

The Perfection Premium

Mathew S. Isaac¹  and Katie Spangenberg² 

Social Psychological and
Personality Science
1-8

© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1948550620944313
journals.sagepub.com/home/spp



Abstract

This research documents a *perfection premium* in evaluative judgments wherein individuals disproportionately reward perfection on an attribute compared to near-perfect values on the same attribute. For example, individuals consider a student who earns a perfect score of 36 on the American College Test to be more intelligent than a student who earns a near-perfect 35, and this difference in perceived intelligence is significantly greater than the difference between students whose scores are 35 versus 34. The authors also show that the perfection premium occurs because people spontaneously place perfect items into a separate mental category than other items. As a result of this categorization process, the perceived evaluative distance between perfect and near-perfect items is exaggerated. Four experiments provide evidence in favor of the perfection premium and support for the proposed underlying mechanism in both social cognition and decision-making contexts.

Keywords

perfection, categorization, numerical cognition, social cognition

Media reports suggest that people disproportionately value perfection. For example, Walnut Hills High School in Cincinnati, OH, was featured in national newspapers after 17 students earned a perfect score of 36 on the American College Test (ACT; Strauss, 2019). However, individuals also seem to value near-perfection. For example, Tesla's Model S received accolades after earning a near-flawless road test score of 99 (out of 100) from *Consumer Reports* (Valdes-Dapena, 2013). In this research, we explore whether preferences and interpersonal judgments meaningfully differ when people consider individuals or items that have achieved perfect versus near-perfect status on the same attribute or rating.

In four experiments, we find that even when the objective numerical gap between two values is equal, people perceive the difference between individuals and items to be greater if one has a perfect attribute value or rating. For example, the perceived difference in intelligence of two students scoring 100% versus 99% on an exam exceeds the perceived gap between students scoring 99% versus 98%, even though the scores differ by 1% in both cases. In addition to documenting this asymmetric “perfection premium,” we provide evidence that it occurs at least in part because of a categorization bias, whereby perfect individuals and items occupy a mental category separate from near-perfect others.

The contexts investigated in this research involve ascending attributes (i.e., higher values are better; Gunasti & Ross, 2010) whose numerical values are located on bounded interval scales with minimum and maximum values. In such contexts, one might presume evaluative equidistance, whereby equivalent differences in attribute values correspond with identical differences in evaluations. However, numerical cognition research

suggests that people may not always exhibit evaluative equidistance. For example, the math skills of a student who placed 10th on a math test are considered much stronger than those of a student who placed 11th, but students who placed 11th and 12th on the same test are perceived as equally skilled (Isaac & Schindler, 2014). Other work has shown that left digits are often perceptually overweighted (Bizer & Schindler, 2005), which can affect how citizens judge public services (such as schools) based on their performance ratings (Olsen, 2013).

Also relevant to our proposed perfection premium is prior research documenting discontinuities in judgments at scale endpoints. Research examining decision making under risk has found evidence of a probability weighting function with an inverse S-shape (Camerer & Ho, 1994; Tversky & Kahneman, 1992; Wu & Gonzalez, 1996). Because decision weights are “not well-behaved near the endpoints” (Kahneman & Tversky, 1979, p. 283), a change in probability near 0% or 100% is weighted more heavily than an identically sized change in the middle of a distribution. For example, in a game of Russian roulette involving a gun with six chambers, people are willing to pay more to reduce the number of bullets from one to zero than from four to three although each option reduces their chance of death by 16.67% (Kahneman & Tversky, 1979).

¹ Seattle University, Seattle, WA, USA

² University of Washington, Seattle, WA, USA

Corresponding Author:

Mathew S. Isaac, Seattle University, Seattle, WA 98122, USA.
Email: isaacm@seattleu.edu

The preference for certain outcomes was first documented in the Allais paradox (1953) and has been labeled the certainty effect (Kahneman & Tversky, 1979) or the 100% effect (Li & Chapman, 2009). Even the *illusion* of certainty affects judgments and choices by eliminating perceived risk. For example, people disproportionately prefer a vaccine whose efficacy claim references 100% probability (e.g., 100% efficacy against 70% of strains) over an equivalent net vaccine efficacy that does not (e.g., 95% efficacy against 74% of strains; Li & Chapman, 2009).

