Time Periods Feel Longer When They Span More Category Boundaries:

Evidence From the Lab and the Field

Five experiments (N = 2,478) 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 hour marks. For example, periods like 1:45pm – 3:15pm (boundary-expanded) feel longer than 1:15pm – 2:45pm (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 hour 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

Consider a consumer who is looking to book a flight to a distant country. She narrows in on two flights that have the same duration but differ slightly in their departure and arrival times. One flight departs at 10:10am and arrives at 11:28pm. The other departs at 9:50am and arrives at 11:08pm. Might their departure and arrival times make one flight *feel* shorter (and thus more appealing) than the other, even though their duration is the same? The present research explores this question, investigating how start and end times affect perceived duration and subsequent consumer decision-making.

We propose that time periods feel longer when they span more category boundaries (e.g., whole hours). For example, a period like 1:45pm – 3:15pm—which spans the 1:00, 2:00, and 3:00 hour markers—encompasses the maximum number of distinct hours possible given its duration. We will call such periods "boundary-expanded". In contrast, a period like 1:15pm – 2:45pm, despite being of equal duration, crosses the minimum number of hour markers (only 1:00 and 2:00); these periods will be hereafter called "boundary-compressed". Based on research showing that stimuli in different categories seem more different from each other than stimuli in the same category (e.g., Goldstone and Hendrickson, 2010), we propose that whole hour markers make points in time feel more or less similar to each other (and thus, closer together or further apart). Specifically, given that boundary-expanded time periods cross more hours, we would expect consumers to perceive those periods to feel longer than equivalent boundary-compressed periods. We test this prediction, the underlying process, and its consequences in five preregistered experiments, together involving over 2,000 participants. Additionally, we test the implications of our theoretical model using field data of over 1.8 million rideshare choices in the Chicago metropolitan area. This analysis reveals that boundary-expansiveness reliably affects choices for independent, non-shared rides over shared rides (e.g., Uber X vs. Uber Pool).

Additionally, we estimate that consumers would be willing to pay around $0.60 per ride to avoid crossing hour boundaries, which suggests that rideshare apps could benefit from incorporating this insight in their pricing strategies. We find that simply charging more for independent rides and less for shared rides, when the former are less likely to be boundary-expanded than the latter, yields over $0.30 in extra expected revenue per ride. This translates to more than 3.5 million dollars of additional revenue in Chicago per year. An additional six studies reported in the online appendix exclude alternative explanations and provide further robustness checks.

Our findings make important theoretical and applied contributions. We are unaware of any previous research that investigates how the number of distinct hours spanned by a given period affect its perceived duration. Yet this question is important for consumer decision-making, as is the study of boundaries and time perception more generally (e.g., Tonietto, Malkoc and Nowlis 2019; Zauberman, Levav, Diehl and Bhargave 2010). 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). The present research provides valuable insight into the ways in which 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.

THEORETICAL BACKGROUND

Boundaries and Perceived Difference

Despite time being a continuous variable, consumers appear to naturally think about time categorically, as if chunks were separated into distinct categories. For example, people often use loose groupings to conceptualize time, such as “morning,” “afternoon,” and “evening”, or even simply “like the present” or “not like the present” (Tu and Soman 2014). Moreover, they naturally organize the narrative of their lives into chapters (Skowronski et al. 2007; Thomsen 2009) and structure their autobiographical memories (Shum 1998) and identity (Peetz and Wilson 2013) around significant temporal landmarks such as birthdays, holidays, or public events. Such demarcations also affect motivation. People are more committed to their goals upon entering a new year, month or week, for example, because the category change signals a new period—a “fresh start” (Dai, Milkman and Riis 2014; Ayers et al. 2014). Thus, while time exists on a continuum, consumers often perceive distinct categories defined by various markers (e.g., the start of a new month).

We propose here that consumers also use hour boundaries to categorize time. For example, 2:00pm marks the beginning of a “2pm” category, much like January 1st begins a new year, or Monday a new week. Thus, times that occur within the same hour (e.g., 2:15pm and 2:45pm) are in the same category, whereas times that occur in different hours (e.g., 2:45pm and 3:15pm) occupy two different categories. As far as we are aware, no research has explicitly treated round hours (e.g., 2:00, 3:00, etc.) as category boundaries. However, round hours have been found to naturally be used as reference points (Allen, Dechow, Pope and Wu 2017), suggesting that they may serve as salient markers of time.

If round hours are indeed perceived as category boundaries, we would expect these boundaries to affect subsequent judgments. Specifically, we would expect time periods crossing more hour markers to feel longer than equivalent periods that cross fewer. This hypothesis is consistent with previous work establishing that across a wide range of domains, the crossing of category boundaries increases perceived distance. For example, geographic locations in different states are perceived to be further apart than equidistant locations within the same state (Burris and Branscombe 2005; Irmak, Naylor and Bearden 2011; Maki 1982; Mishra and Mishra 2010). Temperature changes between two consecutive days seem larger when they fall in different months than when they fall in the same month (Krueger and Clement 1994), the difference in length between two lines seems greater when those lines are given different labels (Tajfel and Wilkes 1963), and people are judged to be more dissimilar when arbitrarily assigned to different social groups (Allen and Wilder 1979; Locksley, Ortiz and Hepburn 1980). Even colors are perceived as more or less similar depending on their categorization. For example, because their language affords two distinct categories for lighter and darker blues, Russian speakers are faster than English speakers to discriminate between such shades (Winawer et al. 2007). Conversely, the Berinmo of Papua New Guinea have difficulty distinguishing between blue and green, as they do not use separate words for each (Roberson, Davies and Davidoff 2000). Linguistic category differences may also affect the perception of shape (Roberson, Davidoff and Shapiro 2002), quantity (Lucy 1992b), and facial expression (Etcoff and Magee 1992).

