Judging a Part by the Size of Its Whole: The Category Size Bias in Probability Judgments

2 authors:

Mathew S. Isaac
Seattle University
35 PUBLICATIONS **969** CITATIONS

SEE PROFILE

Aaron R. Brough
Utah State University
15 PUBLICATIONS **758** CITATIONS

SEE PROFILE

Judging a Part by the Size of Its Whole: The Category Size Bias in Probability Judgments
Author(s): Mathew S. Isaac and Aaron R. Brough
Source: *Journal of Consumer Research*, (–Not available–), pp. 310–325
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/10.1086/676126
Accessed: 27/06/2014 01:51

Judging a Part by the Size of Its Whole: The Category Size Bias in Probability Judgments

MATHEW S. ISAAC
AARON R. BROUGH

Whereas prior research has found that consumers' probability judgments are sensitive to the number of categories into which a set of possible outcomes is grouped, this article demonstrates that categorization can also bias predictions when the number of categories is fixed. Specifically, five experiments document a category size bias in which consumers perceive an outcome as more likely to occur when it is categorized with many rather than few alternative possibilities, even when the grouping criterion is irrelevant and the objective probability of each outcome is identical. For example, participants in one study irrationally predicted being more likely to win a lottery if their ticket color matched many (vs. few) of the other gamblers' tickets—and wagered nearly 25% more as a result. These findings suggest that consumers' perceptions of risk and probability are influenced not only by the number of categories into which possible outcomes are classified but also by category size.

Consumers frequently consider the perceived likelihood of uncertain events when making decisions—they judge the probability that their favorite sports team will win the next game, that a particular stock will increase in value, that they will win a prize in a sweepstakes, or that a certain level of insurance will be sufficient to cover the cost of an unexpected loss. Given how frequently consumers make predictions, one might expect that they would have developed a certain degree of expertise or competency at making them. Nevertheless, a great deal of evidence suggests that egregious errors in probability judgments are quite common (Kahn and Sarin 1988; Ofir and Lynch 1984; Sanbonmatsu, Posavac, and Stasney 1997).

When consumers predict future outcomes, one known source of potential bias is categorization; specifically, prior research has shown that consumers' probability judgments are sensitive to the manner in which information is categorized (Fox and Rottenstreich 2003; Martin and Norton 2009; See, Fox, and Rottenstreich 2006; Tsai and Zhao 2011; Tversky and Koehler 1994). For example, Tversky and Koehler (1994) found that the judged probability of a person dying from unnatural causes differed when alternative possibilities were bundled into a single large category (e.g., natural causes) rather than being presented in three smaller categories (e.g., heart disease, cancer, or some other natural cause). While such biases have been attributed to variation in the number of categories used to classify possible outcomes (Falk and Lann 2008; Fox and Clemen 2005; Fox and Rottenstreich 2003; Teigen 2001), it is not clear whether the coinciding variation in the size of the categories (i.e., the number of possible outcomes in each category) could also contribute to the bias.

Despite the lack of attention to category size in prior research, there are many instances in which consumers' predictions could potentially be influenced by whether a particular outcome is classified into a large or small category —even when category size may not be diagnostic of the likelihood of the outcome. For example, would the perceived probability of a particular Olympic swimmer winning an individual gold medal change if many (vs. few) competitors in the final race hail from the same country? Similarly, would an investor's expectation of a specific stock's performance be affected by whether many (vs. few) other stocks in the portfolio are classified in the same industry? Or would a patient's decision to visit the doctor for a lung cancer

screening be influenced by whether lung cancer is one of many (vs. few) of the diseases listed on a health brochure as being potentially preventable?

In contrast to prior research, in which both the number and size of categories typically change simultaneously, the present research aims to decouple category size from the number of categories, allowing us to determine whether changes in category size alone (holding constant the total number of categories and possible outcomes) will affect the perceived probability of a particular outcome. As an example of how category size can be manipulated independently of a change in the number of categories, consider a scenario in which a gambler is trying to determine the likelihood that a particular horse, whose coat happens to be black, will win a race against five other horses. Would this probability estimate differ if the race was between four black and two chestnut horses versus two black and four chestnut horses? In both cases, the set of six horses can be classified into two categories (i.e., black and chestnut), yet the number of horses in each category differs. If this difference influences perceived probabilities and subsequent wagers placed on the same black horse winning the race, these results may be explained by category size but not by the number of categories (since this did not change).

Our claim is that under certain conditions, category size can impact the perceived likelihood of a specific outcome's occurrence, such that an outcome classified into a large (vs. small) category is perceived as more likely to occur even when the total number of possible outcomes and the number of categories into which they are classified is held constant. We label this phenomenon the "category size bias." Of course, to the extent that the size of a category provides relevant information, it would be logical to factor category size into a probability judgment. Nevertheless, findings from five experiments show that even in cases where each outcome is equally likely to occur and the basis for categorization is nondiagnostic, category size can still exert an influence on consumers' predictions.

We attribute the category size bias to consumers' erroneous belief that individual category members inherit the statistical propensities of the parent category. To illustrate using the racehorse example, a category of four horses is, on average, more likely to contain the eventual winner than a category of two horses (assuming the basis of categorization is nondiagnostic). However, just because the larger category is associated with higher probability does not mean that an individual horse in the larger category has a greater probability of winning than an individual horse in the smaller category. Nevertheless, consumers sometimes behave as though the characteristics of a category apply not only to the category as a whole but also individually to each member of the category. The rationale for our prediction that category size will influence individual-level probability judgments is discussed in more detail in the following sections, as well as the theoretical contributions and practical implications.

THE IMPACT OF IRRELEVANT CATEGORY CUES ON CONSUMER JUDGMENTS

A major premise of this research is that categorization often distorts how consumers process information (Allport 1954; Bruner 1957; Campbell 1956; Rosch 1978). The biasing effects of category membership on judgments of individual category members have been documented across numerous domains, including judgments of physical length (Tajfel and Wilkes 1963), temperature (Krueger and Clement 1994), variety (Kahn and Wansink 2004; Mogilner, Rudnick, and Iyengar 2008), calorie content (Chernev 2011; Chernev and Gal 2010), and monetary value (Brough and Chernev 2012; Thomas and Morwitz 2005). Categorization can even alter the motivation, preferences, choices, and consumption levels of consumers (Cheema and Soman 2008; Fox, Ratner, and Lieb 2005; Isaac and Schindler 2014; Mishra and Mishra 2010; Simonsohn and Gino 2013; Wiltermuth and Gino 2013).

The present research focuses on how categorization impacts consumer judgments related specifically to probability and offers the first systematic test of how irrelevant cues about category size can affect judgments of individual outcomes. The focus of this work can be distinguished from two types of categorization-related biases that have been previously documented in the context of consumers' probability judgments. The first bias, known as the "alternative-outcomes effect," is similar to our research in that certain studies have also examined how irrelevant cues about category size affect probability judgments (Windschitl and Chambers 2004; Windschitl and Wells 1998; Windschitl and Young 2001). To illustrate, participants in one study were asked to judge the likelihood that a person holding a fixed number of tickets would win a raffle when most of the remaining tickets were either evenly or unevenly distributed across five other ticketholders (Windschitl and Young 2001). Whereas research on alternative outcomes solicits judgments about an entire category of possible outcomes (e.g., whether a specific ticketholder will win the raffle), the present research focuses exclusively on judgments of an individual outcome (e.g., whether a specific ticket will be selected) and shows how they are influenced by irrelevant category-level probabilities. Although in this particular example, predicting which ticketholder will win may be more important than predicting which ticket will be selected, there are many situations in which individual-level judgments are likely to be more relevant and important than category-level judgments. For example, to a patient considering a lung cancer screening, predicting the chances of contracting lung cancer may be more relevant than predicting the chances of contracting any potentially preventable disease.

The second categorization-related bias documented in prior work, a phenomenon labeled "partition dependence" (Fox, Bardolet, and Lieb 2005; Fox and Clemen 2005; Fox and Levav 2004; Fox et al. 2005; Fox and Rottenstreich 2003; See et al. 2006) is similar to our work in that it

examines the impact of categorization on individual-level probability judgments. For example, participants estimated that Sunday was less likely to be the hottest day of an upcoming week when the set of possible outcomes was subjectively grouped into seven (e.g., Sunday hottest, Monday hottest, etc.) categories rather than two (e.g., Sunday hottest, another day hottest; Fox and Rottenstreich 2003). Similarly, the estimated probability of Wharton being ranked first in *Business Week*'s next ranking of business schools was reduced when the odds of five competing schools were assessed individually (i.e., as five separate categories) rather than as a group (i.e., one category; Fox and Clemen 2005). As previously discussed, this stream of research differs from ours in that it has focused on the sensitivity of probability judgments to the number of categories used to classify the set of possible outcomes, whereas our research seeks to isolate whether changes in category size uniquely impact individual-level predictions.

