The Influence of Macro-Level Motives on Consideration Set Composition in Novel Purchase Situations

AMITAV CHAKRAVARTI
CHRIS JANISZEWSKI*

Consumers often have to create consideration sets when purchasing goals are not well defined. In these situations, the contents of a consideration set depend on a combination of two motives. First, consumers prefer to create a consideration set of easy-to-compare alternatives. It is easier to compare alternatives that have alignable attributes or alternatives that have overlapping features. Second, consumers prefer to create consideration sets that have a high likelihood of containing their optimal alternative. For example, when the set of available alternatives requires the consumer to make trade-offs between benefits (i.e., to be compensatory), the consumer often delays making a decision about which benefits are preferable, and the consideration set tends to contain a more diverse set of alternatives. We document several factors that influence the relative importance of one or the other motive in consideration set formation and discuss implications for brand managers.

Consideration set formation is a fundamental stage of prechoice decision making (Alba, Hutchinson, and Lynch 1991; Howard and Sheth 1969; Nedungadi 1990; Ratneshwar and Shocker 1991). Consideration set composition can be used to understand consumer choice strategies, brand loyalty, and changes in market share that are independent of brand evaluation (Kardes et al. 1993; Nedungadi 1990). In fact, it is the inherent instability of the composition of a consumer’s consideration set that creates an opportunity for brand managers to increase a brand’s market share (Desai and Hoyer 2000; Hutchinson, Raman, and Mantrala 1994).

Consideration set composition depends on a consumer’s screening criteria and screening process. When consumption goals are well defined, these screening criteria are based on the benefits desired for a particular usage situation (Desai and Hoyer 2000; Nedungadi 1990; Ratneshwar and Shocker 1991). In such routine choice scenarios, consumers use their past experiences to narrow down their awareness set, even at the very initial stages of the decision process. This screening process can depend on a simple heuristic like brand familiarity (Desai and Hoyer 2000; Johnson and Lehman 1997) or a more effortful process described by a disjunctive or elimination-by-aspects (EBA) choice rule (Tversky 1972). Consumers gather information about alternatives from memory or the environment, compare this information to cutoffs, and eliminate alternatives that do not meet these cutoffs (Tversky and Sattath 1979).

There are also purchase situations where consumption goals are poorly defined, as is the case with novel product categories, novel buying situations, and novel consumption contexts (Barsalou 1983; Desai and Hoyer 2000). In these situations, consideration set composition may be influenced by goals other than those created by the consumption context. For example, consumers may prefer to create consideration sets that consist of items that are easy to compare or that increase the likelihood that an optimal alternative is retained for further consideration (Huber and Kline 1991; Wright 1975). Macro-level goals have been well documented in decision making (e.g., Bettman 1979; Johnson and Payne 1985; Ratneshwar and Shocker 1991), but they have usually been used to explain limits on the size of a consideration set as opposed to determining the composition of a consideration set (e.g., Jarvis and Wilcox 1973; Mitra and Lynch 1995; Roberts and Lattin 1991). It is possible that these macro-level goals also make consumers sensitive to the characteristics of the alternatives themselves.

This article investigates the impact of two consideration set formation goals—the motivation to simplify the choice process and the motivation to optimize the choice outcome—on consideration set composition (see fig. 1). The

* Amitav Chakravarti (achakrav@stern.nyu.edu) is assistant professor of marketing, Leonard N. Stern School of Business, New York University, New York, NY 10012-1126. Chris Janiszewski (chris.janiszewski@cba.ufl.edu) is the Jack Faricy Professor of Marketing, Warrington College of Business, University of Florida, P.O. Box 117155, Gainesville, FL 32611-7155. This article is based on the first author’s doctoral dissertation. The authors acknowledge the helpful input of the editor, the assistant editor, and three reviewers.
first section of the article shows that a consumer’s desire to simplify a choice encourages the formation of consideration sets consisting of easily compared alternatives. We argue that both attribute alignability (study 1) and feature overlap (study 2) make alternatives easier to compare, and hence, are fundamental drivers of consideration set composition. The second section of the article shows that a consumer will activate an optimization strategy when necessary, encouraging higher levels of across-category consideration and formation of relatively more heterogeneous consideration sets. Study 3 shows that, if consumers are sensitized to committing a type II error, they are likely to create heterogeneous consideration sets. Study 4 demonstrates that involvement also plays a moderating role in determining the contents of a consideration set. Study 5 shows that people create more heterogeneous consideration sets when confronted with an alternative set having negatively correlated benefits. Finally, study 6 shows that the consideration set becomes less heterogeneous if the consumer is encouraged to determine the relative importance of the competing benefits prior to forming the consideration set. We end with a discussion of marketing implications and avenues for future research.

CONSIDERATION SET FORMATION

Consideration set formation has often been described as a relatively effortless process aimed at simplifying the more burdensome final choice task (e.g., Bettman 1979; Huber and Kline 1991; Johnson and Payne 1985). Consumers often encounter a large set of alternatives and use screening criteria to reduce the number of alternatives that will ultimately be compared. In familiar purchase situations, a simple screening rule might rely on brand familiarity or memory accessibility. In novel purchase situations, especially those that are stimulus based, the consumer is likely to focus on one or more attribute
cutoffs. The screening process is more likely to be a disjunctive or elimination-by-aspects choice rule where (a) each alternative is viewed as a set of aspects or attributes and (b) all alternatives that do not possess a particular desirable aspect are eliminated from further consideration. The advantage of these screening criteria and screening processes is that they allow a person to consider alternatives one at a time, retaining only those alternatives that pass the predetermined cutoffs. Yet, attribute-based, noncompensatory screening rules may not be the only determinants of consideration set composition. The consumer’s motivation to simplify the choice process and optimize the choice outcome may also influence the composition of the consideration set.

Simplify the Choice Process

Past research on consideration set formation is replete with examples showing that the desire to simplify the final choice encourages the consumer to reduce the number of alternatives to a more manageable size (Roberts and Lattin 1991). It is also possible that the desire to simplify the final choice task may encourage consumers to retain alternatives that are easy to compare (Medin, Goldstone, and Markman 1995). Past research from diverse domains provides overwhelming evidence that comparability has a strong influence on individual decision making (see Medin et al. [1995] for an excellent review). This research shows (a) that individuals are more strongly influenced by comparable aspects of a decision context and (b) that individuals attempt to enhance comparability in decision contexts where comparisons are difficult to make. It is quite likely that this desire for comparability also manifests itself at the consideration set formation stage. Retaining comparable items in the consideration set considerably eases information-processing demands during the choice stage since it involves comparing information that is commensurable.