Besides category boundaries affecting perceptions of distance, they have also been shown to have downstream consequences for consumer judgment. For example, a large pricing literature finds that consumers group prices by their left-most digits; $1.99 feels closer to $1.98 than it does to $2.00 (Manning and Sprott 2009; Thomas and Morwitz 2009). As such, consumers disproportionately prefer products that are priced just below the next highest dollar (Schindler and Kibarian 1996). Consumers also spontaneously form various round-number categories (e.g., “top 10”, “top 20”, etc.) when evaluating items in a ranked list (Isaac and Schindler 2014). A 10th ranked restaurant, for instance, feels closer in quality to the 8th ranked restaurant than it does to the 12th, illustrating once more that category boundaries increase perceived difference. This is also true when stimuli are separated by more (vs. fewer) category boundaries—increasing the number of boundaries between two stimuli makes them feel further apart. This was mentioned earlier for color perception (i.e., separate categories for two adjacent shades make them more discriminable), but it also directly applies to a consumer context. When the unit of measurement changes such that more boundaries punctuate a waiting period, consumers perceive it as longer. For example, a delivery time seems longer when described as 7-14 days instead of 1-2 weeks (Pandelaere, Briers and Lembregts 2011; but see Monga and Bagchi 2011 for conditions where this unit effect reverses).

In sum, category boundaries influence judgments across a wide variety of domains. We expect them to also affect consumer judgments of time: Specifically, spanning more hour boundaries may increase the perceived duration of a given period. Just as the distance from San Francisco to Salem, Oregon feels longer than the equivalent distance from San Francisco to San Diego, we propose that periods that span more hours—and thereby cross more category boundaries—feel longer than periods of equal duration. For example, we expect that 2:35- 3:05, which crosses a boundary (3:00), will feel longer than 3:05 – 3:35, which does not. To test this intuition, we ran a series of pilot studies, reported in the online appendix, in which we found that participants indicated that equivalent time periods felt longer when they crossed more whole-hour boundaries. In what follows, we explore the consequences of boundary-expansiveness on consumer judgment and decisions, and examine the process underlying these decisions.

Overview of Studies

Study 1 tests the basic premise of our theory: boundary-expanded times are perceived to last longer than equivalent boundary-compressed periods. Supporting this prediction, study 1 finds that workers estimate that they can complete more tasks when the exact same time period is boundary-expanded rather than boundary-compressed. Since our theory assumes that the effect relies on *perceived* boundaries, changing the perception of those boundaries should affect judgments. This was tested in study 2a and 2b. Study 2a reveals that removing salient end times eliminates the effect. Additionally, changing the salient categories such that the same time period appears to cross more (fewer) boundaries increases (decreases) its perceived duration (study 2b). Thus, the effect appears to stem from the presence and placement of salient hour boundaries.

The next two studies illustrate possible consequences for consumer decision-making. Study 3 explores scheduling preferences and finds that participants prefer to plan 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 4 examines implications for transportation services; namely, we manipulate boundary-expansiveness of long wait times and elicit consumers' willingness to pay to avoid the long wait as well as their required compensation to endure it.

Finally, we examine transportation data from Chicago to study the real-world impact of boundary-expansiveness on consumer behavior. This collection of rideshare choices from over 1.8 million trips suggests that boundary-expansiveness affects the choice between taking an individual ride and sharing the trip with another rider. Specifically, 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 experimental study, trials were presented in random order, and for studies that involved choosing between various time periods, the visual position of options was also randomized. In addition, because the expression of time varies in different countries, we only recruited participants from the U.S. Our pre-registrations, materials, data and code can be found at https://osf.io/dav53/?view_only=71e175b98e024cf7a321d07ac4ea5d24.

STUDY 1: HIT ESTIMATION

After conducting several pilot studies in which participants indicated how long they believed different time periods felt (reported as studies W1 - W4 in the online appendix), we tested for the basic effect of round hour marks on a more behaviorally valid measure—the number of tasks workers anticipated they could complete within certain time periods. To ensure our participants would have the general knowledge required to make reasonably informed judgments, we chose to have workers from Amazon Mechanical Turk predict numbers of Human Intelligence Tasks (HITs) they expect to be able to complete during various time periods. Following the effect observed in our initial studies, we predicted 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

Six hundred and twelve MTurk workers (53.8% female, 44.9% male, 1.8% other, M_{age} = 37) participated in this study. Per our pre-registration, we excluded those who 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 much 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, participants were asked to estimate how many HITs they thought they could accomplish in a given period, answering on a slider that ranged from “0” to “500 or more.” Following our pre-registration, responses at either of these two 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

We performed a mixed-effects negative binomial regression on number of HITs, specifying a fixed effect of boundary type (expanded vs. compressed) and random effects of question and participant. 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 ($M = 58.23$, $SD = 71.44$; $z = 3.16$, $p = .002$, $d = .22$).

Discussion

Consistent with the notion that time-periods that span more hours feel like they last longer than equivalent time-periods that span fewer hours, MTurk workers anticipated that they would be able to complete about 29% more HITs during time periods that were boundary-expanded rather than boundary-compressed; this was a difference of roughly 17 HITs. Before looking at the consequences for consumer decisions, we first investigate the underlying process.

STUDIES 2A AND 2B

In Study 1, and the studies reported in the online appendix, we found that participants judge time periods to last longer and allow them to complete more work when the periods span more hours. These predictions were based on the assumption that whole hours act as natural categories, and that time periods *perceived* to span these category boundaries are thus perceived to last longer. In the previous study, we used the natural variation of time to make equivalent time periods span more or fewer hour boundaries. However, our theoretical model also implies that merely changing the perception of boundaries should affect the perceived length of time periods. Specifically, if boundaries are not salient, and therefore less perceptible, the effect should be attenuated. Moreover, we suggest that the effect is driven by the *number* of boundaries: Put simply, if period X spans more boundaries than period Y, X should feel longer. According to this logic, if a period like 2:00 – 3:30 spans two boundaries (instead of one) and 1:30 – 3:00 spans one period (instead of two), the former should feel longer than the latter. Because we cannot manipulate the actual placement of whole hour boundaries, we introduce different, salient non-hour boundaries to test this prediction. We propose that if the placement of boundaries is changed, resulting in time periods spanning more or fewer boundaries, judgments should reflect this change. The following two studies test these predictions.