The distinction between the focal judgment being an individual outcome versus an entire category is important in distinguishing our research paradigm from previous research. Prior work on partition dependence and the alternative-outcomes effect has frequently examined comparative judgments of singletons (i.e., one-item categories) in which the size of the category to which the focal judgment belongs does not vary, but the size (and number) of other categories does vary (Fox and Clemen 2005; Fox and Rottenstreich 2003; Windschitl and Wells 1998). In contrast, the present research focuses on how consumer judgments are influenced by changes in the size of the category to which the focal outcome belongs. The difference between individual-level and category-level probability judgments is relevant not only to theory but also to practice, since many consumer predictions focus on individual outcomes (e.g., a stock's performance) rather than on an entire category of outcomes (e.g., an industry's performance).

THE INTERPLAY OF CATEGORY-LEVEL AND INDIVIDUAL-LEVEL PROBABILITY JUDGMENTS

The way in which categories influence judgments about individual category members is typically thought to be through an indirect induction process (Gelman 1988; Gelman and Markman 1986), whereby features of either a prototype (i.e., an average category member) or an exemplar (i.e., a representative or salient category member) are used to make inferences about other category members (Medin et al. 2003; Osherson et al. 1990; Sloman 1993). However, other research suggests that a more direct transfer of category features to an individual member is also possible in which "people (at times erroneously) expect units to possess the attributes of the superordinate category" (Mishra and Mishra 2010, 1582–83) without necessarily considering prototypes or exemplars (Sloman 1998; Yamauchi and Markman 2000). For example, participants mistakenly judged the city of Montreal (Canada) to be north of Seattle (US) be-

cause Canada is north of the United States (Stevens and Coupe 1978). We refer to the latter process as categorical inheritance because, unlike indirect induction, in which inferences about an individual are made on the basis of family resemblance (e.g., similarity to other category members), the direct application of category features to individual category members suggests that an individual can also be thought to inherit characteristics directly from the parent category.

In many contexts, the distinction between indirect induction and categorical inheritance is irrelevant because both processes lead to the same judgments about individual category members. However, the two processes make different predictions in the context of probability judgments when each possible outcome has an equal probability of occurring. In such a context, an indirect induction process would not be expected to produce the category size bias because when consumers consider prototypes or exemplars, mental representations of actual or imagined individuals are evoked (Cuddy and Fiske 2004). If each outcome is equally likely to occur, making an inference about the likelihood of one outcome based on the probability of another outcome should lead to an accurate estimate. In contrast, because large categories are overrepresented within a set of possible outcomes, consumers rightly observe a more frequent occurrence of outcomes from large categories than from small categories; thus, large categories are imbued with the property of being likely to occur, whereas small categories are imbued with the property of being unlikely to occur. If consumers directly apply this category-level property to judgments about an individual category member, they will mistakenly conclude that a specific outcome belonging to a large category is more probable than a specific outcome belonging to a small category, even in cases where each outcome is actually equally likely to occur.

To illustrate the different predictions made by these two processes, consider a roulette wheel, where each numbered pocket into which the ball may land is categorized by color. If a gambler was trying to decide whether to bet on the "black two" pocket versus the "green zero" pocket, indirect induction (based on comparison to another individual pocket that is an exemplar or prototype) would lead him to correctly conclude that there is an equal chance that the ball will land on the "black two" pocket versus the "green zero" pocket because the odds that the ball will land on any individual pocket, regardless of its color, are identical. In contrast, since there are more black pockets than green pockets, categorical inheritance would lead to the incorrect conclusion that the "black two" pocket inherits a higher probability of occurrence from the "black" category, whereas the "green zero" pocket inherits a lower probability of occurrence from the "green" category.

As illustrated by the roulette example, the category size bias involves the erroneous influence of category-level probabilities on individual-level judgments. We argue that the extent to which category-level information guides individual judgments may depend on category salience. When category

salience is high, it may seem cognitively efficient to apply a single category-level judgment to all the individual members of that category in order to save time and enable consumers to distinguish between members of different categories very quickly (Heit and Rubinstein 1994; Lassaline 1996; Rosch 2002; Sloutsky 2003). This argument is consistent with prior work, suggesting that activating a mental representation of a category can result in its application to individual-level judgments (Andersen and Baum 1994; Markman and Ross 2003). Although categorization is advantageous for information processing in many situations, it can create a bias when category cues are nondiagnostic for the judgment at hand. Based on this reasoning, we predict that when category-level (vs. individual-level) probabilities are salient, consumers will be more likely to mistakenly infer that probability judgments about an entire category are relevant to predictions about a specific category member, even if the basis for categorization is irrelevant.

OVERVIEW OF EXPERIMENTS

To summarize, we argue that when judging the probability of an individual outcome, consumers may be influenced by category size even when it is irrelevant. This effect, which we refer to as the category size bias, predicts that the judged probability of *any* particular outcome within a fixed set of known outcomes will be higher [lower] when it is classified into a relatively large [small] category. The error in the category size bias stems from the fact that category-level properties are not always inherited by individual category members. Just because an outcome happens to be grouped with many other possible outcomes does not mean that it is any more likely to occur than one that happens to be grouped with few alternative outcomes. Nevertheless, consumers often seem to expect category members to inherit the category's propensity of occurring, such that the perceived probability that any outcome from a category will occur is transferred to each individual category member. We do not claim that consumers will substitute category-level judgments for individual-level judgments, but rather we predict that category size will exert influence on judgments about individual outcomes such that probability judgments for a specific outcome belonging to a large [small] category will be relatively high [low].

Across five experiments, we test the prediction that the size of the category into which a specific outcome is classified can influence consumers' perception that it is likely to occur. Study 1 aims to demonstrate the category size bias by showing that outcomes classified into relatively large categories are perceived as more likely to occur, whereas outcomes classified into relatively small categories are perceived as less likely to occur. Study 2 examines whether the category size bias extends beyond probability estimates to influence other consumer decisions, even when category size is imprecisely known. Study 3 investigates the role of categorical inheritance by testing whether this effect is more likely to occur when category salience is high rather than low. Study 4 further examines the mechanism underlying

the category size bias by pitting categorical inheritance against an anchoring process. Finally, study 5 documents the behavioral consequences of the category size bias in a consumer-relevant context and shows that the bias occurs for subjective as well as objective probability judgments.

In our first three studies, we examine situations in which each outcome, regardless of its category, is equally probable. This approach allows us to test for both indirect induction and categorical inheritance, since these two processes make different predictions about the influence of category size on the judged likelihood of a specific outcome. Specifically, when each outcome is equally likely to occur, an indirect induction process—in which consumers reference a representation of an individual category member (e.g., an exemplar or prototype)—would predict no effect of category size because each outcome is equally probable. In contrast, a categorical inheritance process—in which a representation of an entire category is referenced—would predict the category size bias because the focal judgment would inherit a greater [smaller] likelihood of occurrence from a large [small] parent category. In our last two studies, we shift our attention to cases where the likelihood of every outcome is not necessarily equal and the selection process is not random in order to test the generalizability of the category size bias.

STUDY 1

Study 1 aimed to provide an initial demonstration of the category size bias by showing that variations in category size can impact probability judgments even when the number of categories and possible outcomes is fixed. Based on our theorizing, we expected the judged probability of a particular outcome to be higher when it is classified into a large group and lower when it is classified into a small group due to an erroneous assumption that each individual member of a category is infused with the overall category's characteristics.

Method

A total of 223 participants (76% female; mean age = 39.6 years) were recruited using Amazon Mechanical Turk and evaluated one of two versions of a lottery. In both versions, participants judged the probability of randomly selecting a specific ball from an urn containing 15 balls each labeled with a different number from 1 to 15. In one version, the urn contained two black balls, 11 gray balls, and two white balls, whereas in the other version, the urn contained five balls of each color (see the appendix for stimuli).

Since the target of the probability judgment (i.e., the number 8 ball) happened to be gray, it belonged to a category that was either relatively large (i.e., 11 gray balls) or small (i.e., five gray balls). Nevertheless, this categorization was irrelevant because the mathematical probability of randomly selecting any individual ball was 1/15 (6.7%) regardless of variations in color. Participants in both conditions estimated the probability that the 8 ball would be selected by typing a percentage from 0% to 100% into a text box.