The retention of comparable items in the consideration set may be a less effortful and less active process than a noncompensatory screening process because attribute cutoffs and decision criteria are not necessary (Hogarth 1980). In fact, it is unlikely that the retention of comparable items in the consideration set represents an intentionally controlled, effortful, and conscious endeavor on the part of the decision maker (Medin et al. 1995). Instead, this need to retain comparable items may stem from the consumer’s innate desire to minimize effort at the time of choice (Huber and Kline 1991). Thus, although decision makers are unlikely to articulate that retaining comparable items was an explicit aim of their consideration set, it is highly likely that comparability will influence the composition of the consideration set (Markman and Moreau 2001).

A comparability-based screening process may often occur in conjunction with a noncompensatory screening process. For example, consider a two-stage screening process for homes, movies, or gifts in which consumers screen alternatives using attributes and then screen again using comparability as a guide (Levin and Jasper 1995; Wright and Barbour 1977). In step 1, consumers may actively screen a potential home purchase on the number of bedrooms, price, and location; rental movies on genre; or gifts on price point and product class. This initial screen results in a large number of alternatives that make up the awareness set. In many cases, these initial screening criteria are not related to final choice criteria, yet they are useful because the final choice criteria do not yet exist (Payne, Bettman, and Johnson 1988). In step 2, macro-level choice goals, such as the ease of comparison, lead to a second screen of the alternatives. This second screen can depend on the alignability of the attributes used to describe the alternatives or the feature similarity of the alternatives being compared (i.e., feature overlap). Medin et al. (1995) observe that alternatives having alignable attributes are perceived as easier to compare than alternatives having nonalignable attributes. Likewise, Johnson (1984, 1986) suggests that comparisons among available alternatives are much easier when alternatives have overlapping features.

Attribute Alignability. There is considerable evidence that people naturally assign more weight to alignable attributes and that they experience difficulty comparing alternatives with nonalignable attributes (Medin et al. 1995). For example, Markman and Medin (1995) present consumers with pairs of alternatives consisting of comparable (i.e., brands described on the same attributes) and noncomparable (i.e., brands described on different attributes) brands. They find that people pay more attention to alignable attributes and that people are more likely to mention alignable attributes as a justification for a decision. Thus, brands that have more alignable attributes are perceived as more comparable than brands that have more nonalignable attributes.

The choice literature provides some evidence that attribute alignability influences consideration set formation. When people are asked to select an alternative from a heterogeneous set of alternatives, they often reduce their consideration set to two or more comparable alternatives. For example, Tversky and Sattath (1979) asked subjects to choose a famous person they would like to meet from a set composed of movie stars and politicians. Sets consisted of one person from group A (e.g., Chaplin) and two people from group B (e.g., Gandhi and Churchill). Tversky and Sattath found that subjects would eliminate the sole representative of one group (e.g., Chaplin) and then select from the two alternatives in the other group (e.g., Gandhi and Churchill). Likewise, studies conducted in the domain of noncomparable choice alternatives (e.g., Bettman and Sujan 1987; Johnson 1988, 1989) show that, when choosing among alternatives from different product categories (e.g., two desk lamps and two desk clocks), subjects process information in a hierarchical fashion. Thus, subjects eliminate entire categories, or subcategories, of alternatives during the initial stages of decision making, which, in turn, limits their subsequent choice to members of one category.

It is quite possible that the ease of comparison motive led to the pattern of results observed in the choice studies. In the Tversky and Sattath study, subjects may have chosen to retain the two politicians in their choice set because Gan-
CONSIDERATION SET COMPOSITION

247
dhi and Churchill shared several alignable, easy-to-compare attributes. Similarly, when deciding among noncomparable alternatives, subjects may have chosen to retain alternatives from one category (e.g., lamps) because these alternatives were easier to compare. However, none of these studies establishes an unequivocal link between the ease-of-comparison motive and the composition of the consideration set, because an EBA screening strategy could have generated the same pattern of results. For example, a cutoff that specifies that the famous person must have been a world leader would have eliminated all actors, leading to the retention of Gandhi and Churchill. Thus, although these studies suggest attribute alignability can influence consideration set composition, there have been no direct tests of this hypothesis that are not confounded with an EBA screening strategy.

H1: As the alignability of a pair of alternatives increases, the likelihood the pair will enter the consideration set increases.

Feature Overlap. Consumers also face decision scenarios where all alternatives are described on the same attributes. In such cases, attribute alignability becomes ineffective in guiding screening decisions. Yet, the consumer may still have the motivation to retain easily compared alternatives in order to simplify the choice task. In such scenarios, a strategy for simplifying a choice task is to retain alternatives that have overlapping features (Tversky 1972, 1977). As an example, consider a consumer choosing between different models of automobile radar detectors, described on the presence or absence of four features—360-degree detection, city/highway mode, signal strength detector, and laser detection. Although description on a common set of four attributes ensures that there is perfect attribute alignability between the alternatives, not all comparisons are equally easy to make. Comparisons that have substantially different properties are more difficult to make. It is easier to compare two alternatives that have a fair degree of feature overlap (e.g., option A with 360-degree detection and city/highway mode; option B with 360-degree detection and signal strength detector), than two alternatives that have no overlapping features (e.g., option A with 360-degree detection and city/highway mode; option B with signal strength detector and laser detection). When options have few overlapping features, the consumer has to translate the presence or absence of features into corresponding benefits in order to compare alternatives. This is an effortful process (Medin et al. 1995). It is much easier to compare a set of alternatives that have a fair amount of feature overlap. Thus, to the extent a person is motivated to simplify the choice task, she or he should be sensitive to the amount of feature overlap between alternatives.

Although there is no direct evidence that feature overlap makes alternatives more comparable, there is evidence that feature overlap makes alternatives more similar (Tversky 1977). In addition, there is evidence that similar alternatives are easier to compare (Gentner and Markman 1994), provided that the alternatives can be discriminated (see Moorthy, Ratchford, and Talukdar 1997). Thus, to the extent the consumer can anticipate that underlying benefits will have to be generated at the time of choice, retaining alternatives that have a considerable amount of feature overlap should make this benefit-generation process easier.

H2: As the number of overlapping features of a pair of alternatives increases, the likelihood the pair will enter the consideration set increases.