Study 2a

In study 2a, we manipulated the salience of hour boundaries. Although the previous study presented explicit start and end times, time periods can also be framed in terms of start time and duration without mentioning end time. For example, a colleague may request a 30-minute phone call for 2:00 rather than a call from 2:00 - 2:30. If the effect found in the previous study requires salient boundaries, it should be attenuated when explicit end times are omitted. To test this, study 2a examined boundary expansiveness for time periods without a specified, salient end time.

Participants and procedure. Eight hundred and sixty-five Amazon MTurk workers participated in the study (due to experimenter error, demographics were not collected for study 2a or 2b). As specified in the preregistration, 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.

Participants viewed a series of scenarios that presented time periods as events that they had to attend. The nature of the events was not specified; they were simply "events." Participants were randomly assigned to one of two between-subject conditions that 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) were manipulated to form boundary-compressed and boundary-expanded periods. For example, the duration of 1.5 hours with a start time of 10:30am produces a boundary-expanded period, ending at 12:00pm, while the start time of 11:00am produces a boundary-compressed period, ending at 12:30pm. Thus, half of the time periods were boundary-expanded and half were compressed. Participants rated each period individually for "How long does this event feel?" on a slider from 0 (doesn't feel long at all) -100 (feels extremely long). Each participant completed a total of ten trials (five for each boundary type).

Results. We conducted a 2(boundary-expanded vs. compressed) x 2(salient boundary vs. no salient boundary) mixed-model ANOVA with fixed effects for boundary expansiveness, boundary salience and their interaction, and random effects for participant and trial. We observed an unexpected main effect of boundary salience, such that participants rated time periods as longer when end times were absent (M = 47.94, SD = 23.61) compared to present (M = 44.55, SD = 21.91; F(1,749) = 8.72, p = .003, d = .16). Moreover, boundary-expanded periods were rated as marginally longer than boundary-compressed (F(1,6753) = 3.44, p = .064, d_z = .10).

Crucially, however, we observed the anticipated significant interaction between boundary expansiveness and boundary salience (F(1,6753) = 10.28, p = .001). Analysis of the simple effects confirmed our predictions: 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 2a found that minimizing the salience of hour boundaries eliminates the effect of boundary-expansiveness on perceived duration; that is, boundary-expanded time periods only felt longer than compressed when an end time was salient. A couple of 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 expressed in parentheses. Second, participants were provided with the duration of every period without needing to approximate it themselves.

Study 2b

Our theory proposes that whole hour markers (e.g., 12:00, 1:00, 2:00) serve as natural boundaries, and that the crossing of these boundaries inflates the perceived duration of a given time period. If this is true, then the presence of other boundaries should produce the same effect. That is, we should be able to evoke different kinds of boundaries and observe inflation of periods that span more of them. This is what we tested in study 2b. We employed non-numeric categories—specifically, distinct classes in a student's schedule—and prompted participants to evaluate a period that spanned either two or three classes. Furthermore, the class schedule was manipulated such that the number of distinct hours and distinct classes encompassed by the period did not always align. For example, a period might be boundary-expanded in terms of hours but boundary-compressed in terms of classes, or vice-versa. We should observe that regardless of the number of hour boundaries, time periods spanning more classes feel longer.

Participants and procedure. Sixteen hundred and ten MTurk workers participated in this study, which employed a 2(number of class boundaries: 2 vs. 3) x 2(number of hour boundaries: 2 vs. 3) between-subjects design. As per our pre-registration, no participants were excluded.

Participants viewed a hypothetical class schedule for three 60-minute classes ("Class A," "Class B" and "Class C"). Boundaries occurred at either the hour or half hour mark. Specifically, half of the participants saw a schedule in which a new class 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 that was manipulated between-subject to span either all three classes or only two. Participants in the *3 classes* condition read the following:

> 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.

Those in the *2 classes* condition read 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 in the previous study, participants provided ratings of perceived length on a slider from 0-100. Importantly, for participants who saw a schedule with boundaries occurring on the hour, the period under consideration always spanned the same number of hours as it did classes. That is, if the participant was in the "3 classes" condition, this period was 9:30 - 11:10, which also spans 3 distinct hours. If they were in the "2 classes" condition, this period was 9:00 - 10:40, which also only spans 2 distinct hours.

However, the schedule with boundaries occurring on the half hour produced a mismatch. Now, the time period in the "3 classes" condition was 10:00 - 11:40, which only spans 2 distinct hours. Participants in the "2 classes" condition rated 9:30 - 11:10, which spans 3 distinct hours.

Despite our logic implying that we should expect main effects for both number of classes and number of distinct hours spanned, we worried that the salience of the classes could undermine the basic effect of hours spanned. Thus, we specifically preregistered that we had no predictions about the possible influence of whole hour marks. We reasoned that such predictions require knowledge about how people weigh various boundaries and whether they spontaneously construct time periods when given the information to do so. However, by highlighting the classes contained within the period, we expected to observe our basic effect in this non-numeric domain;

that is, we anticipated that time periods crossing more class boundaries would feel longer than equivalent periods that cross fewer.

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: Participants rated a period as feeling longer when it spanned three classes ($M = 67.94$, $SD = 23.44$) compared to two ($M = 64.53$, $SD = 24.56$; $F(1,1606) = 8.08$, $p < .01$, $d = .14$). There was no significant main effect of boundary-expansiveness according to whole hour markers. That is, collapsing across the number of classes spanned, ratings of perceived duration did not differ between periods that spanned three versus two whole hour markers (*M*s =66.76 and 65.73, *SD*s = 24.27 and 23.84, respectively; $F(1,1606) = .85$, $p = .357$, $d = .04$). We also did not observe an interaction between number of hours and number of classes ($F(1,1606) = .631$, $p = .427$.)

Discussion. The findings of study 2b suggest that boundary-expanded periods feel longer than boundary-compressed because, as our label suggests, they span more boundaries. In the previous studies, whole hour marks provided natural boundaries to conceptualize time. However, we argue that the effect does not result from the whole hour marks themselves, but the fact that they act as boundaries: We demonstrate that the same pattern of results emerges when other boundaries are highlighted—even those that conflict with whole hour marks.