Results

We predicted a category size bias, such that the perceived likelihood of selecting the number 8 ball would be higher when it was part of a large category of 11 gray balls rather than a small category of only five gray balls. Our results were consistent with this prediction; participants judged the probability of selecting ball number 8 to be higher [lower] if it were the same color as many [few] other balls in the urn, even though this color-based categorization was unrelated to the selection likelihood of a particular ball (M_{large} = 18.5% vs. M_{small} = 12.9%; $t(221)$ = 2.55, $p < .02$).

We next examined the number of participants who correctly answered that the probability of selecting the 8 ball was 6.7%. Consistent with prior research showing that people tend to inflate probabilities (Sanbonmatsu et al. 1997; Tversky and Koehler 1994), we found that many participants overestimated the likelihood of drawing the 8 ball, and only 45% reported the correct mathematical probability (rounded to the nearest percent). However, there was no significant difference between the number of accurate responses in the large versus small conditions (42.2% vs. 48.2%; $\chi^2(1)$ = .82, $p > .36$). This suggests that while participants displayed the same general tendency to overestimate probabilities that has been discussed in prior research, category size exerted an independent influence on probability judgments above and beyond the overestimation bias that affected participants in both conditions.

Discussion

Consistent with our predictions, the results of study 1 documented a category size bias. This finding is particularly noteworthy because we designed study 1 to be a strong test of our theory—participants knew the total number of possible outcomes and were aware that the selection process was purely random. These conditions favored the null hypothesis that participants' individual-level probability judgments would not be biased by irrelevant information about category size because the probability that an outcome may randomly occur is a simple mathematical function of the number of possible outcomes. Nevertheless, participants judged the selection of the gray number 8 ball to be more likely when most of the other balls in the set were also gray.

We attribute this bias to a categorical inheritance process, in which participants seem to have assumed that individual outcomes inherit characteristics of the category to which they belong, such that if a gray ball is more likely to be drawn than a ball of another color, the 8 ball (which belongs to the gray category) must also be more likely to be drawn. The observed bias is inconsistent with an indirect induction process because every outcome was equally probable, so regardless of which ball was used as an exemplar or prototype, the probability of selecting that ball would have been identical to the probability of selecting the 8 ball, and no bias would have occurred due to variation in category size.

In the next study, we examined whether the category size bias extends beyond probability estimates to influence other consumer decisions. Our expectation was that when consumers' probability judgments are biased by category size, this bias may carry over to related decisions. In addition, we wanted to test whether the category size bias would be observed when participants had only vague knowledge of the overall set size and each category's exact size.

STUDY 2

Study 2 aimed to show that the category size bias occurs even when category size is approximated rather than explicitly provided and that it can affect not only probability estimates but also wagers placed in a lottery. In order to test the robustness of the category size bias that was observed in study 1, we designed study 2 to be notably different in procedure and execution yet conceptually similar in that the basis for categorization was completely irrelevant.

Method

A total of 81 undergraduate students at Seattle University (42% female; mean age = 21.4 years) participated in a paper-and-pencil study in exchange for course credit. Participants were shown an actual glass urn at the front of their classroom, which contained exactly 90 folded lottery tickets, 81 of which were blue and nine of which were yellow (although no numbers were disclosed to participants). They were told to imagine that a blindfolded and unbiased spectator would soon draw a single ticket from the urn and the winner would receive a cash reward of 100 times his or her wager amount. Subsequently, all other entrants (i.e., the nonwinners) would be asked to pay the amount that they had wagered. Participants were each given a single ticket and instructed to assume that it would be the last entry into the lottery (see the appendix for stimuli).

Category size was manipulated through the color of each participant's ticket—blue or yellow. Based on the relative number of similarly colored tickets already in the urn, participants who received blue tickets were in the large category condition and participants who received yellow tickets were in the small category condition. As in study 1, color was chosen as the arbitrary basis for categorization because it represented an attribute that was meaningless in our experimental context and could not have impacted the random selection process in any way—the mathematical probability of selecting any individual ticket was 1/91 (1.1%) regardless of variations in color.

On their ticket, participants wrote their name, the maximum amount (between $0 and $10) they would wager, and the estimated likelihood that they would win the lottery (between 0% and 100%). Although no money was actually exchanged, participants were directed to act as if real money were at stake. To ensure that neither the total number of tickets in the urn nor the relative color distribution was altered during the experiment, completed tickets were discreetly placed into a separate opaque container rather than being added to the glass urn. Subsequently, to verify that our between-condition manipulation of color did not affect

perceptions of overall set size, on a separate piece of paper participants recalled the color of their ticket and provided an estimate of the total number of tickets that were in the urn at the front of the room.

Results

We predicted a category size bias whereby wagers, as well as the perceived likelihood of winning the lottery, would be higher when a participant's ticket was added to a large category of 81 blue tickets rather than to a small category of nine yellow tickets. Our results were consistent with this prediction; participants were willing to wager 24% more if their ticket was part of a large rather than small category (M_{large} = \$7.65 vs. M_{small} = \$6.18; $t(79)$ = 2.03, $p < .05$) and participants judged the probability of their ticket being selected to be higher [lower] if it was the same color as many [few] other tickets in the urn (M_{large} = 24.2% vs. M_{small} = 10.6%; $t(73)$ = 2.53, $p < .02$). The latter analysis excluded six participants who did not provide a probability estimate. Since participants did not know exactly how many total tickets were in the urn, only 9.3% of participants correctly answered that the probability of winning the lottery was 1% (1/91, rounded to the nearest percent). However, as in study 1, the proportion of accurate responses did not differ for the large versus small conditions (11.1% vs. 7.7%, $\chi^2(1)$ = .26, $p > .61$). Furthermore, our category size manipulation did not distort the perception of the total number of tickets in the urn (M_{large} = 76.9 vs. M_{small} = 84.1; $t(79)$ = .53, $p > .59$).

To test whether wager amounts were mediated by probability estimates, we conducted a bootstrapping analysis (Preacher and Hayes 2008; Zhao, Lynch, and Chen 2010). We found a significant indirect effect of category size on maximum wager amount through probability estimation (β = .33, SE = .26, 95% confidence interval [CI] = .019 to .934), which establishes participants' perceived likelihood of winning the lottery as a mediator. The direct effect of category size on wager amount was not significant (β = .91, SE = .79, $p > .25$), which suggests that the observed mediation may be classified as indirect-only (Zhao et al. 2010, 200). Taken together, this analysis indicates that participants' wagers were guided by their perceived likelihood of winning the lottery, which was in turn influenced by category size.

Discussion

Study 2 highlighted the influence of the category size bias on wagers as a result of its impact on probability judgments. Furthermore, study 2 successfully replicated the finding of study 1 that the purely random outcome of drawing a particular item from an urn was perceived as more likely to occur when the item was classified into a large rather than small category. This replication is meaningful because the two studies differed on a number of dimensions, including the total set size (15 vs. 91 items), the color of the target (gray vs. blue/yellow), the study method (computer-based

vs. pencil/paper), the composition of the participant pool (primarily females from an online subject pool vs. primarily male undergraduates), and the precision of participants' knowledge of the total set size and category sizes. Taken together, the first two studies provide evidence in support of the notion that consumers rely on irrelevant information about category size when judging the probability of a random event and making related decisions.

In addition, study 2 helps rule out the possibility that the category size bias is due to participants' inferences about the information provided in our experimental manipulations. In accordance with Gricean norms (Grice 1975), some participants in study 1 may have inferred that because we specifically mentioned the color of the balls, such information must be relevant to their judgment. However, in study 2 no attempt was made to provide participants with information about category size or draw their attention to the different ticket colors, so the experimental manipulation gave participants no reason to infer that category size would be meaningful or relevant to their judgment. Furthermore, it is not clear from a Gricean norms explanation alone why participants, even if they thought color was relevant, would have assumed that an item from a larger category would be more likely to be drawn—it would have been reasonable to use size information to make the opposite prediction (i.e., that an item from a larger category would be less likely to be drawn, as in a little-fish-big-pond analogy).

In the next study, we conduct a more direct test of our proposition that the observed category size bias is due to a categorical inheritance process. According to our theorizing, categorical inheritance leads consumers to focus on information about an entire category and use category-level probabilities as a guide when making individual-level judgments. In contrast, indirect induction leads consumers to infer the probability of an individual outcome by focusing on information about another individual outcome (e.g., a prototype or exemplar). When category salience is high, consumers are more likely to focus on information about an entire category rather than another individual outcome. Thus, to the extent that categorical inheritance underlies the category size bias, the bias should be more pronounced when category salience is high rather than low. Conversely, when category salience is low and consumers focus more on individual outcomes, an indirect induction process should prevail and the category size bias should be attenuated.

