Journal of Behavioral Decision Making

WILEY

RESEARCH ARTICLE OPEN ACCESS

The Categorization of Continuous Attributes

Yusu Wang

Department of Marketing, School of Business, The University of Chicago Booth, Chicago, Illinois, USA

Correspondence: Yusu Wang (yusuwang@chicagobooth.edu)

Received: 5 December 2020 | **Revised:** 12 February 2024 | **Accepted:** 12 March 2024

Funding: The authors received no specific funding for this work.

Keywords: categorization | multi-attribute decision making | preference and choice | preference shift

ABSTRACT

A continuous attribute (e.g., calorie count) can be classified into separate categories (e.g., high vs. low), and a similar attribute value can fall into different categories depending on where the category boundaries are drawn. This research explores the effect of categorization on judgments of options (e.g., products and incentive-compatible games) with continuous attributes. I predict and find a systematic preference shift between two options that were presented with different categorization criteria: When two options involve a tradeoff between two continuous attributes, people tend to prefer the option with both attributes classified into the favorable categories given the categorization criteria. I further show that this effect is driven by larger perceived differences between attribute values across category boundaries and is moderated by people's tendency to rely on category information. Overall, this effect holds even when people are highly familiar with the attributes and feel confident to make similarity evaluations, when people are cued that the categories provide little informational value, and when people are incentivized to make deliberate decisions. The findings in this research carry both theoretical and practical implications.

1 | Introduction

Continuous attributes can be classified into discrete categories by imposing artificial boundaries. This practice is commonly seen in everyday life; for instance, a consumer can browse products on Neiman Marcus by price categories (e.g., under \$25 and \$25–\$50) and on Farfetch by sale discount categories (e.g., up to 30% off and 30%–50% off). Do these artificial boundaries influence people's perceptions of options in different categories? This research tries to understand the effect of categorization on people's evaluations of continuous attributes.

Imagine a studio renting scenario: Suppose that a consumer is considering renting a studio, and she wants it to be large and close to the city center. After browsing an online rental website, she identified two promising options—studio A is 500 sq ft and 4 miles from the city center, and studio B is 400 sq ft and 2 miles from the city center.

	Size	Distance from city center
Studio A	500 sq ft	4 miles
Studio B	400 sq ft	2 miles

Based on this information alone, the two studios are similarly attractive to her. Now, imagine that the rental website attempts to aid prospective renters by categorizing the properties. Specifically, the website categorizes a studio as "large" if it is larger than 450 sq ft and "small" if it is smaller than 450 sq ft, and it categorizes a studio as "close" if it is less than 5 miles from the city center and "far" if it is more than 5 miles from the city center. Based on these categorization criteria, the two studios are in the same distance category ("close") but different size categories (studio A is "large" while studio B is "small"). Given the categorization information, will the consumer still evaluate the two studios as similarly attractive? What if the rental website had used different categorization criteria such that the two studios were placed to the same size category but different distance categories, making studio B "close" and studio A "far"?

No matter whether the two studios belong to the same size category (e.g., if the size threshold is 350 sq ft, then both studios are "large") or different size categories (e.g., if the size threshold is 450 sq ft, as in the example above), the actual difference in size is constant (i.e., 100 sq ft), as is their difference in distance from the city center (i.e., 1 mile). In contexts where continuous attributes are categorized, the presence of categories may not provide objectively new information to assist people in evaluating their options, especially when most people are familiar with the attributes (e.g., size and distance) and may judge the value of these attributes without referring to the category information. Thus, it is possible that people will ignore the categorization imposed on continuous attributes and form preferences by comparing the concrete numbers.

Nevertheless, I predict and find a categorization effect, in which people's evaluations of continuous attributes are influenced by categorizations. More specifically, when a choice involves a tradeoff between two continuous attributes, people's preference between the two options varies according to how the options are categorized on those two attributes. I propose that people tend to prefer an option when the categorization criteria are set such that the option's superior attribute (i.e., the attribute on which the option is objectively superior than the other option; e.g., size for studio A) falls into a distinct category ("large") from the other option (studio B is "small"), and its inferior attribute (i.e., the attribute on which the option is objectively inferior than the other option; e.g., distance for studio A) falls into the same category as the other option (studios A and B are both "close"). Importantly, I argue that categorization affects people's preferences, even though the categories are known to carry little informational value.

The current research contributes to the literature in several ways. First, this research extends the categorization literature to a context in which people make tradeoff decisions based on the continuous attributes. Unlike previous research that has analyzed perceptions of continuous attributes across natural categories (Donnelly, Compiani, and Evers 2021; Isaac and Schindler 2014) or in people's self-generated mental categories (Chernev and Gal 2010), I investigate contexts in which the categories are imposed by a third party or generated through a random process. Additionally, different from the past research that has focused on the effect of labels on evaluations of continuous attributes (Aydinoğlu and Krishna 2011; Ellison, Lusk, and Davis 2014), this research highlights the effect of categorization while decreasing the additional inferences that people may be making from labels.

Second, this research expands our understanding of categorization effect to contexts where the categories are known to carry little informational value and are likely to be ignored by individuals. People often fail to disregard information that is irrelevant to the focal task (Camerer, Loewenstein, and Weber 1989; Fischhoff 1975; Hell et al. 1988; Jacowitz and Kahneman 1995; Tversky and Kahneman 1974). While past research has explored how categories that carry, or at least assumed to carry, meaningful information assist people's judgments and decisions (Kim and Yoon 2016; Mogilner, Rudnick, and Iyengar 2008), this article found that the categorization effect persists even when people are highly familiar with the attributes and feel confident to make similarity evaluations, and when they receive cues that the categories have little informational value.

Finally, this research makes theoretical contributions to the multi-attribute behavioral research by uncovering preference shifts that are driven by a novel mechanism. The classic preference reversals in the past literature usually are elicited by manipulating the preference scale or solicitation method (Hsee 1996; Hsee and Leclerc 1998; Slovic, Griffin, and Tversky 1990; Tversky, Sattath, and Slovic 1988; Tversky, Slovic, and Kahneman 1990). By contrast, the preference shifts found in this research result from the discontinuity in people's evaluations of continuous attributes. This mechanism is discussed in detail later in this article.

1.1 | Categorization Effect on the Perceived Diversity of an Assortment

Options are often grouped into categories to help people cope with overwhelming amounts and varieties of information (Cohen and Basu 1987; Nosofsky 1986; Rosch 1999). Once categories are established, people tend to incorporate the category information into their attitude-formation or decision-making process (Isaac and Schindler 2014; Sharif and Woolley 2020; Tu and Soman 2014). People perceive a greater similarity between items when they are placed in the same category than when they are placed in different categories (Maki 1982), and they tend to focus on the qualitative aspects of the categories rather than on the quantitative aspects of the categorized items (Chernev and Gal 2010).

Ample research has demonstrated the effect of categorization on people's global impression of an assortment. Mogilner, Rudnick, and Iyengar (2008) showed a mere categorization effect whereby consumers who are unfamiliar with the choice domain perceive that an assortment is more diverse if the items in it are grouped into categories, even though the category labels are uninformative. Redden (2008) found that products or consumption scenarios placed into specific categories (vs. one broad category) are perceived to be more different from each other, resulting in a lower perceived repetitiveness of a consumption experience as a whole. In the same vein, Kim and Yoon (2016) found that consumers, especially those who are not familiar with the categories, perceive a greater variety among the available options when the category labels are abstract (vs. specific or absent).

All the evidence points to the fact that people tend to rely on available category information to assist their similarity judgments of an assortment of objects, especially when they lack adequate knowledge or cognitive resources to evaluate the objects.

1.2 | Categories Imposed on Continuous Attributes

Options in the same category are usually more similar or comparable to each other than are options from different categories

(Goldstone 1994). This generalization is not always accurate, however—categorizations can be imposed on continua such that two objects near the boundaries of two adjacent categories might be more similar to each other (i.e., closer on the continuum) than two objects within the same category.

People tend to see continuous data in a categorical fashion (Fisher and Keil 2018). One consequence of this tendency is a perception of elevated differences between objects in different categories. Past research has analyzed the categorization of continua when the category borderlines naturally exist. For example, Isaac and Schindler (2014) demonstrated a "top-ten" effect in which a ranked list is subjectively divided into smaller subcategories by natural mental boundaries (often using round numbers that end in zero or five; e.g., "0–10" and "11–20"). As a result, a difference in rank is perceived as more pronounced if it falls across the natural mental boundaries (e.g., 10 vs. 11) than if it falls within the same mental boundaries (e.g., 9 vs. 10). Another example of how natural category boundaries interact with judgments is Donnelly, Compiani, and Evers (2021)'s work documenting the impact of natural temporal landmarks (e.g., hour, month, and year) on duration perceptions. Specifically, a period that spans multiple temporal boundaries is estimated to be longer in duration than a period of equivalent length that spans fewer temporal boundaries. In sum, natural boundaries tend to be automatically imposed when people evaluate certain continua such as ranked lists and time periods.