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

STUDY 3: SCHEDULING

For some activities, people may want to minimize time; for example, when getting cavities filled or going to the DMV. People may want to maximize time for other activities, such as exploring a new city or taking a much-needed nap. Thus, when scheduling activities for which they want to maximize time, people may prefer boundary-expanded time periods, and when scheduling activities where they want to minimize time, boundary-compressed periods may be more appealing. Study 3 tested this hypothesis.

Method and Procedure

Participants were 601 workers from Amazon MTurk. A final sample of 600 remained after excluding the one person who failed the attention check. Across eight trials, in a fully within-subjects design, participants indicated which of two time periods they would rather schedule an activity. As before, periods had the same duration, but one period was boundary-expanded and the other boundary-compressed. Participants were told to imagine that both times worked equally well with their schedule. The items were selected to represent four activities that consumers would generally prefer to take as little time as possible, and four activities that they would like to last as long as possible. These activities were pre-tested to make sure they were indeed ones that consumers want to minimize and maximize. Specifically, we asked a separate group of MTurkers ($N = 40$) to rate 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 activities ($M = 6.11$, $SD = 1.14$). 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."

Results and Discussion

We performed a mixed-effects logistic regression on choice (boundary-expanded vs. boundary-compressed), specifying a fixed effect of activity type (maximize vs. minimize) and a random effect of participant. We observed the predicted interaction: Participants' choice of period (boundary-expanded vs. compressed) differed between the two types of activities ($z = -5.53$, $p < .0001$). An analysis of simple effects confirmed our predictions. For the time-maximizing activities, participants disproportionately selected the boundary-expanded periods over the compressed (52% vs. 48%; $z = 2.27$, $p = .023$). The reverse was true for the time-minimizing activities, where participants selected boundary-compressed periods more often (56% vs. 44%; $z = -5.35$, $p < .0001$). Figure 3 displays this pattern.

(Insert Figure 3 about here)

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 4: THE VALUATION OF TIME

Much of consumer behavior involves tradeoffs between time and money. 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, as established in the previous studies, time periods feel longer when they are boundary-expanded compared to boundary-compressed, how might this affect how consumers value their time?

We hypothesized that consumers would be willing to pay more to avoid waiting periods that are boundary-expanded compared to compressed. Moreover, consumers will demand more compensation to endure boundary-expanded waiting periods. Study 4 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 and forty MTurk workers participated in this study. As per our preregistration, we employed strict removal criteria. We removed all trials where participants gave answers that were at $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 211[2].

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 time and departure time were manipulated such that they created either a boundary-compressed or boundary-expanded waiting period. 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. Because participants might generally want to arrive earlier, they were randomly assigned to see one of two sets of times. In one, the boundary-expanded period arrived earlier than the boundary-compressed, and in the other set, it arrived later. Whether the boundary-expanded period was earlier or later than the boundary-compressed made no difference for either scenario; ps > .6. The interested reader may examine these periods in table B5 in the online appendix.

Participants responded on a sliding scale from $0 (I would change my flight for free) to $1000. They were also given space to enter their required amount if it exceeded $1000.

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 sliding 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 provided their WTP on a sliding 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. As before, the order of periods (boundary-expanded vs. compressed) was randomized, as was whether the boundary-expanded period was earlier or later (see table B6 in the online appendix).

Results

For the flight scenario, we performed a mixed-effects regression on required compensation, specifying a fixed effect of waiting period (expanded vs. compressed) and 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 = 271.65, SD = 178.48) compared to compressed (M = 254.44, SD = 177.69, t = 2.46, p = .015, d_z = .18).

The same model was specified for the bus scenario on WTP. Again, as predicted, participants were willing to spend more money to get on an earlier bus when the waiting period was boundary-expanded compared to compressed (M = 68.13, SD = 59.05 for expanded, M = 59.54, SD = 49.61 for compressed; t = 4.04, p < .0001, d_z = .29).

Discussion

The results reveal that boundary-expansiveness may affect consumer-decision making, particularly as it relates to the tradeoff between time and money. Participants reported that they would be willing to pay about 14% 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 boundary-expanded (vs. compressed) waiting times. In the next and final study, we examine the effects of boundary-expansiveness on real-world transportation choices.

STUDY 5: RIDESHARE CHOICES

The results of the previous experimental study suggest that boundary-expanded wait times (as opposed to compressed) 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 around the world every day (Iqbal, 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.

When considering booking a rideshare service like Uber, consumers living in major metropolitan areas often have the option to take a cheaper, longer trip by sharing their ride with other passengers. Additionally, consumers are provided with estimated arrival times for each option. Thus, consumers face one of three possible choice sets: 1) neither the individual ride nor the shared ride is estimated to cross a whole hour boundary, 2) both options are estimated to cross a whole hour boundary, or 3) the shared ride is estimated to cross while the individual ride is not.[3] Unlike the first two choice sets where both rides are equal in the degree to which the trip duration crosses an hour mark, options in the third set are "mixed", such that *only* the shared ride is boundary-expanded while the individual ride is not.

Figure 3 illustrates the difference between equal and mixed choice-sets. We expected that, all else equal, consumers would be more likely to request an independent ride when *only* the shared ride is expected to cross a whole hour boundary (when the choice-set is "mixed") as compared to when neither or both the rides cross a whole hour boundary (when the choice-set is "equal"). In order to test this hypothesis, we first develop a simple model of consumer choice and proceed to estimate this model using a large dataset from the Chicago metropolitan area.

(Insert Figure 3 about here)

Model

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.[4] 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

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

where 1(*ind ride is 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

$$U_{sh} = \alpha_{sh} + \beta_{exp} 1(shared\ ride\ is\ expanded) + \beta_d d_{sh} + \beta_c c_{sh} + \beta_{t,sh} time + \beta_{r,sh} route + e_{sh} \qquad (2)$$

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

$$\alpha + \beta_{exp} 1(mixed\ choice\ set) + \beta_d (d_{sh} - d_{ind}) + \beta_c (c_{sh} - c_{ind}) + \beta_t time + \beta_r route + e > 0 \qquad (3)$$

where $\alpha = \alpha_{ind} - \alpha_{sh}$, 1(*mixed choice set*) = 1(*shared ride is expanded*) – 1(*ind ride is 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.