Given that most research examining whether people perceive meaningful differences near maximum scale values focuses on risk perceptions (for an exception, see Shampanier et al., 2007), it is unclear whether differences will emerge when comparing the performance of individuals or items that do not vary in terms of riskiness. Indeed, the perfection premium may make different predictions from the certainty effect because it relates to performance perceptions rather than risk perceptions. For example, the perfection premium predicts that the evaluative gap will be greater between students earning a 35 versus a perfect 36 on the ACT, as compared to the gap between students earning a 34 versus a 35. In contrast, as there is no difference in the perceived risk associated with any of these scores, the certainty effect would predict these slight differences to be evaluatively imperceptible.

In this research, we propose that the perfection premium arises because people spontaneously categorize perfect items separately from near-perfect items. Prior research has shown that seemingly continuous stimuli may be mentally represented in a categorical manner. For example, Tu and Soman (2014) found that although time elapses continuously, people create temporal boundaries and mentally classify future time events (e.g., deadlines) into “like-the-present” or “unlike-the-present” categories. These category distinctions can exaggerate the perceived difference between adjacent items (e.g., dates) on either side of the boundary (Brenner et al., 1999).

Although numerical values can be ordered on a continuous mental number line, individuals are nevertheless prone to categorize numbers (e.g., odd/even; Laski & Siegler, 2007; Shepard et al., 1975). For example, people spontaneously create mental categories at “round number” (e.g., zero- or five-ending) boundaries, expanding the perceived distance between adjacent numbers in different categories (Isaac & Schindler, 2014). Such effects on perceptions (and thus evaluations) are consistent with Rosch’s (1978) claim that categorizing a stimulus makes it “not only equivalent to other stimuli in the same category but also different from stimuli not in that category” (p. 28).

In sum, we posit that when individuals encounter extremely high attribute values, a natural basis for categorization is whether an item is perfect or not. If an item with a perfect attribute or rating is categorized separately, the evaluative distance between it and a near-perfect item will be exaggerated. Four experiments document this perfection premium in decision making and social perception contexts and implicate categorization as an underlying mechanism for this effect.

Study 1

Study 1 provides an initial demonstration of the perfection premium. Given that companies routinely make claims about their products’ perfect (e.g., “all-natural ingredients,” “100% juice”) or near-perfect (e.g., “99.44% pure,” “98% aloe vera”) performance or attributes, we utilize a decision-making context for this study.

Method

Across studies, we analyzed the data after all responses were collected, and we report all data exclusions, manipulations, and measures related to our hypothesis testing. Manipulations and measures for all studies appear in the Supplementary Online Appendix. Data and code for all studies are available on Open Science Framework (OSF, <https://osf.io/y2p3b/>). Sample sizes were determined in advance based on logistical and financial constraints.

Participants were 450 people (60.2% female, $M_{\text{age}} = 43.35$ years) recruited through a paid online panel (Dynata). The instructions informed participants that they were planning to bake a dessert and were considering two bars of unsweetened baking chocolate, Baking Bar A and Baking Bar B. Bar A was priced 20 cents lower than Bar B (\$3.65 vs. \$3.85). However, Bar B always contained exactly one percentage point more of pure ingredients, creating a price–quality trade-off.

Participants were randomly assigned to one of three conditions. Those in the perfect condition read that Bar A contained 99% pure ingredients, but Bar B contained 100% pure ingredients. In the higher of the two near-perfect conditions, Bars A and B were 98% and 99% pure, respectively. In the lower of the two near-perfect conditions, Bars A and B were 97% and 98% pure, respectively. Thus, in terms of purity, the superior option (i.e., Bar B) attained perfection in the perfect condition but not in the two near-perfect conditions.

First, participants indicated which chocolate bar they would choose when baking the dessert. Subsequently, they reported how strongly they would prefer either Bar A or B on a sliding scale (0 = *strongly prefer Bar A*, 100 = *strongly prefer Bar B*). Finally, participants were asked to explain the rationale for their answers in a text box.

Results

Across conditions, participants exhibited a general preference for Bar B (i.e., superior purity, but more expensive) over Bar A, as reflected by the overall choice shares of 59.3% and 40.7%, respectively, $\chi^2(1) = 15.68, p < .001, \phi = .19$. More germane to our theorizing, we predicted that participants would exhibit a relatively greater preference for the superior Bar B when it was perfect versus when it was near-perfect. Consistent with this prediction, a χ^2 test of independence returned a significant effect of condition on choice share, $\chi^2(2, N = 450) = 8.44, p = .015, \phi = .14$. Bar B was chosen by 67.5% of participants in the perfect condition, as compared to 51.0% of

participants in the higher near-perfect condition, $\chi^2(1, N = 298) = 8.44, p = .004, \phi = .17$, and 59.2% of participants in the lower near-perfect condition, $\chi^2(1, N = 303) = 2.27, p = .13, \phi = .09$.