Past research examining the effect of labels has shed light on the effect of third-party imposed categories. Aydinoğlu and Krishna (2011) found that size labels (e.g., small, medium, and large) can influence consumers' perceptions of size and their consumption behaviors, such that a smaller item labeled as "large" seems bigger than a larger item labeled as "small." However, without concrete numeric size information (e.g., in lbs.), participants in Aydinoğlu and Krishna (2011)'s experiments could only rely on the sensory input (visual and satiation cues) to form size judgments, which might be easily distorted by contextual cues (Coren and Girgus 2020).

As another empirical example in the food domain, Ellison, Lusk, and Davis (2014) ran a field experiment in which they used traffic light labels to categorize food calories: a red label for food items with >800 calories, yellow for 401–800 calories, and green for ≤400 calories. The addition of traffic light labels to numeric calorie labels (vs. numeric labels alone) significantly reduced people's overall calorie intake. Nevertheless, additional factors beyond the mere effect of categorization may influence the observed outcomes. As participants were not informed about the criteria for the food categorization, they may have inferred other attributes, such as the healthiness of the food, from the color labels. This inference is plausible given that traffic light labels have historically been used to indicate the content of health-related nutrients like saturated fat and salt, aiding in the identification of healthier choices (Food Standards Agency 2007). Therefore, it is difficult to attribute the change in behavior solely to the effect of categorization based on calorie content.

1.3 | The Current Research

This research examines how the categorizations of continuous attributes will impact people's evaluations of individual options in a collection and extended the analyses beyond the food consumption context. Based on the prior findings, I propose a categorization effect in which people's preferences between two options can be altered by the categorization of those options' continuous attributes. Suppose that options A and B involve a tradeoff between attributes X and Y such that A is superior on X and B is superior on Y. I propose that people are more likely to prefer A over B in an X-separate condition (in which A and B fall into separate categories on X and the same category on Y) than in a Y-separate condition (in which A and B fall into separate categories on Y and the same category on X). I suggest that this pattern will arise despite individuals' awareness that the categorization merely reflects whether an attribute's value is higher or lower than a specific threshold. Being explicit about the categorization criteria helps diminish the likelihood of people drawing additional inferences beyond assessing the magnitude of the attribute values. Moreover, I propose that this effect will hold even when the attributes are familiar and evaluable to most people, when the categories are known to carry little informational value, and when people are incentivized to spare cognitive resources.

Consistent with past research documenting that individuals often infer similarity from category membership (Irmak, Naylor, and Bearden 2011), I further suggest that the categorization effect occurs because the presence of categories introduces discontinuity in people's evaluations of continuous attribute values across category boundaries, as shown in Figure 1. Specifically, an imposed category boundary diminishes the perceived

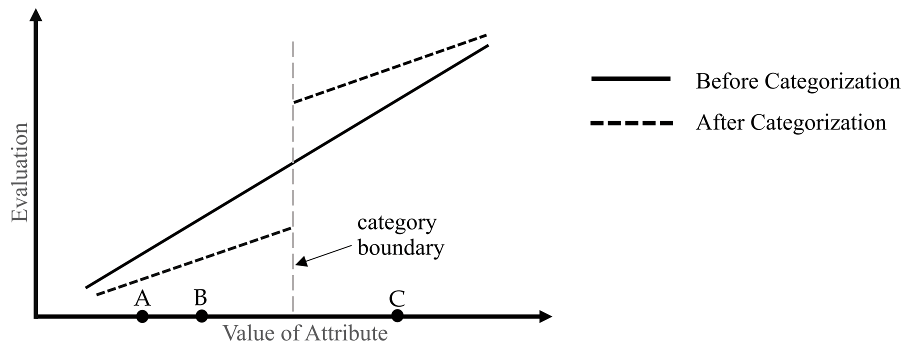

FIGURE 1 | Categorization effect across a category boundary. *Note:* An imposed category boundary diminishes the perceived difference between items in the same category (A and B) and widens the perceived difference between items in different categories (A and C; B and C).

difference between attribute values in the same category and widens the perceived difference between attribute values in different categories, and this discontinuity occurs only near the category boundary. As a result, the superiority of one option on its superior attribute over the other option will be perceived as greater if the two options fall into separate categories on this attribute, and vice versa. When making multi-attribute tradeoff decisions between options, people's preferences are affected by the ways in which the category boundaries make different attributes values seem more or less dissimilar and, consequently, make one option more or less superior to the other.

People draw on category information to judge the distance between two attribute values, especially when they lack adequate knowledge or cognitive resources to make similarity evaluations (Kim and Yoon 2016; Mogilner, Rudnick, and Iyengar 2008). Even though people are generally familiar with and capable of evaluating continuous attributes such as size and distance in the studio-renting scenario, they may rely on the categories as a mental shortcut to save cognitive energy without assessing the utility of categories. I therefore theorize that reducing people's tendency to draw on category information by emphasizing that categories provide no new or useful information (e.g., highlighting that the category boundaries are arbitrarily generated) can attenuate the categorization effect. However, the effect may not be wiped out by the arbitrariness cues. Similar to many to-be-ignored information (Dietvorst and Simonsohn 2019), the category information may continue influencing people's choices and evaluations even when people are cued to disregard the categories.

I tested my propositions across seven experiments (see Table 1). (Preregistrations, materials, and data can be found on OSF at: https://osf.io/yfwgh/?view_only=53db23ac5a1a4353a5be150d93f2755c.) After establishing the basic paradigm in experiment 1, I tested whether the effect persists with noninformative, symbolic category labels in experiment 2. Experiment 3 replicated the categorization effect with an incentive-compatible design, in which participants chose to enter one of two lottery games with a tradeoff between the prize amount and winning probability.

Next, to probe the underlying mechanism, I demonstrated in experiment 4 that the change in perceived differences of attribute values mediated the effect of categorization on choice. Experiment 5 tested whether the categorization effect would be attenuated when the categorization criteria were randomly generated. Experiment 6 extended the effect to a single-attribute evaluation context to further probe the discontinuity mechanism. Finally, in experiment 7, I solicited participants' preferences through willingness to pay, offering them the chance to obtain the item they valued most based on their highest bid.

2 | Experiment 1: Studio Choice

In experiment 1, participants chose between two studios that involved a tradeoff between size and distance. A post-test survey ($N = 57$) showed that participants reported high levels of familiarity with these two attributes, and they felt highly confident in their ability to judge the values of these two attributes. All post-test stimuli and results were reported in Appendix S1. Despite the familiarity and evaluability of the attributes, I predicted that participants' studio choice would still be influenced by the presence of categories, such that a studio with its superior attribute categorized more favorably would be more likely to be preferred by participants.

2.1 | Method

A total of 193 participants recruited through Amazon's Mechanical Turk (MTurk) completed the experiment. Participants were randomly assigned to one of two conditions (size-separate vs. distance-separate). In both conditions, participants imagined that they wanted to rent a studio from an online rental website, and they needed a studio that was as large and as close to the city center as possible. The website provided two available options—studio A (larger) and studio B (closer to the city center)—such that participants faced a tradeoff between the two attributes of interest.

Participants in the [size-separate condition/distance-separate condition] received the following instructions:

> The rental website categorizes a studio as “large” if it is larger than [450 / 350] sq ft, or “small” if it is smaller than [450 / 350] sq ft.

> The rental website categorizes a studio as “close” if it is less than [5 / 3] miles from the city center, or “far” if it is more than [5 / 3] miles from the city center.

> You find two viable options. They are similar in all aspects (including rent) except for size and distance to the city center.

	Size	**Distance from city center**
Studio A	Large (500 sq ft)	[Close/far] (4 miles)
Studio B	[Small/large] (400 sq ft)	Close (2 miles)

Then, all participants indicated which studio they would prefer to rent. Finally, participants answered demographic questions and were debriefed.

2.2 | Results and Discussion

The data were consistent with my prediction. Overall, there was a choice reversal between the two categorization conditions, with the choice share of studio A (i.e., the larger studio) higher in the size-separate condition than in the distance-separate condition (61.0% vs. 29.0%; *Pearson* $\chi^2(1) = 19.85$, $p < 0.001$).

The results of experiment 1 provided initial evidence for the proposed categorization effect. Specifically, participants tended to prefer an option when both of its attributes (vs. only one of its attributes) were categorized favorably relative to the other option, even though the attributes were quantitatively constant across conditions.