We will estimate the coefficients in Equation (3) based on rideshare choice data. Before turning to the estimation results, we describe the dataset.

Data and Empirical Strategy

We acquired rideshare data for November 2018 – December 2019 from the Chicago Transportation Network, accessed from Chicago's open data portal (https://data.cityofchicago.org). 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).

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. 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:[5]

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

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

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).[6] 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. Specifically, 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). We plug the estimated mixed choice-set probability as well as the differences in duration and cost into Equation (3) to 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.

Before turning to that, a few remarks are in order regarding how we chose which trips to include in our dataset. 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, etc.). 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 longer, or travelled further than 30 miles (these exact criteria are somewhat arbitrary, but were preregistered to alleviate concerns of specification-curve hacking, e.g., Leamer 1983). 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%).[7]

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. We control for start time by including fixed effects for day/hour (e.g., January 3rd at 2am) as well as minute. We also control for route by including fixed effects for each pick-up/drop-off location pair.[8]
The results are reported in Table 1.

(Insert Table 1 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. 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.[9]

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.[10] 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.[11] We calculate the expected revenue under this alternative pricing scheme and compare it with that obtained under the pricing policy in the data.[12] Figure 4 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.

The implications from this analysis extend well beyond the rideshare domain. One such area is online commerce, where consumers often choose whether they want to pay more to receive a product sooner. Consider delivery dates, for example (web appendix study W5 finds that the effects of boundary-expansiveness also apply to months). If it is May 25th, and the product would be delivered by May 30th with expedited delivery, but June 2nd with standard delivery, consumers may be more inclined to select the expedited, costlier option. Similarly, food delivery companies with ample competition, such as Grubhub, Postmates, and others, provide their customers with estimated arrival times for orders. If the estimated delivery time for one provider is just under a new hour (e.g., arriving at 6:53pm), but just over a new hour for another provider (e.g., arriving at 7:02pm), ordering from the former provider may be more appealing than it would be if, say, both times fell before or after the hour. In short, the basic choice paradigm faced by rideshare consumers can be found for numerous other decisions consumers make in their day-to-day lives, underscoring the reach and potential impact of these results.

GENERAL DISCUSSION

The present research finds that time periods of equal duration do not always feel equivalent and therefore affects consumer decisions across a variety of domains.

We demonstrate that time periods feel longer when they span more hours because hour units (e.g., the "3" in 3:15) act as categories, and whole hour marks serve as boundaries that increase perceptions of duration. Categories are typically meaningful and function to minimize within-category variance and maximize between-category variance (e.g., Rosch 1998). They reflect similarity—cars are more similar to other cars than to trees, and trees are more similar to other trees than they are to cars, so they are grouped accordingly (Hampton 2001). In that sense, the difference between two objects in different categories *is* often larger than the difference between two objects in the same category. However, in the domain of time, the categories defined by distinct boundaries—such as 3:00—are not meaningful. Rather, hours are simply equally sized units along the abstract continuum of time.

While this is not the case for hours, round-number boundaries *can* signal true categorical shifts and reflect "real" differences. For example, in product versioning, the leading number (e.g., the "7" in version 7.9) only changes when a product has changed more dramatically than usual. Version 7.9 is likely to be more of an incremental improvement from version 7.7 than version 8.1 is from 7.9; the leading digit change may communicate a larger departure from earlier models. In these situations, it is generally appropriate for a consumer to infer greater difference between versions that span different units.

However, steps scale linearly for most numeric category information. The absolute difference between $4.99 and $5.00 is the same as the absolute difference between $4.98 and $4.99 (even if it doesn't feel the same, e.g., Thomas and Morwitz 2009). This is also true for numerically represented product dimensions, such as the display size, battery life, or memory of a laptop—inches, hours, and gigabits are on a ratio scale with equal steps between units. In these situations, although crossing round number boundaries should not affect consumers' perceptions of differences, we expect that it will. A 13.5-inch laptop display might seem more similar to an alternative with a 13-inch display than one with than a 14-inch, just as the time between 3:00 and 3:30 feels shorter than 3:30 to 4:00. That is, even though such round number boundaries do not represent any meaningfully larger shift, people may infer that they do and respond differently to essentially identical situations. This inconsistency is especially salient in study 2b, where arbitrarily imposed boundaries resulted in inconsistent judgments for the exact same time period.

This raises questions about if and how these biased judgments can be corrected, and whether they result in suboptimal outcomes for consumers. 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. In that sense, the bias may be best interpreted as a System 1 type bias (Kahneman 2011). However, such biases are often most pronounced in situations in which consumers give the decision little thought (Kahneman and Frederick 2005), have little experience with the decision situation (Kahneman and Frederick 2002), and have little to lose (Farrell, Goh and White 2014). This is why it is often argued that many biases documented in the lab may not actually affect behavior in the real world, where decision makers are experienced and incentivized to make "correct" decisions (Levitt and List 2007a; Levitt and List 2007b; Levitt, List and Reiley 2010). To this point, the results of study 5 are particularly informative—after all, selecting a rideshare is a common decision for the consumers in our dataset. Moreover, these are incentivized decisions, as the consumer's choice directly affects her wallet. Thus, the rideshare situation has all the characteristics that should reduce or eliminate inconsistencies caused by System 1 decision-making, yet we still find a substantial effect of hour marks 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 it is beyond the scope of the current manuscript, two options seem possible. One, it is possible that consumers are not inaccurate in their forecasts—rather, time periods that cross more hour boundaries *actually* feel longer. This might be true if, for instance, people check the time more often during a boundary-expanded period, thereby drawing their attention to passing more "boundaries." In that case, what appears as inconsistencies in judgments may represent accurate forecasts of experiences. While we do not have data directly testing this possibility, it does not appear likely. In study 1, MTurkers considering boundary-expanded periods estimated being able to complete an average of 17 HITs more than those considering equivalent boundary-compressed periods. Moreover, removing the salience of whole hour marks (as in study 2a) eliminates the effect, suggesting that this bias may be somewhat perceptual in nature. We believe it is more likely that categorical distortion is not like most typical System 1-based heuristics but is more akin to perceptual illusions like the Müller-Lyer and Ebbinghouse illusions (Müller-Lyer 1889; Ebbinghaus 1908; for a discussion on perceptual illusions and their similarity to cognitive illusions, see Kahneman and Tversky 1996). Such illusions are the result of the perceptual system accurately interpreting stimuli within a common, frequently encountered context, but then overapplying that interpretation in contexts where it shouldn't (see, the carpentered world hypothesis: Segall, Campbell and Herskovits 1966). This interpretation also suggests that it will be hard, if not impossible, to avoid these biases, and that any interventions should focus on helping decision-makers recognize situations in which they could fall prey to these biases and actively correct for them.

