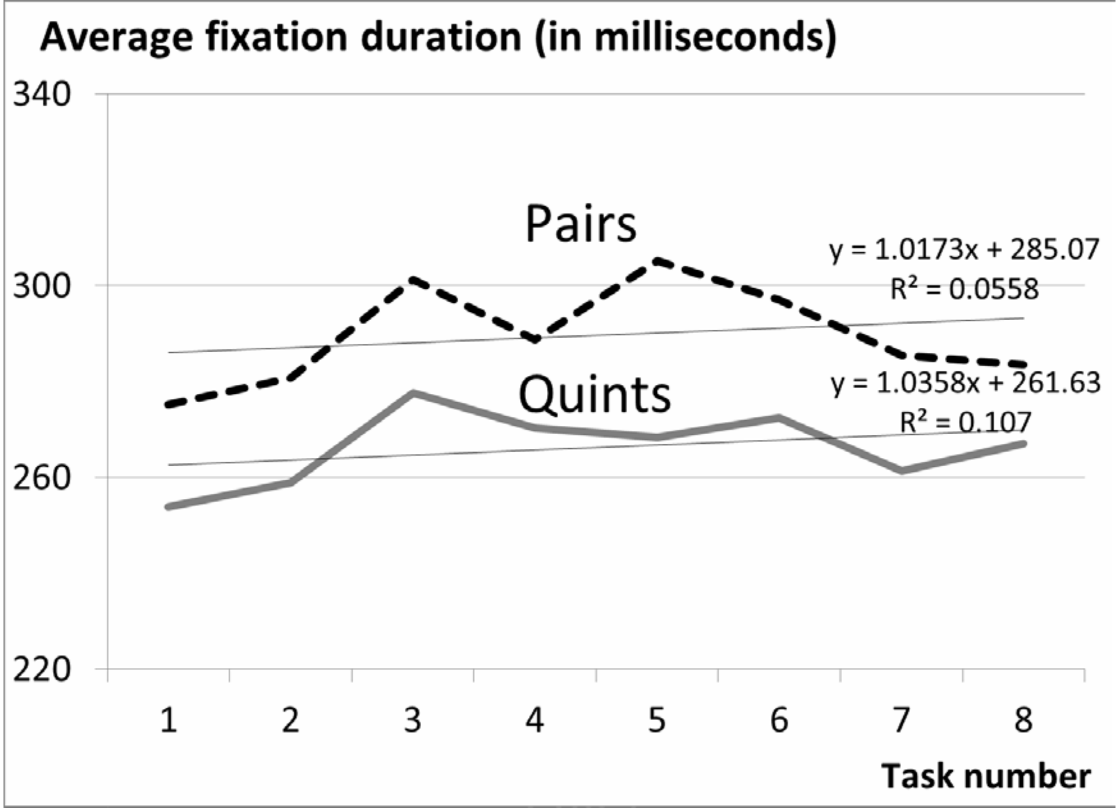
As Figure 7 shows, respondents on average processed the information in the choice tasks more often from the top to the bottom when the tasks included only two instead of five alternatives. This finding is in line with the use of an additive difference strategy in which the respondents compared the alternatives attribute-wise, look at almost all features and process the information from the top to the bottom of the screen. Figure 7 also shows that in later tasks respondents process information less schematically in both the pairs and quints condition.

Figure 7. Top-Down Attribute Attention



Fifth, we consider the average fixation duration of all fixations in a choice task. According to the literature very short fixations taking less than 200 milliseconds are often used for scanning and automatic processes, as for example, to understand the structure of a task when the respondent begins processing the information. By contrast, very long fixations might indicate an increased level of processing in more difficult tasks (see e.g., Velichkovsky *et al.* 2002). The average fixation duration for pairs is 296 ms and for quints it is 267 ms. This difference between pairs and quints is statistically significant ($t=5.3$, $p<.01$). We interpret this difference in fixation durations as evidence for the use of cognitive processes which involve differencing and adding for pairs, a process that is consistent with the idea that comparisons across alternatives are more time consuming than those within alternatives.

