Time Periods Feel Longer When They Span More Category Boundaries:

Evidence From the Lab and the Field

Kristin Donnelly[1]

Giovanni Compiani[2]

Ellen R. K. Evers[1]

[1]Haas School of Business, University of California, Berkeley, [2]Booth School of Business, University of Chicago

Manuscript is under review at *Journal of Marketing Research*

Time Periods Feel Longer When They Span More Category Boundaries:

Evidence From the Lab and the Field

Seven experiments (total *N* = 3,585) and a large field dataset (*N* = 1,820,671) demonstrate that time periods of equal duration are not always perceived as equivalent. We find that periods feel longer when they span more time categories (e.g., hour, month). For example, periods like 1:45pm – 3:15pm and March 31st – April 6th (boundary-expanded) feel longer than, say, 1:15pm – 2:45pm and April 2nd – April 8th (boundary-compressed). Reflecting this, participants anticipated completing more work during boundary-expanded periods than equivalent boundary-compressed periods. This effect appears to result from the salience and placement of time boundaries. As a consequence, participants preferred scheduling pleasant activities for boundary-expanded and unpleasant activities for boundary-compressed periods. Moreover, participants were willing to pay more to avoid—and required more money to endure—a long wait when it was presented as boundary-expanded. Finally, data from over 1.8 million rideshare trips suggest that consumers are more likely to choose independent rides (e.g., UberX) when they are boundary-compressed when the alternative shared option (e.g., UberPool) is boundary-expanded. Together, our studies reveal that time periods feel longer when they span more boundaries, and that this phenomenon shapes consumers' scheduling and purchasing decisions.

Keywords: time perception, scheduling, categories, estimation, field data

In this manuscript, we investigate how the categorization of time affects consumer judgments. We assert that just as they do in other domains, consumers naturally categorize time. For example, the “8” in 8:30am provides a category, as does “March” or the year “2021”. Existing research supports this general idea; for example, people separate the present and future into coarse categories of “like the present” and “not like the present” (Tu and Soman 2014) and organize the narrative of their lives into chapters (Skowronski et al. 2007; Thomsen 2009). Often, consumers categorize time around salient landmarks, such as the beginning of a new age decade (Alter & Hershfield 2014), or new year or month (Dai, Milkman and Riss 2014, Ayers et al. 2014).

Categorization, the process of mentally grouping things and concepts, is fundamental to the human experience (e.g., Smith and Medin 1981; Tversky and Hemenway 1984). Categorization facilitates reasoning and communication (Rosch 1999; Smith and Medin 1981) by organizing stimuli in hierarchical groups defined by salient characteristics (Rosch and Mervis 1975). For example, an iPhone is a member of the category “smartphones” (phones with internet access), which itself is nested in the category “mobile phones”, which is a subcategory of “electronics”, and so on. Categories can describe physical entities (e.g., “apple”, “desk”), their properties (e.g., “green”, “medium-rare”), as well as abstract, non-physical concepts like credit scores (e.g., “good”, “very good”) and hotel ratings (e.g., “3-star”).

If time perception is indeed subject to categorization, we would expect that it should display two core categorization features that lead to a set of novel predictions tested in the present paper. The first feature involves assimilation and contrast: Categories minimize differences within the category and maximize differences across them (e.g., Eiser and Stroebe 1972). In this vein, stimuli in the same category are often judged as more similar than those in different categories, even if category membership is entirely arbitrary (Tajfel 1959, 1969). This leads us to predict the following:

H1: Points in time that occur on different sides of a temporal category boundary (e.g., the start of a new month, hour, or year) are perceived to be more distant from each other than those that do not fall into different temporal categories.

To illustrate, we would expect 2:30 to feel further from 3:00 than 3:00 feels from 3:30, and March 31st to feel further from April 1st than March 29th feels to March 30th.

A large body of research investigates the effects of categorization in non-temporal domains, and consistent with this prediction, finds that the relationship between the exact same stimuli can be judged differently depending on their perceived category membership. For example, people estimate two locations in different states to be further apart than equidistant locations in the same state (Burris and Branscombe 2005; Irmak, Naylor and Bearden 2011; Maki 1982). Consumers automatically categorize items in a ranked list (e.g., into the “top 10”, “top 20”), making a 10th ranked restaurant feel closer in quality to the 8th ranked restaurant than to the 12th (Isaac and Schindler 2014). Temperature changes between two consecutive days seem larger when they fall in different months instead of the same month (Krueger and Clement 1994), the length difference between two lines appears greater when the lines are given different labels (Tajfel and Wilkes 1963), and people seem more dissimilar when arbitrarily assigned to different social groups (Allen and Wilder 1979; Locksley, Ortiz and Hepburn 1980). Even colors appear more or less similar depending on their categorization. Because their language affords distinct categories for light and dark blue, Russian speakers discriminate between such shades faster than English speakers do (Winawer et al. 2007; see also Roberson, Davies and Davidoff 2000). In sum, stimuli across category boundaries are perceived to differ more than stimuli that fall within the same boundaries.

In the case of time, what provides salient category boundaries? We turn now to the second core feature of categorization—hierarchical organization. Events and stimuli are organized into categories based on the most salient difference first, and then subcategorized by smaller differences (e.g., Dougherty 1978). Membership in the most general categories is determined by widely shared, basic qualities (e.g., animal vs. fungi). These categories contain sub-categories (e.g., mammal, amphibian) that in turn subsume others (e.g., cat, dog), each defined by increasingly smaller differences between its members and non-members (e.g., animals differ more from fungi than cats do from dogs; Berlin, Breedlove and Raven 1973).

For time, we would expect this hierarchy to follow a natural size order wherein decades differ more from each another than do years, years differ more than months, months more than days, and so on. As a consequence of this hierarchy, the same differences can appear more or less meaningful depending on the salience of other differences (Dougherty 1978; Gati and Tversky 1984; Tversky 1977). For example, when considering a time period that starts and ends in different months, but the same year, the month category provides the most salient (and useful) way to differentiate start and end time. However, it may be subsumed under the broader “year” category if that period starts and ends in different years.

What does this mean for the perception of time? According to H1, we expect that people perceive, say, 3:09 to 4:02 to last longer than the equivalent 3:03 to 3:56, because the former period crosses a salient category boundary (4:00pm) and the latter does not. This prediction alone is similar to that of Left-Digit Bias, a highly robust phenomenon in which consumers overweight the leftmost unit in price perception. For example, because of Left-Digit Bias, the difference between \$1.98 and \$1.99 feels smaller than the difference between \$1.99 and \$2.00 (Manning and Sprott 2009; Thomas and Morwitz 2009). Applied to time, Left-Digit Bias would also predict, for instance, that 1:58 feels closer to 1:59 than 1:59 feels to 2:00, and 1:30 – 3:00 to feel longer than 1:00 – 2:30.

However, because categories are hierarchical, we would expect that the effect proposed in H1 does not always involve the leftmost digit. Rather, judgments should involve the *largest unit of measurement*—which we conceptualize as the highest-order category—that differentiates start and end time. Consumers may compare prices by their leftmost digits because for prices, the largest unit of measurement *is* leftmost. But when comparing start and end times, the largest unit is sometimes leftmost (e.g., hour), but it can also be rightmost (e.g., year) or in the middle (e.g., month, for non-Americans). Thus, although left-digit bias in pricing is consistent with our theoretical model based on categorization, crucially, our reasoning predicts when we would expect the bias in time perception to involve the left digit, and when we would expect it to occur on other digits. This leads to H2:

H2: The effect of boundary expansiveness results from perceiving differences in the *largest salient unit* that differentiates start and end time.

This means that it is the category information itself—and not the position of that information—that influences judgments. For instance, if a period spans more distinct years than another period of equal duration, we expect it to feel longer, even though year occupies the rightmost position. As an example, a 10-day period should feel longer when it starts in December 2021 and ends in January 2022 rather than, say, starting in January 2022 and ending in February 2022. H2 also assumes a hierarchy. When judging the duration of a time period, consumers consider the largest unit of measurement that *differs* between the start and end time; thus, if the largest unit does not differ, consumers will attend to the largest unit that does.

Importantly, the previous hypotheses involve objective categories (e.g., hour, month), but some temporal categories are defined by events, activities, and experience. For example, if a student has hour-long classes at fifteen minutes past the hour, the most salient boundary will not be the start of a new hour (e.g., 10:00), but the start of a new class (e.g., 10:15). This is because the experiential difference between the end of one class and the start of another is far greater than the experiential difference between 9:59 and 10:00. That is, class is a more diagnostic feature (or a better differentiator) than is hour (e.g., Tversky and Gati 1978). This leads to H3:

H3: The effect proposed in H1 involves the most salient category boundaries. Introducing categories that are more salient than objective ones (e.g., hour, month) will make time periods spanning more of them feel longer.

In sum, we predict that two points in time will feel further apart when they span more distinct temporal categories, such as hour, month, or year (H1). This effect involves attending to the most differentiating categories, not necessarily the leftmost digits (H2). Lastly, categories based on experiences (e.g., events) will show the same boundary-expansiveness effect observed for objective time categories (H3).

The present paper joins a handful of others that study how temporal landmarks affect perceived duration (e.g., Zauberman, Levav, Diehl and Bhargrave 2010; Tonietto, Malkoc and Nowlis 2019), but it examines this topic through an entirely new lens (and with a novel set of "landmarks"). Specifically, we conceptualize objective temporal markers (e.g., 4:00pm, June 1st, etc.) as boundaries that define categories (e.g., "4pm" and "June" categories), and then rely on established principles of categorization to develop and test various predictions. This approach distinguishes the present work from related phenomenon (i.e., Left-Digit Bias), expands our understanding of time perception, and extends the ever-growing umbrella of judgments influenced by categorization.

We are unaware of any previous research that investigates how the number of distinct categories spanned by a given period (e.g., hours, months) affect its perceived duration, but this question is important for consumer decision-making. Consumers routinely choose among time periods when booking or purchasing services and experiences in advance, and online scheduling platforms make such decisions increasingly common (Blaszkiewicz 2018). If the exact same time period can seem longer or shorter depending on how many category boundaries it spans, "boundary-expansiveness" may affect scheduling preferences. For instance, if a long waiting period spans more category boundaries, it should feel longer and thus be less appealing than an equivalent wait time that spans fewer. Similarly, consumers may prefer to schedule pleasant or unpleasant activities during periods that span more or fewer category boundaries, respectively. The present research addresses these possibilities, and more broadly, provides valuable insight into how consumers perceive time, value time, and structure future experiences, which in turn may help companies improve the design of their platforms, apps, and service-based offerings.

Overview of Studies

Study 1 tests the basic premise of our model: Time periods that span more category boundaries are perceived to last longer than equivalent periods spanning fewer; we call such periods "boundary-expanded" and "boundary-compressed", respectively. Consistent with this idea, study 1 finds that workers estimate that they can complete more tasks when the exact same time period is boundary-expanded rather than compressed. Studies 2a and 2b test H2 and reveal that the effect is not due to perceiving changes in the leftmost digit, but in the most differentiating unit. Since our theory assumes that the effect relies on *perceived* boundaries, changing the perception of those boundaries should affect judgments. Study 3 creates alternative categories and finds that when they are made salient, those category boundaries act in the same way as hour- and year-markers did in Studies 1 and 2.

The next three studies focus on the behavioral consequences of boundary expansiveness. Study 4 finds that participants prefer to schedule aversive experiences during boundary-compressed time periods (that feel shorter), but prefer to schedule enjoyable experiences during boundary-expanded time periods (that feel longer). Study 5 manipulates the boundary-expansiveness of long wait times and elicits consumers' willingness to pay to avoid the long wait as well as the compensation they require to endure it. In the final study, we examine archival transportation data from Chicago to study the real-world impact of boundary-expansiveness on consumer behavior. This collection of over 1.8 million rideshare trips shows that consumers are more likely to choose the independent ride when it is boundary-compressed, and the shared ride option is boundary-expanded, compared to when both rides are boundary-compressed or expanded.

The design, hypotheses, sample size and analyses of all experimental studies reported in the paper were pre-registered. The analysis plan and treatment of the rideshare dataset were pre-specified. For all studies, we report all data exclusions, all manipulations, and all measures. We note here that analyzing the data without exclusions does not meaningfully affect the results of any study; those analyses are presented in the online appendix. In each experiment, trials were presented in random order, and for those involving choosing between time periods, the visual position of options was also randomized. Lastly, because the expression of time varies in different countries, we recruited participants from the U.S. in every study except 2a (which intentionally sampled from the UK). Our pre-registrations, materials, data and code are available at https://osf.io/dav53/?view_only=71e175b98e024cf7a321d07ac4ea5d24.

Before discussing study 1, we briefly describe findings from an initial set of pilot studies in which participants indicated how long different time periods felt (studies W1 – W7 in the online appendix). Those studies found that boundary-expanded periods (e.g., 3:30pm – 5:00pm) were selected to "feel longer" than equivalent boundary-compressed ones (e.g., 3:00pm – 4:30pm), even after participants had to calculate each period's duration. They also rated boundary-expanded periods to "feel longer", even when rounding to the nearest hour would make a boundary-expanded period seem shorter than a boundary-compressed one. This effect held when periods spanned months (e.g., January 30th – February 3rd feels longer than January 27th – January 30th), when periods were written out and referenced the same hour categories (e.g., "quarter to two to four o'clock" felt longer than "quarter past two to half past four"), or when periods were displayed on clocks. Full details are available in the online appendix.

STUDY 1: ESTABLISHING THE EFFECT OF BOUNDARY EXPANSIVENESS

Building from our pilot findings, study 1 tests the basic effect of boundary-expansiveness using a behaviorally valid measure. Specifically, we asked MTurkers to predict numbers of Human Intelligence Tasks (HITs) they expect to be able to complete during various time periods. We used a between-subject design; all periods were either expanded or compressed. Per H1, we expected that MTurkers would estimate being able to complete more HITs during boundary-expanded periods compared to boundary-compressed periods of equal length.

Method and Procedure

Our participants were 612 MTurk workers (53.8% female, 44.9% male, 1.8% other, M_{age} = 37). As preregistered, we excluded participants that failed our attention check and/or whose estimated number of HITs for a final attention-check period exceeded their estimates for any of the other periods (which were substantially longer); this left a final sample of 576.

Participants were told to imagine that they had a day completely free to do MTurk work. In each of five trials, they estimated how many HITs they could accomplish in a given period, answering on a slider from “0” to “500 or more.” Following our pre-registration, responses at either of these end points were eliminated. Boundary-expansiveness was manipulated between subject: Participants were randomly assigned to view either expanded or compressed periods. Figure 1 displays these periods and the results.

(Insert Figure 1 about here)

Results and Discussion

A mixed-effects negative binomial model regressed number of HITs on boundary condition (expanded vs. compressed) and specified random effects of question and participant. We opted to use negative binomial regression because it is far less susceptible to producing false positives than Poisson regression (Garner, Mulvey and Shaw 1995; Ryan, Evers and Moore 2018). As predicted, participants considering boundary-expanded periods estimated that they could perform more HITs (M = 75.13, SD = 84.37) than those considering equivalent boundary-compressed periods (M = 58.23, SD = 71.44; z = 3.16, p = .002, d = .22).

Along with the pilot studies, the results of study 1 support H1: Time periods feel longer when they span more category boundaries. Consistent with this hypothesis, MTurk workers anticipated being able to complete about 29% more HITs during time periods that were boundary-expanded rather than boundary-compressed, a difference of roughly 17 HITs. Before looking at the consequences for consumer decisions, we first investigate the underlying process.

STUDIES 2A AND 2B: HIERARCHICAL BOUNDARIES

Based on categories being hierarchical, we expected that the effect observed in the previous studies is a consequence of hours being the most differentiating unit (rather than a consequence of left-digit bias specifically). The next two studies explore whether this is true by varying the position of the differentiating units. To do so, study 2a employs situations in which the differentiating units occur in the middle of the expression rather than on the left. Specifically, Study 2a employs months as boundaries and uses participants in the U.K., where the Day-Month-Year format puts month in the middle position (rather than the leftmost). Study 2b tests whether a period's boundary-expansiveness in terms of months affects judgments when a higher-order category (year) does or does not differentiate start and end time. As a test of generalizability, study 2b also varies the temporal location of the periods, placing some in the past and others in the future.