STUDY 3

The goal of study 3 was to provide converging evidence of the category size bias and to show that, consistent with our theorizing, variation in category salience impacts the magnitude of the bias. Specifically, we expect that when category salience is high, consumers should be more likely to focus on category-level probabilities (leading to a categorical inheritance process) and the category size bias should be more pronounced. However, we expect that when category salience is low, consumers should be more likely to focus on individual-level probabilities (leading to an in-

direct induction process) and the category size bias should be attenuated. The experiment employed a 2 (category size: large vs. small) × 3 (category salience: high, moderate, low) between-participant design.

Method

A total of 175 participants (57% female; mean age = 40.7 years) were recruited using Amazon Mechanical Turk to participate in an online study about probability. Participants were asked to judge the probability of obtaining a particular outcome with the roll of a fair 26-sided die that displayed each letter of the English alphabet on one of its sides. Category size was manipulated by asking participants to evaluate either the probability of rolling the letter A, which is part of a small group of only 5 vowels, or the letter T, which is part of a large group of 21 consonants. Normatively, this categorization is irrelevant because the mathematical probability of rolling any individual letter is 1/26 (3.8%).

Category salience was manipulated by varying the extent to which the focal letter's category membership was highlighted. Specifically, those in the high category salience condition were explicitly informed that there were 5 vowels and 21 consonants on the die and were asked specifically about the probability of rolling either "vowel A" or "consonant T." In the moderate category salience condition, participants were also asked about "vowel A" or "consonant T," but the sizes of each category were not explicitly stated. In contrast, participants in the low salience condition were simply asked to estimate the likelihood that the next roll would be the "letter A" or the "letter T" (see the appendix for stimuli). Each participant recorded his or her response by typing a probability estimate between 0% and 100% for each letter into a text box on their screen.

Results

Based on our proposed inheritance account, we predicted that the category size bias would be most pronounced when category salience was high and attenuated when category salience was diminished. Consistent with this prediction, a 2 (category size: large vs. small) × 3 (category salience: high, moderate, low) ANOVA produced a significant interaction ($F(2, 169) = 4.76$, $p < .02$). When category salience was high, the category size bias was most pronounced (see fig. 1). Participants in this condition estimated the likelihood of rolling a "T" to be significantly greater than the likelihood of rolling an "A" ($M_{large} = 32.2\%$ vs. $M_{small} = 11.2\%$; $t(54) = 3.29$, $p < .01$), even though the chance of rolling either letter is statistically identical. When category salience was moderate, the category size bias was observed but less pronounced ($M_{large} = 24.1\%$ vs. $M_{small} = 12.2\%$; $t(57) = 2.01$, $p < .05$). However, when category salience was low, the category size bias was not observed ($M_{large} = 12.3\%$ vs. $M_{small} = 14.6\%$; $t(58) = .64$, $p > .52$).

The attenuation of the category size bias in study 3 was primarily driven by the large category (consonants). As the

FIGURE 1

ESTIMATED PROBABILITY AS A FUNCTION OF CATEGORY SALIENCE (STUDY 3)
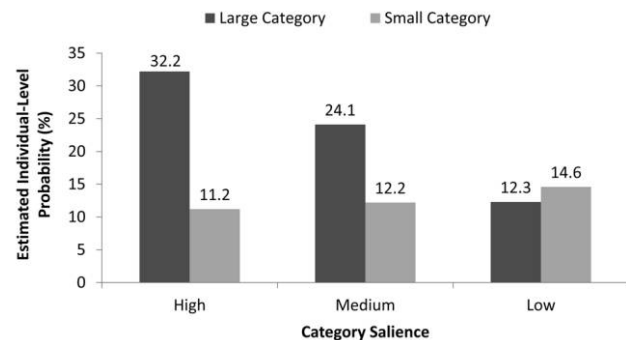
NOTE.—When category salience is high, estimated probabilities are higher for a possible outcome that belongs to a large category than for a possible outcome that belongs to a small category. However, this category size bias is attenuated as category salience diminishes.

salience of the consonant category decreased, so too did the perceived probability of rolling a "T" ($F(2, 78) = 3.96$, $p < .03$). However, as the salience of the vowel category decreased, the perceived probability of rolling an "A" was unchanged ($F(2, 91) = .53$, $p > .59$). Although not hypothesized a priori, we surmise that because there are so few vowels, merely being asked about the letter "A" may spontaneously prompt consideration of the vowel category whereas being asked about the letter "T" does not immediately evoke the consonant category. Another possibility is that there could have been a floor effect, since the vowel category is small relative to the total number of letters in the English alphabet.

Discussion

Consistent with our theorizing, the results of study 3 show that category salience influences the extent to which consumers are biased by category size. This finding lends additional support to our categorical inheritance account, which suggests that when category salience is high, consumers are more likely to rely on category-level probabilities when judging the probability of individual outcomes. In contrast, the observed category size bias runs counter to the predictions of indirect induction, which suggests that participants would have relied on the probability of a prototypical letter when judging the probability of the focal outcomes. Had this been the case, no bias should have been observed as a result of category size even when category salience was high because every letter, including prototypes and exemplars, was equally likely to be rolled. Yet, a category size bias was not only observed, but was also moderated by category salience.

In order to extend the applicability of this research, we transition in the remaining studies to situations where the likelihood of every outcome is not necessarily equal and the selection process is not random. As indicated by the examples described in the introduction of the paper, consumers often make decisions (e.g., to place a wager) based on the perceived likelihood of nonrandom events where factors other than chance (e.g., an athlete's skill) may influence the outcome. Even in these situations, we believe that categorical inheritance may produce the category size bias.

STUDY 4

The purpose of study 4 was to examine a case in which our proposed categorical inheritance account makes opposite predictions to an anchoring account. Specifically, in contrast to our argument that an inheritance process is responsible for the category size bias, one might argue that category size simply acts as an anchor to influence numeric estimates. While related, anchoring and inheritance make different predictions when a low numeric anchor denotes a high likelihood of an event's occurrence (e.g., odds relative to 1). An anchoring account would predict that the values participants provide when making numeric estimates simply mirror the size of the category—outcomes from large [small] categories will be assigned high [low] values. In contrast, categorical inheritance predicts that participants interpret the meaning of a category's size as it relates to probability, such that outcomes from large categories would be perceived as highly probable and therefore assigned low values in an odds-against estimate.

To tease apart these two different explanations for the category size bias, we used a dependent measure of probability in which higher numerical values denote a lower probability (i.e., odds relative to 1). In so doing, we aimed to show that the category size bias is driven by consumers' perception of probability and not merely by anchoring on a category's size when providing numeric estimates.

Method

Eighty-eight undergraduate students at Seattle University (50% female; mean age = 20.7 years) were recruited to participate in an online study regarding their perceptions of the upcoming men's NCAA college basketball tournament. Our experiment used a 2 (category size of focal team: large vs. small) × 2 (focal team: Florida State vs. Wisconsin) between-subjects design.

At the onset of the study, all participants were shown the mascots of an identical set of eight teams and asked to categorize each team into either a large or small category based on certain characteristics of the mascot. Half of the participants categorized the eight teams on the basis of whether the mascot was an animal or human, and the remaining participants categorized the teams on the basis of whether the mascot's face or entire body was visible (see

the appendix for stimuli). Of the eight teams, six mascots were animals and two were humans. Furthermore, six mascot images depicted a body and two depicted a face only. Importantly, both types of categorization (animal/human and face/body) were nondiagnostic of the team's likelihood of success in the college basketball tournament.

Immediately following the mascot categorization task, participants were asked to estimate the odds that either Florida State or Wisconsin would win the tournament by entering a number (with the form XX:1) into an open-ended text box. This primary dependent measure differs from earlier studies in that higher values of XX represent a lower (not higher) probability of winning. These teams were of interest because depending on the basis of categorization ("animal/human" vs. "face/body"), the same team would have been classified into either a large category or a small category (the Florida State Seminoles mascot depicts a human face and the Wisconsin Badgers mascot depicts an animal body). Thus, we expected participants who used the "animal/human" classification scheme to classify Wisconsin into the larger category and Florida State into the smaller category. However, we expected participants who used the "face/body" categorization scheme to place Wisconsin into the smaller category and Florida State into the larger category. This design allowed us to rule out the possibility that the pattern of results could be attributed to either team's expected performance in the tournament. As a nonequivalent dependent variable, all participants also estimated the odds that the Ohio State Buckeyes (which was not among the set of eight teams that participants had previously classified) would win. We expected that estimates provided for Ohio State would not differ based on experimental condition.