Figure 8. Average Fixation Duration



In summary, the average process measures differ strongly between pairs and quints. Pair processing is more thorough, covering proportionately more information, in a more top-down manner, and for greater durations. Quints encourage greater simplification initially as well as over time and lead to deeper processing within a few selected alternatives. Thus pairs fit a model of additive differences while quints reflect a concerted effort to identify a reasonable choice without getting confused by multiple items of available information.

Task Perception

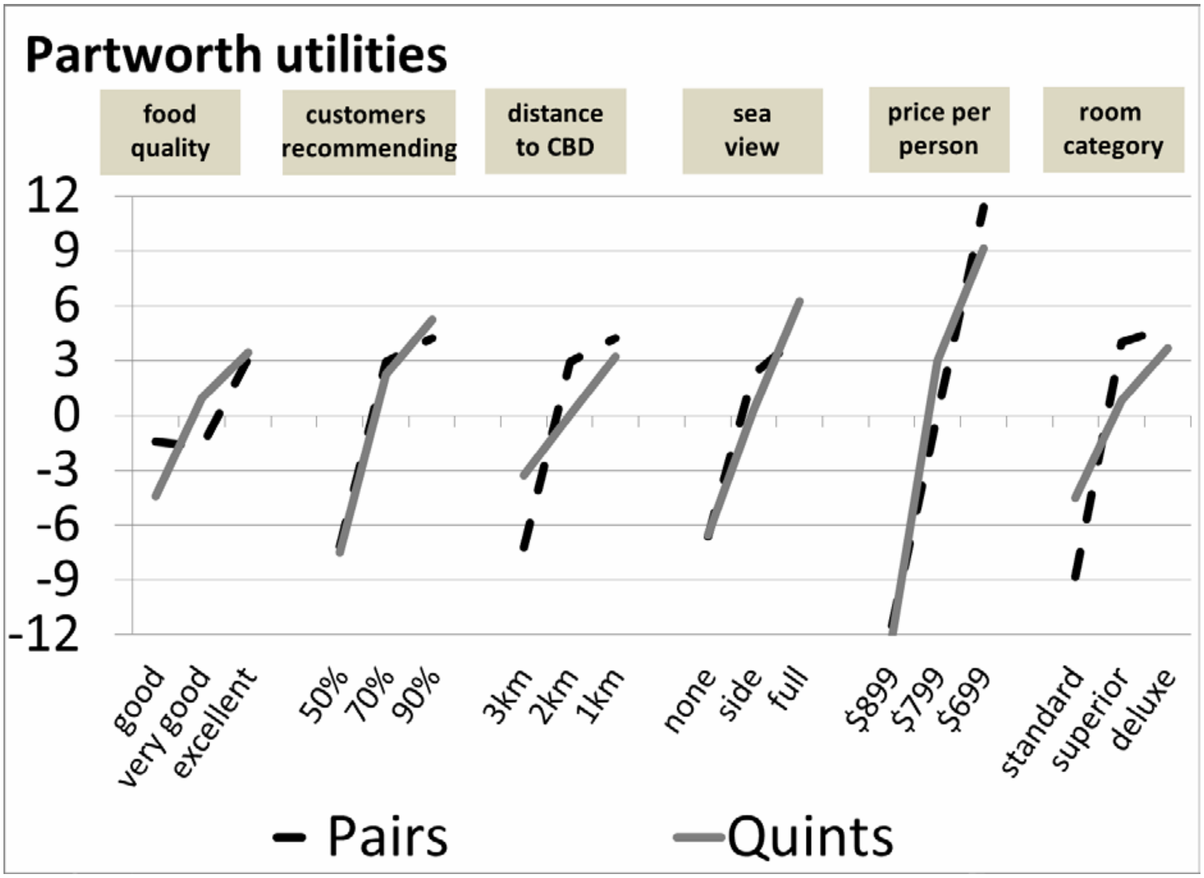
After the initial set of eight choice tasks and before the holdout choice triples, we asked respondents, "How difficult was it for you to choose the vacation package you wanted when last making a choice?" using a 7-point rating scale ranging from "not at all difficult" (-3) to "extremely difficult" (+3). To our surprise, pairs (M=.6, SE=.2) were perceived to be significantly (t=2.9, p<.01) more difficult than quints (M=-.4, SE=.3). Although respondents needed fewer seconds to finish pairs (M=15.0, SE=9.4) compared to quints (M=24.2, SE=14.7), the pairs seem to be cognitively more demanding. Given that most respondents in the pairs condition had to look through most of the information, the tedium of doing that eight times may have made it seem more difficult. By contrast, the goal of finding an acceptable beach vacation was perceived as easier for our respondents in the quints. These differences are also reflected in differences in the patterns of part-worths and predictive accuracy under pairs compared with quints.