Participants' responses to the relative preference measure followed a similar pattern. A one-way analysis of variance (ANOVA) of condition on relative preference returned a significant result, $F(2, 447) = 5.73, p = .003, \eta_p^2 = .03$. Planned contrasts revealed that participants in the perfect condition ($M_{100/99} = 69.66, SD = 25.65, N = 151$) exhibited a stronger preference for Bar B than participants in the higher near-perfect condition, $M_{99/98} = 60.94, SD = 26.33, N = 147, F(1, 447) = 7.88, p = .005, \eta_p^2 = .02$, or the lower near-perfect condition, $M_{98/97} = 60.30, SD = 28.44, N = 152, F(1, 447) = 9.22, p = .003, \eta_p^2 = .02$. Further, the two near-perfect conditions did not differ from each other, $F(1, 447) = .04, p > .83$. These results provide evidence of a perfection premium.

If the perfection premium is driven by a categorization process, participants should attend more to similarities between choice options in the near-perfect conditions and more to differences between options in the perfect condition (because the two products reside in different categories). To test this, two independent coders (blind to condition) coded whether the rationales provided by participants indicated that they viewed the two options as similar (e.g., small, negligible) or different (e.g., large, meaningful). Interrater agreement was an acceptable 93.8% and 76.7% for similarity and difference coding, respectively, and differences were resolved by discussion (Hunt et al., 2015). We found that only 3.3% of participants in the perfect condition indicated that the two options were similar versus 9.7% of participants in the near-perfect conditions, $\chi^2(1, N = 450) = 5.86, p = .015, \phi = -.11$. On the other hand, 28.5% of participants in the perfect condition indicated that the two options were different versus 21.1% of participants in the near-perfect conditions, $\chi^2(1, N = 450) = 3.06, p = .08, \phi = .08$.

Next, we combined the coded similarity and difference scores to create a composite measure of the perceived distance between options. Specifically, we subtracted the similarity code (0 or 1) from the difference code (0 or 1). Thus, a composite distance score between +1 and -1 was assigned to each rationale, with higher values denoting greater perceived distance between the options. Finally, we conducted a mediation analysis using the PROCESS macro (Model 4; Hayes, 2017) to test whether perceived distance mediated the effect of condition on relative preference. This mediation analysis utilized bootstrapping with repeated extraction of 10,000 samples. All conditions were grouped into two cells (*perfect* = 1, *near-perfect* = 0), with the continuous relative preference measure as the dependent variable and the composite perceived distance measure as a potential mediator. Results of the mediation analysis indicated that the indirect effect of condition (perfect vs. near-perfect) through perceived distance was positive ($B = 2.82, SE = 1.04$) and statistically different from zero (95% confidence interval [0.81, 4.87]). This result suggests that perceived distance mediated the relationship between condition

and relative preference. These findings provide preliminary support for our claim that the perfection premium is a categorization bias.

Study 2

In Study 2, we attempt to demonstrate the robustness of the perfection premium in three ways. First, we test whether the perfection premium affects social cognition (i.e., judgments of human characteristics). Second, we examine a condition in which we equate a nonround two-digit number (i.e., 36) with perfection. Because the only three-digit number presented to participants in Study 1 was the perfect value 100, the discontinuous processing of multidigit numerals on the mental number line (Hinrichs et al., 1982; Restle, 1970) may have led to the evaluative gap that we observed. Furthermore, prior work has documented that round numbers such as 100 serve as cognitive reference points, yielding cognitive fluency and enhancing liking (King & Janiszewski, 2011; Pope & Simonsohn, 2011; Rosch, 1975). The design of Study 2 allows us to demonstrate that the perfection premium is not unique to cases in which a three-digit round number represents perfection. Third, we examine a situation where perfection is described as points on a test rather than as a percentage. Prior research has shown that consumers find percentages confusing and often misinterpret them (Kruger & Vargas, 2008); by using points in Study 2, we show that the perfection premium is not limited to contexts where performance or quality is described in percentage terms.

Method

Study 2 was conducted on Amazon Mechanical Turk (MTurk) with 252 U.S. adults (42.1% female, $M_{\text{age}} = 38.59$ years). Participants learned that 20 high school students in a gifted cohort had all earned scores at or above 30 (out of 36) on the ACT. They were also informed that ACT scores are correlated to some degree with intelligence quotient test scores.