Results

To verify our basic assumption that larger categories would be perceived as more likely to produce the tournament's winner than smaller categories, we conducted a pretest with 29 students drawn from the same sample as those in the main experiment. As expected, pretest participants provided lower odds-against estimates that one of the six animal mascot teams ($M = 11.3 : 1$) versus one of the two human mascot teams ($M = 25.9 : 1$) would win the tournament ($t(28) = 3.22$, $p < .01$). Likewise, pretest participants provided lower odds-against estimates that one of the six teams with a mascot body ($M = 10.4 : 1$) versus one of the two teams with a mascot face ($M = 21.0 : 1$) would win the tournament ($t(28) = 3.00$, $p < .01$). Note, however, that the results of the pretest do not necessarily imply that a particular member of a six-team group is any more likely to win the tournament than a member of a two-team group. Yet the category size bias suggests that people may sometimes generalize probabilities in this way.

Our prediction was that participants who had previously classified a team into a large category would perceive a greater probability of that team's winning (and therefore estimate lower "odds against") than participants who had grouped the same team into a small category. To test this

prediction, we first examined participants' focal judgments (i.e., estimates that either Florida State or Wisconsin would win the tournament). Consistent with our theorizing, a 2 (category size of focal team: large vs. small) × 2 (focal team: Florida State vs. Wisconsin) ANOVA revealed a main effect of category size only ($F(1, 84) = 10.77$, $p < .01$). Participants who estimated the likelihood of Florida State's winning the tournament provided a lower odds-against estimate if they had grouped Florida State into a large category ($M = 62.1 : 1$) versus a small category ($M = 78.8 : 1$; $F(1, 84) = 4.27$, $p < .05$). Similarly, participants who estimated the likelihood of Wisconsin's winning the tournament provided a lower odds-against estimate if they had grouped Wisconsin into a large category ($M = 62.9 : 1$) versus a small category ($M = 82.9 : 1$; $F(1, 84) = 6.69$, $p < .02$). Furthermore, as expected, we found no effect of category size on odds estimates for Ohio State ($F(1, 84) = .34$, $p = .56$). The fact that no difference between conditions was observed for this nonequivalent dependent measure provides further indication that the differences in estimated odds across experimental conditions may be attributed to the category size manipulation.

Discussion

The results of study 4 show that participants provided small numeric estimates when assessing probability for members of large categories and large numeric estimates when assessing probability for members of small categories. This data pattern is inconsistent with an anchoring explanation, which suggests that consumers base probability estimates on the magnitude of a numeric anchor (e.g., category size). However, it is consistent with the idea that categorical inheritance leads consumers to erroneously use category size to assess category-level probabilities before estimating the likelihood that a particular outcome will occur.

In each of the studies presented thus far, we measured objective probability (out of 100%), a metric that may have hindered participants who have limited facility with mathematical concepts. Therefore, in the final study, we measured subjective (1 = very unlikely, 7 = very likely) rather than objective probability so that low-numeracy participants would not be disadvantaged. Note, however, that the inability of some participants to accurately compute probabilities in earlier studies cannot explain the category size bias (since there was no difference in accuracy across experimental conditions). In addition, one might argue that some participants in our studies may have been influenced by the fact that we asked them to estimate the likelihood of a single outcome and did not explicitly ask about others. While this explanation would not predict differences across participants who encountered large versus small categories, we asked participants in study 5 to judge the probability of each possible outcome in order to show that the bias is not simply an artifact of the dependent measure used in our experimental design.

STUDY 5

The final study aimed to enhance the external validity of our research by emphasizing the behavioral consequences of the category size bias in a consumer-relevant context. Since consumers routinely estimate the likelihood of a potential threat before deciding whether or not to take precautions, we decided to examine consumers' likelihood to engage in various behaviors that help to prevent different types of information technology (IT) security threats. In addition, to show that the category size bias does not stem from a simple misunderstanding of statistical principles or experimental instructions, we measured subjective instead of objective probability and asked participants to make judgments about each individual member of a category rather than only a subset.

Method

A total of 171 participants (37% female; mean age = 29.2 years) were recruited using Amazon Mechanical Turk to participate in an online study about IT security threats. All participants were shown an identical list of nine behaviors related to IT security and instructed to classify each behavior into one of two categories based on whether it helped to protect against identity theft or against loss of data (see the appendix for stimuli). Category size was manipulated by instructing participants to place seven behaviors into one of the categories (i.e., the large category), thereby creating a second (small) category consisting of the remaining two behaviors. Participants in one condition were asked to classify seven behaviors into the identity theft category, whereas participants in the other condition were asked to classify seven behaviors into the loss of data category. As a dependent measure, participants were asked to indicate the likelihood of performing each of the nine behaviors within the next three months (1 = very unlikely, 7 = very likely).

Consistent with our argument that individual members of a highly probable category will also be judged as highly probable, we predicted that when a specific behavior was classified into a large category of behaviors that protect against the same threat, consumers would perceive a greater risk associated with the threat (i.e., category-level probability) and a correspondingly greater risk of not engaging in the preventative behavior (i.e., individual-level probability). As a result, we expected participants to be more likely to engage in preventative behaviors classified into a large category. In addition, to ensure that there were no a priori differences across conditions in terms of past IT-related behavior, we asked participants to indicate whether or not they had performed each of the nine behaviors during the prior three months. Finally, as a manipulation check, we asked participants to recall whether they had categorized more preventative behaviors into the identity theft category or the loss of data category.

Results

We reasoned that an IT threat associated with more preventative behaviors (i.e., a large category) would be considered more likely to occur than an IT threat associated with fewer preventative behaviors. To verify our assumption, we conducted a pretest with 123 participants (61% female; mean age = 33.6 years) from another online subject pool. After categorizing the same list of behaviors as the main study, pretest participants indicated that a threat associated with seven preventative behaviors was more likely, on a 7-point scale (1 = very unlikely, 7 = very likely), than a threat associated with only two preventative behaviors (M = 4.38 vs. M = 3.87; $t(122)$ = 3.61, $p < .001$).

In the main study, we found no a priori differences across conditions with respect to past IT-related behavior, which suggests that the random assignment was successful. However, we expected, and found, that participants were more likely to engage in a preventative behavior that they had classified into a large (vs. small) category of behaviors. Indeed, participants exhibited a strong category size bias, such that in aggregate the mean likelihood of performing a large-category behavior was higher than the mean likelihood of performing a small-category behavior (M_{large} = 5.74 vs. M_{small} = 4.49; $t(170)$ = 9.86, $p < .001$). This was true regardless of whether participants had classified more behaviors into the identity theft or loss of data category. Specifically, participants who had classified seven behaviors as preventing identity theft exhibited the category size bias (M_{large} = 5.82 vs. M_{small} = 4.45; $t(90)$ = 8.87, $p < .001$), as did their counterparts who had classified seven behaviors as preventing loss of data (M_{large} = 5.65 vs. M_{small} = 4.54; $t(79)$ = 5.39, $p < .001$). This analysis included all 171 participants, regardless of whether they correctly recalled placing more behaviors into the identity theft or loss of data category. However, we also analyzed the subset of 149 participants who correctly recalled their experimental condition and found an identical pattern of results.

As another test of the category size bias, we created a difference score for each participant by subtracting the mean likelihood of performing behaviors classified into a small category from the mean likelihood of performing behaviors classified into a large category. Thus, a positive difference score indicates a greater subjective probability of engaging in behaviors that were classified into large (vs. small) categories. In accordance with the earlier results, difference scores were positive for 71% of participants, which is a proportion greater than chance ($\chi^2(1)$ = 29.48, $p < .001$).

To verify that this participant-level effect was not driven by one or two nonrepresentative behaviors, we analyzed each of the nine preventative behaviors separately. For each behavior, we created another difference score by subtracting the mean likelihood of performing the same behavior when it was classified in a small category by participants in one condition from its mean likelihood when placed in a large category by participants in another condition. We discovered that all difference scores were greater than 0, as shown in the (A)–(D) column of table 1, suggesting that the likelihood of performing each of the nine behaviors was higher when it had been classified into a large versus a small category.