Study 2a

Participants and Procedure. Participants were 203 Prolific Academic workers (56.2% female, 39.9% male, 3.9% did not answer, M_{age} = 37) recruited from the United Kingdom. As before, participants were excluded if their rating of a much shorter attention-check period exceeded their rating for any other period, leaving a final sample of 152.

Participants evaluated various time periods for "how long does this feel?" on a 100-point scale from 0 (doesn't feel long at all) to 100 (feels extremely long). Each of five different durations were expressed in two forms: once as a boundary-expanded period (e.g., October 30th to December 10th), and once as boundary-compressed (spanning fewer month categories, e.g., November 1st to December 11th). Thus, participants provided ratings for a total of 10 periods. Importantly, adopting the format used in the UK (and most countries outside of the U.S.), dates were expressed as Day-Month-Year (see Figure 2 for these periods and the results).

(Insert Figure 2 about here)

Results and Discussion. A mixed-effects model regressed ratings of perceived length on boundary type (expanded vs. compressed), specifying random effects of participant and question. As predicted, boundary-expanded periods (M = 30.80, SD = 23.38) were rated to feel longer than compressed (M = 25.09, SD = 20.94; t = 9.00, p < .001, d = .26). These results demonstrate that judgments are influenced by the most differentiating unit being boundary-expanded or compressed. This also suggests that what is known as the left-digit effect in pricing may actually be a consequence of a more general phenomenon—the effect of boundary-expansiveness on most differentiating units, which, in the case of prices, happen to be leftmost. In the next study, instead of relying on cross-cultural differences in the expression of dates, we examine the effect of differentiating units by manipulating whether same units are or are not the most differentiating.

Study 2b

Study 2b tests the hierarchical assumption discussed earlier—that consumers judge duration according to the largest unit of measurement (or highest-level category) that differentiates a period's start and end time. Specifically, we examine how the month category affects judgments when the higher-order year category does or does not differentiate start and end time. Study 2b also tests generalizability by manipulating the temporal location of the periods (past vs. future).

Participants and Procedure. Participants were 257 Prolific Academic workers (45.1% female, 53.7% male, 1.2% did not answer, M_{age} = 36). Due to a coding error, no attention check was employed in this study.

Participants rated the perceived duration of time periods on a 100-point slider. This study involved four periods, all with the same duration, each occurring in different times of the year. We manipulated temporal location within-subject. Two periods were randomly selected to occur in the past, and the other two, in the future—the year in which the period started was selected at random from the past 20 or the next 20 years, respectively. We manipulated boundary-expansiveness in terms of months: Each period was presented once as boundary-expanded (spanning more distinct months) and once as boundary-compressed (spanning fewer) for a total of 8 judgments.

For half of the participants, the time periods always started and ended in the same year (e.g., 10-20-2021 to 11-30-2021). For the other half, periods were exactly one year longer, starting and ended in different (adjacent) years (e.g., 10-20-2021 to 11-30-2022). Thus, this study used a 2(boundary: expanded vs. compressed) x 2(temporal location: past vs. future) x 2(start and end year: same vs. adjacent) mixed design.

Results. A mixed-effects ANOVA regressed ratings of perceived length on boundary type (expanded vs. compressed), temporal location (past vs. future), and start and end year (same vs. different), allowing for the interaction of these three factors as well as random effects of participant and question. Unsurprisingly, participants assigned to see periods that span adjacent years rated them as longer (*M* = 48.50, *SD* = 25.27) than participants evaluating periods that started and ended in the same year (*M* = 26.84, *SD* = 23.52, *F*(1,255.43) = 68.67, *p* < .001, *d* = .89). Replicating our basic effect, boundary-expanded periods (*M* = 39.48, *SD* = 26.34) felt longer than boundary-compressed (*M* = 35.39, *SD* = 26.87; *F*(1,1779.94) = 48.94, *p* < .001, *d* = .15). We also observed a significant effect of temporal location, such that periods in the future (*M* = 39.04, *SD* = 27.30) were rated to feel longer than those in the past (*M* = 35.65, *SD* = 25.86; *F*(1,1787.09) = 36.11, *p* < .001, *d* = .13. But the effect of boundary-expansiveness on perceived duration did not differ by temporal location—future and past periods showed the effect to the same degree, *F*(1,1779.94) = 0.616, *p* = .433.

However, our primary term of interest was the interaction between boundary-type and year, which emerged as the only significant interaction in the model, *F*(1,1779.94) = 11.08, *p* < .001. The results reflect the hierarchy we anticipated. For participants who saw periods that always started and ended in the same year, boundary-expanded periods (*M* = 29.85, *SD* = 23.61) felt longer than compressed (*M* = 23.85, *SD* = 23.06; *t* = 7.38, *p* < .001). But for participants who saw periods that started and ended in different (adjacent) years, this difference was only marginally significant (*M*s = 49.56 and 47.45, *SD*s = 25.27, 25.23; *t* = 2.57, *p* = .051).

Discussion. Based on the hierarchical nature of categories, we predicted that the effect of boundary expansiveness depends on whether it occurs on the most differentiating unit. Consistent with this prediction, we found that the effect of boundary expansiveness on months strongly attenuated when the time periods spanned multiple years, and months thus were no longer the most differentiating factor. That is, when the higher-order year category did not differentiate start and end time, consumers attended to the month category. Conversely, they relied *less* on month category differences when every period spanned two adjacent years.

Combined, studies 2a and 2b provide support for our second hypothesis (H2), which posits that the effect of boundary-expansiveness results from consumers perceiving differences in the highest-order category that differentiates start and end time.

STUDY 3: MANIPULATING CATEGORIES

The previous studies rely on the natural variation of time to make equivalent time periods span more or fewer category boundaries. Our theoretical model predicts that boundary expansiveness mostly affects judgments when it involves the most salient categories. In study 2b we tested this by varying the degree to which months were more or less differentiating (H2). We next turn to testing this prediction by creating new, salient categories (H3).

According to our third hypothesis, we should be able to evoke different kinds of time categories—beyond the "natural" ones provided by, say, hour, month or year—and observe inflation of periods that span more of them. To test this, study 3 prompted participants to evaluate a period that spanned different classes in a student's schedule. That is, the period was either boundary-compressed (spanning Classes B and C) or expanded (spanning Classes A, B and C) in terms of classes. Orthogonally, we manipulated the class schedule such that the period was either boundary-expanded or boundary-compressed in terms of hours. Because this study presents class categories as more salient than hour categories, we predict that the time period feels longer when it spans more classes, regardless of the number of hour categories it spans.

Participants and procedure. Sixteen hundred ten MTurk workers participated; as per our pre-registration, none were excluded. Demographics were not collected for studies 3 or 4.

This study employed a 2(number of class boundaries: 2 vs. 3) x 2(number of hour boundaries: 2 vs. 3) between-subjects design. Participants saw a hypothetical class schedule for three 60-minute classes (“Class A,” “Class B” and “Class C”). Half saw a schedule in which classes started every hour on the hour:

Class A: 9:00 - 10:00
Class B: 10:00 - 11:00
Class C: 11:00 - 12:00

For the other half of participants, a new class started every hour on the *half hour*:

Class A: 9:30 - 10:30
Class B: 10:30 - 11:30
Class C: 11:30 - 12:30

Participants were then asked to consider a 1 hour, 40-minute period, manipulated between-subject to span either all three classes or only two. Those assigned to a *3 classes* condition read:

> “Shelly charged her cell phone for the last thirty minutes of Class A, all sixty minutes of Class B, and the first ten minutes of Class C.”

Participants in the *2 classes* condition saw a slightly different version:

> “Shelly charged her cell phone for all sixty minutes of Class A and the first forty minutes of Class B.”

Just as before, participants rated the period for “how long does it feel?” on a 100-point slider.

Note that for participants presented with the schedule that has classes start on the hour (i.e., Class A at 9:00, Class B at 10:00, and Class C at 11:00), the period under consideration always spanned the same number of hours as it did classes. Thus, if the participant was assigned to the "3 classes" condition, this period was 9:30 - 11:10, which also spans 3 distinct hours. This meant that the period was boundary-expanded in terms of both classes *and* hours. Similarly, for participants in the “2 classes” condition, the period was 9:00 – 10:40, which is boundary-compressed on both dimensions (spanning only 2 distinct hours and 2 distinct classes).

However, for participants presented with the schedule where boundaries occurred on the half hour (i.e., Class A at 9:30, Class B at 10:30, and Class C at 11:30), the period spanned different numbers of classes and hours. Specifically, the period in the "3 classes" condition was 10:00 - 11:40, which only spans 2 distinct hours; this made the period boundary-expanded in terms of classes but compressed in terms of hours. The reverse was true for participants assigned to the "2 classes" condition—they rated 9:30 - 11:10, which is expanded in terms of hours but compressed in terms of classes (it spans 3 distinct hours and 2 distinct classes).

Based on H3, we expected that boundary-expansiveness in terms of classes would affect judgments of the period's duration, but we did not know whether the hour category would continue to exert any influence. That is, if hours are strong natural categories, they might affect judgments in addition to classes. Thus, our pre-registration specified that that we were agnostic about the influence of whole hour marks. Our confirmatory hypothesis was simply that salient new categories (classes) would show the same boundary-expansiveness effect that we observed with natural categories in previous studies.

Results. We performed a 2(number of classes: two vs. three) x 2(number of hours: two vs. three) ANOVA. As predicted, we observed a significant main effect of number of classes spanned. Specifically, the period was rated as feeling longer when, in terms of classes, it was boundary-expanded (spanning three classes, $M = 67.94$, $SD = 23.44$) rather than boundary-compressed (spanning two, $M = 64.53$, $SD = 24.56$; $F(1,1606) = 8.08$, $p < .01$, $d = .14$).

Conversely, perceived duration did not differ when, in terms of the hour category, the period was boundary-expanded (spanning three distinct hours) compared to compressed (spanning two; *M*s =66.76 and 65.73, *SD*s = 24.27 and 23.84, respectively; $F(1,1606) = .85$, $p = .357$, $d = .04$). There was no interaction between number of hours and number of classes ($F(1,1606) = .631$, $p = .427$). In short, the time period felt longer when it spanned more classes, regardless of whether it spanned more or fewer hours. This suggests that when periods span categories of different types, the effect of boundary-expansiveness is limited to the category that is most salient.

Discussion. These results support our third hypothesis—that a period's boundary-expansiveness according to idiosyncratic categories (i.e., events) affected perceived duration in the same way as the standard categories provided by, say, hour or month. People appear to rely on those standard categories when no other categories are salient; however, when we introduced more salient, alternative categories, those provided the boundaries that affected judgments. Interestingly, participants only responded to the newly-imposed categories and their boundaries, without any additional effect remaining for hour categories. This could be because judgments only rely on the most salient categories, though it possibly the result of how explicitly boundaries were provided. Specifically, the time period presented in this study was explicitly defined in terms of classes, while hours were merely implied. This suggests category boundaries may need to be salient in order to affect perceived duration.

Given that time periods that span more category boundaries are estimated to last longer, boundary-expansiveness should affect consumer decision-making across a variety of contexts. The next three studies examine its effects on scheduling decisions, valuation of delays and time savings, and finally, on rideshare choices within a large set of real-world transportation data.

STUDY 4: SCHEDULING

Consumers may want to minimize the amount of time taken by unpleasant activities, like getting cavities filled or going to the DMV. On the other hand, consumers may want to maximize time spent on enjoyable activities, such as exploring a new city or taking a much-needed nap. Thus, when scheduling activities for which they want to maximize time, consumers may prefer boundary-expanded time periods, and when scheduling activities where they want to minimize time, boundary-compressed periods may be more appealing. Study 4 tested this hypothesis.

Method and Procedure

Participants were 601 Amazon MTurk workers, with 600 remaining after excluding the one person who failed the attention check. Across eight trials, in a fully within-subjects design, participants indicated which of two equal-duration time periods they would rather schedule an activity. One period was boundary-expanded and the other boundary-compressed.

Participants were told to imagine that both time slots worked equally well with their schedule. The items represented four activities that consumers would prefer to expedite, and four that they would like to savor. To confirm this, a separate group of MTurkers (N = 40) rated each activity on a scale from 1 (“I would want to feel like I got it over with as fast as possible”) to 7 (“I would want to feel like it lasted as long as possible”). The time-minimizing activities received a much lower score (M = 1.48, SD = .88) than the time-maximizing ones (M = 6.11, SD = 1.14).

Time-minimizing activities were “going to the doctor to get blood drawn”, “lunch with someone you really dislike”, “visiting the dentist to have cavities filled”, and “going to the DMV”. Time-maximizing activities were “watching the finale of your favorite show”, “taking a much-needed nap”, “on a work trip and getting some time to explore”, and “free time in the middle of the day to do whatever you want.” For each of the time-minimizing activities, the boundary-compressed period was randomly presented to occur either before or after the boundary-expanded option. For example, if the boundary-compressed option was 12:00pm – 1:00pm, the expanded choice would either be 12:30pm – 2:00pm or 11:30pm – 1:00pm. The reverse was true for the time-maximizing activities, where we randomized whether the boundary-compressed option was earlier or later than the expanded. This design choice reduced the influence of time-of-day preferences (e.g., if an activity is generally preferred earlier or later in the day).

Results and Discussion

We performed a mixed-effects logistic regression of choice (boundary-expanded vs. compressed) on activity type (maximize vs. minimize), with a random effect of participant. As predicted, participants' choice of period differed between the two types of activities ($z = -5.53$, $p < .001$). For the time-maximizing activities, participants selected boundary-expanded periods over compressed ones (52% vs. 48%; $z = 2.27$, $p = .023$). But for the time-minimizing activities, participants selected boundary-compressed periods more often (56% vs. 44%; $z = -5.35$, $p < .001$).

In sum, we found that boundary-compressed periods were more attractive when scheduling activities that most people want to minimize, like going to the DMV. In contrast, participants preferred boundary-expanded periods for activities that most people want to maximize, like exploring a new city. These results suggest that consumer scheduling preferences may be affected by the boundary-expansiveness of the available time periods. In the next study, we explore consequences for how consumers value their time—specifically, how much they are willing to pay to avoid a long wait and how much compensation they require to endure it.

STUDY 5: THE VALUATION OF TIME

Consumers frequently decide whether they want to spend money to save time, or spend time to save money. This is particularly true in the domain of transportation, which is rife with long (and often painful) waiting periods (e.g., for Uber X, consumers pay more to arrive sooner; for Uber Pool, they pay less but arrive later). If time periods feel longer when they are boundary-expanded, how might this affect how consumers value their time?

We hypothesized that consumers would be willing to pay more to avoid a long wait when it is boundary-expanded compared to compressed, and will demand more compensation to endure it. Study 5 tested these predictions, examining changes in willingness to pay (WTP) and required minimum compensation in two scenarios involving waiting for transportation.

Method and Procedure

Three hundred two workers from Prolific Academic (45.1% female, 53.7% male, 1.2% did not answer, M_{age} = 36) participated. As per our preregistration, we removed trials where participants answered $0 or above the scale maximum for that question. We further removed participants who failed an attention check[1]. These exclusions left a final sample size of 260.

In a 2 x 2 within-subject design, every participant responded to two slightly different versions for each of two scenarios, one involving flight times and one involving bus times, respectively eliciting required compensation and willingness-to-pay. For the flight scenario, participants read the following:

> "Imagine that you are waiting to board a plane, and the flight is overbooked. It is {*present time*} right now. The airline is offering to pay you to take a later flight that leaves at {*departure time*}. What is the lowest amount of money you'd have to receive in order to take the later flight?"

Present and departure times formed either an expanded or compressed wait. For example, participants would view a present time of 10:47am [11:24am] and a departure time of 3:12pm [3:49pm], thereby creating a boundary-expanded [boundary-compressed] waiting period.

It is possible that arrival time influences responses, as people may generally want to arrive earlier. To address this, participants were randomly assigned to see a set of times where the boundary-expanded period began (and ended) either earlier or later than the equivalent compressed period. For example, participants would consider 10:47am – 3:12pm for the expanded period and 11:24am – 3:49pm for the compressed, or 9:47am – 2:12pm and 9:24am – 1:49pm (expanded and compressed, respectively). Whether the boundary-expanded period was earlier or later did not affect our pattern of results, $p = .89$.