Similarity of the Part-Worths

The part-worth utilities were computed on the individual level by applying Sawtooth Software's Hierarchical Bayes (HB) multinomial logit (MNL) estimation. As shown in Figure 9, the average part-worth utilities are similar, with a correlation of r=.92. Contrary to our expectations, visually it appears that pairs demonstrate greater non-linearity in valuations within attributes. This result suggests that it is more likely that non-linear cutoff values were used when respondents answered the pairs questions.

Next, we also analyzed attribute importance weights. We calculated attribute importances by calculating the ratio of the range of an attribute's utilities against the sum of the ranges across all attributes. The correlation of the average importance weights is also high (r=.86). Importantly, pairs elevate unimportant attributes. The mean of the standard deviation of importances across respondents is 20% less for pairs than for quints (M(pairs)=.097, M(quints)=.117; t=-3.1, p<.01). That finding is consistent with pairs generating a focus on all attributes. The increased attention on less important attributes increases the relative importance of these attributes in the decision process.

Figure 9. Comparison of the Part-Worth Utilities in Pairs and Quints



Predictive Performance

To evaluate the predictive performance, we first looked at the internal hit rates and the output from Sawtooth Software's Hierarchical Bayes estimation. The key measures are included in Table 2.

It does not make much sense to directly compare the internal hit rates for pairs and quints, because the probability of correctly predicting a pair at random is 50%, but is only 20% for quints. The average hit rate is 72% for pairs and therefore is 22% above chance level. For quints, the improvement above chance level is 35%, given a hit rate of 55%.

Because it is hard to correct hit rates for the number of alternatives in the choice set, percent certainty, or another likelihood-based statistic, is a more appropriate indicator of model fit. It is an information-theoretic measure that compares the information explained by the model to the total uncertainty of the system (Hauser 1978). The measures included in Table 2 show that the internal model fit is better for pairs than for quints. The percent certainty for pairs is .90 whereas it is only .69 for quints. One explanation for this result might be that respondents in the pairs condition more consistently applied the same (additive difference) strategy, but in case of quints used all kinds of different decision strategies. As a consequence the internal consistency might be lowered for quints.

Table 2: Predictive Performance Measures

Measure	Pairs	Quints
Percent Certainty	.90	.69
RLH: Root Likelihood	.93	.60
Internal hit rate	72%	55%
Hit rate from cross-validation	75%	53%
Hit rate predicting holdout triples	76%	57%

Consistent with the low error as indicated by the Percent Certainty, pairs more consistently predicted the holdout triples shown at the end of the survey. The hit rate for pairs is 87% (69%) in the first (second) holdout task whereas it is only 56% (58%) for quints. This difference is significant for the first holdout task (=6.4, p=.01), and is directionally consistent for the second holdout task (=1.1, p=.30). There are two possible reasons why pairs predicted the holdout choices better than quints. First, pairs are more similar to triples than quints, meaning that respondents who have frequently used an additive difference strategy in a sequence of pairs might continue to do so in the consecutive triples. Second, the error around pairs to predict holdouts may simply be sufficiently smaller for quints enabling pairs to overcome their 30% deficit in statistical efficiency.