Next, we created four additional difference scores for each behavior that allowed us to compare all possible combinations of category size (large vs. small) and category type (identity theft vs. loss of data). For example, the (B)–(E) column of table 1 compares participants in different conditions who categorized the same behavior into the large (vs. small) identity theft category. Similarly, column (C)–(F) makes the same comparison for behaviors classified into the large (vs. small) loss of data category. In contrast, column (B)–(F) compares behaviors that were placed into differently sized categories by participants who were in the same condition, as does column (C)–(E). Because 7/9ths of all behaviors were by design classified into the large category condition, the cell sizes of small category behaviors were too small to report statistical tests at the individual-behavior level. Nevertheless, 29 of the 34 difference scores that we calculated were positive, which suggests that the category size bias exhibited in study 5 is robust and was not driven by one or two nonrepresentative behaviors.

Discussion

The results of study 5 provide evidence that the category size bias influences decisions in a consumer-relevant context and that participants were more likely to engage in a preventative behavior that they had classified into a large (vs. small) category of behaviors. The fact that the category size bias was robust across all behaviors and observed even when the dependent measure was subjective rather than objective probability judgments provides further support for our theoretical account. One caveat, however, is that the design of study 5 intentionally forced participants to categorize a certain number of behaviors into each group rather than choosing how many behaviors to classify into each group. The rationale for this design is that if respondents were to self-categorize, it is possible that those who were more concerned with one category of IT security threats (and therefore more likely to take behaviors that would prevent it) would place more items into that category. In such a design, self-selection could provide an alternative explanation. However, by manipulating category size, we were able to rule out self-selection as a possible explanation for the effect.

Nevertheless, in order to show that the results generalize to situations where respondents determine category size without guidance, we also ran a separate test with 37 separate respondents (30% female; mean age = 31.8 years) that was nearly identical to study 5, except that we allowed respondents to categorize without any constraints. On average, participants placed 2.35 more items into one group than another, suggesting that differently sized categories occur naturally. Moreover, participants' intention to engage in behaviors that were classified into the larger group was higher on average than their intention to engage in behaviors that were classified into the smaller group (M_{large} = 5.54 vs. M_{small} = 5.03; $t(36)$ = 2.44, $p < .02$). These results suggest that people often categorize items into differently

TABLE 1

LIKELIHOOD OF PERFORMING PREVENTATIVE BEHAVIORS AS A FUNCTION OF CATEGORY SIZE (STUDY 5)

Preventative Behavior	Mean likelihood of performing each behavior during the next three months (1 = very unlikely, 7 = very likely)					
	(A) Large Category	(B) Large Identity Theft Category	(C) Large Loss of Data Category	(D) Small Category	(E) Small Identity Theft Category	(F) Small Loss of Data Category
(1) Change password	5.24	5.31	5.13	4.28	4.28	*
(2) Encrypt sensitive file	3.76	3.91	3.60	3.58	3.65	3.54
(3) Use pop-up blocker or firewall	6.36	6.45	6.21	6.32	6.33	6.29
(4) Use password-protected screensaver	4.89	4.85	4.94	3.22	3.11	3.32
(5) Install (or maintain) security software with automatic updates	6.07	6.13	6.01	5.37	5.63	5.18
(6) Verify the publisher before downloading and installing software	5.94	6.10	5.74	5.52	5.47	5.67
(7) Use antivirus or antispy software	6.46	6.56	6.36	5.00	3.50	5.60
(8) Avoid opening unsolicited e-mail attachments	6.65	6.60	6.71	6.32	6.07	7.00
(9) Back up data	5.53	5.37	5.58	4.91	4.94	4.90
ALL	*5.74*	*5.82*	*5.65*	*4.49*	*4.54*	*4.45*

	Mean difference score for each behavior				
	(A)–(D)	(B)–(E)	(C)–(F)	(B)–(F)	(C)–(E)
(1) Change password	.96	1.03	*	*	.85
(2) Encrypt sensitive file	.18	.26	.06	.37	−.04
(3) Use pop-up blocker or firewall	.03	.12	−.07	.17	−.12
(4) Use password-protected screensaver	1.67	1.74	1.62	1.53	1.83
(5) Install (or maintain) security software with automatic updates	.70	.50	.83	.94	.39
(6) Verify the publisher before downloading and installing software	.42	.63	.07	.43	.27
(7) Use antivirus or antispy software	1.46	3.06	.76	.96	2.86
(8) Avoid opening unsolicited e-mail attachments	.34	.53	−.29	−.40	.64
(9) Back up data	.62	.43	.68	.47	.64
ALL	*1.25*	*1.28*	*1.20*	*1.37*	*1.11*

NOTE.—Positive difference scores indicate that a consumer was more likely to engage in behaviors that were classified into large (vs. small) categories. Column (A)–(D) represents an overall difference score for each behavior, regardless of the experimental condition. Column (B)–(E) compares participants in different conditions who categorized the same behavior into the Large (vs. Small) Identity Theft category. Similarly, column (C)–(F) makes the same comparison for behaviors classified into the Large (vs. Small) Loss of Data category. In contrast, column (B)–(F) compares behaviors that were placed into differently sized categories by participants who were in the same condition, as does column (C)–(E).

*No participants classified "Change password" into the Small Loss of Data Category.

sized groups without being directed to do so and that such categorization can affect their perceptions of risk and associated behaviors even when their attention is not drawn to differences in category size. Indeed, the fact that this pattern of results was obtained without category size instructions seems to suggest that the results of study 5 are not an artifact of the experimental design.

This post-test also provided insight into the natural categorization of each behavior and therefore lends credibility to the reasonableness of the categorization instructions used in the main study. Specifically, two of the behaviors ("change password regularly" and "use a password-protected screensaver") were classified by more than 60% of respondents into the Identity Theft category, whereas two

other behaviors ("back up data regularly" and "encrypt sensitive files") were classified by more than 60% of respondents into the Loss of Data category. None of the remaining five behaviors achieved more than a 60% majority, suggesting that respondents found them to be more ambiguous and open to classification into either category.

GENERAL DISCUSSION

This research examines how consumers' probability judgments of a particular outcome are influenced by the way in which the set of all possible outcomes is grouped. In contrast to prior research, which focuses on how the number of categories can impact judgments, our results show that cat-

egory size exerts an independent effect on probability judgments above and beyond any influence exerted by the number of categories, such that outcomes classified into large [small] categories are perceived as more [less] likely to occur. We attribute this bias to a categorical inheritance process whereby individual outcomes are thought to inherit properties of the overall category.

Five studies document a category size bias and are consistent with a categorical inheritance explanation. Study 1 shows that outcomes classified into relatively large [small] categories are perceived as more [less] likely to occur. Study 2 demonstrates that the category size bias influences wager amounts in addition to probability judgments, even when category size is only approximated. Study 3 shows that category salience moderates the effect, such that the bias is more pronounced when the category is highly salient. Study 4 rules out an anchoring explanation by showing that perceptions of probability rather than magnitude produce the effect. Finally, study 5 highlights the implications of the category size bias for consumer behavior by assessing behavioral intentions related to the subjective probability of future behaviors classified into large or small categories.

At a conceptual level, this research contributes to the literature on judgment under uncertainty by providing the first systematic empirical demonstration that irrelevant cues about the size of a focal outcome's category can independently affect judgments of its likelihood to occur. This is a novel contribution that is distinct from two related findings in prior research on probability judgments: the alternative-outcomes effect and partition dependence. Although the alternative-outcomes effect is similar to our work in that it also examines irrelevant cues about category size, it differs from our work by soliciting judgments about an entire category of possible outcomes (e.g., the likelihood that someone holding multiple tickets to a lottery will win) rather than a specific individual outcome (e.g., the likelihood that a specific ticket will be selected). Likewise, whereas partition dependence research resembles our work in that it examines individual-level probability judgments, it differs from our work by focusing on how changes in the number of categories rather than category size affect probability judgments. Further, most work on partition dependence has simultaneously manipulated the number of categories along with category size, making it difficult to assess the unique role of category size. The identification of these confounds, together with our empirical demonstration of the category size bias when they are controlled for, represents a significant advance in our understanding of judgment under uncertainty.

Our research also contributes to the literature on categorization by showing how the salience of a category can impact the way in which category size influences judgments of individual category members. Specifically, we show that when information about a category is salient, the category size bias is more likely to occur. Consistent with the predictions made by categorical inheritance, we document a category size bias even when the basis of categorization is nondiagnostic and add to prior categorization research sug-

gesting that indirect induction is not the only way in which category membership can influence individual-level judgments. Indeed, although the potential for category-level information to transfer directly to category members (without the use of individual exemplars or prototypes) has been implied by previous research (Mishra and Mishra 2010; Sloman 1998; Yamauchi and Markman 2000), our findings suggest that this process can apply not only to concrete and well-established features (e.g., geographic location of countries and cities) but also to context-dependent features (e.g., likelihood of occurrence). This adds to prior work examining the consequences of categorization on consumer judgment and decision making and suggests yet another context in which consumers seem to rely on a categorical judgment when making a numeric estimate (Brough and Chernev 2012; Chernev and Gal 2010; Krueger and Clement 1994; Mishra and Mishra 2010; Tajfel and Wilkes 1963).