TABLE 1 | Overview and main results of all studies.

Study	Stimulus	Attributes	Objectives	Results			
1	Studios	Size vs. distance	Tests the proposed effect	% choosing the larger studio	Size-separate	61.0%	Pearson $\chi^2(1) = 19.85, p < 0.001$
					Distance-separate	29.0%	
2	Yogurts	Sugar vs. fat	Replicates the effect using symbolic category designations	% choosing the lower-fat yogurt	Control	35.8%	Pearson $\chi^2(1) = 28.23, p < 0.001$
					Fat-separate	64.2%	
					Sugar-separate	19.1%	
3 (incentive-compatible)	Lottery games	Prize amount vs. winning probability	Replicates the effect in an incentive-compatible decision	% choosing the larger-prize game	Prize-separate	71.7%	Pearson $\chi^2(1) = 19.73, p < 0.001$
					Probability-separate	41.2%	
4	Smartphones	Battery life vs. storage capacity	Tests the perceived differences in attribute values as a mediator	% choosing the option with a longer battery life	Noise-separate	71.1%	Pearson $\chi^2(1) = 25.46, p < 0.001$
					Energy-separate	35.6%	
5	Window ACs	Noise vs. energy consumption	Tests the tendency to rely on category information as a moderator	% choosing the more energy-efficient AC model	(Given criteria) Energy-separate	65.2%	Interaction effect: $\beta = 2.06$, Wald $\chi^2(1) = 6.17, p = 0.013$
					(Given criteria) Noise-separate	2.2%	
					(Random criteria) Energy-separate	59.1%	
					(Random criteria) Noise-separate	11.8%	
6	Cupcakes	Calories	Extends the context to a single-attribute evaluation task	Purchase likelihood of the focal cupcake (7-point scale)	High-cutoff	4.95	$t(298) = 3.73, p < 0.001, d = 0.43$
					Low-cutoff	4.24	
7 (incentive-compatible)	Flashlights	Battery life vs. weight	Solicits preferences using willingness to pay in an incentive-compatible design	The amount that one is willing to pay for each flashlight	Battery-separate (option A)	\$21.98	Interaction effect: $F(1, 199) = 22.03, p < 0.001, \eta^2 = 0.100$
					Battery-separate (option B)	\$17.01	
					Weight-separate (option A)	\$20.27	
					Weight-separate (option B)	\$21.82	

Nevertheless, participants likely experienced a greater decision fluency when reading semantic labels that were congruent with the stated goals of the hypothetical scenario (i.e., to find an apartment that was "large" and "close" to the city center). This may have increased participants' liking toward the option for which both category labels matched the stated goals (Labroo, Dhar, and Schwarz 2008). Moreover, as there was no control condition and options were categorized in both conditions, it is unclear whether the categorizations influenced the evaluations of one of the attributes or both. I addressed these possibilities in the following experiments.

3 | Experiment 2: Color Labels Yogurt Choice

Experiment 1 found that people's evaluations of familiar attributes were influenced by categorization information, but the effect may have been caused by the semantic nature of the labels rather than the mere presence of categories. The objective of experiment 2 was to test whether the effect could replicate when categories were designated with nonsemantic, neutral symbols—specifically, different colors. Participants chose between two yogurts that involved a tradeoff between sugar content and fat content. In this case, a yogurt is "superior" if it contains *less* of the nutrient, as sugar and fat are often considered undesirable in excess. A posttest survey ($N=57$) showed that participants reported high levels of familiarity with these two attributes, and they felt highly confident in their ability to judge the values of these two attributes (see details in Appendix S1). I avoided colors that usually signal behavioral norms (e.g., green = "go!"/"good," red = "stop!"/"bad") or healthiness (e.g., green = "healthy"/"organic," red = "unhealthy"/"fast food"); instead, I used dark blue and light blue as neutral symbols.

Besides, to test whether categorization affects people's evaluations of both attributes, I added a control group in which no category information was given, and I compared the choice shares between the two categorization conditions and the control condition. I expected to observe differences in choice across all three conditions.

3.1 | Method

A total of 202 participants recruited through MTurk completed the experiment. Participants were randomly assigned to one of three conditions (control vs. fat-separate vs. sugar-separate).

All participants imagined that they were purchasing yogurt from an online store, and they wanted a yogurt that contained as little sugar and as little fat as possible. A post-test survey found that this instruction was consistent with people's general preference (see details in Appendix S1). They learned that there were only two options available: yogurt A (lower in fat) and yogurt B (lower in sugar).

In the control condition, participants read:

	Sugar per 100 g	**Fat per 100 g**
Yogurt A	10 g	3 g
Yogurt B	8 g	5 g

In the [fat-separate condition/sugar-separate condition], participants saw the following instructions:

> The online store uses **DARK BLUE** to highlight a sugar level of over [11 g / 9 g] per 100 g; otherwise, it uses LIGHT BLUE to highlight the sugar level.

> Likewise, the online store uses **DARK BLUE** to highlight a fat level of over [4 g / 6 g] per 100 g; otherwise, it uses LIGHT BLUE to highlight the fat level.

	Sugar per 100 g	**Fat per 100 g**
Yogurt A	[10 g/**10 g**]	3 g
Yogurt B	8 g	[**5 g**/5 g]

In the original stimuli, words that are in bold were highlighted with a dark blue background, while those that are underlined were marked with a light blue background. Then, all participants indicated which yogurt they would buy. Finally, participants answered demographic questions and were debriefed.

3.2 | Results and Discussion

The data provided support for my proposition. The choice share of yogurt A (i.e., the lower-fat yogurt) was higher in the fat-separate condition than in the control condition (64.2% vs. 35.8%; Pearson $\chi^2(1)=10.78$, $p=0.001$), while the choice share of yogurt B (i.e., the lower-sugar yogurt) was higher in the sugar-separate condition than in the control condition (80.9% vs. 64.2%; Pearson $\chi^2(1)=4.73$, $p=0.030$). Overall, there was a significant choice reversal between the fat-separate condition and the sugar-separate condition (Pearson $\chi^2(1)=28.33$, $p<0.001$; see Figure 2).

In experiment 2, I adjusted the basic paradigm from experiment 1 by adding a control group and using symbolic (rather than semantic), neutral category designations. Replicating the findings in experiment 1, I found a choice reversal between the two

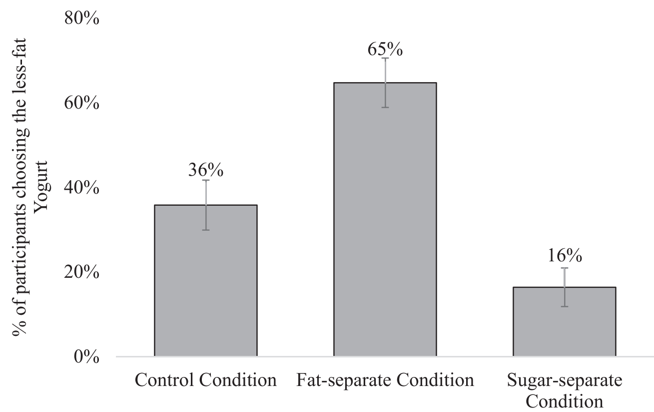

FIGURE 2 | Results of experiment 2. *Note:* The choice share of the lower-fat yogurt (i.e., Yogurt A) was highest in the fat-separate condition, followed by the control condition and then the sugar-separate condition.

categorization conditions, and I demonstrated preference shifts between the control condition and either of the categorization conditions. Besides, the results showed that categorization has a major impact on people's choice even when the category designations are symbolic and neutral. Thus, this study provided further evidence for the categorization effect.

4 | Experiment 3: Incentive-Compatible Lottery Game

Experiment 3 adopted an incentive-compatible setting and examined the categorization effect in a risk–return tradeoff decision. Participants were instructed to choose between two lottery games that involved a tradeoff between two ubiquitous attributes: stakes and winning probability. Participants' payoffs were directly related to their choices: They were entered into the lottery game of their choice and were paid if they won. The incentive-compatible design ensured that participants were fully engaged and were motivated to make decisions that reflected their true preferences. Importantly, participants had frequent interactions with both attributes, especially with monetary values, in their daily lives. A post-test survey ($N = 60$) found that participants reported high levels of familiarity with the two attributes. Besides, they revealed high levels of confidence in their ability to judge monetary values and probabilities (see details in Appendix S1). Therefore, participants should be capable of evaluating the attribute values based on their past experiences and personal preferences. However, I predicted that they would still make choices relying on categorization information.