We employed a within-participant perfection manipulation by having participants indicate the relative level of intelligence of two students (drawn from the cohort without replacement) on 10 separate occasions, each of which constituted a separate decision. Each time, we randomly varied the numerical information that participants saw. Specifically, the score of the higher scoring student in each pair was between 32 and 36 (i.e., a perfect ACT score), inclusive. The score of the lower scoring student was either 1 or 2 points lower than the higher scoring student in each pair. Participants rated which student was more intelligent on a 100-point scale (0 = *definitely [the lower scoring student]*, 100 = *definitely [the higher scoring student]*). Based on this design, all participants encountered the same 10 pairs of students (presentation order randomized), of which two pairs involved a perfect score (i.e., 36 vs. 35 and 36 vs. 34).

Results

In accordance with our presumption that participants would expect the higher scoring student to be more intelligent, the mean relative intelligence rating was 72.56 ($SD = 17.18$), which is significantly greater than the scale midpoint of 50, $t(251) = 20.85, p < .001$. Our main prediction is that this difference will be more pronounced if the higher scoring student had earned a perfect ACT score (i.e., 36). To test this prediction, we classified each of the 10 student pairs into one of the four within-participant conditions and calculated the mean relative intelligence rating of each condition.

A 2 (perfection: perfect, near-perfect) \times 2 (score difference: 1 point, 2 points) repeated-measures ANOVA on relative intelligence ratings returned a significant main effect of ACT score difference, $F(1, 251) = 8.55, p = .004, \eta_p^2 = .033$, indicating that the difference in perceived intelligence between each pair of students was greater when the higher scoring student had scored 2 points (vs. 1 point) higher than the lower scoring student.

In support of the perfection premium, we observed a significant main effect of perfection, $F(1, 251) = 22.32, p < .001, \eta_p^2 = .082$. When paired students' ACT scores were 1 point apart, relative intelligence ratings were more disparate when the higher scoring student achieved a perfect score ($M_{36/35} = 74.27, SD = 20.41$) as compared to a near-perfect score ($M_{\text{other}} = 70.64, SD = 17.78$), $F(1, 251) = 23.81, p < .001, \eta_p^2 = .087$. Similarly, when paired students' ACT scores were 2 points apart, relative intelligence ratings were more disparate when the higher scoring student achieved a perfect score ($M_{36/34} = 75.49, SD = 22.30$) as compared to a near-perfect score ($M_{\text{other}} = 73.32, SD = 17.61$), $F(1, 251) = 4.91, p = .028, \eta_p^2 = .019$. The interaction between perfection and score difference was nonsignificant, $F(1, 251) = 1.41, p = .24, \eta_p^2 = .006$.

We also conducted a 5 (higher score: 36, 35, 34, 33, 32) \times 2 (score difference: 1 point, 2 points) repeated-measures ANOVA on relative intelligence ratings. These results are presented visually in Figure 1 and the complete table of contrasts is provided in the Supplementary Online Appendix. This analysis provides further evidence of a discontinuity between perfect and near-perfect values.

Study 3

Based on our theorizing, if a near-perfect and perfect numerical value are placed in the same evaluative category, the perfection premium should be attenuated. In Study 3, we test this prediction in a different choice context.

Method

Study 3 was conducted on MTurk with 322 U.S. adults (63.0% female, $M_{\text{age}} = 37.34$ years). Participants were randomly assigned to one of the four conditions in a 2 (perfection: perfect, near-perfect) \times 2 (categorization cue: present, absent) between-participants design.

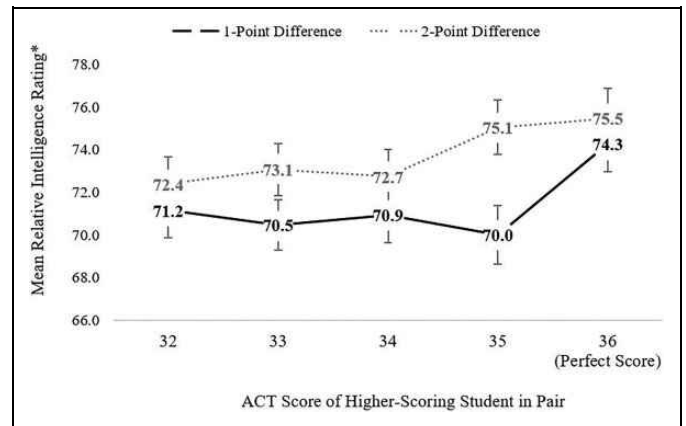


Figure 1. Perceived difference in intelligence of students whose American College Test scores differed by either 1 or 2 points (Study 2). Note: Error bars represent ± 1 Standard Error. *0 = definitely [the lower scoring student] is more intelligent, 100 = definitely [the higher scoring student] is more intelligent.