Our findings contribute to the body of evidence that consumers are often poor judges of probability (Kahn and Sarin 1988; Ofir and Lynch 1984; Sanbonmatsu, Posavac, and Stasney 1997), but they reflect more than a simple lack of statistical ability. Indeed, the category size bias cannot be attributed to low numeracy alone because the percentage of accurate responses did not differ across experimental conditions, yet both the direction and magnitude of errors were systematically influenced by category size. This evidence provides a nuanced view of how reliance on categorical information can erroneously influence individual-level probability estimates. Moreover, our finding that the bias can be attenuated by reducing the salience of category-level information improves our understanding of the way in which categorization affects judgments under uncertainty and suggests a boundary condition that can potentially improve the accuracy of consumers' predictions.

The notion that category size can bias probability judgments has important public policy implications and suggests that categorization may be used strategically to promote healthy and safe behaviors. For example, when crafting health-related messages for consumers, policy makers might group a highly preventable disease such as lung cancer with a large (vs. small) number of other potential health risks. Our findings suggest that because a person may believe he is more likely to contract a disease from a large (vs. small) group of health risks, classifying lung cancer into a large category in this way may increase the perceived risk of contracting lung cancer. Through this technique, policy makers may encourage consumers to visit their doctor for regular screenings and avoid risky behaviors such as smoking. Similarly, to the extent that a fatality report increases perceived risk by categorizing automobile accidents with a large (vs. small) number of avoidable causes of death, drivers may be more inclined to wear a seat belt. Managing the categorical presentation of risk-related information in these ways may increase early detection of preventable diseases and encourage socially beneficial conduct. By identifying category size as a factor that can influence consumers' probability judgments, our research informs the

literature on consumer health-risk perceptions (Keller, Lipkus, and Rimer 2002; Yan and Sengupta 2013).

This research also suggests several areas for further investigation. First, future research might test whether a similar bias may occur when consumers rather than possible outcomes are the objects of categorization. For example, would consumers feel more likely to win a sweepstakes if they are part of a majority versus minority group of participants (e.g., based on race, gender, etc.)? Second, future research may investigate the stage at which categorization influences probability judgments. For example, does category-level information impact probability judgments during encoding, at retrieval, or both? Third, although our results provide robust evidence that category size influences probability judgments, further work is required to identify when consumers will be more sensitive to small categories, more sensitive to large categories, or equally sensitive to both. We suspect that sensitivity to category size may depend upon the degree of deviation from the expected category size in a particular context. Finally, future work can also explore the influence of absolute category size (e.g., number of outcomes) versus relative category size (e.g., number of outcomes compared to other categories). This article lays the groundwork for a formal investigation of these issues relating to how consumers may be influenced by category size.

DATA COLLECTION

The first author supervised the collection of data for studies 2 and 4. The second author supervised the collection of data for studies 1, 3, and 5. Study 2 was conducted by paper and pencil with students at Seattle University in May 2013 in exchange for course credit. Study 4 was conducted with students at Seattle University in May 2012 in exchange for course credit. Studies 1, 3, and 5 were conducted using paid online participants living in the United States who were recruited using Amazon Mechanical Turk in March 2011, March 2011, and July 2013. The first and second authors shared responsibility for data analysis for studies 1, 3, and 5. The first author assumed primary responsibility for the analysis of data from studies 2 and 4. For all studies, data were discussed and results reviewed on multiple occasions by both authors.

APPENDIX
Study 1 Stimuli

An urn contains 15 balls that are numbered from 1 to 15.

A. Small Category Size Condition

Balls 1-5 are black, 6-10 are gray, and 11-15 are white.

B. Large Category Size Condition

Balls 1-2 are black, 3-13 are gray, and 14-15 are white.

If one ball is drawn from the urn, what is the probability that it will be ball 8? ____%

Study 2 Stimuli

In the front of the room is a jar containing the tickets of those who have already entered the lottery. All entrants were instructed to write their names on their tickets before placing them in the jar. In this lottery, only one entry is allowed per person. Imagine that one (and only one) ticket will soon be drawn randomly from the jar (by a blindfolded and unbiased spectator), and the winner will receive a cash reward.

Now suppose you have the option to buy the very LAST ticket to this lottery, which is the colored slip of paper that you have been given.

Step 1: Write your name on the ticket.

Step 2: Below your name, write down the maximum amount (between $0 and $10) that you would be willing to pay to enter this lottery.

If you write "$0" or any amount greater than $10, you will NOT be entered into the lottery. If you instead write an amount between $0 and $10 and your entry is the one that is drawn from the jar, you will receive 100 times the amount that you wrote down on your lottery ticket. In other words, if you decide to spend one penny to enter, you will get back one dollar if you win. If you don't win, however, you will lose the penny. If you spend $10 (the maximum amount), you will get $1,000 in return if you win the lottery. If you don't win, you will lose $10.

Step 3: Below the maximum dollar amount that you would be willing to pay, please also provide an estimate of the likelihood that you will be the lottery winner by entering a percentage between 0% (very unlikely) and 100% (very likely).

Study 3 Stimuli

Condition 1: High Category Salience. Consider a die with twenty-six faces. On each face is a different letter of the alphabet, such that there are 5 vowels and 21 consonants. How likely do you think it is that the next roll will be the vowel A [consonant T]? ____%

Condition 2: Moderate Category Salience. Consider a die with twenty-six faces. On each face is a different letter

of the alphabet, such that all 26 letters are represented. How likely do you think it is that the next roll will be the vowel A [consonant T]? ____%

Condition 3: Low Category Salience. Consider a die with twenty-six faces. On each face is a different letter of the alphabet, such that all 26 letters are represented. How likely do you think it is that the next roll will be the letter A [T]? ____%

Study 4 Stimuli

Below are the names and mascots of teams that are likely to be in next year's NCAA Men's Basketball Tournament.

Condition 1: Face/Body Categorization. Please classify each team mascot either as "face" or "body" by dragging each logo and dropping it into the appropriate box.

Condition 2. Animal/Human Categorization. Please classify each team mascot either as "animal" or "human" by dragging each logo and dropping it into the appropriate box.

Study 5 Stimuli

The risk of IT security threats may be minimized by taking various precautions. Below is a list of preventative behaviors that help to protect against two different kinds of IT security threats. Each behavior helps prevent either Identity Theft or Loss of Data. Please drag each behavior listed below into the category where you think it belongs.

- Change password frequently.
- Encrypt sensitive files.
- Use a pop-up blocker and firewall.
- Use a password-protected screensaver.
- Install security software with automatic updates.
- Verify publisher before downloading or installing software.
- Use antivirus and antispy software.
- Avoid opening unsolicited e-mail attachments.
- Back up data regularly.

Condition 1: Large Identity Theft Category/Small Loss of Data Category. Helps to prevent Identity Theft (select seven items)

Helps to prevent Loss of Data (select two items])

Condition 2: Small Identity Theft Category/Large Loss of Data Category. Helps to prevent Identity Theft (select two items)

Helps to prevent Loss of Data (select seven items)

REFERENCES

Allport, Gordon W. (1954), *The Nature of Prejudice*, Reading, MA: Addison-Wesley.

Andersen, Susan M., and Alana Baum (1994), "Transference in Interpersonal Relations: Inferences and Affect Based on Significant-Other Representations," *Journal of Personality*, 62 (4), 459–97.

Brough, Aaron R., and Alexander Chernev (2012), "When Opposites Detract: Categorical Reasoning and Subtractive Valuations of Product Combinations," *Journal of Consumer Research*, 39 (2), 399–414.

Bruner, Jerome S. (1957), "On Perceptual Readiness," *Psychological Review*, 64 (2), 123–52.

Campbell, Donald T. (1956), "Enhancement of Contrast as Composite Habit," *Journal of Abnormal and Social Psychology*, 53 (3), 350–55.

Cheema, Amar, and Dilip Soman (2008), "The Effect of Partitions on Controlling Consumption," *Journal of Marketing Research*, 45 (6), 665–75.

Chernev, Alexander (2011), "The Dieter's Paradox," *Journal of Consumer Psychology*, 21 (2), 178–83.

Chernev, Alexander, and David Gal (2010), "Categorization Effects in Value Judgments: Averaging Bias in Evaluating Combinations of Vices and Virtues," *Journal of Marketing Research*, 47 (4), 738–47.