Participants responded on a slider scale from $0 (I would change my flight for free) to $1000 or more. Those selecting the latter option were asked to enter their required amount; as pre-registered, responses larger than the scale maximum (in this case, $1000) were excluded.

On the next screen, the scenario was updated with a different set of times: "Now imagine that it is {*present time*} right now, and the airline is offering to pay you to take a later flight that leaves at {*departure time*}." Rating order was randomized, such that those who first rated a scenario with a boundary-expanded period now saw a compressed period of equal duration, and vice versa. Participants again indicated their required compensation on the same slider scale.

For the bus scenario, participants read:

> "Imagine that you get to the Greyhound bus station and learn that the tickets for the next bus are sold out, so you have to buy a seat on the bus that leaves at {*departure time*}. It is {*present time*} right now. The man next to you has a ticket for the bus that leaves in half an hour. What is the MOST you'd be willing to pay to switch tickets?"

Just as in the flight scenario, present and departure times formed either an expanded or compressed waiting period. Participants indicated their WTP on a slider scale from $0 to $300. They then saw a modified version of the bus scenario ("Now imagine that it is {*present time*} right now, and the only available seats are on a bus that leaves at {*departure time*}"), and again answered how much they would be willing to pay to switch to an earlier bus. The order of periods (boundary-expanded vs. compressed) was randomized. Just as before, we randomized whether the expanded period was earlier or later than the compressed; again, this factor did not affect our results, $p = .55$. The interested reader may examine tables B5 and B6 in the online appendix for means broken down by period for the flight and bus scenarios, respectively.

Results and Discussion

For the flight scenario, a mixed-effects model regressed required compensation on waiting period (expanded vs. compressed), specifying a random effect of participant and rating order. As predicted, the amount of money that participants required to take a later flight was higher when the waiting period was boundary-expanded ($M = 251.98$, $SD = 188.49$) compared to compressed ($M = 234.81$, $SD = 175.78$, $t = 4.44$, $p < .001$, $d_z = .29$).

We ran the same model for the bus scenario on WTP. As predicted, participants were willing to spend more to take an earlier bus when the waiting period was expanded ($M = 64.70$, $SD = 51.83$) compared to compressed ($M = 60.93$, $SD = 51.83$; $t = 2.97$, $p = .003$, $d_z = .20$).

These results reveal that boundary-expansiveness may affect consumer decisions involving the tradeoff between time and money. Participants reported being willing to pay about 6% more to avoid a waiting period when it was boundary-expanded compared to boundary-compressed. In addition, participants indicated that they required about 7% more compensation to endure a boundary-expanded (vs. compressed) waiting time. In the next and final study, we examine the effects of boundary-expansiveness on real-world transportation choices.

STUDY 6: RIDESHARE CHOICES

The results of the previous experimental study suggest that boundary-expanded wait times may make consumers more likely to upgrade to a faster, more expensive option. We now investigate whether this finding continues to hold in a real-world setting, specifically rideshare choices. There are two primary reasons for focusing on this market. First, rideshares are pervasive. Uber and Lyft have respectively surpassed 10 billion and 1 billion rides since their inception (Arevalo, 2020). Booking a rideshare is a common consumer decision in the modern world: Uber alone provides over 14 million rides every day (Srivastav, 2019), 20% of which are UberPool, and is the most frequently expensed vendor for business travelers (Hagen, 2019). Second, the availability of very large datasets on rideshares provides the statistical power needed to detect our hypothesized effects in consumer choice under noisy real-world conditions.

Importantly, rideshare decisions sometimes involve choosing between boundary-expanded and boundary-compressed options. Consumers living in major metropolitan areas often have the option to take a cheaper, longer trip by sharing their ride with other passengers (instead of taking an independent trip). They are provided with estimated arrival times for both options, and because the shared ride is always estimated to take longer, consumers sometimes face a choice set where the shared ride option crosses into a new hour, and the independent does not. Figure 3 illustrates the difference between same and different-boundary choice sets.

(Insert Figure 3 about here)

If the estimated arrival time crosses into a new hour for the shared option, but not for the independent, the shared trip may seem longer, increasing the likelihood of choosing the faster and more expensive solo option. Thus, we propose that if consumers face a “mixed” choice set—where the shared ride is expanded and the independent ride is compressed—they would be less likely to request the shared ride option than they would if the two options did not differ in boundary-expansiveness (that is, if neither or both rides crossed into the next hour).[2] We test this hypothesis on a large dataset from the Chicago metropolitan area. We describe the data in the next subsection and then develop and estimate a model of consumer choice.

Data and Variables Construction

We acquired rideshare choice data from the Chicago Transportation Network via the city's open data portal (https://data.cityofchicago.org). To our knowledge, Chicago is the only city that requires rideshare companies to publicly document, for all trips, pick-up and drop-off date, time, and location, trip duration, fare, miles travelled, whether the rider requested a shared or independent ride, and, if they requested a shared ride, whether they were matched with another rider (Table 1 provides summary statistics; see Web Appendix for more details). Following the emerging standard procedure for large archival data, we employed a split-half analysis. That is, exploratory analyses on roughly half the available data informed our treatment of the remaining half, for which data cleaning procedures and statistical analyses were pre-registered to reduce researcher degrees of freedom (Simonsohn, Simmons and Nelson 2019). We examined trips taken between November 2018 and May 2019 (N = 1,559,675) for the exploratory half and trips taken between June 2019 and December 2019 for the confirmatory half (N = 1,820,671).

(Insert Table 1 about here)

We pre-registered various restrictions. One obvious reason to select an independent ride, particularly when the shared ride is boundary-expanded, is to arrive on time to events that begin on the hour (e.g., work at 9am, a dinner party at 7pm). To mitigate this concern, we only examined trips that that began on weekdays between 1am and 5am. Further, trips were excluded if they took less than 3 minutes, 60 minutes or more, or travelled further than 30 miles. Trips that were missing location coordinates, or those that contained a coding error wherein the rider was recorded as receiving a shared ride when they had requested independent, were removed (0.3%). Lastly, to provide more precise estimates, trips with fewer than 100 “similar” trips present in the dataset—in this case, those with the same route, start hour, and ride type—were excluded (0.9%), as were those that were two SDs above or below the average duration of their similar trips (3.9%).

After all exclusions, a final sample of N = 1,820,671 remained (80.7% independent, 19.3% shared). Of course, although exact cut-offs for exclusions are always somewhat arbitrary, we explicitly preregistered them to reduce concerns of specification-curve hacking (e.g., Leamer 1983). Importantly, alternate analyses in the web appendix show that the results are robust to all exclusions and also hold for rides taken during the day (i.e., between 8am and 6pm).

The data only records cost and expected duration for the ride type that was actually chosen by the consumer. For instance, if a consumer chose the independent ride, we do not observe the cost and expected duration for the shared ride that they could have chosen instead. In addition, while the dataset provides precise estimates of trip duration, it rounds the start and end times for each trip to the nearest 15 minutes. Therefore, some work is needed to approximate the choice menus faced by consumers when they requested their rides. We now describe the steps we took.

First, we computed an interval of possible start times for each trip. The earliest and latest possible start times were determined, respectively:[3]

$$b_k = max(y_k - d_k, t_k) - 7.5$$

$$B_k = min(y_k - d_k,\ t_k) + 7.5$$

where y_k is the end time provided in the dataset, t_k is the provided start time, d_k is the duration of the trip in minutes, and k indexes the trip. An interval of possible start times S_k was computed for each trip as

$$S_k = \{x \in \mathbb{N} | b_k \leq x \leq B_k\}$$

We then estimated the probability that the independent and shared ride options would have been boundary-expanded. First, for each trip, we found the set of similar trips—that is, those with same start hour and route as the trip; route was a coarse grouping of pickup and drop-off location coordinates (i.e., latitude and longitude rounded to the nearest tenth).[4] For similar trips of each ride type, we calculated the proportion that, if they had started at the same time as the trip, would have crossed into a new hour. Proportions were calculated for each time within the trip's interval of possible start times and averaged. Expressed mathematically, the probability that the independent and shared rides are boundary-expanded for trip k were constructed as follows:

$$P_{ind}(k) = \frac{1}{|S_k|} \sum_{s_k \in S_k} \left(\frac{1}{|N_k|} \sum_{n \in N_k} 1\,(d_n + s_k \geq 60, ride\ n\ is\ ind)\right)$$

$$P_{shared}(k) = \frac{1}{|S_k|} \sum_{s_k \in S_k} \left(\frac{1}{|N_k|} \sum_{n \in N_k} 1\,(d_n + s_k \geq 60, ride\ n\ is\ shared)\right)$$

where N_k denotes the set of trips similar to k (as defined above), and $|N_k|$ denotes the number of elements in N_k and similarly for $|S_k|$. We then created our variable of interest,

$$P_{diff}(k) = P_{shared}(k) - P_{ind}(k)$$

which is simply the difference between the two proportions. This value captures the "mixed" boundary-expansiveness probability mentioned earlier—that is, the likelihood that the consumer would have been choosing between a boundary-expanded shared ride and a boundary-compressed independent ride. A difference score of 1 indicates that for all other similar trips, the shared ride would have crossed into a new hour, but the independent ride would not; a score of 0 implies that both or neither of the rides would have crossed.

Next, we followed a similar procedure to approximate the cost and expected duration for each ride type. For each trip, we found the difference between the average duration of all other shared rides and all other independent rides that had the same start hour and route as that trip:

$$D_{diff}(k) = \frac{1}{|N_k|} \sum_{n \in N_k} d_n 1(ride\ n\ is\ shared) - \frac{1}{|N_k|} \sum_{n \in N_k} d_n 1(ride\ n\ is\ ind)$$

The same procedure was performed for cost, subtracting the average cost of the analogous independent ride from the average cost of the shared one:

$$C_{diff}(k) = \frac{1}{|N_k|} \sum_{n \in N_k} c_n 1(ride\ n\ is\ shared) - \frac{1}{|N_k|} \sum_{n \in N_k} c_n 1(ride\ n\ is\ ind)$$

where c_n is the cost of the trip in dollars (provided in the dataset rounded to the nearest $2.50). The variables $P_{diff}(k), D_{diff}(k), C_{diff}(k)$ will be used as explanatory variables in our empirical model, which we now turn to.

Model

We now outline the simple model of consumer choice which will be estimated on the rideshare data. We focus on a consumer who is in the process of booking a ride on an app and is faced with the choice between an independent and a shared ride.[5] It is natural to assume that the decision will be affected by the expected duration and cost of each ride type, as well as the date and time in which the decision is made. In addition, we consider the possibility that the consumer might be influenced by whether each ride type crosses an hour mark (boundary-expanded) or not (boundary-compressed). Accordingly, we specify the utility that a consumer gets from choosing the independent ride as

$$U_{ind} = \alpha_{ind} + \beta_{exp} ind_{expanded} + \beta_d d_{ind} + \beta_c c_{ind} + \beta_{t,ind} time + \beta_{r,ind} route + e_{ind}, \quad (1)$$

where $ind_{expanded}$ denotes the dummy variable indicating that the independent ride is boundary-expanded, d_{ind}, c_{ind} are the duration and cost of the independent ride, respectively, *time* denotes the time when the choice is made (this term is composed of date, hour, and minute) and *route* denotes the pick-up and drop-off locations. Any other factors (e.g. the consumer's idiosyncratic preference for independent rides) are captured by the consumer-specific term e_{ind}. Similarly, we let the utility that a consumer gets from choosing the shared ride be

$$U_{sh} = \alpha_{sh} + \beta_{exp} shared_{expanded} + \beta_d d_{sh} + \beta_c c_{sh} + \beta_{t,sh} time + \beta_{r,sh} route + e_{sh} \quad (2)$$

Our main hypothesis is that the coefficient β_{exp} is negative, i.e. that, all else equal, consumers prefer a boundary-compressed ride relative to a boundary-expanded one. Further, we would expect that β_d and β_c are both negative, reflecting the fact that consumers tend to dislike spending time in transit and dislike spending money.

A consumer will choose the shared ride whenever $U_{sh} > U_{ind}$, i.e. whenever

$$\alpha + \beta_{exp} mixed + \beta_d (d_{sh} - d_{ind}) + \beta_c (c_{sh} - c_{ind}) + \beta_t time + \beta_r route + e > 0 \quad (3)$$

where $\alpha = \alpha_{ind} - \alpha_{sh}$, $mixed = shared_{expanded} - ind_{expanded}$, $\beta_t = \beta_{t,sh} - \beta_{t,ind}$, $\beta_r = \beta_{r,sh} - \beta_{r,ind}$, and $e = e_{sh} - e_{ind}$. Note that consumer choices are allowed to be affected not just by the features of the two ride types, but also by time and route.

As discussed above, the data only contains information on the ride type that was actually chosen by the consumer. Therefore, we do not observe some of the variables in Equation (3). To address this, for a given trip k, we approximate $mixed$, $d_{sh} - d_{ind}$, $c_{sh} - c_{ind}$ with $P_{diff}(k), D_{diff}(k), C_{diff}(k)$, respectively. These variables were defined in the previous subsection. The idea is that if $P_{diff}(k)$ is large then it is more likely that trip k has a mixed choice set, and similarly $d_{sh} - d_{ind}$, $c_{sh} - c_{ind}$ are more likely to be large when $D_{diff}(k), C_{diff}(k)$ are large, respectively. Thus, from Equation (3) we obtain, for trip k,

$$\alpha + \beta_{exp} P_{diff}(k) + \beta_d D_{diff}(k) + \beta_c C_{diff}(k) + \beta_t time + \beta_r route + e > 0 \qquad (4)$$

This is the equation we will use to estimate the effects of interest.

Results and Discussion

We estimate the coefficients in Equation (4) via logistic regression, where the dependent variable is a dummy equal to one when the consumer chooses the shared ride. As shown in Equation (4), we control for the expected differences in duration and fare; we also control for route and start time by respectively including fixed effects for each pick-up/drop-off location pair[6] and day/hour (e.g., January 3rd at 2am) as well as minute. Table 2 reports these results. Table 2 also presents results from a linear regression with robust standard errors.

(Insert Table 2 about here)

As expected, all estimated coefficients are negative and statistically significant. In particular, the negative coefficient on the mixed choice set probability indicates the following: As the probability of a mixed choice set increases—that is, choice sets where the shared ride is boundary-expanded and the independent is boundary-compressed—the consumer is less likely to choose the shared ride, all else equal.[7] It is important to emphasize that this holds conditional on the duration and cost differences between the two ride types. In other words, the mixed choice set probability has a negative impact on the probability of choosing the shared ride that is distinct from the sheer effect of differences in cost and, importantly, duration between the two ride types.

In the online appendix, we show that the estimate of the effect of interest is robust across several alternate specifications. In particular, we estimate a least-squares linear regression based on Equation (4). The fact that our main effect continues to be statistically significant provides some reassurance regarding measurement error. Under the standard assumption that measurement error is independent of the mis-measured variable, the estimates would be biased towards zero. Thus, the fact that we still find significant effects suggests that the true coefficients might in fact be even larger.

We now turn to quantifying the effect of a mixed choice set on consumer behavior. One standard way to proceed is to consider odds ratios. Define the odds of a shared ride as the probability that a consumer chooses the shared ride divided by the probability that she chooses the independent ride. Then, our results imply that, all else equal, the odds of selecting a shared ride when the choice set is mixed are 41% lower than when the choice set is not mixed.[8]

Another way to assess this is to look at how much money consumers would be willing to pay in order to avoid crossing the hour boundaries. To this end, we proceed in three steps. First, for every choice in the dataset, we calculate the utility the consumer expects to derive from the choice set she is facing. Second, we repeat the same calculation for the hypothetical scenario in which neither ride type crosses a boundary. Given the estimated negative coefficient on boundary crossing, the expected consumer utility computed for this hypothetical scenario will be higher relative to that in step one. Finally, we use the quantities in the previous two steps to compute the amount of money that consumers would be willing to pay to go from the status quo to the hypothetical scenario.[9] This is a measure of consumers' willingness to pay to avoid crossing the hour boundaries. On average, we find that consumers in our data would be willing to pay $0.60 per trip, or around 5% of the fare for independent ride.