4.1 | Method

A total of 208 participants recruited through MTurk completed this experiment. Participants were randomly assigned to one of two conditions (prize-separate vs. probability-separate). They were informed that they could choose to play one of two available lottery games—game A (the larger-prize game) and game B (the higher-probability game)—and they would actually get the prize if they won.

All participants received the following instructions:

> In this study, you have the opportunity to play a lottery game, which gives a certain chance to win a certain prize. We have many such games in our repertoire. These games entail different winning chances and different prizes.

Participants in the [prize-separate condition/probability-separate condition] then read:

> We categorize the winning chance of a game as "high" if it is over [30% / 20%]; otherwise, we categorize it as "low." We categorize the prize of a game as "high" if it is over [40 cents / 60 cents]; otherwise, we categorize it as "low."
>
> Each worker is allowed to play only one game, but they are given two games to choose from. The following are our options:

	Your chance to win	**The prize you will get if you win**
Game A	Low (16%)	[High/low] (50 cents)
Game B	[Low/high] (26%)	Low (25 cents)

Then, all participants made their choices and were entered into the corresponding lottery game. Winners of the game were paid in the form of a bonus 3 days after they completed the study.

4.2 | Results and Discussion

Overall, there was a choice reversal between the two conditions, with the choice share of game A (i.e., the larger-prize game) higher in the prize-separate condition than in the probability-separate condition (71.7% vs. 41.2%; Pearson $\chi^2(1) = 19.73$, $p < 0.001$).

Experiment 3 replicated the findings of previous studies in a design that held real consequences for participants and used two attributes (i.e., monetary value and probability) that were highly ubiquitous. My manipulation of the categorization criteria once again yielded a robust categorization effect: Participants' choices were skewed toward the option that was favored by the given categorization criteria, even though the attributes were familiar to the participants and participants were incentivized to maximize their payoff.

5 | Experiment 4: Categorization Influences Choices by Changing Perceived Difference Between Attribute Values

Thus far, I have found that the categorization effect persists when category designations are symbolic and neutral (experiment 2) and when people are incentivized to maximize the payoff of their choices (experiment 3). Experiment 4 tested the underlying mechanism that the effect occurs because categorizations change the perceived similarity between continuous attribute values. In this experiment, participants made a choice between two smartphones that involved a tradeoff between battery life and storage capacity. Results from a post-test survey ($N = 61$) showed that participants were familiar with these two attributed and reported a high level of confidence in their ability to evaluate the values of these two attributes (see details in Appendix S1).

To measure the perceived difference between attribute values, I measured the perceived superiority of each option on their superior attribute. The larger the perceived superiority of one option relative to the other option on one attribute, the greater the perceived difference between the two options along that attribute. I tested whether the perceived differences in attribute values mediated the effect of categorization on choice. Furthermore, I tested the generalizability of the categorization effect using category labels that had a smaller qualitative contrast: The less favorable category label was neutral in valance (i.e., average).

5.1 | Method

A total of 201 participants recruited through MTurk completed this experiment (preregistered at AsPredicted https://aspredicted.org/75W_Q8K). Participants were randomly assigned to one of two conditions (separated attribute: battery-separate vs. storage-separate). They were instructed to imagine choosing between two smartphone options: option A (the smartphone with a longer battery life) and option B (the smartphone with a larger storage capacity).

All participants read the following instructions:

> "When buying a smartphone, battery life and storage are two of the most important factors to consider. Battery life affects how long a phone can last throughout the day without needing a recharge. Storage affects the number of applications that can be installed on the phone as well as the performance of the operating system."

Participants in the [battery-separate condition/storage-separate condition] then read:

> A retailer categorizes a smartphone as LONG in battery life if its battery life is longer than [14 / 6] hours of use, or else categorizes it as AVERAGE in battery life.

> The retailer categorizes a smartphone as LARGE in storage if it has more than [32 / 96] GB of storage, or else categorizes it as AVERAGE in storage.

Next, participants answered two comprehension questions by filling the values for categorization in the text boxes. For example, they answered: "A smartphone is categorized as LONG in battery life by this retailer if its battery life is longer than __ hours of use?" Participants were allowed to proceed to the next page only after entering a correct value for each question.

Then, participants in the [battery-separate condition/storage-separate condition] saw the following two options.

Option A	LONG battery life (18 h of use)	[LARGE/AVERAGE] storage (64 GB)
Option B	[AVERAGE/LONG] battery life (10 h of use)	LARGE storage (128 GB)

Participants then indicated which smartphone they would prefer to choose (option A or option B). On the next page, they indicated their perceived superiority of option A's battery life relative to option B's battery life ("How long do you think option A's battery life is relative to option B's battery life?" 1 = *slightly longer*, 4 = *moderately longer*, 7 = *significantly longer*) and their perceived superiority of option B's storage capacity relative to option A's storage capacity ("How large do you think option B's storage is relative to option A's storage?" 1 = *slightly larger*, 4 = *moderately larger*, 7 = *significantly larger*). Finally, participants answered demographic questions and were debriefed.

5.2 | Results and Discussion

Consistent with my prediction, there was a choice reversal between conditions. A larger percentage of participants in the battery-separate (vs. storage-separate) condition chose to buy option A, which was the smartphone with a longer battery life (71.1% vs. 35.6%; Pearson $\chi^2(1) = 25.46$, $p < 0.001$).

I then ran a repeated measures ANOVA with the perceptions of the relative battery life and storage capacity levels between the two smartphone options as two within-subjects measures and the separation condition as the between-subjects factor. As predicted, there was a significant interaction between the relative

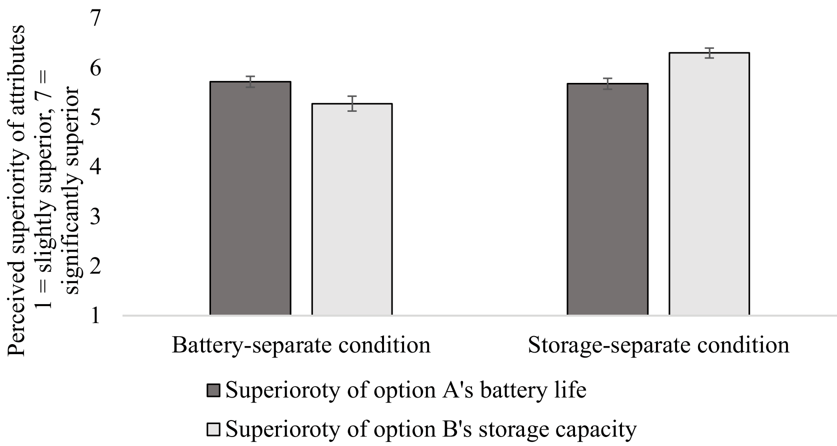

FIGURE 3 | Results of experiment 4. *Note:* The superiority of option A's battery life (relative to option B's battery life) was perceived as larger than the superiority of option B's storage capacity (relative to option A's storage capacity) in the battery-separate condition. The superiority of option B's storage capacity was perceived as lager than the superiority of option A's battery life in the storage-separate condition. Overall, there was a significant interaction between the relative perceptions of attribute values and separation conditions.

perceptions of attribute values and separation conditions (F(1, 199) = 33.69, $p < 0.001$, $\eta^2 = 0.145$; see Figure 3). Specifically, within the battery-separate condition, the superiority of option A's battery life relative to option B's battery life (M = 5.72, SD = 1.06) was perceived as larger than the superiority of option B's storage capacity relative to option A' storage capacity (M = 5.28, SD = 1.33; F(1, 96) = 10.90, $p = 0.001$, $\eta^2 = 0.102$). Conversely, within the storage-separate condition, the superiority of option B's storage capacity (M = 6.31, SD = 0.97) was perceived as lager than the superiority of option A's battery life (M = 5.68, SD = 1.12; F(1, 103) = 24.56, $p < 0.001$, $\eta^2 = 0.193$). These pattern suggested that categorization changed the perceived similarity between attribute values by lowering the perceived difference between two attribute values in the same category and enlarging the perceived difference between two attribute values in separate categories, therefore influencing consumers choices that involve a tradeoff between two attributes.

I then took the difference between the perceived superiority of option A relative to option B along battery life and the perceived superiority of option B relative to option A along storage capacity to form a single measure of the relative attribute value perception between the two options, with a higher number of the relative attribute value perception reflecting a stronger perception of option A's superiority on battery life than the perception of option B's superiority on storage capacity. I ran a mediation analysis using Hayes's (2017) PROCESS macro (Model 4) with the categorization condition as the independent variable (coded as −1 for storage-separation and 1 for battery-separation), the relative attribute value perception as a mediator (mean-centered), and choice as the dependent variable (coded as 1 for option A and 0 for option B). The results revealed that the relative attribute value perception mediated the impact of categorization on choice ($b = 0.39$, $SE = 0.12$, CI: [0.20, 0.67]). Specifically, separating the battery life into long and average contributed to a stronger perception of option A's superiority on battery life ($b = 0.53$, $SE = 0.09$, $t(199) = 5.80$, $p < 0.001$), with was further associated with a greater likelihood of choosing option A ($b = 0.72$, $SE = 0.17$, $Z = 4.22$, $p < 0.001$).