Optimize the Choice Outcome

A second macro goal that may be active during consideration set formation is the desire to optimize the choice outcome by avoiding type II errors. Huber and Kline (1991) describe consideration set formation as a trade-off between imposing cutoffs that simplify a decision and taking a chance that an optimal alternative will be discarded. This concern about errors of exclusion is likely to become more salient as the risk of neglecting an optimal alternative increases. Bettman (1979) argues that this risk increases when (a) the consumer lacks experience, (b) the available brand information is ambiguous, and (c) the consumer is not confident about the brand information. Risk also increases when the consumer is uncertain about his or her preferences (Kardes et al. 1993; Simonson 1990). First, the consumer can be uncertain about the combination of desired features (Kardes et al. 1993). First-time buyers or buyers of new products and innovations will not have well-established preference structures. Second, consumers can be uncertain about the weights to assign to desired benefits when making an overall evaluation. Alternatives having negatively correlated benefit structures (Huber and Kline 1991) create uncertainty about the reliability of the choice. Third, the consumer can have a difficult time anticipating the benefit weights at the time of choice. The further into the future a choice will be made, the more uncertain the consumer becomes about the stability of currently known benefit weights (Simonson 1990).

Consumers are likely to perceive more risk of a type II error when purchasing in novel product categories, novel buying situations, and novel consumption contexts. To hedge against the risk of committing a type II error, the consumer is likely to retain relatively more dissimilar items in the consideration set. Consumers may choose to retain maximally dissimilar items, preferably with uncorrelated errors, to ensure that they do not miss out on a brand that could have otherwise been optimal (Roberts and Lattin 1997):

H3: As the concern for committing a type II error increases, the likelihood a diverse pair of alternatives will enter the consideration set increases. The concern for committing a type II error is more likely when (a) benefits are negatively correlated, and (b) there is uncertainty about the relative importance of benefits.
STUDY 1

The goal of study 1 was to assess whether consideration set composition was sensitive to the comparability of alternatives. To investigate this issue, we manipulated comparability by altering the degree of attribute alignability between two subcategories of homogeneous brands. We reasoned that increasing the alignability of the attributes describing the alternatives in these two subcategories of brands should encourage people to form consideration sets consisting of brands from both subcategories (hypothesis 1).

Design and Stimuli

Four replicates (e.g., video games, cars, restaurants, and vacations) were used to investigate the hypothesis that consideration set composition is sensitive to the alignability of the attributes describing the alternatives. Each replicate was made up of 10 alternatives, five from one subcategory (e.g., baseball video games) and five from a second subcategory (e.g., football video games; see fig. 2). Alternatives within a subcategory were always described on a common set of attributes (i.e., they were perfectly alignable). The comparability of the alternatives across the two subcategories was manipulated by varying the number of attributes common to the two subcategories. In the low alignability condition, there were no alignable attributes across the two subcategories. The five alternatives in subcategory 1 were described on one set of six attributes (A1–A6), and the five alternatives in subcategory 2 were described on a second set of six attributes (A7–A12). In the moderate alignability condition, there were two alignable attributes across the two subcategories (A6–A7) plus five unique attributes per subcategory (A1–A5, A8–A12). In the high alignability condition, there were four alignable attributes across the two subcategories (A5–A8) and four unique attributes per subcategory (A1–A4, A9–A12).

Additional moderate alignability and high alignability conditions were created because of a potential confound created by the alignability manipulation. The number of attributes describing the alternatives in the original low, moderate, and high alignability conditions were six, seven, and eight, respectively. This confound was created because of the desire to keep the total number of attributes used to describe the entire set of alternatives constant at 12. Alternatively, we could have created stimuli that kept the number of attributes describing any one alternative constant at six. Therefore, an additional moderate alignability condition was created in which there were two alignable attributes across the two subcategories (A6–A7) plus four unique attributes per subcategory (A1–A4, A9–A12). The additional high alignability condition had four alignable attributes across the two subcategories (A5–A8) and two unique attributes per subcategory (A1–A2, A11–A12).

Based on these manipulations, attribute descriptions were created for the 10 alternatives in each replicate. For example,
CONSIDERATION SET COMPOSITION

in the low alignability condition the description for each alternative consisted of an introductory sentence, a concluding sentence, and six attribute description sentences. The attribute description sentences were based on actual descriptions available on product packages and web sites.

Procedure

The procedure was designed to mimic a stimulus-based screening task in which people gather information as part of the screening process. Subjects were seated at computers and told how to perform a two-step decision-making task. In step 1, they would review the alternatives available in a category and form a “short list” of the brands for further consideration. In step 2, they would choose one alternative from the short list. Subjects were asked to base their decisions strictly on the information that was provided and to discount any previous experiences they may have had with the brands or the categories.

Subjects were then asked to consider the first replicate. On the first screen, subjects saw the names of 10 brands displayed as links. Below these links were instructions asking the subjects to form a short list of four alternatives and a text box where the subjects could list the names of the alternatives. Clicking on the link for a particular brand displayed the corresponding description. Subjects were required to click on each brand link at least once to ensure that no alternative was unintentionally overlooked and that there was no difference in the awareness set across the different conditions. There was no restriction on the amount of time the subjects could spend creating the consideration set, though we expected that subjects would spend less than five minutes forming a consideration set (the equivalent of 30 seconds per product description). This expectation was confirmed.

After the subjects had formed their consideration set, they clicked to a second screen and provided a brief explanation about how they created their consideration set. On the following screen, subjects were asked to make their final choice. After making their final choice, subjects were required to provide a brief explanation about how they made their final choice. Subjects then repeated this procedure for the other three replicates. The presentation order of the replicates was random. All four of the replicates represented the same experimental condition for a given subject.

After considering all four replicates, a manipulation check was administered. Subjects were asked to judge the comparability of the two subcategories within each replicate. For example, for the video game replicate they rated how comparable they thought the baseball games were to the football games. If the alignability manipulation was successful, the alternatives in the two subcategories should become more comparable as they shared more alignable features.

Dependent Measure

The four alternatives in a consideration set could be decomposed into six pairs of alternatives. The key dependent measure was the number of pairs consisting of an alternative from each subcategory. For example, a consideration set consisting solely of baseball video games would have no cross-category pairs (i.e., the minimum number of cross-category pairs), whereas a consideration set consisting of two baseball video games and two football video games would have four cross-category pairs (i.e., the maximum number of cross-category pairs).

Results

Sixty-six subjects from an introductory marketing course subject pool participated in the study for extra credit. The manipulation check verified that alternatives in different clusters became more comparable as the amount of alignability increased ($F(1, 61) = 37.7$, $p < .01$). It was easier to compare alternatives from different subcategories in the moderate alignability conditions ($M = 5.64$) than in the low alignability conditions ($M = 3.75$; $F(1, 61) = 44.53$, $p < .01$). It was also easier to compare alternatives in the high alignability conditions ($M = 6.25$) than in the moderate alignability conditions ($M = 5.64$; $F(1, 61) = 6.52$, $p < .05$). The manipulation check was not sensitive to the replicate ($F(12, 183) = 0.78$, $p > .10$).