Round numbers are not the only boundaries that may alter perceptions of duration. Everyday events may provide additional boundaries. For a person who normally has lunch at noon, 12:00pm signals a new hour *and* a new activity, and thus may be a stronger boundary than, say, 11:00am. Relatedly, crossing into 12:00pm might signal a coarser shift from "morning" to "afternoon"; concepts of morning, afternoon, and evening may be their own categories and therefore provide additional boundaries. Consistent with both previous research as well as study 2b in the present paper, the same duration may also feel longer if it contains more punctuating events. A past point in time, such as the year 1995, feels further away when people consider more (vs. fewer) relevant intervening events (Zauberman et al. 2010). Moreover, an interval of time feels longer when it contains many (vs. few) shifts between events (the "filled duration illusion"; Ornstein 1969). While these studies involve retrospective and subjective judgments of time, their conclusions match our own—that shifts exaggerate duration. Applying an "events-as-boundaries" perspective to the present research, people may anticipate periods earlier in the day, regardless of their boundary-expansiveness, to feel longer than equivalent later periods because the former usually contain more events (e.g., lunch, meetings, etc.) and the latter fewer (e.g., dinner, free time). Examining some periods within study W4 in the online appendix offers an initial test of this idea, as it has a few trials where periods have the same duration but occur at different times of the day (in either the afternoon or evening; time periods are marked in the web appendix). Indeed, we find that the afternoon period was rated as feeling longer than the equal-duration evening period, $F(1, 850.55) = 45.47$, $p < .001$. However, the effect of boundary-expansiveness on perceived duration did not differ between these two periods.

At first glance, our findings may appear to be inconsistent with recent work by Tonietto et al. (2019), which demonstrates that time periods feel shorter when they are bounded by another upcoming task compared to unbounded (i.e., when the time afterward is free and unaccounted for). What might explain this apparent discrepancy? A crucial difference rests in the nature of the tasks in each paper. Participants in Tonietto et al. (2019) estimate *only* the time that falls between boundaries—which are either absent or defined by a subsequent event—and the paper finds that the inclusion of a boundary makes 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 may appear contradictory on the surface, but are actually very compatible, and combined illustrate the different ways in which boundaries can both make the time between boundaries feel shorter (Tonietto et al. 2019), and the time across boundaries feel longer (this paper). Interestingly, the design of study 2a 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 studies underscore the importance of boundaries in time estimation and their implications for consumer behavior.

We argue that the number of boundaries crossed inflates the estimation of time. If true, two logical extensions of our theory emerge. First, increasing the number of boundaries within a given time period should increase perceptions of its length. We show that our effect relies on the presence and placement of boundaries, but do not manipulate the size and meaning of the space between them. That is, we do not alter the scale itself, but simply take advantage of the fact that some time periods naturally span more boundaries than equivalent others (and manipulate the nature of those boundaries in study 2b). However, if time periods feel longer because they span more boundaries, increasing the resolution of the step size (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 a quantity seems larger when it is expressed in units that increase the number of boundaries spanned (e.g., for a product warranty, 84 months feels longer than 7 years; Pandelaere, Briers and Lembregts 2011). Second, because the difference of one extra unit is proportionally larger, one could expect that boundary-expansiveness affects shorter periods more than longer. For example, crossing two boundaries instead of one should feel like a larger difference than crossing eight boundaries instead of seven. Although not designed to test for this, three of our studies involve manipulations of boundary-expansiveness for varying lengths of time and thus allow us to examine whether such a pattern emerges (see the online appendix). Study 1 finds the anticipated pattern—that the perceived difference between boundary-expanded and compressed periods decreases as their duration increases, $z = -5.48$, $p < .0001$. The other two studies (W2 and W4 in the online appendix) show non-significant patterns in the same direction as study 1; with roughly one-sixth the number of participants, we suspect those studies are underpowered.

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. This prescription varies by the type of service offered—namely, whether it is one that consumers want to prolong or expedite. For example, dental offices may attract more patients if they offer boundary-compressed appointment times, and boundary-expanded appointments may be more appealing to clientele of massage parlors. In the same vein, given that most people prefer to minimize the wait that comes with traveling to one's destination, transportation companies may have greater success attracting customers for trips that are boundary-compressed instead of expanded. For example, if a boundary-compressed flight is more appealing than an expanded flight of equal duration, it might fill up more quickly.

Second, when travel options differ in boundary-expansiveness, businesses may be able to steer consumers toward choosing one option over the other. Our analysis of real-world transportation data reveals a compelling relationship between boundary-expansiveness and a routine and ubiquitous rideshare decision. Our findings imply that when choosing between a boundary-expanded shared ride and a boundary-compressed independent ride, consumers are more likely to pick the latter, all else equal. Indeed, we find 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. Therefore, by making small alterations in the estimated time of arrival, rideshare companies may nudge consumers toward choosing a faster, more expensive solo ride or a cheaper, longer shared ride. To facilitate more independent ride bookings, rideshare companies may consider biasing estimated arrival times upward for shared ride options. On the other hand, companies may want to prioritize pooled ridesharing—UberPool, for example, has not been nearly as profitable as anticipated (Sherman 2019). In that case, the present research suggests making the projected arrival time for shared trips slightly shorter, and independent trips slightly longer, when either ride has a non-zero probability of crossing into the next hour.