Participants read that they needed to choose between two pairs of wool socks to buy. In all conditions, the more expensive pair of socks cost USD \$4.15, whereas the less expensive pair cost USD \$3.65. The more expensive pair contained a higher percentage of Merino wool, thereby creating a tradeoff between price and wool quality. Participants in the perfect (near-perfect) condition learned that the superior socks contained 100% (96%) Merino wool, whereas the inferior socks contained 98% (94%) Merino wool. Participants in the categorization-cue-absent condition were given no other information. In the absence of a common categorization cue, we predicted that participants in the perfect condition would spontaneously place the perfect and near-perfect socks into different categories, thereby inflating the evaluative distance between them. Participants in the categorization-cue-present condition also learned that each pair of socks had received a five-star rating on Amazon from the same number of reviewers. A posttest confirmed that the categorization cue (i.e., five-star ratings) increased the extent to which participants placed the perfect 100% Merino wool socks in the same evaluative category as the near-perfect socks.¹

Next, participants designated which pair of socks they would select. Subsequently, they indicated their relative preference on a sliding scale from 0 to 100, with higher numbers indicating greater preference for the more expensive pair of socks with higher wool content.²

Results

We performed a binary logistic regression with participants' choice (0 = less expensive, lower wool quality socks, 1 = more expensive, higher wool quality socks) as our dependent variable. Choice share was regressed on perfection (0 = near-perfect, 1 = perfect), categorization cue (0 = absent, 1 = present), and the interaction of perfection and categorization cue. We observed a significant effect of perfection ($B = 1.07$,

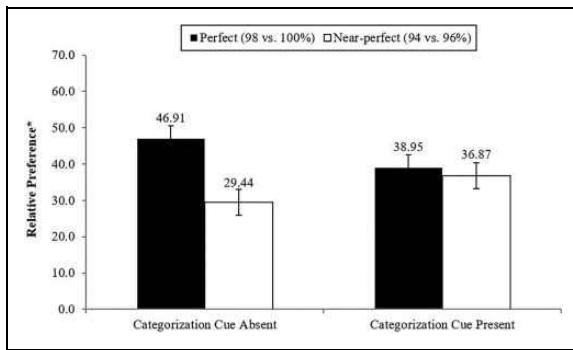


Figure 2. Relative preference for higher quality but more expensive socks (Study 3). *0 = strongly prefer the less expensive, but lower wool content option, 100 = strongly prefer the more expensive, but higher wool content option.

$SE = 0.37$, $Wald = 8.44$, $p = .004$), but no effect of categorization cue ($B = 0.25$, $SE = 0.39$, $Wald = 0.42$, $p = .52$). More germane to our theorizing, we obtained a marginally significant interaction between perfection and categorization cue ($B = -0.87$, $SE = 0.52$, $Wald = 2.86$, $p = .091$), which suggests that the categorization cue weakened the perfection premium.

In the absence of a common categorization cue, a χ^2 test of independence returned a significant effect of perfection condition on choice share, $\chi^2(1, N = 158) = 8.75$, $p = .003$, $\phi = .24$. The more expensive but higher quality socks were chosen by 40.5% of participants in the perfect condition versus 19.0% of participants in the near-perfect condition. However, in the presence of a categorization cue, the effect of perfection condition on choice share was nonsignificant, $\chi^2(1, N = 164) = 0.29$, $p = .59$, $\phi = .04$. The more expensive, higher quality socks were chosen by 26.8% of participants in the perfect condition versus 23.2% of participants in the near-perfect condition.

We also conducted a 2 (perfection: perfect, near-perfect) \times 2 (categorization cue: present, absent) between-participants ANOVA on relative sock preference. There was a significant main effect of perfection, $F(1, 318) = 7.37$, $p = .007$, $\eta_p^2 = .02$, but no main effect of categorization cue, $F(1, 318) = .01$, $p = .94$, $\eta_p^2 < .001$. Importantly, this main effect was qualified by a significant interaction between perfection and categorization cue, $F(1, 318) = 4.56$, $p = .033$, $\eta_p^2 = .014$. Replicating the perfection premium, planned contrasts revealed that when the categorization cue was absent, participants in the perfect condition ($M_{100/98} = 46.91$, $SD = 33.92$, $N = 79$) exhibited a stronger relative preference for the superior socks compared to participants in the near-perfect condition ($M_{96/94} = 29.44$, $SD = 30.39$, $N = 79$), $F(1, 318) = 11.54$, $p = .001$, $\eta_p^2 = .035$. However, when the common categorization cue was present, the relative preference of participants in the perfect condition ($M_{100/98} = 38.95$, $SD = 33.57$, $N = 82$) versus participants in the near-perfect condition ($M_{96/94} = 36.87$, $SD = 31.22$, $N = 82$) did not differ, $F(1, 318) = .17$, $p = .68$, $\eta_p^2 = .001$. Figure 2 shows these results.