The critical test involved the number of cross-category pairs in the consideration sets at each level of alignability. The alignability manipulation significantly influenced the number of cross-category pairs included in the consideration set ($F(4, 61) = 23.61$, $p < .01$, $\eta^2 = .28$). There were more cross-category pairs in the consideration sets in the moderate alignability condition ($M = 2.78$) than in the low alignability condition ($M = 1.29$; $F(1, 61) = 42.70$, $p < .05$, $\eta^2 = .46$). There were also more cross-category pairs in the consideration sets in the high alignability condition ($M = 3.50$) than in the moderate alignability condition ($M = 2.78$; $F(1, 61) = 15.64$, $p < .05$, $\eta^2 = .20$). We also confirmed that the two moderate alignability conditions ($F(1, 61) = 0.31$, $p > .10$) and the two high alignability conditions ($F(1, 61) = 2.53$, $p > .10$) did not significantly differ and that the replicate did not interact with the alignability manipulation ($F(12, 183) = 0.15$, $p > .10$).

These results show that as the comparability of two subcategories (e.g., football and baseball) of alternatives increased, people were more likely to include alternatives from both subcategories in their consideration set.

Discussion

The results of the first study are consistent with the hypothesis that consideration set composition is sensitive to the ease of comparing alternatives. As the alignability of the attributes describing two homogeneous groups of alternatives increased, people became more willing to include alternatives from both groups. As predicted by hypothesis 1, alignability promoted comparability and allowed people to consider a more diverse set of alternatives without a substantial increase in the effort required to compare the alternatives. Additionally, the use of a between-subjects manipulation of alignability with equivalent stimulus labels across
all conditions ensures that the effects are not attributable to the use of an EBA strategy. Since subjects were randomly assigned to an experimental condition, subjects should have had the same attribute cutoffs in all three alignability conditions. Therefore, the use of an EBA screening strategy cannot be used to explain the differences in across-category consideration that were observed across the low, moderate, and high alignability conditions. The only systematic difference across the three conditions was the degree of attribute alignability, and consequently, the ease of making comparisons. Thus the ease of comparison motive appears to play an important role in determining the content of consideration sets.

STUDY 2

Previously, we argued that two factors increase the comparability of alternatives. First, we argued that the ease of comparing alternatives depends on the alignability of the attributes used to describe the alternatives, as was shown in study 1. Second, we argued that, when alternatives are completely aligned, comparability could be further enhanced by an overlap of the features describing the alternatives. In study 2, we tested hypothesis 2 by manipulating the feature overlap of two sets of completely aligned alternatives.

Design, Stimuli, and Procedure

The experimental stimuli consisted of four replicates (automobile radar detectors, vacuum cleaners, microwave ovens, and air conditioners) that had low, moderate, or high feature overlap. A replicate consisted of 30 alternatives, all described on a common set of 12 features (i.e., the alternatives were completely aligned). For any one alternative, each of the 12 features was listed as present or absent. Having more alternatives than study 1, with simpler attribute descriptions, meant that the information load was comparable across the two studies.

Feature overlap was manipulated by varying the number of shared features between a first cluster (alternatives 1–15) and a second cluster (alternatives 16–30). In the low overlap condition, each member of the first cluster of alternatives was described as having a different set of four of the F1–F6 features but none of the F7–F12 features. Each member of the second cluster of alternatives was described as having a different set of four of the F7–F12 features but none of the F1–F6 features. Thus, the first cluster and second cluster of alternatives shared no features. Stimuli in the moderate feature overlap condition had the same F1–F5 and F8–F12 values as the low feature overlap condition. Moderate feature overlap was created by having F6 and F7 present for six alternatives from the first cluster and six alternatives from the second cluster. Stimuli in the high feature overlap condition had the same F1–F3 and F10–F12 values as the low feature overlap condition. High overlap was created by having a different set of the four F4–F9 features present for each of the 30 alternatives. Thus, as feature overlap went from low to moderate to high, the chance that alternatives from cluster 1 and cluster 2 were described as having a common pair of attributes increased, while the feature overlap of alternatives within a cluster remained constant (i.e., across-cluster overlap varied, within-cluster overlap was invariant).

The experimental procedure was similar to study 1. The primary difference was that people were asked to rank features in the order of their importance at the beginning of the experimental session. The feature rank information was retained and used to determine the feature labels used in the stimulus matrix. Specifically, the features were assigned to the stimulus matrix in order to accomplish two things. First, the features were distributed across the matrix so that the two clusters of alternatives were roughly equally preferred (e.g., in the low feature overlap condition, features ranked first, third, fifth, seventh, ninth, and eleventh were assigned to the first cluster, while the features ranked second, fourth, sixth, eighth, tenth, and twelfth were assigned to the second cluster). Second, the overlapping features were the least important features so that the feature overlap manipulation was insensitive to the use of a reverse-EBA strategy. For example, in the moderate feature overlap condition, the two overlapping features (F6–F7) were the two least important features and, in the high feature overlap condition, the six overlapping features (F4–F9) were the six least important features. If subjects followed a reverse-EBA strategy and retained alternatives with the most important features, this strategy would not be sensitive to a manipulation that relied on feature overlap among the least important features. The study ended with a manipulation check for each of the four replicates. Subjects were shown the original stimuli matrix and asked how comparable they found alternatives 1–15 to alternatives 16–30.

The only other difference from study 1 was that people were asked to create consideration sets of size eight, a reasonable size given the 30 alternatives. Like study 1, the dependent variable was the number of cross-cluster pairs included in the consideration set. For example, a consideration set consisting solely of alternatives from one cluster would have no cross-cluster pairs (i.e., the minimum number of cross-cluster pairs), whereas a consideration set consisting of four alternatives from each cluster would have 16 cross-cluster pairs (i.e., the maximum number of cross-cluster pairs).

Results

Sixty-seven subjects from an introductory marketing course subject pool were given extra credit to participate in the study. The manipulation check verified that alternatives in different clusters became more comparable as the amount of feature overlap increased ($F(1, 64) = 14.42, p < .01$). Alternatives from different clusters were easier to compare in the moderate feature overlap condition ($M = 4.54$) than in the low feature overlap condition ($M = 3.42; F(1, 64) = 7.51, p < .05$). Alternatives from different clusters were easier to compare in the high feature overlap condition.
CONSIDERATION SET COMPOSITION

(M = 5.53) than in the moderate feature overlap condition (M = 4.54; F(1, 64) = 6.60, p < .05).

The results were generally consistent with the predictions of hypothesis 2. The feature overlap manipulation significantly influenced the number of cross-cluster pairs included in the consideration set \( F(2, 64) = 32.19, p < .05, \eta^2 = .33 \). The number of cross-category pairs was larger in the moderate feature overlap condition \( M = 6.47 \) than in the low feature overlap condition \( M = 4.51 \), but the difference was only marginally significant \( F(1, 64) = 2.95, \) one-tail, \( p < .05, \eta^2 = .04 \). The number of cross-category pairs was significantly larger in the high feature overlap condition \( M = 13.93 \) than in the moderate feature overlap condition \( M = 6.47; F(1, 64) = 33.64, p < .05, \eta^2 = .34 \).