This amount may scale up massively; in Chicago alone, approximately 10.9 million trips were taken in 2019 (https://data.cityofchicago.org). As such, platforms might substantially increase their revenues by incorporating these insights in their pricing strategy. To further explore this, we consider the following change in prices. For every trip with $P_{diff} > 0$, we increase the price of the independent ride and simultaneously decrease the price of the shared ride by the same amount. This policy increases the price for the ride type that is less likely to be boundary expanded, while leaving the average price faced by each consumer constant. We consider this type of pricing policy to account for the fact that each rideshare app might not want to increase the overall price levels for fear of losing customers to competing apps or other modes of transportation.[10] We calculate the expected revenue under this alternative pricing scheme and compare it with that obtained under the pricing policy in the data.[11] Web Appendix figure D1 shows that, for a range of price changes, expected revenue per trip would increase. In particular, by increasing the price of the independent ride by about $1.8 (and decreasing the price of the shared ride by the same amount), our estimates suggest that rideshare apps could increase their revenue per trip by more than $0.30 on average. Scaling this by the number of annual rides in Chicago yields an increase in expected revenues by more than $3.5 million dollars. Of course, changing the pricing policy could alter consumer choices in ways that are not captured by our model (e.g., consumers might be more willing to bike or walk when independent rides become more expensive, even if the average price stays constant) and our estimates do not reflect this. Nonetheless, the strong evidence that boundary crossing affects consumer choice does support the notion that rideshare apps could increase their revenues by incorporating this insight in their pricing strategies.

Lastly, based on a reviewer suggestion, we consider possible consequences of boundary-crossing for customer satisfaction. If, according to their original choice menu, the trip was supposed to be boundary-expanded but ended up being boundary-compressed, the consumer may feel particularly satisfied and tip the driver more. Controlling for ride duration, choice (independent vs. shared), fare, route and time of day, we find evidence of such correlation. Specifically, customers gave higher tips as it became more likely that the ride had been forecasted to be boundary-expanded but was actually boundary-compressed. These results suggest that expectations around boundary-expansiveness may have downstream consequences for consumer satisfaction.

GENERAL DISCUSSION

The present research finds that time periods of equal duration do not always feel equivalent and therefore affect consumer decisions across a variety of domains. We demonstrate that time periods feel longer when they span more distinct time categories (e.g., the "3" in 3:15, the "March" in March 3rd).

We believe that the bias documented here results from a simple but fundamental psychological process whereby category boundaries, which typically denote larger differences, are interpreted as doing so in situations where this is not applicable. It may be best interpreted as a System 1 type bias (Kahneman 2011). However, such biases are often most pronounced in decisions for which consumers lack experience and incentive (Kahneman and Frederick 2002; Farrell, Goh and White 2014), factors which also fuel real-world generalizability concerns about many biases found in the lab (Levitt and List 2007a; Levitt, List and Reiley 2010). To this point, the results of study 6 are particularly informative. For the consumers in our dataset, selecting a rideshare is a common decision that directly affects their wallets, yet we still found a substantial effect of hour categories on choices, with consumers being willing to pay about $0.60 more just to avoid crossing a salient hour mark.

The fact that this bias has such a strong effect on routine consumer decisions leads to questions about its origins. While beyond the scope of the current manuscript, two options seem possible. First, boundary-expanded time periods might *actually* feel longer while one is experiencing them. This might be true if, for instance, people check the time more often during a boundary-expanded period, thus drawing their attention to passing more "boundaries." In that case, what appears as inconsistencies in judgments may represent accurate forecasts of experiences. We do not have data testing this possibility, but it does not appear likely. In study 1, MTurkers considering boundary-expanded periods estimated being able to complete an average of 17 more HITs than those considering equivalent compressed periods. We believe that categorical distortion is likely more similar to perceptual illusions than it is to typical System 1 heuristics (for a discussion on the similarity between perceptual and cognitive illusions, see Kahneman and Tversky 1996). Such illusions result from the perceptual system accurately interpreting stimuli within a frequently encountered context, but overapplying that interpretation in contexts where it shouldn't (see, the carpentered world hypothesis: Segall, Campbell and Herskovits 1966). In the same vein, even though time categories are arbitrary segments along an abstract continuum, they are somewhat informative; categories like 1pm and March are respectively closer to 2pm and April than they are to 3pm and May. Relying on categories when considering duration may be an adaptive, effort-saving heuristic (Gigerenzer & Gaissmaier 2011)— on average, a period that starts and ends in different hours, for example, is likely to be longer than one that starts and ends in the same hour. But our results imply that people may overapply that interpretation. This suggests that these biases will be hard, if not impossible, to avoid, and any interventions should aim to help decision-makers recognize situations in which they could fall prey to these biases and actively correct for them.

Over and above the ones provided by time (e.g., hour, month), everyday events may provide additional category boundaries. For many people, 12:00pm signals a new hour *and* a new activity, lunch, and thus may be a stronger boundary than, say, 11:00am. Consistent with study 3, the same duration may also feel longer if it contains more punctuating events. A past point in time feels further away when people consider more (vs. fewer) relevant intervening events (Zauberman et al. 2010). Moreover, a time interval feels longer when it has many (vs. few) shifts between events (Ornstein 1969). Although these studies involve retrospective and subjective judgments of time, their conclusions match our own—that shifts exaggerate duration. If a period feels longer when it contains more events, people may anticipate periods earlier in the day to feel longer than equivalent later periods because the former usually contain more events (e.g., lunch, meetings) and the latter fewer (e.g., dinner). For an initial test of this idea, some periods in study W4 in the online appendix have the same duration, but some are in the afternoon (i.e., 11:30am – 3:00pm, 12:00pm – 3:30pm) and others in the evening (i.e., 6:30pm – 10:00pm, 7:00pm – 10:30pm). We indeed find that the afternoon period was rated to feel longer than the equal-duration evening period, $F(1,850.55) = 45.47$, $p < .001$. But the effect of boundary-expansiveness on perceived duration did not differ between these two periods.

At first glance, our findings may appear to contradict recent work by Tonietto et al. (2019), which shows that time periods feel shorter when they are bounded by an upcoming task (e.g., a meeting) compared to unbounded (i.e., when the time afterward is unaccounted for). However, participants in Tonietto et al. (2019) estimate the time that falls *between* boundaries—which are either absent or defined by a subsequent event—and the inclusion of a boundary made the intermittent time feel shorter. In contrast, we ask participants to estimate time periods that *span* boundaries, and, critically, find that it is the boundaries themselves that make the period feel longer. Thus, these results are actually very compatible, and combined illustrate the different ways in which category boundaries can both make the time between them feel shorter (Tonietto et al. 2019) but across them feel longer (this paper). Interestingly, the design of study W8 in the web appendix allows us to test both effects within the same study. In line with Tonietto et al. (2019), we found that the unbounded periods (i.e., those without salient end times) felt longer than the bounded (i.e., those with salient end times). Together, both papers underscore the importance of boundaries in time estimation and their implications for consumer behavior.

We argue that the number of category boundaries crossed inflates perceived duration. If true, increasing the number of boundaries within a given time period (e.g., by making salient markers occur every half hour or 15 minutes) should produce similar effects. This idea closely relates the unit effect, wherein quantities seem larger when expressed in more (smaller) units (e.g., for a warranty, 84 months feels longer than 7 years; Pandelaere, Briers and Lembregts 2011). Moreover, because the difference of one extra unit is proportionally larger, one could expect that boundary-expansiveness affects shorter periods more than longer (e.g., spanning two categories instead of one should feel like a larger difference than spanning eight instead of seven). Although not designed to test for such a pattern, three of our studies have enough variability in duration to provide an initial look (see the online appendix). Of these three studies, only study 1 has a large enough sample to reliable test this prediction. For this study, we do find the anticipated pattern—the perceived difference between boundary-expanded and compressed periods decreased as their duration increased, $z = -5.48$, $p < .001$. The other two studies (W2 and W4 in the online appendix) have roughly one-sixth the number of participants as study 1, but show (non-significant) patterns in the same direction.

Practical Implications

These results have numerous implications for businesses looking to optimally position and price their services. First, our findings suggest that altering the start and end time of appointment slots may facilitate booking. Although doing so may be impossible for busy service providers with urgent customer needs, such a practice may be useful for businesses with flexible scheduling. Our prescription varies by the type of service offered—namely, whether it is one that consumers want to prolong or expedite. For example, boundary-expanded appointments may be more appealing to clients of a travelling masseuse. In the same vein, given that most people prefer to minimize their travel time, transportation companies may have greater success attracting customers for trips that are boundary-compressed instead of expanded. Indeed, our rideshare analysis finds that the odds of selecting the shared ride were over 40% lower when the shared ride was likely to cross into a new hour, but the independent ride was not.

Second, businesses may be able to sustain a slight increase in the price of certain offerings. We found a 6% increase in willingness-to-pay for a bus ride when it was boundary-compressed instead of expanded (study 5). Were consumers willing to pay even 1% more for, say, boundary-compressed flights, this amount scales up massively. For instance, with a carrying capacity of roughly 200 on a Boeing 737, pricing a boundary-compressed flight at $505 (instead of $500) would net an additional $1,000 per flight. The difference reaches substantial significance when one considers that over 38 million flights were taken in 2019 alone (Statista 2020). Similarly, our rideshare analysis determined that customers were willing to pay about 5% of the cost of the independent ride—roughly $0.60 more—to avoid crossing hour boundaries. Thus, rideshare platforms may observe sizable gains in revenue if they consider the influence of boundary-expansiveness when pricing options. Upon doing so, by our estimates, rideshare companies would increase expected revenue by more than 3.5 million dollars per year in Chicago alone.

The basic choice paradigm faced by rideshare consumers exists for many other routine decisions. For example, consider shipping and product delivery. Food and grocery delivery services provide estimated arrival times. If the estimate is before a new hour from one provider (e.g., arriving at 6:53pm), but after a new hour for another (e.g., 7:02pm), consumers may be more inclined to order from the former. Similarly, consumers may be more likely to buy from a seller who estimates delivery on, say, May 31st rather than June 1st. Thus, businesses may want to prioritize making deliveries boundary-compressed when possible. We note that this should be true for both estimated *and* actual delivery time, because any delays in delivery may seem more substantial if they cross into the next time category (i.e., hour, month).

This raises implications for customer satisfaction across several domains. Whether waiting for deliveries, appointments, or transportation, customers may be more unhappy with delays that broach a new time category (e.g., hour or month) compared to those that do not. They may also be especially satisfied when an arrival time was projected to fall after a new time category but came *before* it instead. The exploratory tipping analysis mentioned in the rideshare study supports this idea (see table D2 in the web appendix). In short, differences between the expected and actual boundary-expansiveness of a waiting period may influence customer satisfaction.

Beyond a consumer context, our findings have implications for recruitment, hiring, and employee satisfaction. Managers may increase attendance for non-mandatory workplace trainings if they offer them during boundary-compressed periods. This also applies when recruiting for volunteer work, favors or unpaid labor—people may be more willing to work during a boundary-compressed period, rather than expanded, because it seems like less of a commitment. Moreover, companies and institutions with fixed vacation periods may want to make them boundary-expanded when possible. For instance, summer break at UC Berkeley currently goes from May 14th to August 19th, but to the extent that prospective students or staff value long summer breaks, attendance or employment may be (somewhat) more appealing if the break were from May 28th to September 2nd instead.

From an entirely different angle, our work highlights an important factor currently biasing consumer decision-making. It is unlikely that most consumers spontaneously consider the way in which boundary-expansiveness influences their perception of time. Knowledge of the bias presented in the present paper may prevent consumers from falling prey to it, potentially improving consumer welfare.

Conclusion

Together, our studies suggest that time periods feel longer when they span more boundaries, and that this phenomenon may shape the scheduling and purchasing decisions consumers make in everyday life. Broadly, this research provides novel insight into the ways in which consumers perceive time and anticipate the duration of future experiences.

REFERENCES

Allen, Vernon L. and David A. Wilder (1979), “Group Categorization and Attribution of Belief Similarity,” *Small Group Behavior* 10(1)73–80.

Alter, Adam L., and Hal E. Hershfield (2014), "People search for meaning when they approach a new decade in chronological age." *Proceedings of the National Academy of Sciences* 111(48)17066-17070.

Ayers, John W., Benjamin M. Althouse, Morgan Johnson, and Joanna E. Cohen (2014), "Circaseptan (weekly) rhythms in smoking cessation considerations," *JAMA internal medicine* 174(1)146-148.

Berlin, Brent, Dennis E. Breedlove, and Peter H. Raven (1973), "General principles of classification and nomenclature in folk biology." *American anthropologist* 75(1)214-242.

Blaszkiewicz, Suzie. (2018, March 13). “Online booking options can get you more clients.” Retrieved from https://lab.getapp.com/research-online-booking-importance-of-appointment-scheduling/

Burris, Christopher T. and Nyla R. Branscombe (2005), “Distorted Distance Estimation Induced by a Self-Relevant National Boundary,” *Journal of Experimental Social Psychology* 41(3)305–12.

Dai, Hengchen, Katherine L. Milkman, and Jason Riis (2014),"The fresh start effect: Temporal landmarks motivate aspirational behavior." *Management Science* 60, no. 10: 2563-2582.

Dougherty, Janet WD (1978), "Salience and relativity in classification," *American ethnologist* 5.1 (1978): 66-80.

Eiser, J. Richard, and Wolfgang Stroebe. "Categorization and social judgment." (1972).

Farrell, Anne M., Joshua O. Goh, and Brian J. White (2014), "The effect of performance-based incentive contracts on system 1 and system 2 processing in affective decision contexts: fMRI and behavioral evidence," *The Accounting Review* 89(6)1979-2010.

Gardner, William, Edward P. Mulvey, and Esther C. Shaw (2010), "Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models," *Psychological bulletin* 118 (3)392.

Gati, Itamar, and Amos Tversky (1984), "Weighting common and distinctive features in perceptual and conceptual judgments." *Cognitive Psychology* 16, no. 3: 341-370.

Gigerenzer, Gerd, and Wolfgang Gaissmaier. "Heuristic decision making." *Annual review of psychology* 62 (2011): 451-482.

Hagen, Shelly (2019, July 25). “Uber, Lyft and Scooters Are Winning Over Corporate Travelers.” Retrieved from https://www.bloomberg.com/news/articles/2019-07-25/uber-lyft-and-starbucks-among-most-popular-business-expenses

Irmak, Caglar, Rebecca Walker Naylor, and William O. Bearden (2010), “The out-of-Region Bias: Distance Estimations Based on Geographic Category Membership,” *Marketing Letters* 22(2)181–96.

Isaac, Mathew S. and Robert M. Schindler (2014), “The Top-Ten Effect: Consumers Subjective Categorization of Ranked Lists,” *Journal of Consumer Research* 40(6)1181–1202.

Kahneman, Daniel (2011), *Thinking fast and slow*, New York: Farrar, Straus and Giroux.

Kahneman, Daniel and Shane Frederick (2002), “Representativeness revisited: Attribute substitution in intuitive judgment,” *Heuristics and biases: The psychology of intuitive judgment*, *49*, 81.

Kahneman, Daniel and Amos Tversky (1996), "On the reality of cognitive illusions,” *Psychological Review*, 103, 582 - 591.

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.

Leamer, Edward E. (1983), "Let's take the con out of econometrics," *The American Economic Review* 73(1)31-43.

Levitt, Steven D., and John A. List (2007a), *"*What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives*, *21*(2)153-174.

Levitt, Steven D., John A. List, and David H. Reiley (2010), "What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments," *Econometrica* 78.4 (2010): 1413-1434.

Locksley, Anne, Vilma Ortiz, and Christine Hepburn (1980), “Social Categorization and Discriminatory Behavior: Extinguishing the Minimal Intergroup Discrimination Effect.,” *Journal of Personality and Social Psychology* 39(5) 773–83.

Maki, Ruth H. (1982), “Why Do Categorization Effects Occur in Comparative Judgment Tasks?,” *Memory & Cognition* 10(3)252–64.

Manning, Kenneth C., and David E. Sprott (2009), "Price endings, left-digit effects, and choice", *Journal of Consumer Research* 36(2) 328-335.

Pandelaere, Mario, Barbara Briers, and Christophe Lembregts (2011), "How to make a 29% increase look bigger: The unit effect in option comparisons." *Journal of Consumer Research* 38(2)308-322.

Roberson, Debi, Ian Davies, and Jules Davidoff (2000), “Color Categories Are Not Universal: Replications and New Evidence from a Stone-Age Culture.,” *Journal of Experimental Psychology: General* 129(3) 369–98.