Experiment 4 provided support for the proposed underlying mechanism that categories exert their effect by changing the perceived difference between attribute values either across or within category boundaries. I measured the perceived superiority of each option on their superior attribute as a proxy for the perceived difference between the options along each attribute. The results showed that the imposition of category boundaries influences the perceived differences between attribute values, leading to a shift in choice. Importantly, the categorization effect holds even when people are familiar with the attributes and are confident in their ability to compare and evaluate the attribute levels.

6 | Experiment 5: Attenuating the Categorization Effect by Decreasing the Tendency to Rely on Category Information

The objective of experiment 5 was two-fold. First, it tested whether changing people's tendency to rely on category information will moderate the effect. People tend to incorporate category information into their judgment and decisions to simplify their decision making and conserve cognitive effort. Although people may fail to evaluate the informational value of categories and overweight category information in their decisions, they can be cued to notice that categories may not always provide useful insights and therefore rely less on the category information. Thus, I predicted that decreasing participants' tendency to rely on category information—by highlighting the limited utility of the categories—would attenuate the categorization effect.

Second, experiment 5 investigated whether the categorization effect would persist when participants were prompted to recognize that categories lacked meaningful insight. Specifically, participants randomly chose two letters from two sets of meaningless letters to define the category boundaries for categorizing product attributes. This procedure highlighted that the category borderlines were arbitrary and lacked informative value. However, the categorization effect might remain because individuals might fail to evaluate the informational value of categories and continue to rely on available category information at the moment of choice. To this point, this design provided a conservative test of the robustness of categorization effect.

In this experiment, participants made a choice between two models of window AC units that involved a tradeoff between noise and electricity consumption. Again, the less favorable category label was neutral in valance (i.e., average). A post-test survey ($N = 60$) found that participants were generally familiar with these two attributes and reported feeling confident in their ability to compare and judge these attribute values (see details in Appendix S1).

6.1 | Method

A total of 402 participants, recruited through Prolific, completed this experiment (preregistered at AsPredicted https://aspredicted.org/NTY_HR3). Participants were randomly assigned to one of four conditions in a 2 (separated attribute: energy-separate vs. noise-separate) by 2 (categorization criteria: given vs. random) between-subjects design.

All participants were instructed to imagine shopping for a window AC that made as little noise and consumes as little electricity as possible. The post-test survey found that this instruction was consistent with people's general preference (see Appendix S1). Participants who were assigned to the random criteria condition then read that there were many ways to categorize the noise level and the electricity consumption of a window AC such that a window AC could be categorized as LOW or AVERAGE in noise and LOW or AVERAGE in electricity consumption. They were told that each of the letters—"E," "G," "N," "P," and "Z"—represented a unique number for the categorization of noise level and randomly picked one of the letters. Similarly, they were told that each of the following letters—"H," "L," "X," "V," "Q"—represented a unique number for the categorization of electricity consumption and randomly picked one of the letters.

On the next page, participants in the random criteria condition read the following ([energy-separate condition/noise-separate condition]):

> You picked the letter {choice of the participant} for the categorization of noise level, and it represents the number [50 / 35]. Thus, a window AC is categorized as LOW in noise if its noise level is lower than [50 /35] dB, or else it is categorized as AVERAGE in noise.

> You picked the letter {choice of the participant} for the categorization of electricity consumption, and it represents the number [650 / 700]. Thus, a window AC is categorized as LOW in electricity consumption if it uses fewer than [650 / 700] watts of electricity, or else it is categorized as AVERAGE in electricity consumption.

Participants in the given criteria condition read the same categorization information without knowing how the category boundaries were formed.

Next, participants answered two comprehension questions in the same format as experiment 4. They could only proceed after responding to each question correctly.

Then, participants saw two images with one describing each AC unit, listed side by side: one labeled as model X that consumed less electricity (625 watts) but was noisier (42 dB), and the other one labeled as model Y that consumed more electricity (675 watts) but was less noisy (28 dB). In the energy-separate condition, model X was categorized as low in noise and low in electricity consumption, whereas model Y was categorized as low in noise and average in electricity consumption. In the noise-separate condition, model X was categorized as average in noise and low in electricity consumption, whereas model Y was categorized as low in noise and low in electricity consumption. Participants then indicated which window AC model they would buy (model X or model Y). Finally, participants answered demographic questions and were debriefed.

6.2 | Results and Discussion

Overall, there were significant choice reversals as a function of the separated attribute in both given and random criteria conditions. Specifically, the choice share of the more energy-efficient model (i.e., model X) was larger in the energy-separate condition than in the noise-separate condition both when the categorization criteria were given (65.2% vs. 2.2%; Pearson $\chi^2(1) = 86.24$, $p < 0.001$) and when the categorization criteria were random (59.1% vs. 11.8%; Pearson $\chi^2(1) = 49.55$, $p < 0.001$).

I then ran a logistic regression with choice between the two models as the dependent variable (coded as 1 for choice of model X and 0 for choice model Y), and with separation condition (coded as 1 for energy-separation and 0 for noise-separation), categorization criteria (coded as 1 for given and 0 for random), and their interaction as the independent variables. As expected, there was a main effect of separation condition on choice, with the separation of electricity level associated with a greater likelihood of choosing the more energy-efficient model (i.e., model X; $\beta = 2.38$, Wald $\chi^2(1) = 42.11$, $p < 0.001$). There was also a significant main effect of categorization criteria ($\beta = -1.80$, Wald

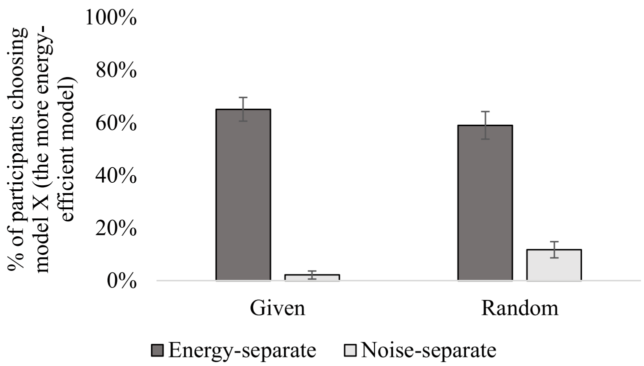

FIGURE 4 | Results of experiment 5. *Note:* Overall, participants were more likely to choose model X (the more energy efficient model) in the energy-separate conditions than in the noise-separate conditions. The effect of categorization was attenuated, but remained significant, when the categorization criteria were generated through a random process, compared to when the criteria were given.

$\chi^2(1) = 5.40$, $p = 0.020$), with a given (vs. random) set of criteria associated with smaller likelihood of choosing model X.

More central to my interest, there was a significant interaction effect between separation condition and categorization criteria ($\beta = 2.06$, Wald $\chi^2(1) = 6.17$, $p = 0.013$; see Figure 4), suggesting that the effect of categorization on choice was moderated by the way through which the categorization criteria were generated. Specifically, the effect of categorization was stronger when the criteria were given than when the criteria were generated through a random and arbitrary procedure.

The results in experiment 5 suggested that although people may fail to evaluate the informational value of categories before incorporating categories in their decision-making, the effect of categorization could be attenuated by highlighting the lack of meaningful insight provided by categories. Despite this, the results also demonstrated the robustness of the categorization effect, which remained significant even if participants were made aware that the categorization criteria were randomly assigned. If in the prior experiments, participants perceived the categories as carrying important information about the evaluative standards that should be used in the specific evaluation context, this possibility was diminished when the random nature of categories was made salient. Thus, the findings in this experiment suggested that highlighting the limited utility of categories beforehand had limited impact on individuals' decision-making at the moment of choice. Once attribute values were grouped into categories, the category information became difficult to ignore and continued to influence people's judgments and decisions.