Discussion

Study 2 provides evidence that people prefer to create consideration sets of comparable alternatives, even when all of the alternatives are described on the same attributes (i.e., attributes were alignable). Consistent with hypothesis 2, as the number of overlapping features between two clusters of alternatives increased, people became more willing to consider alternatives from both clusters. We believe that this desire for feature overlap can be attributed to consumers wanting to avoid the need to create a common currency in order to compare alternatives. When options have very few overlapping features, the consumer has to translate features into benefits in order to compare alternatives. This is an effortful process. It is much easier to compare a set of alternatives that have a fair amount of feature overlap because differences between alternatives are limited to a small number of features.

The data from study 2 cannot be attributed to the consumer’s use of a reverse-EBA selection strategy. Again, since subjects were randomly assigned to an experimental condition, the use of a reverse-EBA strategy should have been equally prevalent in all three types of feature overlap conditions. Therefore, the use of a reverse-EBA screening strategy cannot be used to explain the observed differences in cross-cluster consideration across the conditions. The only systematic difference across the three conditions was the extent of feature overlap and, consequently, the ease of making comparisons. Thus the studies presented so far show that people have a preference for creating consideration sets consisting of alternatives that are easy to compare. We argue that the ease of comparison can occur because there is attribute alignability across alternatives or because alternatives have overlapping features.

One of the major remaining issues is our assumption that consumers have an ease-of-comparison motive. It is difficult to directly test for this motive because we expect it is a default motive. People are unlikely to be able to articulate the relationship between alignability or feature overlap and the ease of comparison. However, we can imagine situations in which the ease of comparison motive is likely to be less active. For example, when consumers are made sensitive to the possibility of committing a type II error (i.e., erroneously excluding a good alternative), they may become less motivated to have easily compared alternatives. To the extent that we can place the motive to avoid a type II error in opposition to the motive to have easily compared alternatives, we should be able to observe the independent influence of the two motives. We demonstrate this in study 3 using instructions to manipulate consumer sensitivity to committing errors of exclusion.

STUDY 3

Although individuals may have a predisposition toward simplifying a choice task, decision makers are also sensitive to committing type II errors. Sensitivity to errors of exclusion has been found to influence decision makers’ categorization width (Pettigrew 1958), trial of new products (Donnelly, Etzel, and Roeth 1973), reaction to change (Bruner and Tajfel 1961), and search for variety of experience (Taylor and Levitt 1967). In the case of consideration set composition, increased concern about committing a type II error should manifest by the retention of a more diverse set of alternatives. Consequently, the risk of committing a type II error is minimized.

Past research has shown that the sensitivity to committing a type II error not only can vary chronically across individuals but also may be induced in momentary situations (Liberman et al. 2001). So, in this study we instructionally manipulated the motivation to simplify and the motivation to avoid type II errors. In the simplifying-motive condition, subjects were told that the primary objective of the short-listing task was to make their final choice task easier. In the optimizing-motive condition, subjects were told that the main aim of the short-listing task was to make sure that they picked the best possible alternative during the final choice stage. In addition, subjects in a control condition were not given any instructions regarding the objective of the short-listing task. It was expected that consideration sets would be significantly more diverse in the optimizing condition than in the other two experimental conditions.

Stimuli and Procedure

Four replicates (video games, cars, restaurants, and vacations) were used to investigate our hypothesis. Each replicate was made up of 16 alternatives, and each alternative was described on six attributes. The nature of the stimuli can be illustrated with the video game replicate (see fig. 3). There were eight sports games and eight combat games. The eight sports games included four baseball games and four football games. The eight combat games included four aerial combat games and four individual combat games. Games of a particular type were highly alignable alternatives because they were described on the same six attributes (e.g., baseball games were described on A1–A6). Games within a subcategory but from two different types were moderately alignable alternatives because they were described on three common attributes and three unique attributes (e.g., baseball games were described on A1–A6 and football games were
Finally, games from two different subcategories were nonalignable alternatives because they were described by two unique sets of six attributes (e.g., baseball games were described on A1–A6, and aerial combat games were described on A10–A15).

The procedure was identical to that of study 1, barring the motivation instructions. In the simplifying-motive condition, subjects were told the following: “The main aim of this shortlisting task is to simplify your decision making during the final choice phase, so that you can pick your final choice easily, without too much effort.” Subjects in the optimizing-motive condition were told the following: “The main aim of this short-listing task is to help you pick the best possible alternative during the final choice phase. Create your short list very carefully, so as to make sure that you do not regret leaving out an alternative that could have otherwise been optimal.” Subjects in the control condition did not receive motivation instructions, and they had the same set of instructions as in study 1.

Dependent Measure

The 16 alternatives in the awareness set varied in their alignability to each other: some were highly alignable, some were moderately alignable, and some were nonalignable. Subjects selected four of these alternatives for further consideration. Thus, we can describe each subject’s consideration set by the number of high, moderate, and low align-
ability pairs in the set. Sensitivity to the ease of choice should be reflected by relatively more high-alignability pairs, whereas sensitivity to committing a type II error should be reflected by relatively more low-alignability pairs.

Results

Seventy-one subjects from an introductory marketing course subject pool participated in the study for extra credit. The data were collapsed across the four replicates. As predicted, there was no significant difference between the simplifying motive condition (\(M_{\text{low}} = 1.3, M_{\text{moderate}} = 1.5, M_{\text{high}} = 3.2\)) and the control condition (\(M_{\text{low}} = 1.4, M_{\text{moderate}} = 1.4, M_{\text{high}} = 3.2\); Mann-Whitney \(U = 240.0, p > .10\)). As predicted, there was a significant difference between the optimizing-motive condition (\(M_{\text{low}} = 3.2, M_{\text{moderate}} = 1.4, M_{\text{high}} = 1.4\)) and the control condition (\(M_{\text{low}} = 1.4, M_{\text{moderate}} = 1.4, M_{\text{high}} = 3.2\); Mann-Whitney \(U = 45.5, p < .01\)). We note that there were more non-alignable pairs in the optimizing condition than in the control condition (\(F(1, 68) = 50.75, p < .01, \eta^2 = .43\)) and that there were fewer high alignability pairs in the optimizing condition than in the control condition (\(F(1, 68) = 52.42, p < .01, \eta^2 = .44\)).