Table 1. Proportion of Participants Placing Each Student in a Singleton Category (Study 4).

Condition	Student Q (Lowest Score, %)	Student R (Middle Score, %)	Student S (Highest Score, %)	N
Perfect condition				
Round perfect value (i.e., 100)	34.9	3.6	61.4	83
Nonround perfect value (i.e., 88)	30.6	5.9	63.5	85
Near-perfect condition				
Round perfect value (i.e., 100)	54.1	2.4	43.5	85
Nonround perfect value (i.e., 88)	58.8	9.4	31.8	85

Study 4

In Study 4, we attempt to provide more direct evidence for our mechanism using categorization itself as a dependent variable in a social cognition context. We predict that participants should be more likely to place an individual associated with a perfect numerical value in a separate category from those with near-perfect numerical values.

Method

Study 4 was conducted on MTurk with 338 U.S. participants (39.6% female, $M_{age} = 35.13$). Participants were randomly assigned to one of four conditions in a 2 (perfection: perfect, near-perfect) \times 2 (perfect value: round, nonround) between-participants design.

Participants were informed that three students (i.e., Students Q, R, and S) had taken a test in which a perfect score was either 100 points (round perfect value condition) or 88 points (non-round perfect value condition). Those in the perfect condition learned that the three students had earned 98, 99, and 100 points (out of 100) or 86, 87, and 88 points (out of 88). In contrast, those in the near-perfect condition learned that the three students had earned 97, 98, and 99 points (out of 100) or 85, 86, and 87 points (out of 88).

Subsequently, we instructed participants to sort the students into two groups (with at least one student in each group) based on any criteria of their choosing. Sorting was done using the cursor to drag and drop labels of each student (i.e., Q, R, and S) into two boxes that were named "Group A" and "Group B." Participants were not allowed to proceed until they had placed all three students into one of the two groups.

Results

Following the instructions, all participants placed one student into a singleton category (i.e., a category of its own; Brenner et al., 1999) and the other two students into a shared category. Table 1 shows the proportion of participants in each condition

who placed Student Q, R, or S into a singleton category. Unsurprisingly, the student who received the middle score (i.e., Student R) was always least likely to be in a singleton category. However, across conditions, the placement of either the lowest scoring student (Student Q) or the highest scoring student (Student S) in a singleton category varied.

We predicted that participants would be more likely to place the highest scoring student into a singleton category if this student earned a perfect (vs. near-perfect) score. Supporting this prediction, a binary logistic regression showed a significant main effect of perfection ($B = 1.32$, $SE = 0.32$, Wald = 16.58, $p < .001$) on participants' likelihood of placing Student S into a singleton category. Among participants in the perfect conditions, 62.5% placed Student S into a singleton category. In contrast, only 37.6% of participants in the near-perfect conditions placed Student S into a singleton category, $\chi^2(1, N = 338) = 20.88$; $p < .001$, $\phi = .25$. Neither the main effect of the perfect value condition (i.e., round vs. nonround) nor the interaction effect was significant ($ps > .11$).

General Discussion

This research shows that people disproportionately value individuals and items when they are perfect on certain attributes versus near-perfect on those same attributes. A single-paper meta-analysis validated the robustness of the perfection premium (estimate = 3.21, $SE = 1.22$, $z = 2.64$, $p < .01$; McShane & Böckenholt, 2017). Further, we find that this perfection premium occurs because individuals spontaneously categorize perfect items separately from near-perfect items. We provide evidence for the perfection premium in both decision-making (Studies 1 and 3) and social cognition (Studies 2 and 4) contexts.

Demonstrating the perfection premium in interpersonal evaluations is particularly valuable because it helps rule out the alternative explanations of physical contamination (Rozin & Nemeroff, 1990) and certainty (Kahneman & Tversky, 1979). According to a physical contamination account, adding *any* amount of a nonnatural entity destroys perceptions of an item's naturalness. Although a physical contamination explanation could potentially explain a perfection premium in quality judgments (Studies 1 and 3), it cannot explain the perfection premium obtained in judgments of intelligence (Study 2) since physical contamination concerns do not apply to this interpersonal context. Similarly, although perfect options may sometimes be viewed as less risky, it is unlikely that students whose test scores are 1 point apart (Study 2) have disparate risk profiles. Thus, our proposed categorization process explains the results of our four studies more parsimoniously than the certainty effect.