CONCLUSIONS

Summary of Empirical Findings

The important lesson is that the decision making process is very different when choosing between two versus higher multiples of alternatives. Those processing differences lead to different patterns of part-worths and predictive accuracy, and suggest contexts in which either task is more appropriate.

For pairs the pattern of fixations and saccades is consistent with an additive difference strategy. That strategy assesses the relative benefit one attribute at a time and sums those differences across attributes to identify the most preferred option. This within-attribute processing has the advantage of enabling an assessment of each attribute independently from the other attributes. We find that an additive difference strategy leads to greater use of available information, with 92% of the pair information fixated on compared with 69% of the quint information.

The process with more alternatives is quite different. For quints, the need for simplification across 30 pieces of information encourages the use of an important attribute to rule out less promising options. The process of finding a good option from many alternatives can best be described as a search that over time gets more effective at focusing on less information to identify a satisfactory choice. The variability in the search strategy for quints contrasts a relatively mechanistic and stable choice process on pairs.

In terms of efficiency, compared with quints, pairs took 40% less time, were 31% less statistically efficient, but generated 33% more accurate predictions of holdout triples and were more consistent internally. Our results therefore contradict Pinnell and Englert (1997) who find that pairs are no better at predicting holdouts. Perhaps because of the need for accuracy, pairs are perceived to be substantially more difficult. We suggest that this difference can be explained with the cognitive process in pairs which seems to be more demanding.

The average part-worths from the pairs and the quints seem similar, with a correlation of .92. However, differences in details matter. Pairs provide more discrimination of levels within attributes while quints reveal greater discrimination across attributes. In particular, visual inspection suggests greater non-linearity within attributes for pairs, e.g., revealing a large difference between poor and good food compared with good to excellent. This finding is in line with the results by Pinnell and Englert (1997) who observe a “non-linear relationship indicating a loss aversion effect” (p. 150) for pairs. By contrast, for quints the relationship between the three levels is far more linear. That said, quints reveal 20% greater separation in the relative importance of attributes; it appears the complexity of having more attributes reveals respondents’ strategy to focus attention on more important attributes. Here, our results contradict the earlier findings by Pinnell and Englert (1997) who found that “important attributes are more important in pairs” (p. 150).

How Many Options Should You Use in Your Conjoint Study?

We suggest that pairs are appropriate if one wants efficient measures of how people use many attributes to make choices. Pairs also make sense when the choice is difficult or highly emotional. When patients make choices that involve trading off substantial loss of income and

hospital time against longer expected life, having just two options makes the decisions less overwhelming.

However, where the goal is to simulate choices where there are many alternatives and relatively few attributes, then a multi-alternative CBC is appropriate. A good example would be shelf studies that explore consumer ability to find preferred brands in a complex display and respond to different promotional efforts. Put differently, if the decision process involves substantial simplification to find a reasonable option from a large set, then there are advantages to showing a greater number of alternatives per conjoint choice set.

We remain surprised at finding that fewer than 10% of the studies reported in the past four Sawtooth Software Proceedings used pairs. One reason for the lack of use of pairs may stem from the well-known finding that having many alternatives improves the technical statistical efficiency of the design. A second and more reasonable problem with pairs arises from the process revealed by eye-tracking. The additive difference process may be more effective at revealing consistent tradeoffs, but may be even farther removed from what happens in the marketplace with many attributes and many alternatives.

That said, we believe that pairs are underutilized. Apart from greater efficiency, pairs are appropriate when decisions are sufficiently important that simplification to a few attributes makes little normative sense. Furthermore, pairs will be more efficient at assessing consumer reaction to changes in all attributes, in cases where decisions are very important thus justifying consideration of all attributes. Pairs also are reasonable when the attributes are novel, or where when respondents have deep emotional reactions. In the latter situations pairs might help respondents because a weighted additive process facilitates the thoughtful integration of all the attributes of a decision.


Martin Meissner


Harmen Oppewal


Joel Huber

REFERENCES

Bettman, J., Johnson, E. J., & Payne, J. W. (1991). Consumer Decision Making. In T. Robertson & H. Kassarjian (Eds.), Handbook of Consumer Behavior (50–84). NJ: Prentice-Hall: Englewood Cliffs.