Third, businesses may be able to sustain a slight increase in the price of certain offerings without deterring sales. For example, we found a 14% increase in willingness-to-pay for a bus ride that was boundary-compressed compared to one of equal duration that was expanded (study 4). If consumers were willing to pay even 5% more for, say, boundary-compressed flights, the difference would scale up massively. For instance, with a carrying capacity of roughly 200 on a Boeing 737, pricing tickets on a boundary-compressed flight at $525 (instead of $500) would net airlines an additional $5,000 per flight. The difference reaches substantial significance when one considers that over thirty-eight million flights were performed in 2018 alone (Statista 2019).

Moreover, as determined in our rideshare analysis, customers are 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 the 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.

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, Eric. J., Patricia. M. Dechow, Devin G. Pope, & George Wu (2017), "Reference-dependent preferences: Evidence from marathon runners," *Management Science*, *63*(6), 1657-1672.

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

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.

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.

Ebbinghaus, Hermann (1908), *Psychology: An elementary textbook*. DC Heath.

Etcoff, Nancy L. and John J. Magee (1992), "Categorical Perception of Facial Expressions," *Cognition* 44(3)227–40.

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.

Goldstone, Robert L., and Andrew T. Hendrickson (2010), "Categorical perception," *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1) 69-78.

Hampton, James A. (2001), The role of similarity in natural categorization.

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 Shane Frederick (2005), “A model of heuristic judgment,” In *Thinking and reasoning*, eds. K. Holyoak & R. G. Morrison, New York: Cambridge University Press, 267-293.

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. and John A. List (2007b), "Viewpoint: On the generalizability of lab behavior to the field," *Canadian Journal of Economics,* 40: 347-370.

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.

Lucy, John A. (1992), *Language diversity and thought: A reformulation of the linguistic relativity hypothesis* (Vol. 12). Cambridge University Press.

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.

Mishra, Arul and Himanshu Mishra (2010), "Border Bias," *Psychological Science* 21(11) 1582–86.

Müller-Lyer, F. C. (1889) *Arch. Anat. Physiol. Physiol. Abt*., 2, 263–270.

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.

Peetz, Johanna, and Anne E. Wilson (2013), "The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation," *Journal of Personality and Social Psychology* 104(2), 249.

Roberson, Debi, Jules Davidoff, and Laura Shapiro (2002), "Squaring the circle: The cultural relativity of 'good' shape", *Journal of Cognition and Culture 2*(1), 29-51.

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

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.

Shum, Michael S. (1998), “The role of temporal landmarks in autobiographical memory processes,” *Psychological bulletin* 124(3), 423.

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

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.

Sherman, Len. “Are Uber’s Short-Term Paths to Profitability All Dead Ends?” (2019), Retrieved 16 March 2020 from https://www.forbes.com/sites/lensherman/2019/06/04/are-ubers-short-term-paths-to-profitability-all-dead-ends/#420568f97c08

Statista. “Number of flights performed by the global airline industry from 2004 to 2019 (in millions)”. (2019), Retrieved 05 June 2019 from https://www.statista.com/statistics/564769/airline-industry-number-of-flights/

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

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.

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 Kirkegaard (2009), "There is more to life stories than memories", *Memory,* 17(4) 445–457.

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 a version of the bus scenario wherein the wait was 2 hours (instead of ~6). Following our pre-registration, participants whose WTP on this trial was greater than their WTP for either bus trial were excluded.

[2] This high dropout rate appears to result from confusion about the text entry box where participants could enter values exceeding the scale maximum, as many entered numbers that were available on the slider.

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

[4] 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.).

[5] 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 $y_k + 7.5 - d_k$, respectively. A similar argument applies to b_t.

[6] 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.

[7] Such exclusions do not meaningfully change the estimate of the effect of interest.

[8] 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).

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

[10] 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.

[11] 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.

[12] 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. Coefficient estimates, standard errors and p-values. The dependent variable is a dummy equal to 1 if the consumer chooses the shared ride.

	Estimate	Standard Error
P_{Diff}	-0.525***	0.035
D_{Diff}	-0.035***	0.002
C_{Diff}	-0.220***	0.002

*** p < .0001

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.

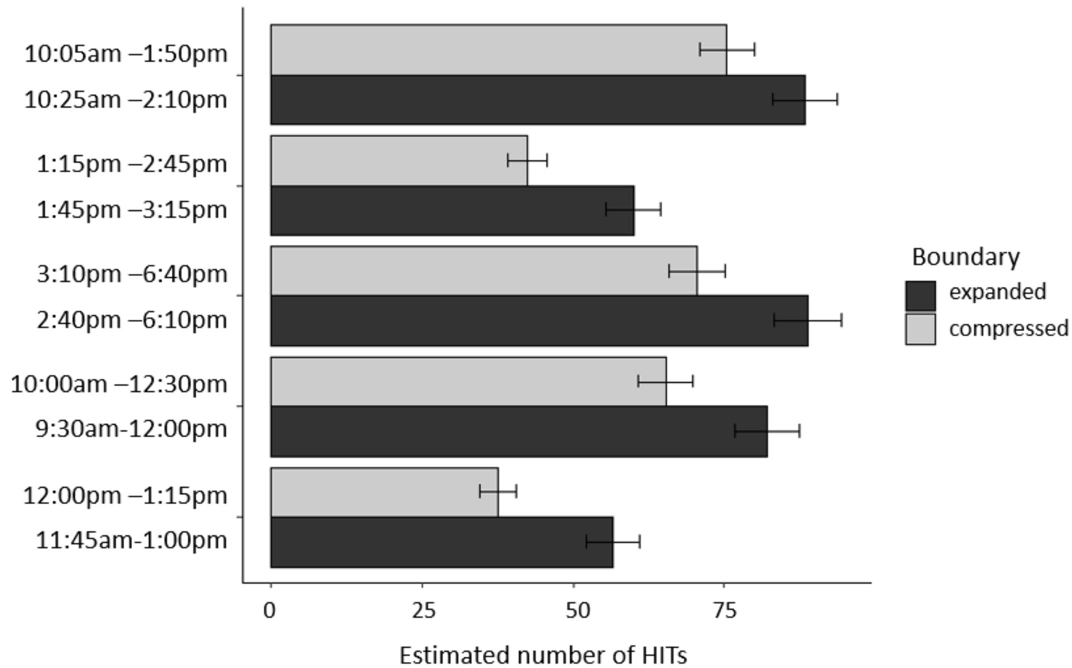


Figure 2. Boundary-expanded periods were selected more often for time-maximizing activities. Boundary-compressed periods were selected more often for time-minimizing activities.