This article contributes to four streams of literature. First, we add to the literature on categorization by showing that perfect items are spontaneously placed into separate mental categories from other items. This work complements research showing that quantitative judgments sometimes lead participants to group options into qualitative categories (e.g., calorie

estimates of food affect grouping as vices or virtues; Chernev & Gal, 2010). We document a novel basis for categorization (i.e., perfection vs. near-perfection) that extends the previously established good/bad dichotomy (Rozin et al., 1996).

Second, our work relates to research on authenticity and psychological essentialism. Newman (2016) argued that items are categorized based on their underlying attributes, and these categorizations drive demand. For example, consumers categorize a tape measure owned by President Kennedy (vs. a generic tape measure) as authentic and therefore value it more (Newman et al., 2011). Similarly, we find that people categorize perfect items separately from near-perfect items and thus value them to a greater extent.

Third, we add to the well-established literature on prospect theory and decision making, which has shown nonlinearity in probability judgments and discontinuities in distributed weights near the endpoints of a scale (Shampanier et al., 2007; Tversky & Kahneman, 1992). Whereas prior research has primarily focused on risk perceptions (e.g., Li & Chapman, 2009), we find that even in relatively low-risk choice contexts or when judging personality traits, people disproportionately value items at the maximum boundary of a scale (i.e., perfection) even though they are not less risky than near-perfect items. Furthermore, this research provides the first demonstration that categorization plays a critical role in causing endpoint discontinuities.

Finally, we demonstrate that biases in numerical cognition can exert influence on social cognition judgments. Prior research has shown that numerical values associated with an individual can affect human motivation. For example, high school students are particularly motivated to retake the Scholastic Aptitude Test if their initial score is just short of a round number (Pope & Simonsohn, 2011). We show that perfect versus near-perfect numerical values associated with an individual can also disproportionately affect the inferences that others make about the individual's personality (e.g., intelligence). There is considerable practical and theoretical benefit in understanding when and how numerical values affect judgments about other human beings. For instance, Olenski and Colleagues (2020) recently found that physicians were more likely to perform coronary-artery bypass grafting on patients admitted in the 2 weeks prior to their 80th birthday as compared to patients admitted in the 2 weeks after their 80th birthday. This difference may arise because physicians more readily apply labels such as "fragile" or "elderly" to patients they have placed into the "over 80" category. We encourage future research to explore how the assignment of numerical values to specific individuals affects judgments about their personality or well-being (e.g., extraversion, healthiness). From a theoretical perspective, marrying numerical cognition with social cognition can help disentangle multiple potential explanations for an effect, as was the case in the present research.

Acknowledgments

The first author gratefully acknowledges organizers and attendees of the 2017 IDEA Conference for helping develop this research idea.

Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Mathew S. Isaac  <https://orcid.org/0000-0002-3182-1341>

Katie Spangenberg  <https://orcid.org/0000-0003-3443-9225>

Supplemental Material

The supplemental material is available in the online version of the article.

Notes

1. The posttest was conducted with 119 participants recruited on Amazon Mechanical Turk (41.2% female, $M_{\text{age}} = 38.99$ years). Participants evaluated the extent to which they considered the perfect and near-perfect socks to be part of the same category or in two separate categories on a 100-point scale (0 = *one category*, 100 = *two categories*). They were more likely to evaluate the socks as part of the same category when the categorization cue was present ($M = 53.77$, $SD = 33.79$) versus absent ($M = 56.61$, $SD = 33.63$), $t(118) = -2.08$, $p = .040$.
2. Participants indicated how believable they found the wool content claims (1 = *extremely unbelievable*, 7 = *extremely believable*). Believability ratings did not differ significantly by condition and are therefore not discussed any further.