Rosch, Eleanor (1999), Principles of categorization. *Concepts: core readings*, *189*.

Rosch, Eleanor, and Carolyn B. Mervis (1975), "Family resemblances: Studies in the internal structure of categories," *Cognitive psychology* 7, no. 4: 573-605.

Ryan, William, Ellen Evers, and Don A. Moore (2018), "False Positive Poisson." *Available at SSRN 3270063*.

Schindler, Robert M. and Thomas M. Kibarian (1996), “Increased Consumer Sales Response Though Use of 99-Ending Prices,” *Journal of Retailing* 72(2)187–99.

Segall, Marshall H., Donald Thomas Campbell, and Melville Jean Herskovits (1966), *The influence of culture on visual perception*. Indianapolis: Bobbs-Merrill.

Simonsohn, Uri, Joseph P. Simmons, and Leif D. Nelson (2009), "Specification curve: Descriptive and inferential statistics on all reasonable specifications," *Available at SSRN 2694998*.

Medin, Douglas L., and Edward E. Smith (1984), "Concepts and concept formation," *Annual review of psychology* 3(1)113-138.

Skowronski, John J., Timothy D. Ritchie, W. Richard Walker, Andrew L. Betz, Constantine Sedikides, Leslie A. Bethencourt, and Amy L. Martin (2007), "Ordering our world: The quest for traces of temporal organization in autobiographical memory," *Journal of Experimental Social Psychology* 43(5)850-856.

Srivastav, Sudeep (2019, July 27). “35 best Uber Statistics to Know (2019-2020 Updated).” Retrieved from https://appinventiv.com/blog/uber-statistics/#A16

Statista. “Number of flights performed by the global airline industry from 2004 to 2021.” Retrieved 27 November 2020 from https://www.statista.com/statistics/564769/airline-industry-number-of-flights/

Tajfel, Henri (1969), "Cognitive aspects of prejudice." *Journal of biosocial science* 1(1)173-191.

Tajfel, Henri (1959), "Quantitative judgement in social perception,” *British journal of psychology* 50(1)16-29.

Tajfel, Henri and Alan L. Wilkes (1963), "Classification and quantitative judgement," *British journal of psychology* 54(2)101-114.

Thomas, Manoj and Vicki Morwitz (2005), “Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition,” *Journal of Consumer Research* 32(1)54–64.

Thomsen, Dorthe (2009), “There is more to life stories than memories”, *Memory,* 17(4)445–457.

Tonietto, Gabriela N., Selin A. Malkoc, and Stephen M. Nowlis (2018), “When an Hour Feels Shorter: Future Boundary Tasks Alter Consumption by Contracting Time,” *Journal of Consumer Research* 45(5)1085–1102.

Tu, Yanping, and Dilip Soman (2014), "The categorization of time and its impact on task initiation." *Journal of Consumer Research* 41(3) 810-822.

Tversky, Barbara, and Kathleen Hemenway (1984), "Objects, parts, and categories," *Journal of experimental psychology: General* 113(2)169.

Winawer, Jonathan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky (2007), “Russian Blues Reveal Effects of Language on Color Discrimination,” *Proceedings of the National Academy of Sciences* 104(19)7780–85.

Zauberman, Gal, Jonathan Levav, Kristin Diehl, and Rajesh Bhargave (2009), “1995 Feels So Close Yet So Far,” *Psychological Science* 21(1)133–39.

FOOTNOTES

[1] This attention check was similar to the bus scenario, but the wait was 4 hours shorter. As pre-registered, we removed participants who had a higher WTP here than they did on either bus trial.

[2] As discussed below, we only consider trips that last at most 60 minutes.

[3] B_k is the latest possible start time because, given the provided start and end times, the trip must have started before both $t_k + 7.5$ and $y_k + 7.5 - d_k$. A similar argument applies to b_t.

[4] This corresponds to partitioning the city into cells roughly of size 7 miles (north-south) by 5 miles (east-west). This yields 95 unique pick-up and drop-off location pairs.

[5] Note that we do not model the choice of which app to use (e.g., Uber or Lyft) or the choice between using a rideshare app and other means of transportation (driving, biking, etc.).

[6] Specifically, we partition Chicago into four roughly equally-sized quadrants and include fixed effects for each combination of pick-up and drop-off location (e.g., from southwest to northeast).

[7] P_{diff} improves the model likelihood from -831721.2 to -831609.6, $\chi^2(1) = 223.28$, $p < .001$.

[8] This follows from the fact that the ratio between the odds of the shared ride when the probability of a mixed set is 1 and the odds when that probability is zero equals $e^{-0.525} = 0.592$.

[9] Note that, under the assumption that the terms e_{ind}, e_{sh} are independent type-I extreme value random variables, expected utility can be conveniently computed in closed form using the logsum formula.

[10] Our model does not capture this type of substitution since the data does not contain information on the customers who considered booking a ride but eventually chose not to.

[11] We focus on changes in revenues as opposed to profits since assessing the latter would require a measure of the costs incurred by rideshare apps for the different types of rides, which is not in our data.

Table 1. Summary statistics of the rideshare data.

	Mean	S.D.	Min	Max
Proportion selecting shared ride	0.19	0.39	0	1
Probability that trip was boundary-expanded	0.05	0.08	-0.15	0.61
Duration (in minutes)	14.92	9.25	3.02	59.92
Fare (in dollars, rounded to nearest $2.50)	12.40	8.65	0	392.50

Table 2. Coefficient estimates, standard errors and p-values under logistic regression and linear regression with robust standard errors. The dependent variable is a dummy equal to 1 if the consumer chooses the shared ride.

	Logistic regression		*Linear regression with robust s.e.*	
	Estimate	Standard Error	Estimate	Standard Error
P_{Diff}	-0.525***	0.035	-0.068***	0.005
D_{Diff}	-0.035***	0.002	-0.008***	0.0003
C_{Diff}	-0.220***	0.002	-0.032***	0.0003

*** p < .001

Figure 1. MTurk workers estimated that they could complete more HITs during boundary-expanded periods compared to boundary-compressed. Error bars represent the standard error of the mean.

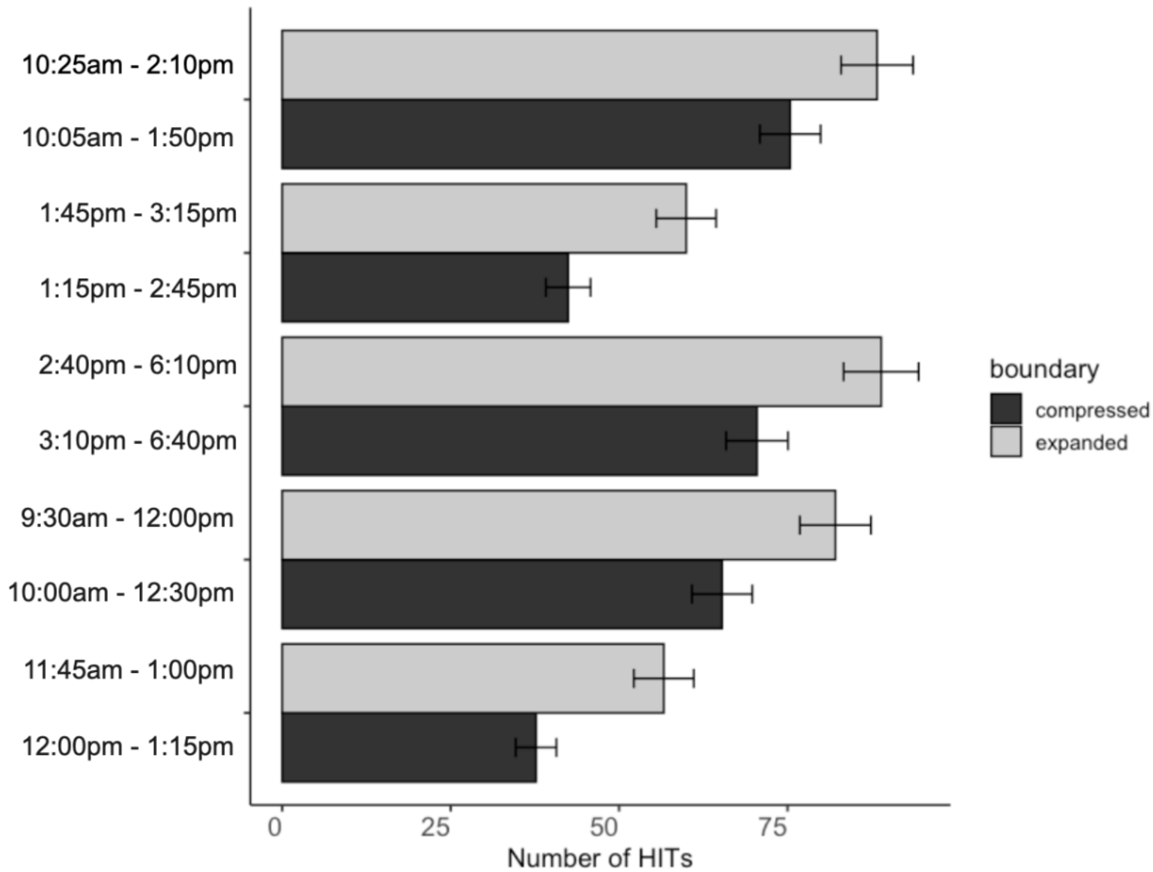


Figure 2. Time periods spanning more month categories felt longer. Periods are expressed as Day-Month-Year and rated by participants in the United Kingdom. Error bars represent the standard error of the mean.

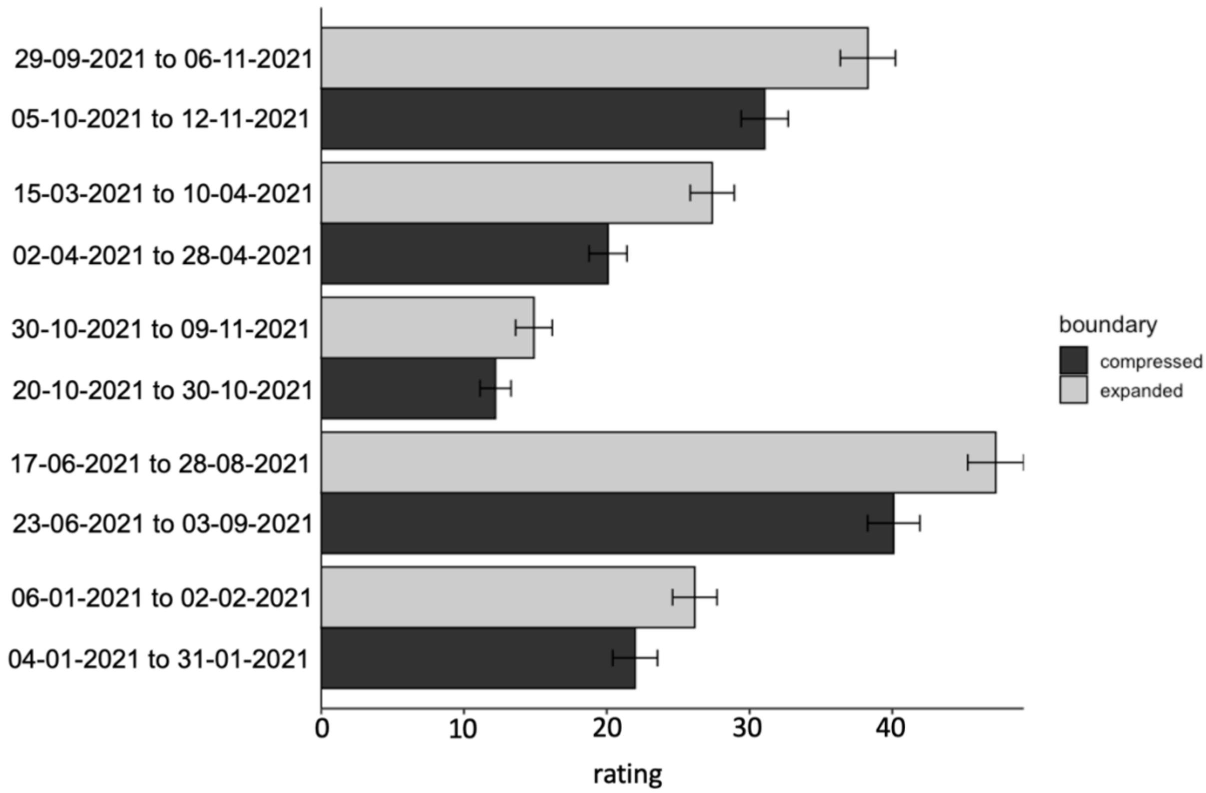
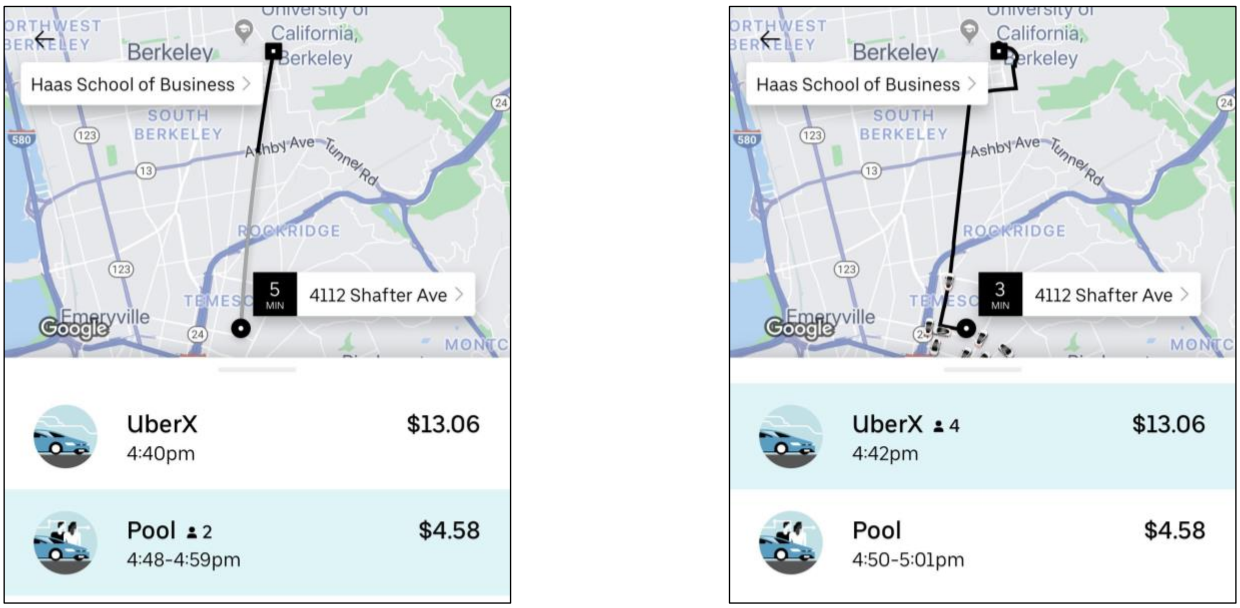


Figure 3. Example rideshare choice sets. Left image: Neither ride option is expected to cross into the next hour. Right image: Shared ride option is expected to cross, independent is not.



WEB APPENDIX

WEB APPENDIX A: ADDITIONAL STUDIES

Studies W1 and W2

Studies W1 and W2 tested the basic effect of boundary-expansiveness on perceived duration. We predicted that time periods feel longer when they span more hours; that is, participants will judge boundary-expanded periods to feel longer than boundary-compressed.

Study W1

Participants and procedure. One hundred and four workers from Amazon Mechanical Turk completed this study. All workers passed a language quality check, which involved correctly identifying that an image depicted a dog playing piano. Participants were told that we were interested in “whether some time periods *feel* longer than others, even if they are technically the same.” Participants completed five trials in which they were asked to indicate which of two time periods “feel longer” (see appendix for the time periods used in these trials and all other studies). Each pair of periods had the same duration, but one period was boundary-expanded (e.g., 3:30pm – 5:00pm) and one was boundary-compressed (e.g., 3:00pm – 4:30pm).

Results. We performed a logistic mixed-effects regression on choice (expanded vs. compressed), specifying a random effect for participant and a random effect for trial. Participants selected boundary-expanded times to feel longer more often than boundary-compressed (68% vs. 32%, respectively; $z = 4.94$, $p < .001$, $d_z = .54$).