Discussion

The results demonstrate that priming the optimizing motive leads to formation of relatively more diverse consideration sets. Consumers were willing to include more pairs of nonalignable alternatives and fewer pairs of alignable alternatives in their consideration set when the optimization motive was primed. Additionally, the lack of a significant difference between the control and simplifying-motive conditions indicates that the ease-of-comparison motive is perhaps the default motivational perspective. This conforms to the general cognitive miser paradigm observed elsewhere in the literature (e.g., see Bettman 1979).

STUDY 4

In the previous study, we observed that, when left to themselves, subjects preferred to retain the more alignable alternatives in their consideration sets (see control condition of study 3). Although, in theory, consumers can compare two nonalignable items by creating common currency measures (i.e., by translating the nonalignable attributes into their corresponding benefits and then comparing the benefits), it is an effortful and time-consuming process. However, if, by some means, subjects were encouraged to put in the extra effort and time required to create these common currencies, it is quite likely that the consideration set would consist of more heterogeneous alternatives. The main aim of the current study was to explore this possibility.

Although there is overwhelming evidence (see Medin et al. [1995] for details) that decision makers tend to attend to alignable information and neglect nonalignable information, Zhang and Markman (2001) found that, when subjects are more involved, they do attend to nonalignable information. In the context of consideration sets, this would imply that subjects with higher involvement might spend more time and effort in creating their consideration sets and consequently be more likely to create the common currency measures needed to compare dissimilar items. For example, more involved consumers who spend more time and effort creating a consideration set of video games are more likely to come up with a common currency measure like “excitement” in order to compare a baseball video game that has “120 batting stances” to an aerial combat video game that has “98 flying missions.” Thus, subjects who are more involved should create more heterogeneous consideration sets relative to their less involved counterparts. The current study was designed specifically to test this proposition.

Design, Stimuli, and Procedure

Our hypothesis was tested in a 3 (task involvement) × 4 (replicate) mixed-design experiment. Fifty-four subjects were randomly assigned to one of the three conditions: lower involvement, higher involvement, or control. Similar to Zhang and Markman (2001), we used task instructions to manipulate involvement. Prior to forming their short lists, subjects in the lower involvement condition were told the following: “You are among 10,000 respondents participating in this study, which is being conducted in several cities in the United States. Your responses will be combined with those of other people to get a sense of the average consumer.” Subjects in the higher involvement condition were told the following: “You are among a very small and select group chosen to participate in the study, and your responses are very important. A local retailer, who wishes to remain unidentified, will be using your responses to decide on the products to be marketed.” Subjects in the control condition did not receive any task-involvement-related instructions.

The stimuli and dependent measure used were identical to that used in study 3. The only other difference between the procedures was the use of an involvement manipulation check. Adapted from Zhang and Markman (2001), these questions were nine-point scales anchored by (a) were not interested/were very interested, (b) were not absorbed/were very absorbed, (c) skimmed the descriptions quickly/read the descriptions thoroughly, (d) unimportant/important, (e) irrelevant/of concern to you, (f) worthless/valuable, (g) boring/interesting, and (h) uninvolving/involving. While the first three scales were about the subject’s involvement with the short-listing task, the last four scales were about the product descriptions provided in the study.

Results

The seven items used for measuring involvement showed high reliability (Cronbach \(\alpha = .87\)), so the mean of these seven items was used for further analysis. Analysis of these manipulation check measures showed that our instructions were successful in manipulating task involvement (\(F(2, 51) = 4.88, p < .05\)). As intended, subjects in the higher
involvement condition were more highly motivated ($M = 6.48$) than subjects in both the lower involvement condition ($M = 5.53$; $F(1, 51) = 6.96$, $p < .05$) and the control condition ($M = 5.51$; $F(1, 51) = 7.29$, $p < .05$). Subjects in the control and lower involvement conditions did not differ in their self-rated involvement ($F(1, 51) = 0.0$, $p > .10$).

The data were collapsed across the four replicates. As predicted, there was a significant difference between the higher involvement condition ($M_{\text{high}} = 2.6$, $M_{\text{moderate}} = 1.4$, $M_{\text{low}} = 2.0$) and the lower involvement condition ($M_{\text{low}} = 1.9$, $M_{\text{moderate}} = 1.6$, $M_{\text{high}} = 2.5$; Mann-Whitney $U = 86.5$, $p < .05$). There were more nonaligned pairs in the higher involvement condition ($M = 2.6$) than in the lower involvement condition ($M = 1.9$; $F(1, 51) = 7.89$, $p < .01$, $\eta^2 = .13$), and there were marginally fewer high alignability pairs in the higher involvement condition ($M = 2.0$) than in the lower involvement condition ($M = 2.5$; $F(1, 51) = 2.75$, one-tail, $p = .05$, $\eta^2 = .05$).

Discussion

These results demonstrate the moderating role of involvement in consideration set composition. Decision makers who were more involved used more analytic decision strategies and spent more time and effort creating their consideration sets. As a result they were able to compare and include relatively more nonalignable alternatives in their consideration sets. Although both the involvement manipulation (study 4) and the motivational manipulation (study 3) had similar effects on consideration set composition, it is important to point out the differences underlying these effects. Clearly, both higher involvement and a heightened concern about committing type II errors increase consideration set heterogeneity. Yet, the underlying reasons for this increased heterogeneity differ. When sensitized to committing a type II error (study 3), subjects retain items from different subcategories as a convenient hedge against erroneously leaving out an optimal alternative. In contrast, subjects with higher task involvement (study 4) retain items from different subcategories because of their willingness to put in greater effort. This increased effort may lead to more careful decisions or it may encourage the creation of the common currency measures necessary to compare items from different subcategories that are otherwise difficult to compare.

STUDY 5

In study 3, we used an instruction to increase subjects’ concerns about excluding an optimal alternative from the consideration set. The characteristics of the alternatives in the awareness set could also create concern about excluding an optimal alternative even in the absence of explicit instructions. For example, when competing clusters of alternatives have a negatively correlated benefit structure, selecting an item from one or the other cluster increases the probability of a type II error. To illustrate this idea, reconsider the video game stimuli used in study 3. All 16 alternatives had two equally salient underlying benefits (e.g., realism and challenge). This meant that subjects could create consideration sets of easily comparable alternatives without worrying that they might have missed a more challenging or realistic game. Only our prompt to avoid type II errors increased the heterogeneity of the consideration set. In contrast, consider an alternative set where all sports games are high on challenge but low on realism and all combat games are low on challenge but high on realism. Under these conditions, consumers are likely to become sensitive to the fact that relying solely on the simplification motive could result in a consideration set that lacked alternatives having one of these two important benefit dimensions. Thus, consumers become more likely to retain heterogeneous items in their consideration set when subcategories of alternatives have a negatively correlated benefit structure (hypothesis 3a).