7 | Experiment 6: The Categorization of a Single Attribute

So far, I examined people's multi-attribute choice between two options. In experiment 6 (preregistered at AsPredicted https://aspredicted.org/R38_J8Y), I furthered the understanding of categorization effect in several ways. First, experiment 6 adopted a different paradigm in which participants evaluated five hypothetical options, side by side, varying on one continuous attribute. It is possible that an increase in the number of options would attenuate the categorization effect by making within-category comparisons more salient. Besides, an increase in the number of options may provide valuable insights into the range of possible attribute values (Janiszewski and Lichtenstein 1999) and may interact with the effect of category boundaries. To address these possibilities, in experiment 6, I presented participants with five cupcakes in ascending order of calorie content and asked them to evaluate each cupcake. A post-test survey ($N=60$) showed that participants reported high levels of familiarity with the concept of calorie, and they felt highly confident in their ability to judge calorie amounts in food (see details in Appendix S1).

Participants were randomly assigned to either a high-cutoff or a low-cutoff condition. Given the cutoff for categorization, a cupcake with its calorie content higher (lower) than the cutoff was labeled with a red (green) sign. The color labels were used to minimize the semantic effect. Different from Ellison, Lusk, and Davis (2014)'s traffic light experiment, it was unlikely that participants would infer information beyond the level of calories (e.g., healthiness) from green versus red colors in the current experiment, as there existed little variances in healthiness levels across different cupcakes, and because it was made clear to participants that the categories were formed purely based on calorie levels. I predicted that the imposition of category boundaries would influence people's evaluations of multiple options varying on one continuous attribute.

Importantly, experiment 6 probed the effect of categorization on options either within or across category boundaries. I expected to see a discontinuity in participants' evaluations of options with continuous attribute values, whereby the perceived differences would be elevated only between options across category boundaries. Essentially of my interest was that the cupcake containing 150 calories (i.e., the focal item) was categorized with a green sign in the high-cutoff condition and a red sign in the low-cutoff condition. If category boundaries impact people's evaluation of objects along a continuous attribute, I predicted that participants' purchase likelihood of the focal item would differ across conditions.

Additionally, experiment 6 extended the categorization effect to a context in which the decision did not involve a tradeoff between different attributes, and it adopted purchase likelihood as the dependent variable. This design relieved the tension caused by opportunity cost considerations.

Finally, to test the persistence of categorization effect, I informed all participants that there were two different sets of categorization criteria, with one set being applied for a reason that was irrelevant to the evaluation task. I predicted that even participants were cued to notice that the categories might be uninformative, their evaluations of continuous attribute would still be impacted by categories.

7.1 | Method

A total of 300 participants were recruited through Prolific. They were randomly assigned to one of two conditions (low-cutoff or high-cutoff). They were asked to imagine that they wanted to purchase some cupcakes, and they were trying to limit their calorie intake. Then, they learnt that two local bakeries, Amber Cupcakes and Divine Cupcakes, categorized a cupcake as either low in calorie (green) or high in calorie (red) using different cutoff points. Amber (Divine) Cupcakes gives a cupcake a green label if it contains fewer than 140 (160) calories, and gives a cupcake a red label if it contains at least 140 (160) calories.

On the next page, participants assigned to the low- (high-) cutoff condition were told that they happened to be near Amber (Divine) Cupcakes, and participants in both conditions saw a menu that listed five cupcakes with various calorie levels (110, 130, 150, 170, and 190 calories). The cupcakes with fewer than 140 calories were labeled with green signs and otherwise red signs on Amber's menu (low-cutoff), whereas the cupcakes with fewer than 160 calories were labeled with green signs and otherwise red signs on Divine's menu (high-cutoff). The focal item—the cupcake that had 150 calories—was labeled with a red (green) sign in the low- (high-) cutoff condition. Then, participants indicated their purchase likelihood for each cupcake on 7-point scales (1 = *very unlikely*, 7 = *very likely*). Finally, participants answered demographic questions and were debriefed.

7.2 | Results and Discussion

A repeated measures ANOVA analysis with categorization condition as a between-subjects factor and five calorie levels as the within-subject measures revealed a significant effect of calorie level on purchase likelihood ($F(295)=68.95$, $p<0.001$, $\eta^2=0.483$). As predicted, there was a significant interaction effect between categorization condition and the calorie level ($F(295)=9.50$, $p<0.001$, $\eta^2=0.114$). Consistent with my theory, the interaction effect was driven by the difference in purchase likelihood of the focal item (the cupcake with 150 calories). Participants in the high-cutoff condition were more likely to purchase the focal item than participants in the low-cutoff condition ($M_{\text{low-cutoff}}=4.24$, SD = 1.69; $M_{\text{high-cutoff}}=4.95$, SD = 1.60; $F(298)=13.90$, $p<0.001$, $\eta^2=0.045$; see Figure 5).

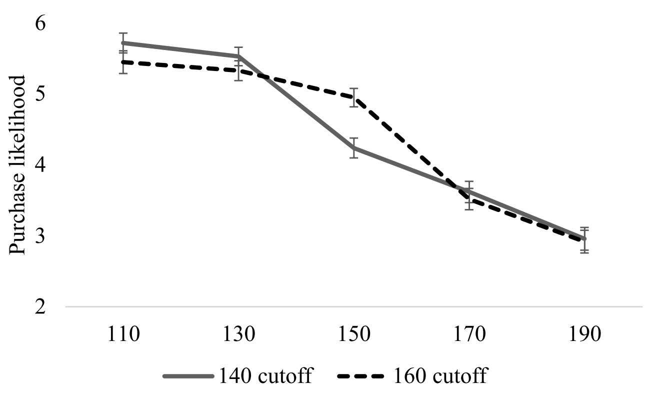

FIGURE 5 | Results of experiment 6. *Note:* The purchase likelihood for the focal item (the cupcake containing 150 calories) was higher in the high cut-off condition (with 140 calories as the category boundary) than in the low cut-off condition (with 160 calories as the category boundary).

Experiment 6 extended the categorization effect from multi-attribute choices to a single-attribute evaluation task, and the findings suggested that the pattern of categorization effect is generalizable to situations where no tradeoffs are involved in the judgment. I demonstrated that the purchase likelihood of the same product varied depending on how it was categorized along one continuous attribute, and this phenomenon was driven by the larger perceived differences across the category boundary. The results also showed that categorization effect persisted when participants were aware that there existed multiple categorization criteria and the categories might thus be uninformative.

8 | Experiment 7: Soliciting Preferences Through Willingness to Pay

So far, I have tested the effect of categorization of continuous attributes on choice and purchase likelihood. Experiment 7 tested whether the categorization effect could be extended to impact people's willingness to pay for each option. Willingness to pay has been shown to increase deliberations relative to choice, and consequently, the preference elicited through willingness to pay may systematically differ from the preference expressed through choice (O'Donnell and Evers 2019). Therefore, it is essential to test the persistence of the categorization effect soliciting preferences through willingness to pay.

In this experiment, participants chose between two flashlights that involved a tradeoff between battery life and weight. A post-test survey ($N=60$) found that participants were generally familiar with these two attributes and reported feeling confident in their ability to compare and judge these attribute values (see details in Appendix S1). Importantly, experiment 7 adopted an incentive-compatible deign, which could further rule out the possibility that participants relied on the category information because they did not care about the decision in a hypothetical scenario or because they made nonconsequential decisions that they perceived as desired by the researcher.

8.1 | Method

A total of 201 participants, recruited through MTurk, completed this experiment (preregistered at AsPredicted https://aspredicted.org/1ZB_QT1). Participants were randomly assigned to one of two conditions—battery-separate or weight-separate—and considered two flashlights that involved a tradeoff between battery life and weight—option A (the flashlight with a longer battery life) and option B (the flashlight with a lighter weight). All participants read the following instructions:

> "When buying a flashlight for hiking and camping, battery life and weight are two of the most important factors to consider. Battery life affects how long the flashlight can be used without needing a recharge, and weight determines how easy it is to carry the flashlight around."

Participants in the [battery-separate condition/weight-separate condition] then read:

> A retailer categorizes a flashlight as LONG in battery life if its battery life is longer than [10.5 / 7.5] hours, or else categorizes it as AVERAGE in battery life.

> The retailer categorizes a flashlight as LIGHT in weight if it weighs less than [23 / 11] ounces, or else categorizes it as AVERAGE in weight.

Participants then answered two comprehension questions in the same format as experiments 4 and 5. They could only proceed after responding to each question correctly.

Next, participants imaged purchasing a flashlight for hiking and camping. They saw the following two options that were described to be equivalent in all aspects except for battery life and weight.

	Battery life	**The prize you will get if you win**
Option A	LONG (12 h of use)	[LIGHT/AVERAGE] (17 ounces)
Option B	[AVERAGE/LONG] (9 h of use)	LIGHT (5 ounces)

Participant also saw one generic image of flashlight alongside each option. They were informed that one participant would be randomly selected to receive the option for which they indicated a higher willingness-to-pay. After entering the amount they would be willing to pay for each option, participants answered several demographic questions and were debriefed. As promised, after the study, I randomly selected one participant and sent them the instructions to receive the flashlight option they had a higher willingness to pay through Amazon Wishlist.