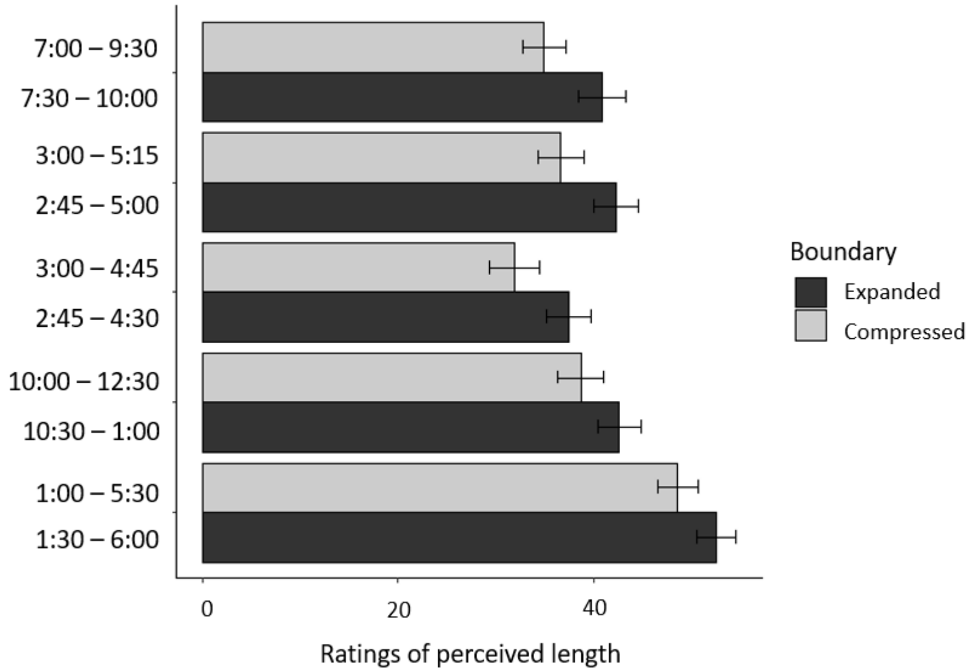
Study W2

In study W1, participants made a forced choice. To ensure that the effect holds outside of this binary situation, we had participants rate their perceived duration of each period on a continuous slider.

Participants and Procedure. One hundred and four MTurk workers participated in this study. Participants viewed twelve time periods one at a time and rated each for "How long does this feel?" on a scale from 0 (doesn't feel long at all) – 100 (feels extremely long). The first two trials were practice trials composed of one very short time period and one very long one to give participants some idea of how to interpret the scale. For the next ten trials, the trials of interest, half of the periods were boundary-expanded (e.g., 2:30pm – 3:00pm), and the other half were of equal length but boundary-compressed (e.g., 3:00pm – 3:30pm). We expected that participants would indicate that the expanded trials feel longer than the compressed trials.

Results. The results are illustrated in figure A1. We performed a mixed-effects regression on ratings, specifying a fixed effect for period type (expanded vs. compressed) and random effects for participant and trial. Participants indicated that boundary-expanded time periods felt longer (M = 43.14, SD = 23.31) than boundary-compressed (M = 38.15, SD = 24.04; t = 4.84, p < .001).

Figure A1. Boundary-expanded periods were rated as “feeling longer” than boundary-compressed. Error bars represent standard error of the mean.



Discussion. The results of study W1 and W2 revealed that time periods felt longer when they spanned more hours. In study W1, participants selected boundary-expanded periods to feel longer more often than boundary-compressed. This finding was echoed by study W2: When time periods were evaluated individually and rated on a slider for perceived length, boundary-expanded periods felt longer than boundary-compressed.

These studies provide initial support for our hypothesis. However, they cannot rule out mundane explanations for the effect; for example, that participants are bad at math in some systematic way, or rounding times in a way that exaggerates the duration of boundary-expanded periods. Studies W3 and W4 investigate these possibilities.

Study W3

Study W3 investigated whether the effect results from participants relying on bad math. When evaluating time periods, participants might rely on a quick and erroneous estimate based on the first hour unit, encoding times like 12:30pm – 2:00pm and 12:00pm – 1:30pm as 12-2 and 12-1, respectively. To address this concern, study W3 presented participants with pairs of equivalent time periods (one boundary-expanded, one boundary-compressed) and required them to calculate the duration of both periods before selecting which felt longer. If the effect observed in the previous studies results from some participants not paying attention, then it should be eliminated when they must explicitly calculate duration before making their choice.

Participants and Procedure. Participants were 154 MTurk workers. As specified in the pre-registration, we removed all trials where participants incorrectly calculated duration for one or both time periods. Further, we eliminated any participant who provided incorrect calculations on more than two trials[1]. These exclusions resulted in a final sample of 137.

Just as in study W3, participants completed five trials in which they evaluated pairs of equivalent time periods—one was boundary-expanded and one was boundary-compressed. Before selecting which period "feels longer," participants had to calculate the duration of each. They indicated their answers on a menu of possible responses that were set in 15-minute increments (from 15 minutes to 6 hours).

Results. A binomial logistic mixed-effects model regressed random effects of participant and question on choice (boundary-expanded vs. compressed). We observed the predicted

[1] Participants were slightly worse when calculating the expanded duration; on average, participants answered incorrectly for .53 out of 5 expanded trials, and .35 out of 5 compressed trials.

significant intercept: Participants selected the boundary-expanded periods to "feel longer" more often than the boundary-compressed (75% vs. 25%, respectively; $z = 6.93$, $p < .001$).

Study W4

It is possible that people round the start and end times of the periods under consideration. They might, for example, round 2:30 down to 2:00 or up to 3:00, doing this for both start and end times, or for only one or the other. Some forms of rounding might produce the effect demonstrated in the previous studies. For instance, if people round down start and end times, periods like 1:30 – 3:00 and 2:00 – 3:30 might be respectively rounded to 1:00 – 3:00 and 2:00 – 3:00, making the latter (boundary-compressed) period feel shorter.

In the previous studies, most periods had times ending in :30; if, for example, participants rounded down times ending in :30, they should be even more likely to round down times that end in :05. Similarly, if they were rounding up times that end in :30, they should be even more likely to round up times ending in :45. In this study, we used time periods that are likely to be rounded in a way that runs counter to our predictions—specifically, where rounding renders boundary-expanded periods *shorter* than compressed. In other words, we manipulated periods such that rounding would always make boundary-compressed periods longer than expanded (see Table B12 for a full list of periods). For example, we contrasted a boundary-compressed period like 3:15 pm – 4:45 pm with an equivalent boundary-expanded counterpart like 2:45pm – 4:15pm. In this case, if participants round the times, the boundary compressed period would be rounded to 3pm – 5pm, while the boundary expanded time would be rounded to 3pm – 4pm. Thus, if rounding explains the previous results, we should find the reverse of our hypothesis in these situations. However, if the effect relies on the number of boundaries crossed, participants in the rounding conditions should still evaluate the boundary-expanded times as longer than the boundary-compressed.

In sum, in study W4, we varied both boundary-expansiveness (expanded vs. compressed) and the consequences of rounding: For half the periods, rounding would make the boundary-expanded period shorter than the boundary-compressed.

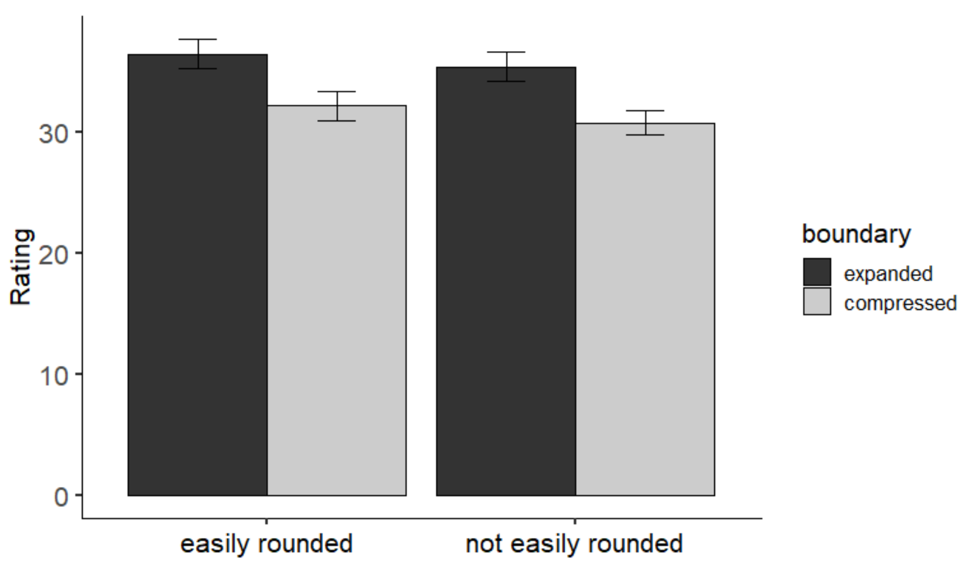
Participants and Procedure. One hundred and thirty-four MTurk workers participated in this study. Just as in study W2, boundary-expanded and boundary-compressed time periods were rated one at a time for "how long they feel" on a sliding scale from 0-100. However, in this study, we also included trials in which rounding would result in a pattern opposite from our predictions. That is, half the periods would produce the opposite effect if rounding plays a role. Thus, this 2(boundary expanded vs. compressed) x 2(more vs. less susceptible to rounding) within-subject design produced four possible types of time periods: the basic expanded and compressed trials and a new set of expanded and compressed trials designed in such a way that rounding would result in the opposite of our predictions. Each participant rated a total of 20 time periods (displayed in table A1).

Table A1. Time periods rated in Appendix Study W4.

	Boundary-expanded	Boundary-compressed
Less likely to be rounded	6:30pm – 10:00pm 9:30pm – 10:00pm 9:30am – 12:00pm 11:30am – 3:00pm 1:30pm – 3:00pm	7:00pm – 10:30pm 9:00pm – 9:30pm 9:00am – 11:30am 12:00pm – 3:30pm 2:00pm – 3:30pm
More likely to be rounded	6:40pm – 10:10pm 9:40pm – 10:10pm 9:45am – 12:15pm 11:45am – 3:15pm 2:45pm – 4:15pm	7:20pm – 10:50pm 9:15pm – 9:45pm 10:10am – 12:40pm 12:25pm – 3:55pm 3:15pm – 4:45pm

Results. A linear mixed-effects regression specified fixed effects for time period type (boundary-expanded vs. compressed), rounding type (more vs. less susceptible), and their interaction, and a random effect for participant. Overall ratings did not differ between periods that were more vs. less susceptible to rounding (F(1, 2307.2) = 2.70, p = .10). As before, participants rated boundary-expanded time periods to feel longer than boundary-compressed (*M*s = 35.96 and 31.51; *SD*s = 23.97 and 24.14, respectively; F(1, 2308.1) = 33.14, p < .001). Crucially, there was no interaction with rounding type; this effect held to the same degree for periods that were more versus less susceptible to rounding (F(1,2307.1) = .035, p = .85). That is, boundary-expanded periods felt longer than boundary-compressed even for periods where rounding would produce the opposite pattern. Figure A2 displays these results (see also table B12).

Figure A2. Boundary-expansiveness versus susceptibility to rounding. Error bars represent the standard error of the mean.



Discussion. Studies W3 and W4 ruled out two possible explanations for the effect of boundary-expansiveness on perceived duration. The results of study W3 suggest that the effect is not driven by participants not paying attention and miscalculating duration. Even after calculating and explicitly indicating that both periods had the same duration, participants disproportionately selected boundary-expanded periods to "feel longer" than boundary-compressed. Study W4 suggests that it is not likely that the previously documented effects of boundary expansiveness result from participants rounding the times in specific ways. Even in situations where sensible rounding would reduce duration (e.g., 3:55 – 5:05 would be rounded to 4-5), participants still felt that boundary-expanded time periods lasted longer.

Study W5

While the previous studies focused on boundaries occurring on the hour, the present study examined whether months also provide boundaries (e.g., January, February, etc.).

Participants and procedure. One hundred and forty-one MTurk workers participated in this study. Just as in study W1, participants were presented with pairs of time periods and selected which one felt longer. However, time periods spanned months instead of hours. For example, a period like 03-11 to 03-31 would be boundary-expanded, while 03-14 to 04-04 would be boundary-compressed. On half the trials, the expanded period was earlier than the compressed; on the other half, it was later. Participants were informed of the format (Month-Day) and given an example before starting. Participants completed a total of 14 trials, although we removed one trial due to experimenter error.

Results. We subjected the data to a logistic mixed-effects model, regressing choice on a random effect of participant and trial. Time periods were selected to feel longer more often when they were boundary-expanded compared to boundary-compressed, $z = 6.91$, $p < .001$. Expanded periods were selected approximately 73% of the time. Table B13 displays the results broken down by each period.

Study W6

Time periods were presented numerically in all of our studies, although study 3a used a non-numeric category. Can the effect emerge when the presentation is fully visual, such as when one reads a clock? Appendix study W6 contrasts time periods presented a digital (i.e., numeric) format from time periods presented in an analog (i.e., clock-based) format. Because telling time requires the perceiver to mentally convert the visual time to a numeric time, we predicted that the effect—boundary-expanded times feeling longer than boundary-compressed—would emerge for both mediums.

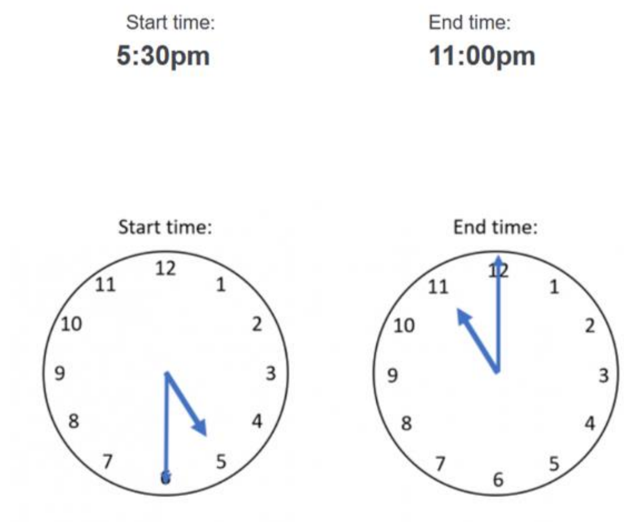
Participants and procedure. Three hundred and three workers from MTurk participated in this study. Following our pre-registered criteria, we excluded participants who failed the attention check as well as those in the analog condition who failed to correctly tell time on an initial test clock. We also eliminated participants whose rating of a short practice trial exceeded their rating of the long practice trial. This left a final sample of 288.

Just as in study W2 (and others), participants completed a series of trials in which they rated their perceived length of various time periods on a sliding scale. Participants were randomly assigned to view time periods that were expressed in either an analog or digital format. That is, half of the participants viewed periods in numeric format, and the other half viewed periods that were depicted by two clocks (one indicating start time and one indicating end time). Importantly, these analog periods always specified that both times were in either am or pm (the numeric times matched this categorization). Boundary-expansiveness varied within participant, just as before, such that half the periods were expanded, and half were compressed. Participants were given a short and a long practice trial before starting to give them a sense of the scale. Figure A3 displays example stimuli.

Figure A3. Example stimuli, digital (A) and analog (B) formats.