Design, Stimuli, and Procedure

A 3 (benefits trade-off) × 2 (replicates) mixed design was used to test the hypothesis. In the no trade-off condition, all alternatives contained features consistent with both benefits. However, in the two benefits trade-off conditions, one subcategory of alternatives was high on one benefit and low on the other, whereas the reverse was true for the other subcategory. For example, in the video game replicate, we varied the attribute values such that the sports video games were rich in terms of challenge but poor in regard to realism, while the combat video games were rich in realism but poor in terms of challenge. The second benefits trade-off condition simply reversed this pattern of benefits in order to counterbalance the stimuli.

The stimuli were identical to those used in study 3 for the video games and vacations replicates, barring the attribute value manipulations in the benefits trade-off conditions. Consider the following example using the video game replicate. For the football video games, an attribute like the number of offensive/defensive formations was related to challenge, while an attribute like stadium graphics was related to realism. For the aerial combat video games, an attribute like the number of flying missions was related to challenge, while an attribute like physical impact (e.g., sound effects of ricocheting bullets and shattering glass) was related to realism. In the condition without any benefits trade-off, both football and aerial combat games had high values on these attributes (e.g., 25 offensive/defensive formations, 12 meticulously recreated NFL stadiums, 30 flying missions, a lot of physical impact special effects). In a condition with a benefits trade-off, the sports games were poor in challenge (e.g., only five offensive/defensive formations) but rich in realism (e.g., 12 meticulously recreated NFL stadiums), and the combat games were rich in challenge (e.g., 30 flying missions) but poor in realism (e.g., no physical impact special effects). Likewise, in the other benefits trade-off condition, the sports games were rich in challenge but poor in realism (e.g., only three NFL stadiums), and the combat games were poor in challenge (e.g., only five flying missions) but rich in realism.

The procedure used was identical to that of study 3, bar-
ing a manipulation check that consisted of a single question asking subjects to indicate the extent to which they were aware of a trade-off between the two benefit dimensions.

Results

Fifty-one subjects from an introductory marketing course subject pool were given extra credit to participate in the study. Across all measures, there was no difference in the two negatively correlated benefits conditions, so data were collapsed from these two conditions. The manipulation check confirmed that subjects perceived a greater trade-off between the two benefit dimensions in the benefits trade-off conditions ($M = 6.23$) than in the no trade-off condition ($M = 5.12; F(1, 49) = 5.97, p < .05$). As predicted, there was a significant difference between the benefit trade-off conditions ($M_{\text{low}} = 3.8, M_{\text{moderate}} = 1.4, M_{\text{high}} = 0.8$) and the no benefit trade-off condition ($M_{\text{low}} = 1.9, M_{\text{moderate}} = 0.8, M_{\text{high}} = 3.3$; Mann-Whitney $U = 26.50, p < .01$). There were more nonaligned pairs in the benefit trade-off condition ($M = 3.8$) than in the no benefit trade-off condition ($M = 1.9; F(1, 48) = 67.44, p < .01, \eta^2 = .58$), and there were fewer high alignability pairs in the benefit trade-off condition ($M = 0.8$) than in the no benefit trade-off condition ($M = 3.3; F(1, 48) = 78.08, p < .01, \eta^2 = .62$).

Discussion

Study 5 provides additional evidence that people create more heterogeneous consideration sets when they become concerned about leaving an optimal alternative out of their consideration set. When benefits are negatively correlated across subcategories of alternatives, people recognize that the optimal alternative could be in either subcategory. If people were willing to make a decision about the relative importance of each benefit, they could create a comparable set of alternatives consisting of alternatives from primarily one category. Yet, decisions about the relative importance of benefits is effortful, so subjects create heterogeneous consideration sets and leave the more effortful task to the final choice stage.

STUDY 6

In study 5 we showed that people create heterogeneous consideration sets when they become concerned about leaving a potentially optimal alternative out of their consideration set. When benefits are negatively correlated across alternatives, people form more heterogeneous consideration sets and delay the decision about the relative importance of each benefit to the final choice stage. If our reasoning is correct, then the relative heterogeneity of the consideration set should be sensitive to the consumer’s anticipation of the time available to form the consideration set and the time available to make the choice. For example, when the consumer anticipates having very little time for consideration but a lot of time to make the final choice, the consumer should delay determining the relative importance of the benefits to the final choice stage, and the consideration set should be more heterogeneous. However, when a consumer anticipates having very little time during the final choice stage but a lot of time for consideration, the consumer should determine the relative importance of the benefits during consideration set formation and be able to create a set of more comparable alternatives. Thus, study 6 was a test of hypothesis 3b.

Design, Stimuli, and Procedure

The hypothesis was tested using a 2 (anticipated time pressure) $\times$ 2 (replicates) mixed design. Anticipated time pressure was manipulated using instructions as well as countdown clocks embedded in the computerized procedure. In one condition, subjects were told that, although they would have very little time (e.g., two minutes) for creating their consideration set, they could have as much time as they wanted to make their final choice. In another condition, subjects were told that, although they could have as much time as they wanted for creating their consideration set, they would have very little time (e.g., two minutes) to make their final choice.

The stimuli were the video game and vacation replicate from study 5. Only the counterbalanced benefit trade-off conditions were used. The procedure was also identical to the experimental procedures used in study 5. Subjects selected four alternatives to include in their consideration set.

Results

Twenty-six subjects from an introductory marketing course subject pool were given extra credit to participate in the study. Across all measures, there was no difference in the two negatively correlated benefits conditions, so data were collapsed from these two conditions. As predicted, there was a significant difference between the low time pressure at choice ($M_{\text{low}} = 3.7, M_{\text{moderate}} = 0.9, M_{\text{high}} = 1.4$) and low time pressure at consideration ($M_{\text{low}} = 1.2, M_{\text{moderate}} = 2.6, M_{\text{high}} = 2.2$; Mann-Whitney $U = 5.0, p < .01$). There were more nonaligned pairs when there was low time pressure at choice ($M = 3.7$) than when there was low time pressure at consideration ($M = 1.2; F(1, 23) = 55.66, p < .01, \eta^2 = .71$), and there were fewer high alignability pairs when there was low time pressure at choice ($M = 1.4$) than when there was low time pressure at consideration ($M = 2.2; F(1, 23) = 8.15, p < .01, \eta^2 = .26$).

Discussion

Manipulating whether the consumer experienced time pressure at consideration set formation or anticipated there would be time pressure at choice had a significant influence on how consumers reacted to a benefit trade-off in the set of alternatives. When consumers anticipated that they would have a lot of time at choice, they created heterogeneous consideration sets and delayed the difficult task of resolving which benefit was more important. When consumers antic-
ipated that they would have little time at choice, they made decisions about which benefit was more significant prior to forming the consideration set and, thus, they created consideration sets of more comparable alternatives.