8.2 | Results and Discussion

A repeated measures ANOVA with categorization condition as a between-subjects factor and the willingness to pay for each option as a within-subject measure revealed a significant difference between the valuations for option A and option B, with option A (i.e., the flashlight with a longer battery life; $M=21.13$, $SD=21.10$) yielding an overall higher willingness to pay than option B (i.e., the flashlight that was lighter in weight; $M=19.41$, $SD=24.23$; $F(1, 199)=6.06$, $p=0.015$, $\eta^2=0.030$). The valuations did not vary as a function of categorization condition ($F(1, 199)=0.25$, $p=0.621$).

Importantly, the categorization changed the valuation of option A relative to option B, as revealed by a significant interaction effect between the willingness to pay for each option and categorization condition ($F(1, 199)=22.03$, $p<0.001$, $\eta^2=0.100$). Specifically, participants were willing to pay a larger amount for option A ($M=21.98$, $SD=15.71$) than option B ($M=17.01$, $SD=11.71$, $F(1, 100)=22.57$, $p<0.001$, $\eta^2=0.184$) in the battery-separate condition, where both attributes of option A were categorized favorably. In contrast, participants were willing to pay marginally significantly more for option B ($M=21.82$, $SD=32.18$) than option A ($M=20.27$, $SD=25.47$, $F(1, 99)=2.89$, $p=0.092$, $\eta^2=0.028$) in the weight-separate condition. To summarize, the results in experiment 6 further supported the hypothesis that categorization changes people's valuations between options that involve a tradeoff between two continuous attributes.

Using an incentive-compatible method, experiment 7 showed that category influenced participants' relative valuations between product options soliciting preferences through willingness to pay. The results suggested that the categorization effect is generalizable to decisions that require more deliberations.

9 | General Discussion

Categorization can influence how people both view an assortment of items and approach individual items in an assortment. When items are grouped into distinct categories, people tend to perceive two items as more similar to each other if they fall into the same category than if they fall into different categories (Isaac and Schindler 2014; Maki 1982; Mogilner, Rudnick, and Iyengar 2008). Building on prior findings, this research examines how categorization affects people's evaluations of objects with continuous attributes. In the present research, I demonstrated preference shifts between two product options or two incentive-compatible tasks that were presented with different categorization criteria, and I showed that this phenomenon occurs because category boundaries introduce discontinuities in the evaluations of continuous attribute values across category boundaries.

Across seven experiments, I demonstrated that people tend to prefer an option when it is categorized relatively more favorably than the other option, whether the preferences are solicited through choice (experiments 1–5), purchase likelihood (experiment 6), or willingness to pay (experiment 7), when the categorization criteria are generated through an arbitrary process (experiments 5 and 6), and when people face real financial consequences and are thus motivated to reveal their true preferences (experiments 2 and 7). Importantly, in each of the experiments conducted, participants were explicitly informed that the categories merely reflected whether an attribute value was higher or lower than a specific value. This clarification was intended to limit any extraneous inferences that participants might otherwise draw from the category labels. Additionally, the attributes used in the experiments were familiar and evaluable to most people. These designs helped reduce the perceived differences in relevant background information that could arise from the categorization, thus diminishing the potential information leakage (Sher and McKenzie 2006).

9.1 | Boundary Conditions

The categorization effect should happen when the following conditions are satisfied. First, the categorization effect requires that comparisons within a given category are not salient. I suggest that within-category rankings increase the salience of comparisons within a given category, thereby increasing perceived differences between items within the same category, as documented by the ranking effect (Leclerc, Hsee, and Nunes 2005). If within-category comparisons are not salient, however, then the difference between categories becomes more noticeable, and the categorization effect analyzed in the current research is more likely to emerge. This notion is consistent with the findings in experiment 6, where participants' evaluations of product items farther away from the category boundaries were less impacted by the categorization. As people move away from the boundaries, the within-category comparisons become more predominant, and this is when the ranking effect begins to trickle in, gradually diminishing the previously dominant categorization effect.

Besides, the categorization effect will occur only if people can tell that, for the particular attribute, one category is more favorable than the other category (e.g., people know that a "large" studio is better than a "small" studio). In the studies conducted by Leclerc, Hsee, and Nunes (2005), the favorability of the categories was not salient to participants, so participants did not systematically prefer one category over another. In experiment 2, I find that the pattern of categorization effect persists without semantic labels or evident signs that point to the superiority of one category over the other as long as people can easily distinguish the more favorable category from the less favorable one. However, if people do not generally agree on which side of the category boundary is better, then the imposition of category boundaries will not skew people's preferences systematically in one direction or the other. For example, an ice-cream categorized as "high" (vs. "low") in sugar level may be desired more by sweet-toothed people, but the preference is reversed for people who are trying to control their sugar intake. Therefore, we may not expect a systematic preference shift by categorizing the sugar content of an ice-cream.

9.2 | Theoretical Implications

The categorization effect introduced in this article contributes to several streams of research. First of all, the current research contributes to the categorization literature. Judgments and decisions are influenced by the context in which people make those evaluations (Lewis, Gaertig, and Simmons 2019; Mislavsky and Gaertig 2019; Prelec, Wernerfelt, and Zettelmeyer 1997; Sussman 2017). While research has extensively examined the contextual effects of categorization on variety perceptions (Kim and Yoon 2016; Mogilner, Rudnick, and Iyengar 2008; Redden 2008), goal progress perceptions (Sharif and Woolley 2020), and task implementation (Tu and Soman 2014), this research demonstrated that categorization affects people's multi-attribute decisions. Unlike past research that analyzed the perceptions of continuous attributes across natural boundaries (Donnelly, Compiani, and Evers 2021; Isaac and Schindler 2014) or categories generated by the evaluators themselves (Chernev and Gal 2010), the current research showed that third-party categorizations, instead of third-party labels (Aydinoğlu and Krishna 2011; Ellison, Lusk, and Davis 2014), also have a major impact on preferences, and the impact persists even when the categorizations are known to carry minimal informational value.

What is more, this research extends the body of literature on human biases that involve the use of to-be-ignored information.

People oftentimes fail to disregard information that is irrelevant to the focal task (e.g., the anchoring effect; see Tversky and Kahneman 1974; the hindsight bias; see Fischhoff 1975; the curse of knowledge; see Camerer, Loewenstein, and Weber 1989) even if prompted to ignore the information (Steblay et al. 2006). While past research has examined the effect of categories that carry, or at least assumed to carry, meaningful information that may assist decisions (Mogilner, Rudnick, and Iyengar 2008), this research highlights that the categorization effect persists even when people receive heavy-handed cues suggesting that the categories carry little informational value.

Finally, the key findings of this article add to the research on multi-attribute choices by uncovering a novel mechanism that drives preference shifts. While the classic preference reversals documented in the literature usually involve the manipulations of the preference scale or solicitation method (Hsee 1996; Hsee and Leclerc 1998; Slovic, Griffin, and Tversky 1990; Tversky, Sattath, and Slovic 1988; Tversky, Slovic, and Kahneman 1990), the preference reversals found in this research are instead caused by a discontinuity in people's evaluations of continuous attributes.

9.3 | Managerial Implications

The findings in this article have practical implications for marketers and policymakers who wish to affect decision makers' judgments and choices of the products or services that they provide and who are concerned about the effectiveness of their marketing messages and policies. Based on the findings in this research, the classification of continuous attributes into discrete categories can increase people's valuations of or skew their preferences toward the option that is favored by the categorization criteria, even when the criteria generating process is arbitrary. For example, simply changing the price category boundary from \$50 to \$100 may increase people's preference for a product originally in the "over \$50" category and now in the "below \$100" category compared to another product that has a superior (i.e., lower) price. As another example, categorizing the traffic volume on a highway as "heavy" rather than "normal" can effectively mitigate congestion by rerouting vehicles to less congested roads.

Besides, marketers and policymakers may consider relocating resources to best fit the discontinuous utility that people derive from categorization. For example, a highway that is newly categorized as "fast" (rather than "slow") may experience a significant increase in traffic. Thus, such a change should be accompanied by the deployment of sufficient traffic officers.

9.4 | Limitations and Future Directions

The current research provides an initial examination of the effect of categorization on people's judgments of continuous attributes, leaving many avenues for future research. This research shows that the objects farther away from the category boundary are less impacted by the boundary, but it remains unexplored how distant an attribute value from the boundary needs to be to diminish the impact of the category information, and what inferences people might draw from the available options. It is possible that people view the attribute values of available options as natural category boundaries, such that two attribute values are seen as more distinct from each other when there is a greater number of available options with attribute values lying between them. Furthermore, future studies might examine the effect of categorization on choice difficulty and post-choice evaluations. When does categorization make choices either easier or more difficult? Will people feel more or less satisfied about their decisions when the options are categorized along the continuous attributes? What factors may interact with these relationships? Although the current research did not test these possibilities, future research might investigate the categorization effect in more complicated contexts.