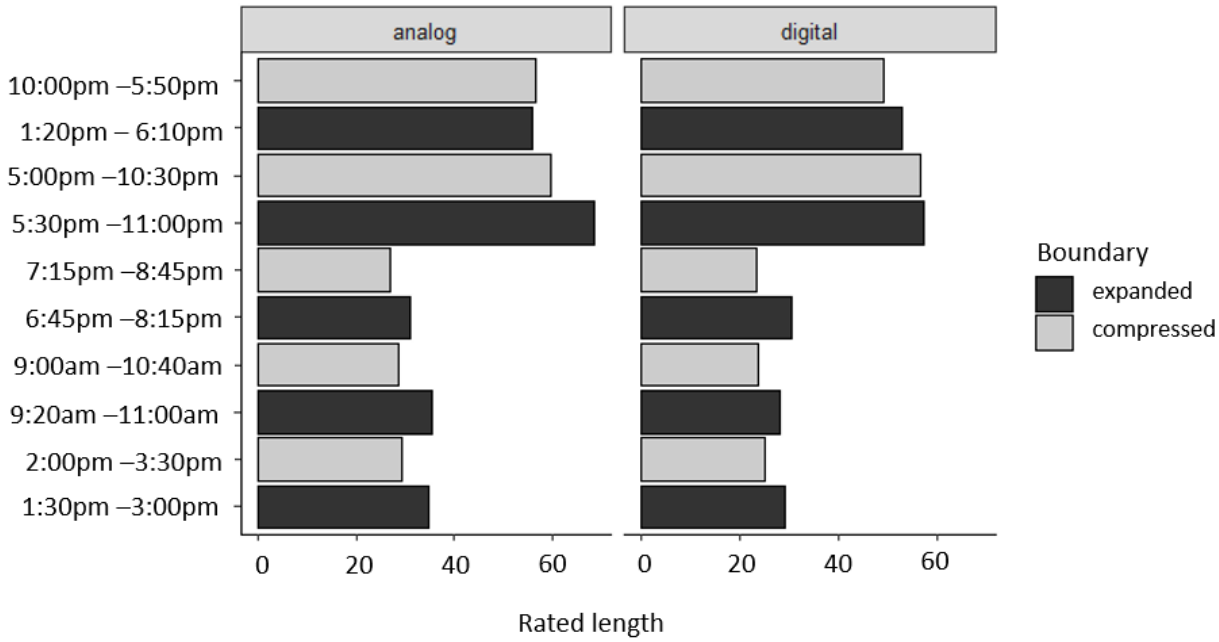
A)

B)



Results. We conducted an ANOVA on a mixed-effects model that regressed ratings of perceived length on a fixed effect of format (digital vs. analog), a fixed effect of boundary type (expanded vs. compressed), the format by boundary type interaction, a fixed effect of trial, and a random effect of participant. We observe a significant effect of format, such that periods were rated as longer when they were presented on clocks, *M*s = 42.68 and 37.69, *SD*s = 24.09 and 24.58 for analog and digital, respectively; $F(1,286) = 7.96$, $p = .005$. An effect of boundary-type emerged such that expanded periods ($M = 42.61$, $SD = 24.47$) were rated as feeling longer than compressed ($M = 38.08$, $SD = 24.21$), $F(1,2586) = 79.13$, $p < .001$. The effect appears to be just as strong with numbers as it is with clocks, as we found no interaction between format and expansiveness, $F(1,2586) = .908$, $p = .34$. The results broken down by period are displayed in figure A4.

Figure A4. Ratings of "how long does it feel" for time periods shown in analog versus digital format.



Study W7

This study describes time periods using words instead of digits. We used wording that removed the influence of left-to-right processing. Specifically, each type of period referenced the same hour categories—for example, "quarter to two to four o'clock" would be a boundary-compressed period, and "quarter past two to half past four" its compressed counterpart. If perceivers were simply anchoring on the largest value they saw when characterizing the period, we would not expect boundary-expansiveness to affect perceived duration. Moreover, much like studies 2a and 2b, this study tests for the influence of left-to-right processing; we would not expect to observe the effect if perceivers anchor on the leftmost information.

Participants and procedure. Four hundred seventy-nine participants from Prolific Academic participated in this study. Due to an error in the pre-registered attention check question, none were excluded.

Participants saw four distinct periods, once in boundary-expanded form and once in boundary compressed for a total of 8 ratings. Just as they did in several other studies, participants rated each period on a 100-point slider for “how long does this period feel?”. A full list of periods is displayed in Table A5.

Table A5. Time periods rated in Appendix Study W7.

Boundary-expanded	Boundary-compressed
Quarter to two to four o’clock	Quarter past two to half past four
Half past eight to quarter past nine	Eight o’clock to quarter to nine
Quarter to noon to three o’clock	Quarter past noon to half past three
Half past three to six o’clock	Quarter past three to quarter to six

Results. We conducted a mixed-effects model that regressed ratings of perceived length on a fixed effect of boundary type (expanded vs. compressed), specifying random effects of participants and question. The results indicate that once again, boundary-expanded periods ($\underline{M}$ = 45.30, *SD* = 25.21) were rated to feel longer than boundary-compressed ($\underline{M}$ = 43.73, *SD* = 25.46), *t* = 3.16, *p* = .002. Thus, it does not appear to be the case that neither left-to-right processing nor only paying attention to the largest value explains the effect.

Study W8

Although the previous studies presented explicit start and end times, periods can also be framed in terms of start time and duration without mentioning end time (e.g., a 30-minute meeting at 2:00). Hypothesis 4 asserts that the effect found in the previous studies requires salient boundaries. If so, it should be attenuated when explicit end times are omitted. To test this, study 3b examined boundary expansiveness for periods without a specified, salient end time.

Participants and procedure. Eight hundred sixty-five Amazon MTurk workers participated in the study. As preregistered, we eliminated any participants who rated the short (10 minute) practice period as feeling longer than the long (8 hour) practice period (N = 84) and/or who failed our attention check (N = 30), leaving a final sample of 751.

Five unique durations, framed as unspecified events that participants had to attend, were presented once as boundary-expanded and once as boundary-compressed for a total of ten trials. Participants rated each period on a 100-point slider for "How long does this event feel?". Between-subjects, we manipulated the salience of hour boundaries. Those in the *salient boundary* condition saw an end time in parentheses:

> "Imagine that you have to attend an event. It starts at {*start time*} and ends {*duration*} hours later (at {*end time*})."

This information was absent for those in the *non-salient boundary* condition:

> "Imagine that you have to attend an event. It starts at {*start time*} and ends {*duration*} hours later."

Start time and duration (italicized above) formed boundary-compressed and boundary-expanded periods. For example, a duration of 1.5 hours produces an expanded period when it starts at 10:30 (ending at 12:00pm), and a compressed period when it starts at 11:00 (ending at 12:30pm).

Results. A mixed-model ANOVA specified fixed effects for boundary expansiveness, boundary salience and their interaction, with random effects for participant and trial. A main effect of boundary salience found that periods were rated as longer when end times were absent (M = 47.94, SD = 23.61) versus present (M = 44.55, SD = 21.91; F(1,749) = 8.72, p = .003, d =

.16). Moreover, boundary-expanded periods (*M* = 46.44, *SD* = 22.87) were rated as marginally longer than compressed (*M* = 45.84, *SD* = 22.69; $F(1,6753) = 3.44$, $p = .064$, $d_z = .10$).

Crucially, we observed the anticipated significant interaction between boundary expansiveness and boundary salience ($F(1,6753) = 10.28$, $p = .001$). As predicted, when boundaries were salient, boundary-expanded periods (*M* = 45.29, *SD* = 22.09) were rated as feeling longer than compressed (*M* = 43.81, *SD* = 21.70; $t = 3.63$, $p = .0003$, $d_z = .23$), but there was no difference in ratings when boundaries were not salient ($t = -.948$, $p = .343$, $d_z = -.07$).

Discussion. Study 3b found that minimizing the salience of hour boundaries eliminated the effect of boundary-expansiveness on perceived duration. That is, boundary-expanded time periods only felt longer when an end time was salient. Two design features underscore the strength of these results. First, the boundary salience manipulation was extremely subtle: the presence or absence of a small amount of information shown in parentheses. Second, participants were provided with the duration of every period without needing to approximate it themselves.

WEB APPENDIX B: TRIAL-LEVEL INFORMATION

Study 1

Table B1. Estimates of how many HITs can be completed, broken down by period.

Expanded	*HITs*	*Compressed*	*HITs*
11:45am – 1:00pm	*M* = 56.63 *SD* = 73.83	12:00pm – 1:15pm	*M* = 37.68 *SD* = 52.34
9:30am – 12:00pm	*M* = 82.11 *SD* = 87.28	10:00am – 12:30pm	*M* = 65.29 *SD* = 77.40
2:40pm – 6:10pm	*M* = 88.89 *SD* = 91.92	3:10pm – 6:40pm	*M* = 70.47 *SD* = 79.34
1:45pm – 3:15pm	*M* = 59.96 *SD* = 73.66	1:15pm – 2:45pm	*M* = 42.45 *SD* = 57.55
10:25am – 2:10pm	*M* = 88.29 *SD* = 88.24	10:05am – 1:50pm	*M* = 75.40 *SD* = 78.06

Study 2a

Table B2. Ratings of perceived length of periods spanning months from UK participants, broken down by period.

Expanded	*Rating*	*Compressed*	*Rating*
06-01-2021 to 02-02-2021	*M* = 26.16 *SD* = 19.18	04-01-2021 to 31-01-2021	*M* = 21.99 *SD* = 19.29
23-06-2021 to 03-09-2021	*M* = 47.24 *SD* = 24.17	17-06-2021 to 28-08-2021	*M* = 40.10 *SD* = 22.67
30-10-2021 to 09-11-2021	*M* = 14.90 *SD* = 15.83	20-10-2021 to 30-10-2021	*M* = 12.21 *SD* = 13.40
15-03-2021 to 10-04-2021	*M* = 27.39 *SD* = 19.03	02-04-2021 to 28-04-2021	*M* = 20.09 *SD* = 16.37

29-09-2021 to 06-11-2021	*M* = 38.29 *SD* = 23.78	05-10-2021 to 12-11-2021	*M* = 31.07 *SD* = 20.29

Study 2b

Table B3. Ratings of perceived length broken down by period and condition (same vs. adjacent start and end year). Note that start and end year were drawn from either the past 20 or future 20 years.

Expanded	Rating *(same start and end year*	Rating *(adjacent start and end year)*	Compressed	Rating *(same start and end year)*	Rating *(adjacent start and end year)*
10-30-{*start year*} to 12-10-{*end year*}	*M* = 28.96 *SD* = 22.85	*M* = 50.12 *SD* = 25.81	10-20-{*start year*} to 11-30-{*end year*}	*M* = 22.79 *SD* = 22.41	*M* = 51.18 *SD* = 25.83
03-31-{*start year*} to 05-11-{*end year*}	*M* = 29.85 *SD* = 23.80	*M* = 48.30 *SD* = 24.82	03-19-{*start year*} to 04-29-{*end year*}	*M* = 24.23 *SD* = 23.85	*M* = 45.10 *SD* = 25.50
07-27-{*start year*} to 09-06-{*end year*}	*M* = 29.28 *SD* = 23.93	*M* = 50.98 *SD* = 25.53	07-17-{*start year*} to 08-27-{*end year*}	*M* = 24.24 *SD* = 23.19	*M* = 47.80 *SD* = 24.37
05-29-{*start year*} to 07-09-{*end year*}	*M* = 31.30 *SD* = 24.03	*M* = 48.85 *SD* = 25.13	06-01-{*start year*} to 07-12-{*end year*}	*M* = 24.12 *SD* = 23.07	*M* = 45.69 *SD* = 25.07

Study 3a

Table B4. Ratings broken down by number of classes and hours spanned.

	Spans 2 classes	*Spans 3 classes*
Spans 2 hours	*M* = 64.46	*M* = 66.92

	SD = 23.74	*SD* = 23.90
Spans 3 hours	*M* = 64.60 *SD* = 25.35	*M* = 68.98 *SD* = 22.94

Study 3b

Table B5. Ratings of perceived length broken down by trial and condition (boundaries salient vs. not).

Condition	*Expanded*	*Rating*	*Compressed*	*Rating*
Salient end time	10:30am – 1:00pm	*M* = 38.62 *SD* = 17.98	10:00am – 12:30pm	*M* = 36.31 *SD* = 18.37
	2:45pm – 4:30pm	*M* = 32.87 *SD* = 19.09	3:00pm – 4:45pm	*M* = 32.51 *SD* = 19.61
	1:30pm – 6:00pm	*M* = 64.30 *SD* = 16.59	1:00pm – 5:30pm	*M* = 62.22 *SD* = 17.09
	7:30pm – 10:00pm	*M* = 49.20 *SD* = 23.20	7:00pm – 9:30pm	*M* = 46.52 *SD* = 21.84
	2:45pm – 5:00pm	*M* = 41.47 *SD* = 18.78	3:00pm – 5:15pm	*M* = 41.47 *SD* = 18.22
No salient end time	10:30am – 1:00pm	*M* = 39.04 *SD* = 20.23	10:00am – 12:30pm	*M* = 38.54 *SD* = 20.38
	2:45pm – 4:30pm	*M* = 35.32 *SD* = 20.22	3:00pm – 4:45pm	*M* = 36.11 *SD* = 20.70
	1:30pm – 6:00pm	*M* = 68.10 *SD* = 17.89	1:00pm – 5:30pm	*M* = 68.64 *SD* = 17.78
	7:30pm – 10:00pm	*M* = 52.88 *SD* = 23.44	7:00pm – 9:30pm	*M* = 52.10 *SD* = 23.13
	2:45pm – 5:00pm	*M* = 43.41 *SD* = 20.63	3:00pm – 5:15pm	*M* = 45.33 *SD* = 20.11

Study 4

Table B6. Frequency of selecting boundary-expanded and compressed periods as a function of activity.

Activity Type	*Item*	*Expanded*	*N*	*Compressed*	*N*
Time-minimizing	Lunch with someone you dislike	11:30am – 1:00pm	196	12:00pm – 1:30pm	101
		12:30pm – 2:00pm	80	12:00pm – 1:30pm	223
		Total:	**276**	*Total*:	**324**
	Getting blood drawn	9:45am – 10:00am	183	10:00am – 10:15am	95
		10:45am – 11:00am	86	10:00am – 10:15am	236
		Total:	**269**	*Total*:	**331**
	Dental cleaning	12:30pm – 2:00pm	171	1:00pm – 2:30pm	145
		1:30pm – 3:00pm	90	1:00pm – 2:30pm	194
		Total:	**261**	*Total*:	**339**
	DMV appointment	12:30pm – 1:00pm	157	1:00pm – 1:30pm	146
		1:30pm – 2:00pm	102	1:00pm – 1:30pm	195
		Total:	**259**	*Total*:	**341**
Time-maximizing	Free time	2:30pm – 4:00pm	172	2:00pm – 3:30pm	133
		2:30pm – 4:00pm	155	3:00pm – 4:30pm	140
		Total:	**327**	*Total*:	**273**
	Exploring on a work trip	12:30pm – 2:00pm	202	12:00pm – 1:30pm	93
		12:30pm – 2:00pm	114	1:00pm – 2:30pm	191
		Total:	**316**	*Total*:	**284**
	Watching favorite show	7:30pm – 9:00pm	176	7:00pm – 8:30pm	122
		7:30pm – 9:00pm	109	8:00pm – 9:30pm	193
		Total:	**285**	*Total*:	**315**
	Taking a nap	2:15pm – 3:00pm	136	2:00pm – 2:45pm	170
		2:15pm – 3:00pm	192	3:00pm – 3:45pm	102
		Total:	**328**	*Total*:	**272**

Study 5

Table B7. Flight Scenario: Compensation required to endure boundary-expanded and compressed waiting periods broken down by period.

Expanded period	*Compensation*	*Compressed period*	*Compensation*
9:47am – 2:12pm	*M* = 245.14 *SD* = 187.33	9:24am – 1:49pm	*M* = 227.55 *SD* = 171.69
10:47am – 3:12pm	*M* = 258.77 *SD* = 190.14	11:24am – 3:49pm	*M* = 2.42.07 *SD* = 180.17

Table B8. Bus Scenario: Willingness-to-pay to avoid boundary-expanded and compressed waiting periods broken down by period.

Expanded period	*WTP*	*Compressed period*	*WTP*
10:40am – 5:05pm	*M* = 62.07 *SD* = 49.23	11:15am – 5:40pm	*M* = 59.36 *SD* = 53.24
12:40pm – 7:05pm	*M* = 66.94 *SD* = 55.15	12:15pm – 6:40pm	*M* = 62.27 *SD* = 50.78

Supplemental Studies

Study W1

Table B9. Frequency of selecting each period as feeling longer by trial.

Expanded	*N*	*Compressed*	*N*
1:30pm – 6:00pm	76	1:00pm – 5:30pm	28
10:30am – 1:00pm	73	10:00am – 12:30pm	31
2:45pm - 4:30pm	60	3:00pm - 4:45pm	44
2:45pm – 5:00pm	65	3:00pm – 5:15pm	39

7:30pm – 10:00pm	79	7:00pm – 9:30pm	25

Study W2

Table B10. Ratings of perceived length broken down by period.