Cuddy, Amy J. C., and Susan T. Fiske (2004), "Doddering but Dear: Process, Content, and Function in Stereotyping of Older Persons," in *Ageism: Stereotyping and Prejudice against Older Persons*, ed. Todd Nelson, Cambridge, MA: MIT Press, 3–26.

Falk, Ruma, and Avital Lann (2008), "The Allure of Equality:

Uniformity in Probabilistic and Statistical Judgment," *Cognitive Psychology*, 57 (4), 293–334.

Fox, Craig R., David Bardolet, and Daniel S. Lieb (2005), "Partition Dependence in Decision Analysis, Resource Allocation, and Consumer Choice," in *Experimental Business Research*, Vol. 3, ed. R. Zwick and A. Rapoport, Amsterdam: Kluwer, 229–51.

Fox, Craig R., and Robert T. Clemen (2005), "Subjective Probability Assessment in Decision Analysis: Partition Dependence and Bias toward the Ignorance Prior," *Management Science*, 51 (9), 1417–32.

Fox, Craig R., and Jonathan Levav (2004), "Partition-Edit-Count: Naive Extensional Reasoning in Judgment of Conditional Probability," *Journal of Experimental Psychology: General*, 133 (4), 626–42.

Fox, Craig R., Rebecca K. Ratner, and Daniel S. Lieb (2005), "How Subjective Grouping of Options Influences Choice and Allocation: Diversification Bias and the Phenomenon of Partition Dependence," *Journal of Experimental Psychology: General*, 134 (4), 538–51.

Fox, Craig R., and Yuval Rottenstreich (2003), "Partition Priming in Judgment under Uncertainty," *Psychological Science*, 14 (3), 195–200.

Gelman, Susan A. (1988), "The Development of Induction within Natural Kind and Artifact Categories," *Cognitive Psychology*, 20 (1), 65–95.

Gelman, Susan A., and Ellen M. Markman (1986), "Categories and Induction in Young Children," *Cognition*, 23 (3), 183–209.

Grice, H. Paul (1975), "Logic and Conversation," in *Syntax and Semantics 3: Speech Acts*, ed. Peter Cole and Jerry L. Morgan, New York: Academic Press.

Heit, Evan, and Joshua Rubinstein (1994), "Similarity and Property Effects in Inductive Reasoning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (2), 411–22.

Isaac, Mathew S., and Robert M. Schindler (2014), "The Top-Ten Effect: Consumers' Subjective Categorization of Ranked Lists," *Journal of Consumer Research*, 40 (2), 1181–202.

Kahn, Barbara E., and Rakesh K. Sarin (1988), "Modeling Ambiguity in Decisions under Uncertainty," *Journal of Consumer Research*, 15 (2), 265–72.

Kahn, Barbara E., and Brian Wansink (2004), "The Influence of Assortment Structure on Perceived Variety and Consumption Quantities," *Journal of Consumer Research*, 30 (4), 519–33.

Keller, Punam Anand, Isaac M. Lipkus, and Barbara K. Rimer (2002), "Depressive Realism and Health Risk Accuracy: The Negative Consequences of Positive Mood," *Journal of Consumer Research*, 29 (1), 57–69.

Krueger, Joachim, and Russell W. Clement (1994), "Memory-Based Judgments about Multiple Categories: A Revision and Extension of Tajfel's Accentuation Theory," *Journal of Personality and Social Psychology*, 67 (1), 35–47.

Lassaline, Mary E. (1996), "Structural Alignment in Induction and Similarity," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 754–70.

Markman, Arthur B., and Brian H. Ross (2003), "Category Use and Category Learning," *Psychological Bulletin*, 129 (4), 592–613.

Martin, Jolie M., and Michael I. Norton (2009), "Shaping Online Consumer Choice by Partitioning the Web," *Psychology and Marketing*, 26 (10), 908–26.

Medin, Douglas L., John D. Coley, Gert Storms, and Brett L. Hayes (2003), "A Relevance Theory of Induction," *Psychonomic Bulletin and Review*, 10 (3), 517–32.

Mishra, Arul, and Himanshu Mishra (2010), "Border Bias: The Belief That State Borders Can Protect against Disasters," *Psychological Science*, 21 (11), 1582–86.

Mogilner, Cassie, Tamar Rudnick, and Sheena 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 (2), 202–15.

Ofir, Chezy, and John G. Lynch (1984), "Context Effects on Judgment under Uncertainty," *Journal of Consumer Research*, 11 (2), 668–79.

Osherson, Daniel N., Edward E. Smith, Ormond Wilkie, Alejandro Lopez, and Eldar Shafir (1990), "Category-Based Induction," *Psychological Review*, 97 (2), 185–200.

Preacher, Kristopher, and Andrew Hayes (2008), "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models," *Behavior Research Methods*, 40 (3), 879–91.

Rosch, Eleanor (1978), "Principles of Categorization," in *Cognition and Categorization*, ed. Eleanor Rosch and Barbara B. Lloyd, Hillsdale, NJ: Erlbaum, 27–47.

——— (2002), "Principles of Categorization," in *Cognitive Psychology: Core Readings*, ed. Daniel Levitin, Cambridge, MA: MIT Press, 251–70.

Sanbonmatsu, David M., Steven S. Posavac, and Randon Stasney (1997), "The Subjective Beliefs Underlying Probability Overestimation," *Journal of Experimental Social Psychology*, 33 (3), 276–95.

See, Kelly E., Craig R. Fox, and Yuval S. Rottenstreich (2006), "Between Ignorance and Truth: Partition Dependence and Learning in Judgment under Uncertainty," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32 (6), 1385–402.

Simonsohn, Uri, and Francesca Gino (2013), "Daily Horizons: Evidence of Narrow Bracketing in Judgment from 10 Years of MBA Admissions Interviews," *Psychological Science*, 24 (2), 219–24.

Sloman, Steven A. (1993), "Feature-Based Induction," *Cognitive Psychology*, 25 (2), 231–80.

——— (1998), "Categorical Inference Is Not a Tree: The Myth of Inheritance Hierarchies," *Cognitive Psychology*, 35 (1), 1–33.

Sloutsky, Vladimir M. (2003), "The Role of Similarity in the Development of Categorization," *Trends in Cognitive Sciences*, 7 (6), 246–51.

Stevens, Albert, and Patty Coupe (1978), "Distortions in Judged Spatial Relations," *Cognitive Psychology*, 10 (4), 422–37.

Tajfel, Henri, and Alan L. Wilkes (1963), "Classification and Quantitative Judgement," *British Journal of Psychology*, 54 (2), 101–14.

Teigen, Karl Halvor (2001), "When Equal Chances = Good Chances: Verbal Probabilities and the Equiprobability Effect," *Organizational Behavior and Human Decision Processes*, 85 (1), 77–108.

Thomas, Manoj, and Vicki G. Morwitz (2005), "Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition," *Journal of Consumer Research*, 32 (1), 154–64.

Tsai, Claire I., and Min Zhao (2011), "Predicting Consumption Time: The Role of Event Valence and Unpacking," *Journal of Consumer Research*, 38 (3), 459–73.

Tversky, Amos, and Derek J. Koehler (1994), "Support Theory: A Nonextensional Representation of Subjective Probability," *Psychological Review*, 101 (4), 547–67.

Wiltermuth, Scott S., and Francesca Gino (2013), "I'll Have One of Each: How Separating Rewards into (Meaningless) Cate-

gories Increases Motivation," *Journal of Personality and Social Psychology*, 104 (1), 1–13.

Windschitl, Paul D., and John R. Chambers (2004), "The Dud-Alternative Effect in Likelihood Judgment," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30 (1), 198–215.

Windschitl, Paul D., and Gary L. Wells (1998), "The Alternative-Outcomes Effect," *Journal of Personality and Social Psychology*, 75 (6), 1411–23.

Windschitl, Paul D., and Michael E. Young (2001), "The Influence of Alternative Outcomes on Gut-Level Perceptions of Certainty," *Organizational Behavior and Human Decision Processes*, 85 (1), 109–34.

Yamauchi, Takashi, and Arthur B. Markman (2000), "Inference Using Categories," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (3), 776–95.

Yan, Dengfeng, and Jaideep Sengupta (2013), "The Influence of Base Rate and Case Information on Health-Risk Perceptions: A Unified Model of Self-Positivity and Self-Negativity," *Journal of Consumer Research*, 39 (5), 931–46.

Zhao, Xinshu, John G. Lynch, and Qimei Chen (2010), "Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis," *Journal of Consumer Research*, 37 (2), 197–206.