Böckenholt, U., & Hynan, L. S. (1994). Caveats on a Process-Tracing Measure and a Remedy. Journal of Behavioral Decision Making, 7(2), 103–117.

Bunch, D., Louviere, J., & Anderson, D. (1983). "A Comparison of Experimental Design Strategies for Multinomial Logit Models: The Case of Generic Attributes," Working Paper,

Graduate School of Business, University of California, Davis (available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.196.4913&rep=rep1&type=pdf).

Eggers, F., Hauser, J. R., & Selove, M. (2016). The Effects of Incentive Alignment, Realistic Images, Video Instructions, and Ceteris Paribus Instructions on Willingness to Pay and Price Equilibria. Sawtooth Software Conference Proceedings 2016.

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process Tracing Methods: Contributions, Problems, and Neglected Research Questions. Organizational Behavior and Human Decision Processes, 43(1), 75–117.

Hauser, J. R. (1978). Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information-Theoretic Approach. Operations Research, 26(3), 406–421.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press: Oxford.

Louviere, J. J., & Woodworth, G. (1983). Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. Journal of Marketing Research, 20(4), 350–367.

Meissner, M., & Decker, R. (2010). Eye-Tracking Information Processing in Choice-Based Conjoint Analysis. International Journal of Market Research, 52(5), 593–610.

Meissner, M., Musalem, A., & Huber, J. (2016). Eye-Tracking Reveals a Process of Conjoint Choice That Is Quick, Efficient and Largely Free from Contextual Biases. Journal of Marketing Research, 53(1), 1–17.

Orme, B. (2013). The CBC System for Choice-Based Conjoint Analysis—Version 8. Sawtooth Software Inc., Orem, Utah, (accessed October 10, 2016, available at http://www.sawtoothsoftware.com/support/technical-papers/cbc-related-papers/cbc-technical-paper-2013).

Orquin, J. L., Bagger, M. P., & Mueller Loose, S. (2013). Learning Affects Top Down and Bottom Up Modulation of Eye Movements in Decision Making. Judgment and Decision Making, 8(6), 700–716.

Payne, J. W. (1976). Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. Organizational Behavior and Human Performance, 16(2), 366–387.

Payne, J. W., Bettman, J. R., Coupey, E., & Johnson, E. J. (1992). A Constructive Process View of Decision Making: Multiple Strategies in Judgment and Choice. Acta Psychologica, 80 (1–3), 107–141.

Pinnell, J., & Englert, S. (1997). The Number of Choice Alternatives in Discrete Choice Modeling. Sawtooth Software Conference Proceedings 1997, 121–153, (available at https://www.sawtoothsoftware.com/download/techpap/1997Proceedings.pdf#page=135).

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. Psychological Bulletin, 124(3), 372–422.

Russo, J. E., & Dosher, B. A. (1983). Strategies for Multiattribute Binary Choice. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9(4), 676–696.

Schulte-Mecklenbeck, M., Kühberger, A., & Ranyard, R. (2011). The Role of Process Data in the Development and Testing of Process Models of Judgment and Decision Making. Judgment and Decision Making, 6(8), 733–739.

Stüttgen, P., Boatwright, P., & Monroe, R. T. (2012). A Satisficing Choice Model. Marketing Science, 31(6), 878–899.

Tobii Software (2016). Tobii Studio 3.4.5 User Manual. (accessed October 10, 2016, available at http://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf).

Todd, P. M. (2007). How Much Information Do We Need? European Journal of Operational Research, 177 (3), 1317–1332.

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an Express-Diagnostics for Level of Processing and Hazard Perception. Transportation Research Part F: Traffic Psychology and Behaviour, 5(2), 145–156.

Yang, L., Toubia, O., & De Jong, M. G. (2015). A Bounded Rationality Model of Information Search and Choice in Preference Measurement. Journal of Marketing Research, 52(2), 166–183.