In conclusion, through seven experiments, including two incentive-compatible experiments, this research finds a categorization effect in multi-attribute decisions, documenting systematic preference shifts between two options that involve a tradeoff between two continuous attributes. In specific, the choice shares of the two options varied with the categorization criteria—people preferred the option whose attributes were categorized more favorably. This phenomenon occurs because categorization introduces discontinuity in people's judgments of continuous attributes across category boundaries.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

Pre-registrations, materials, and data for experiments 4 to 7 can be found on OSF at: https://osf.io/yfwgh/?view_only=53db23ac5a1a4353a5be150d93f2755c.

References

Aydinoğlu, N. Z., and A. Krishna. 2011. "Guiltless Gluttony: The Asymmetric Effect of Size Labels on Size Perceptions and Consumption." *Journal of Consumer Research* 37, no. 6: 1095–1112.

Camerer, C., G. Loewenstein, and M. Weber. 1989. "The Curse of Knowledge in Economic Settings: An Experimental Analysis." *Journal of Political Economy* 97, no. 5: 1232–1254.

Chernev, A., and D. Gal. 2010. "Categorization Effects in Value Judgments: Averaging Bias in Evaluating Combinations of Vices and Virtues." *Journal of Marketing Research* 47, no. 4: 738–747.

Cohen, J. B., and K. Basu. 1987. "Alternative Models of Categorization: Toward a Contingent Processing Framework." *Journal of Consumer Research* 13, no. 4: 455–472.

Coren, S., and J. S. Girgus. 2020. *Seeing Is Deceiving: The Psychology of Visual Illusions*. London: Routledge.

Dietvorst, B. J., and U. Simonsohn. 2019. "Intentionally "Biased": People Purposely Use to-Be-Ignored Information, but Can Be Persuaded Not to." *Journal of Experimental Psychology: General* 148, no. 7: 1228–1238.

Donnelly, K., G. Compiani, and E. R. Evers. 2021. "Time Periods Feel Longer When They Span More Category Boundaries: Evidence From the Lab and the Field." *Journal of Marketing Research* 59, no. 4: 821–839.

Ellison, B., J. L. Lusk, and D. Davis. 2014. "The Impact of Restaurant Calorie Labels on Food Choice: Results From a Field Experiment." *Economic Inquiry* 52, no. 2: 666–681.

Fischhoff, B. 1975. “Hindsight Is Not Equal to Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty.” *Journal of Experimental Psychology: Human Perception and Performance* 1, no. 3: 288.

Fisher, M., and F. C. Keil. 2018. “The Binary Bias: A Systematic Distortion in the Integration of Information.” *Psychological Science* 29, no. 11: 1846–1858.

Food Standards Agency. 2007. “Front-of-Pack Traffic Light Signpost Labelling Technical Guidance,” accessed February 21, 2013. http://www.food.gov.uk/multimedia/pdfs/frontofpackguidance2.pdf.

Goldstone, R. L. 1994. “The Role of Similarity in Categorization: Providing a Groundwork.” *Cognition* 52, no. 2: 125–157.

Hayes, A. F. 2017. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford publications.

Hell, W., G. Gigerenzer, S. Gauggel, M. Mall, and M. Müller. 1988. “Hindsight Bias: An Interaction of Automatic and Motivational Factors?” *Memory & Cognition* 16, no. 6: 533–538.

Hsee, C. K. 1996. “The Evaluability Hypothesis: An Explanation for Preference Reversals Between Joint and Separate Evaluations of Alternatives.” *Organizational Behavior and Human Decision Processes* 67, no. 3: 247–257.

Hsee, C. K., and F. Leclerc. 1998. “Will Products Look More Attractive When Presented Separately or Together?” *Journal of Consumer Research* 25, no. 2: 175–186.

Irmak, C., R. W. Naylor, and W. O. Bearden. 2011. “The Out-of-Region Bias: Distance Estimations Based on Geographic Category Membership.” *Marketing Letters* 22, no. 2: 181–196.

Isaac, M. S., and R. M. Schindler. 2014. “The Top-Ten Effect: Consumers' Subjective Categorization of Ranked Lists.” *Journal of Consumer Research* 40, no. 6: 1181–1202.

Jacowitz, K. E., and D. Kahneman. 1995. “Measures of Anchoring in Estimation Tasks.” *Personality and Social Psychology Bulletin* 21, no. 11: 1161–1166.

Janiszewski, C., and D. R. Lichtenstein. 1999. “A Range Theory Account of Price Perception.” *Journal of Consumer Research* 25, no. 4: 353–368.

Kim, H. J., and S. O. Yoon. 2016. “The Effect of Category Label Specificity on Consumer Choice.” *Marketing Letters* 27, no. 4: 765–777.

Labroo, A. A., R. Dhar, and N. Schwarz. 2008. “Of Frog Wines and Frowning Watches: Semantic Priming, Perceptual Fluency, and Brand Evaluation.” *Journal of Consumer Research* 34, no. 6: 819–831.

Leclerc, F., C. K. Hsee, and J. C. Nunes. 2005. “Narrow Focusing: Why the Relative Position of a Good in Its Category Matters More Than It Should.” *Marketing Science* 24, no. 2: 194–205.

Lewis, J., C. Gaertig, and J. P. Simmons. 2019. “Extremeness Aversion Is a Cause of Anchoring.” *Psychological Science* 30, no. 2: 159–173.

Maki, R. H. 1982. “Why Do Categorization Effects Occur in Comparative Judgment Tasks?” *Memory & Cognition* 10, no. 3: 252–264.

Mislavsky, R., and C. Gaertig. 2019. “Combining Probability Forecasts: 60% and 60% Is 60%, but Likely and Likely Is Very Likely.” Available at SSRN 3454796.

Mogilner, C., T. Rudnick, and S. S. Iyengar. 2008. “The Mere Categorization Effect: How the Presence of Categories Increases Choosers' Perceptions of Assortment Variety and Outcome Satisfaction.” *Journal of Consumer Research* 35, no. 2: 202–215.

Nosofsky, R. M. 1986. “Attention, Similarity, and the Identification–Categorization Relationship.” *Journal of Experimental Psychology: General* 115, no. 1: 39–57.

O'Donnell, M., and E. R. Evers. 2019. “Preference Reversals in Willingness to Pay and Choice.” *Journal of Consumer Research* 45, no. 6: 1315–1330.

Prelec, D., B. Wernerfelt, and F. Zettelmeyer. 1997. “The Role of Inference in Context Effects: Inferring What You Want From What Is Available.” *Journal of Consumer Research* 24, no. 1: 118–125.

Redden, J. P. 2008. “Reducing Satiation: The Role of Categorization Level.” *Journal of Consumer Research* 34, no. 5: 624–634.

Rosch, E. 1999. “Principles of Categorization. Concepts: Core Readings.” 189.

Sharif, M. A., and K. Woolley. 2020. “The Effect of Categorization on Goal Progress Perceptions and Motivation.” *Journal of Consumer Research* 47: 608–630.

Sher, S., and C. R. McKenzie. 2006. “Information Leakage From Logically Equivalent Frames.” *Cognition* 101, no. 3: 467–494.

Slovic, P., D. Griffin, and A. Tversky. 1990. “Compatibility Effects in Judgment and Choice.”

Steblay, N., H. M. Hosch, S. E. Culhane, and A. McWethy. 2006. “The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis.” *Law and Human Behavior* 30, no. 4: 469–492.

Sussman, A. B. 2017. “Valence in Context: Asymmetric Reactions to Realized Gains and Losses.” *Journal of Experimental Psychology: General* 146, no. 3: 376–394.

Tu, Y., and D. Soman. 2014. “The Categorization of Time and Its Impact on Task Initiation.” *Journal of Consumer Research* 41, no. 3: 810–822.

Tversky, A., and D. Kahneman. 1974. “Judgment Under Uncertainty: Heuristics and Biases.” *Science* 185, no. 4157: 1124–1131.

Tversky, A., S. Sattath, and P. Slovic. 1988. “Contingent Weighting in Judgment and Choice.” *Psychological Review* 95, no. 3: 371–384.

Tversky, A., P. Slovic, and D. Kahneman. 1990. “The Causes of Preference Reversal.” *The American Economic Review* 80, no. 1: 204–217.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.