Expanded	*Rating*	*Compressed*	*Rating*
1:30pm – 6:00pm	*M* = 52.49 *SD* = 19.92	1:00pm – 5:30pm	*M* = 48.57 *SD* = 21.20
10:30am – 1:00pm	*M* = 42.64 *SD* = 22.89	10:00am – 12:30pm	*M* = 38.66 *SD* = 24.45
2:45pm - 4:30pm	*M* = 37.42 *SD* = 23.21	3:00pm - 4:45pm	*M* = 31.86 *SD* = 26.05
2:45pm – 5:00pm	*M* = 42.27 *SD* = 23.60	3:00pm – 5:15pm	*M* = 36.66 *SD* = 23.40
7:30pm – 10:00pm	*M* = 40.86 *SD* = 24.36	7:00pm – 9:30pm	*M* = 34.94 *SD* = 21.83

Study W3

Table B11. After correctly computing duration, frequency of selecting each period as feeling longer by trial.

Expanded	*N*	*Compressed*	*N*
1:30pm – 6:00pm	106	1:00pm – 5:30pm	29
10:30am – 1:00pm	101	10:00am – 12:30pm	32
2:45pm - 4:30pm	91	3:00pm - 4:45pm	39
2:45pm – 5:00pm	87	3:00pm – 5:15pm	44
7:30pm – 10:00pm	112	7:00pm – 9:00pm	22

Study W4

Table B12. Ratings of perceived length broken down by period. Times marked with an asterisk were used in the exploratory analysis reported in the Discussion.

	Boundary-expanded		*Boundary-compressed*	
Type	*Time*	*Rating*	*Time*	*Rating*
Opposing roundability	2:45pm – 4:15pm	*M* = 35.30 *SD* = 23.69	3:15pm – 4:45pm	*M* = 31.01 *SD* = 23.45
	11:45am – 3:15pm*	*M* = 48.69 *SD* = 20.56	12:25pm – 3:55pm*	*M* = 44.16 *SD* = 22.00
	9:45am – 12:15pm	*M* = 37.93 *SD* = 20.71	10:10am – 12:40pm	*M* = 35.17 *SD* = 20.98
	9:40pm – 10:10pm	*M* = 18.54 *SD* = 21.80	9:15pm – 9:45pm	*M* = 13.37 *SD* = 19.26
	6:40pm – 10:10pm*	*M* = 41.90 *SD* = 22.12	7:20pm – 10:50pm*	*M* = 37.59 *SD* = 22.99
“Regular” roundability	1:30pm – 3:00pm	*M* = 34.23 *SD* = 22.07	2:00pm – 3:30pm	*M* = 28.64 *SD* = 22.45
	11:30am – 3:00pm*	*M* = 49.29 *SD* = 20.54	12:00pm – 3:30pm*	*M* = 43.00 *SD* = 21.87
	9:30am – 12:00pm	*M* = 37.65 *SD* = 20.11	9:00am – 11:30am	*M* = 30.76 *SD* = 22.73
	9:30pm – 10:00pm	*M* = 14.51 *SD* = 19.14	9:00pm – 9:30pm	*M* = 14.87 *SD* = 22.28
	6:30pm – 10:00pm*	*M* = 41.45 *SD* = 23.34	7:00pm – 10:30pm*	*M* = 36.66 *SD* = 22.58

Study W5

Table B13. For periods that spanned months instead of hours, frequency of selecting each period as feeling longer by trial.

Temporal order	*Expanded*	*N*	*Compressed*	*N*
Expanded before compressed	5/31 – 6/1	102	6/7 – 6/8	31
	11/29 – 12/1	97	12/6 – 12/8	35
	1/31 – 2/3	91	2/7 – 2/10	43
	4/26 – 5/1	99	5/3 – 5/8	33
	7/27 – 8/2	99	8/3 – 8/9	36
	3/28 – 4/4	90	4/11 – 4/18	42
Expanded after compressed	5/31 – 6/1	104	5/24 – 5/25	30
	11/29 – 12/1	103	11/22 – 11/24	31
	1/31 – 2/3	97	1/24 – 1/27	36
	4/26 – 5/1	91	4/19 – 4/24	42
	7/27 – 8/2	95	7/20 – 7/26	40
	3/28 – 4/4	89	3/14 – 3/21	45
	9/30 – 10/4	102	9/23 – 9/27	31

Study W7

Table B14. Ratings of perceived length broken down by period.

Expanded	*Rating*	*Compressed*	*Rating*
Quarter to two to four o'clock	*M* = 47.37 *SD* = 23.40	Quarter past two to half past four	*M* = 46.38 *SD* = 22.81
Half past eight to quarter past nine	*M* = 26.12 *SD* = 20.98	Eight o'clock to quarter to nine	*M* = 23.06 *SD* = 20.23
Quarter to noon to three o'clock	*M* = 55.81 *SD* = 22.72	Quarter past noon to half past three	*M* = 53.86 *SD* = 22.83

Half past three to six o'clock	*M* = 51.89 *SD* = 22.68	Quarter past three to quarter to six	*M* = 51.62 *SD* = 23.36

Study W8

Table B15. Ratings of perceived length broken down by period.

Salient Boundary	*Expanded*	*Rating*	*Compressed*	*Rating*
End time salient	10:30am – 1:00pm	*M* = 38.62 *SD* = 17.98	10:00am - 12:30pm	*M* = 36.31 *SD* = 18.37
	2:45pm – 4:30pm	*M* = 32.87 *SD* = 19.09	3:00pm – 4:45pm	*M* = 32.51 *SD* = 19.62
	1:30pm – 6:00pm	*M* = 64.30 *SD* = 16.59	1:00pm – 5:30pm	*M* = 62.22 *SD* = 17.09
	7:30pm – 10:00pm	*M* = 49.20 *SD* = 23.20	7:00pm – 9:30pm	*M* = 46.52 *SD* = 21.84
	2:45pm – 5:00pm	*M* = 41.47 *SD* = 18.78	3:00pm – 5:15pm	*M* = 41.47 *SD* = 18.22
No salient end time	10:30am – 1:00pm	*M* = 39.04 *SD* = 20.23	10:00am - 12:30pm	*M* = 38.54 *SD* = 20.37
	2:45pm – 4:30pm	*M* = 35.32 *SD* = 20.22	3:00pm – 4:45pm	*M* = 36.11 *SD* = 20.70
	1:30pm – 6:00pm	*M* = 68.10 *SD* = 17.89	1:00pm – 5:30pm	*M* = 68.65 *SD* = 17.78
	7:30pm – 10:00pm	*M* = 52.88 *SD* = 23.44	7:00pm – 9:30pm	*M* = 52.10 *SD* = 23.13
	2:45pm – 5:00pm	*M* = 43.41 *SD* = 20.63	3:00pm – 5:15pm	*M* = 45.33 *SD* = 20.11

WEB APPENDIX C: ALTERNATE ANALYSES

Treating trial as a fixed effect

The following analyses specify trial as a fixed effect (instead of random) for studies where it is appropriate to do so.

Study 1. Participants estimated that they would be able to complete more hits during boundary-expanded periods than boundary-compressed: $z = 3.06$, $p = .002$.

Study 2a. Participants rated boundary-expanded periods to feel longer than boundary-compressed, $t = 8.00$, $p < .001$.

Study 2b. Participants rated boundary-expanded periods to feel longer than boundary-compressed, $F(1,1779.94) = 48.95$, $p < .001$. The effect of boundary-expansiveness on perceived duration did not differ by temporal location (future vs. past), $F(1,1779.94) = 0.617$, $p = .432$. There only significant interaction was between boundary type and year condition, $F(1,1779.94) = 11.09$, $p = .0009$.

Analysis without exclusions

Study 1. MTurk workers considering boundary-expanded time periods estimated that they could complete more HITs ($M = 75.63$, $SD = 87.53$) than those considering boundary-compressed periods ($M = 59.64$, $SD = 77.58$); $z = 3.12$, $p = .002$.

Study 2a. Boundary-expanded periods ($M = 31.98$, $SD = 23.73$) were rated to feel longer than compressed ($M = 26.48$, $SD = 22.04$; $t = 9.32$, $p < .001$.

Studies 2b and *3* had no exclusions.

Study 4. The predicted interaction emerges such that choice of period (boundary-expanded vs. compressed) differed between the two types of activities; $z = -5.35$, $p < .001$. For time-maximizing activities, participants disproportionately selected the boundary-expanded periods over the compressed (52% vs. 48%; $z = 2.23$, $p = .026$). For time-minimizing activities, they chose boundary-compressed periods more often (56% vs. 44%; $z = -5.35$, $p < .001$).

Study 5. In the bus scenario, WTP for avoiding the wait was higher when it was boundary-expanded ($M = 86.56$, $SD = 80.02$) compared to boundary-compressed ($M = 78.59$, $SD = 75.14$); $t = 3.27$, $p = .001$. In the flight scenario, required compensation to endure the wait was higher for boundary-expanded waits ($M = 334.42$, $SD = 244.23$) compared to boundary-compressed ($M = 320.95$, $SD = 254.77$); $t = 2.18$, $p = .030$.

Does the Effect Differ Based on Duration?

Below, we present the results of exploratory analyses designed to test whether the effect of boundary-expansiveness differs by duration. That is, does the effect differ for longer and shorter periods? To answer this question, we re-run the models reported in the main text. We enter duration into an interaction with boundary type (expanded vs. compressed) for each study where such an analysis is possible.

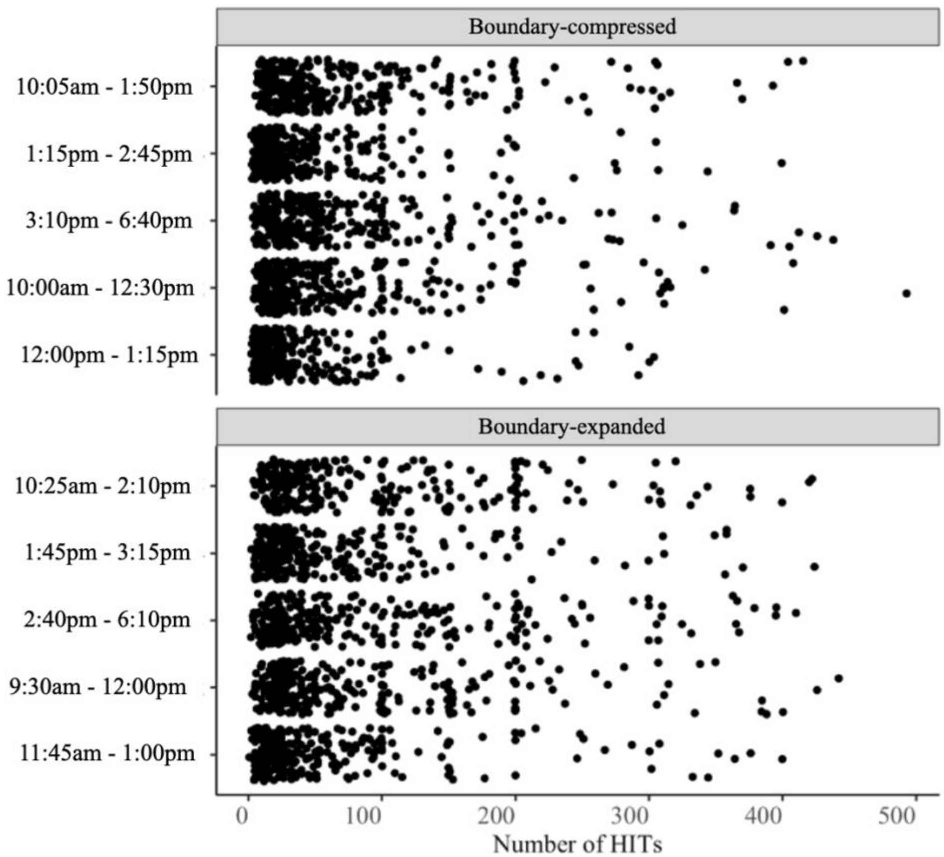
Table C1. Effects of boundary-expansiveness by duration.

Study	*Levels of duration (in hours)*	*Interaction between boundary-type and duration*	*Conclusion*
A2	1.75, 2.25, 2.5, 4.5	$F(1,917.17) = .32$, $p = .57$	No difference
A4	.5, 1.5, 2.5, 3.5	$F(1,2304.2) = .90$, $p = .44$	No difference
1	1.25, 1.5, 2.5, 3.5, 3.75	$z = -5.48$, $p < .001$	The difference between boundary-expanded and compressed decreases as duration increases.

Alternate graph for study 1

Study 1 employs negative binomial regression to analyze how the number of HITs varied by boundary-expansiveness. This decision was informed, in part, by the overdispersion present in the response distribution (shown in Figure C1).

Figure C1. Distribution of responses in study 1.



WEB APPENDIX D: MORE INFORMATION ABOUT THE RIDESHARE STUDY

Alternate analyses and specifications

We present three alternate analyses of the rideshare data. In the first, we run the same model presented in the paper with linear regression instead of logistic. In the second, we set any unrealistic values—specifically, negative values for "mixed" probability or duration difference, or positive values for fare difference—to zero. The reason these values might arise in the first place is that we are approximating the characteristics of each ride type, as discussed in the main text. The third analysis shows results without any of our exclusion criteria. This means that a trip could have any duration (instead of more than 3 minutes and less than 60) and span any distance (instead of a maximum of 30 miles). Moreover, trips were included if they had at least one other trip in the dataset with the same route, start hour, and ride type (as opposed to at least 100 other trips; a minimum of one is required for modeling purposes). Lastly, trips that were two SDs above or below the average duration of their similar trips were not removed. In the fourth analysis, we examine trips taken between 8am and 6pm (rather in the middle of the night). The fifth analysis controls for holidays and sporting events; specifically, Cubs games and Bears games.. Results are presented in table D1.

Table D1. Alternate analyses of rideshare data in Study 6.

		P_{Diff}	D_{Diff}	C_{Diff}
Coefficient estimates under linear regression	Estimate	-0.068***	-0.036***	-0.006***
	Standard Error	0.0048	0.0003	0.0002
Coefficient estimates with corrections to zero (logistic regression)	Estimate	-0.538***	-0. 041***	-0. 235***
	Standard Error	0.0351	0.0017	0.0022

Coefficient estimates with no exclusions	Estimate	-0.358***	-0.027***	-0.221***
	Standard Error	0.0335	0.0014	0.0016
Coefficient estimates during a different time of day (rides between 8am and 6pm)	Estimate	-0.199***	0.080***	-0.068***
	Standard Error	0.0399	0.0017	0.0022
Coefficient estimates accounting for holidays	Estimate	-0.520***	-0.0350***	-0.220***
	Standard Error	0.0349	0.0017	0.0019

***$p < 0.001$

More detail about variables

Below, we list all variables present in the dataset. Full information can be found at

https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

- Start and end time (date, hour, minute) rounded to the nearest 15 minutes
- Trip duration (in seconds)
- Distance travelled (in miles)
- Whether or not the rider requested to share their ride
- If they did request to share, whether they were actually matched with another rider
- The census tracts and community areas in which the trip began and ended
- The start and end location coordinates (as the centroid of the pickup census tract)
- Fare (in dollars, rounded to the nearest $2.50)
- Tip (in dollars, rounded to the nearest $1.00)
- Additional charges (taxes, fees and any other charges for the trip)
- Trip total (sum of fare, additional charges, and tip)

Information about the ride the consumer actually chose

If the consumer had selected the independent ride, the probability that they had faced a mixed choice set was .05, and the probability that their ride actually crossed into a new hour was .34. If they had selected the shared ride, the probability that they had faced a mixed choice set was .04, and the probability that the ride had actually crossed into a new hour was .35.

Additional Figures and Tables

Table D2. Coefficient estimates, standard errors and p-values for tip analysis. The “Diff.Expectations” term reflects the likelihood that the trip was forecasted to be boundary-expanded, but was actually-boundary-compressed.

	Estimate	Standard Error
Diff.Expectations	0.4061***	0.00845
Duration	0.0234***	0.00025
Fare	0.0341***	0.00022

***$p < .001$.

Figure D1. Change in expected revenues under alternative pricing policies.