**GENERAL DISCUSSION**

Relative to most previous research on consideration sets that tended to either focus on factors affecting consideration set size (e.g., Jarvis and Wilcox 1973; Mitra and Lynch 1995) or study consumption goal influences on consideration set formation (e.g., Ratneshwar, Pechmann, and Shocker 1996), we offer a more comprehensive framework for anticipating the composition of consideration sets. We show that priming different macro-level motives predisposes consumers to employ different types of screening strategies. These screening strategies interact with characteristics of the stimuli, consumption goals, and situational variables to determine the content of consideration sets.

Studies 1 and 2 provide strong evidence that consideration set formation not only involves reducing the number of alternatives to a more manageable size but also involves retaining alternatives that are easy to compare during the final choice stage. This is an important conceptual point, one overlooked by previous research where the “simplifying” role of screening is restricted to paring down the number of alternatives to a more manageable set size. Studies 3–6 show that there are also situations that motivate people to create consideration sets of heterogeneous alternatives. First, when people want to avoid type II errors, they often include a more heterogeneous set of alternatives in their consideration sets (study 3). Second, when people are able to create common currencies (e.g., benefits) that allow them to compare heterogeneous alternatives (study 4) or when different clusters of alternatives pose a benefits trade-off (study 5), people become more likely to create heterogeneous consideration sets. Finally, consumers anticipating limited time at the choice stage resolved their conflict about which benefit they preferred and created less heterogeneous consideration sets (study 6).

The role of comparability, as highlighted in these studies, may help resolve some conflicting findings reported in past literature. When modeling consideration sets, Hauser and Wenerfelt (1989) reported evidence supporting heterogeneous consideration sets, while Lattin and Roberts (1992) reported evidence supporting homogeneous consideration sets. It is possible that this conflicting pattern of results may be explained by differences in alignability across the two data sets. Hauser and Wenerfelt (1989) used data from the plastic wraps market, while Lattin and Roberts (1992) used data from the ready-to-eat cereals market. Whereas the plastic wraps market had a relatively simple two-segment structure made up of national and store brands, the ready-to-eat cereal category had four clearly identified clusters (e.g., healthful, artificial, interesting, and nonadult). Thus, it is quite likely that the degree of alignability between alternatives was much higher in the plastic wraps market in comparison to the highly segmented ready-to-eat cereal market. This is consistent with the findings in our studies where consideration set heterogeneity was found to covary with the degree of alignability between alternatives.

Our results also show that the manner in which consumers partition the choice task may influence the consideration set composition. Both in study 5 and in study 6, subjects had two tasks that needed to be accomplished: (a) compare scale values and (b) resolve the benefit weights problem. The former task involved comparing the attribute values of various alternatives and selecting those alternatives that had high attribute values. The latter task involved resolving the trade-off between the two benefit dimensions. The pattern of results observed in both studies seems to arise from the fact that subjects chose to perform the more effortful task (i.e., resolving the benefits trade-off) when the time constraint was the least. Thus the nature of task partitioning may be an important driver of consideration set composition.

Our results underscore the role of alignability in promoting or dampening competition. Manufacturers may actively reduce the alignability between their products and competing alternatives and, at the same time, create alignability between their own branded variants. For example, manufacturers use irrelevant attributes (e.g., “mountain grown” for Folgers Coffee) to reduce alignability and discourage consumers from considering competing alternatives. Manufacturers of appliances, computers, and toys create product lines that encourage comparison of alignable alternatives within the product line instead of between competing lines. This strategy may be particularly effective when novice consumers are gathering product attribute information in order to screen alternatives at the time of purchase.

From the retailer’s point of view, adding alignable attributes to existing product descriptions may encourage heterogeneity in consumer considerations sets and, as a consequence, ensure relatively fast turnovers for different types of inventory and stock. For example, a category such as wines is characterized by nonalignable descriptions. It is becoming more common to find retailers (e.g., the internet-based “wine.com”) and information sources (e.g., Wine Spectator) that provide tasting charts describing even the most dissimilar wines on a common set of attributes (e.g., acidity, tannin). These descriptions increase benefit alignability and have the potential to encourage inclusion of otherwise diverse alternatives in the consumer’s consideration set. Identification of additional acceptable alternatives can promote additional consumption in product categories where variety seeking is common.

**Limitations**

Although this article takes an important first step toward understanding the determinants of consideration set composition, there are several characteristics of the experiments that limit the generalizability of the findings. First, all the studies were restricted to situations in which consideration
CONSIDERATION SET COMPOSITION

set formation was primarily stimulus based. However, in reality consideration sets can be stimulus based or memory based, and consumers may use a combination of these methods to form consideration sets. The second drawback of our studies is that the procedures and accompanying stimuli are more relevant to novel purchase scenarios. This limits the extent to which we can generalize our findings to more routine, repeat purchase situations. A third limiting characteristic of these studies is that information on the alternatives was made available to the subjects simultaneously. In contrast to such a scenario, it is fairly common practice for consumers to examine alternatives in a sequential fashion and to use only a partial amount of all available information in order to create a consideration set. A final limitation is that subjects were asked to form consideration sets of a given size, as opposed to forming consideration sets of variable sizes.

Future Research

Two extensions of this research have the potential to be useful. First, it would be useful to understand the influence of the alignability and feature overlap of the alternative set on the size of a consumer’s consideration set. As alternatives become more alignable or similar, they become easier to compare. As alternatives become easier to compare, consumers should be willing to compare more alternatives. Thus, we might expect that there is more brand-switching behavior, an indicator of larger consideration sets, in highly alignable product categories. We might also find larger product lines, an attempt to block competing products from being considered, in product categories having more alignable product attributes.

Second, it would be useful to understand the potential malleability of consideration sets. Desai and Hoyer (2000) find that consideration sets become more stable as consumers gain more experience (see also Johnson and Lehmann 1997). Our research shows that consideration set composition is a function of the characteristics of the set of alternatives. We have little insight into how individual differences related to product experience will interact with characteristics of the alternative set during consideration set formation. For example, it may be that experienced consumers have well-established preferences that lead to a small, focused consideration set (selecting a movie). It may also be that experienced consumers can easily create a common currency that allows them to compare a heterogeneous set of alternatives (e.g., selecting a dinner entrée, buying wine). The factors that encourage a consumer to rely on preexisting preferences or alternative set characteristics during consideration set formation are worthy of further study.

[David Glen Mick served as editor and Wayne D. Hoyer served as associate editor for this article.]

REFERENCES


Ratneshwar, S., Cornelia Pechmann, and Allan D. Shocker (1996), “Goal-Derived Categories and the Antecedents of Across-Cat-