References

- Bizer, G. Y., & Schindler, R. M. (2005). Direct evidence of ending-digit drop-off in price information processing. *Psychology & Marketing*, *22*, 771–783.
- Brenner, L., Rottenstreich, Y., & Sood, S. (1999). Comparison, grouping, and preference. *Psychological Science*, *10*, 225–229.
- Camerer, C. F., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, *8*(2), 167–196.
- Chernev, A., & Gal, D. (2010). Categorization effects in value judgments: Averaging bias in evaluating combinations of vices and virtues. *Journal of Marketing Research*, *47*, 738–747.
- Gunasti, K., & Ross, W. T. Jr. (2010). How and when alphanumeric brand names affect consumer preferences. *Journal of Marketing Research*, *47*, 1177–1192.
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis*. Guilford Press.
- Hinrichs, J. V., Berie, J. L., & Mosell, M. K. (1982). Place information in multidigit number comparison. *Memory & Cognition*, *10*, 487–495.
- Hunt, L. L., Eastwick, P. W., & Finkel, E. J. (2015). Leveling the playing field: Longer acquaintance predicts reduced assortative mating on attractiveness. *Psychological Science*, *26*, 1046–1053.
- Isaac, M. S., & Schindler, R. M. (2014). The top-ten effect: Consumers' subjective categorization of ranked lists. *Journal of Consumer Research*, *40*, 1181–1202.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
- King, D., & Janiszewski, C. (2011). The sources and consequences of the fluent processing of numbers. *Journal of Marketing Research*, *48*, 327–341.
- Kruger, J., & Vargas, P. (2008). Consumer confusion of percent differences. *Journal of Consumer Psychology*, *18* (1), 49–61.
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, *78*, 1723–1743.
- Li, M., & Chapman, G. B. (2009). "100% of anything looks good": The appeal of one hundred percent. *Psychonomic Bulletin & Review*, *16*, 156–162.
- McShane, B. B., & Böckenholt, U. (2017). Single-paper meta-analysis: Benefits for study summary, theory testing, and replicability. *Journal of Consumer Research*, *43*(6), 1048–1063.
- Newman, G. E. (2016). An essentialist account of authenticity. *Journal of Cognition and Culture*, *16*, 294–321.
- Newman, G. E., Diesendruck, G., & Bloom, P. (2011). Celebrity contagion and the value of objects. *Journal of Consumer Research*, *38*, 215–228.
- Olenksi, A. R., Zimmerman, A., Coussens, S., & Jena, A. B. (2020). Behavioral heuristics in coronary-artery bypass graft surgery. *New England Journal of Medicine*, *382*(8), 778–779.
- Olsen, A. L. (2013). Leftmost-digit-bias in an enumerated public sector? An experiment on citizens' judgment of performance information. *Judgment and Decision Making*, *8*(3), 365–371.
- Pope, D., & Simonsohn, U. (2011). Round numbers as goals: Evidence from baseball, SAT takers, and the lab. *Psychological Science*, *22*(1), 71–79.
- Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*, 274–278.
- Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, *7*, 532–547.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum Associates.
- Rozin, P., Ashmore, M., & Markwith, M. (1996). Lay American conceptions of nutrition: Dose insensitivity, categorical thinking, contagion, and the monotonic mind. *Health Psychology*, *15*, 438–447.
- Rozin, P., & Nemeroff, C. (1990). The laws of sympathetic magic: A psychological analysis of similarity and contagion. In J. Stigler, G. Herdy, & R. A. Shweder (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 205–223). Cambridge University Press.
- Shampanier, K., Mazar, N., & Ariely, D. (2007). Zero as a special price: The true value of free products. *Marketing Science*, *26*, 742–757.
- Shepard, R. N., Kilpatrick, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology*, *7*, 82–138.

- Strauss, V. (2019). 17 Students from one high school get perfect scores on the ACT. *The Washington Post*. WP Company. Retrieved April 25, 2019 from www.washingtonpost.com/education/2019/04/25/students-one-high-school-get-perfect-scores-act/
- Tu, Y., & Soman, D. (2014). The categorization of time and its impact on task initiation. *Journal of Consumer Research*, *41*, 810–822.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
- Valdes-Dapena, Peter. Tesla gets near-perfect score from consumer reports. *CNNMoney*. Cable News Network. Retrieved May 9, 2013 from www.money.cnn.com/2013/05/09/autos/tesla-model-s-consumer-reports/index.html
- Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, *42*, 1676–1690.

Author Biographies

Mathew S. Isaac is the Genevieve Albers Professor of marketing at the Albers School of Business and Economics, Seattle University, Seattle, WA. His research on consumer judgment and decision making, which examines how contextual and motivational factors influence product evaluations and purchase intentions, has been published in a number of leading scientific journals including *Journal of Consumer Research*, *Journal of Marketing Research*, *Journal of Marketing*, and *Journal of Consumer Psychology*.

Katie Spangenberg is a doctoral candidate in marketing at the University of Washington Foster School of Business, Seattle, WA. Her research on implicit cognition, self-identity, brand perception, and advertising has been published in journals such as *Journal of Advertising*, and presented at top marketing and psychology conferences.

Handling Editor: Eva Walther