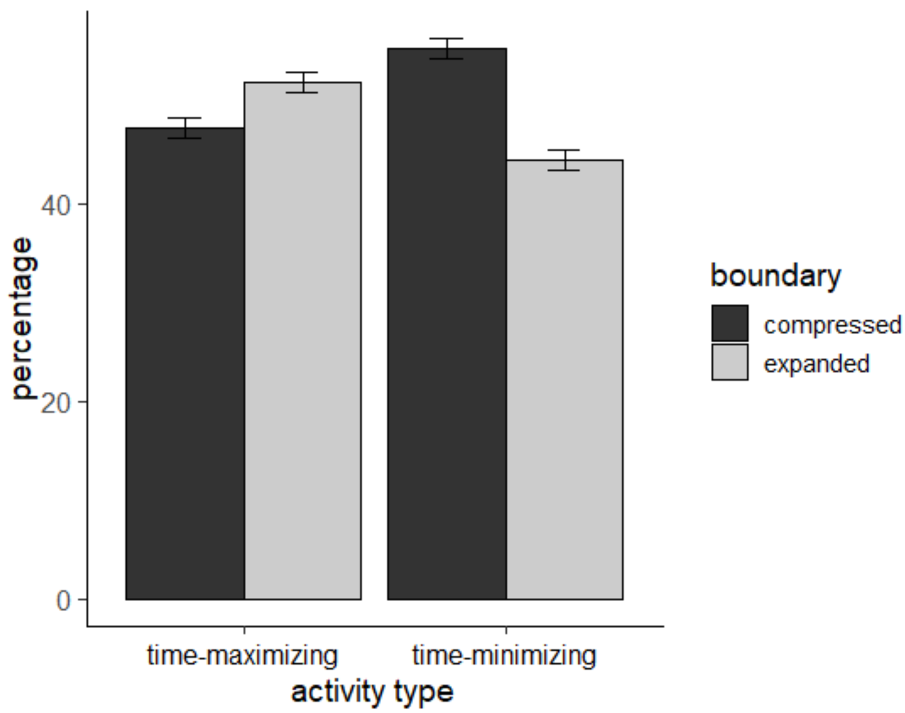


Figure 3. Example rideshare choice sets. Left image: Neither ride option is expected to cross into the next hour. Right image: Shared ride option is expected to cross, independent is not.

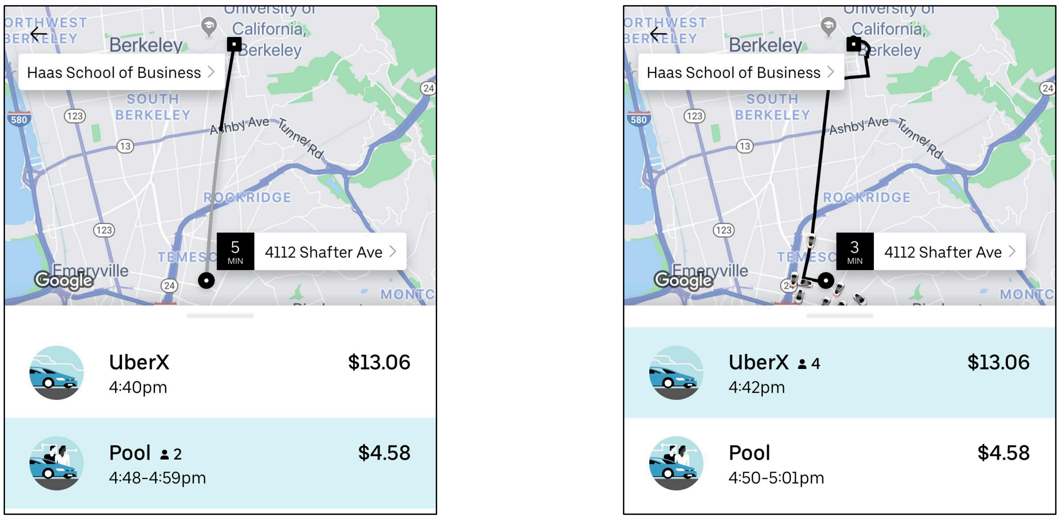
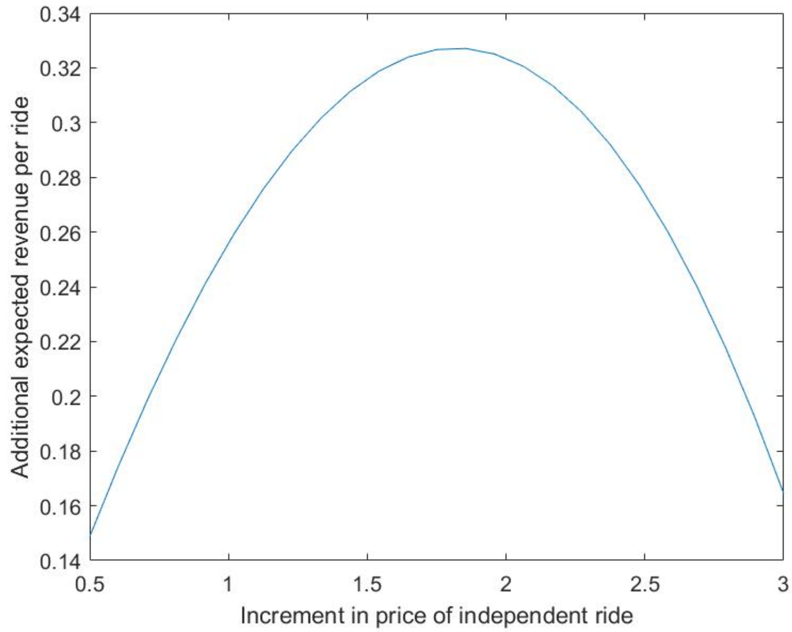


Figure 4. Change in expected revenues under alternative pricing policies.



APPENDIX

Studies A1 and A2

Studies A1 and A2 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 A1

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, $p < .0001$, $d_z = .54$).

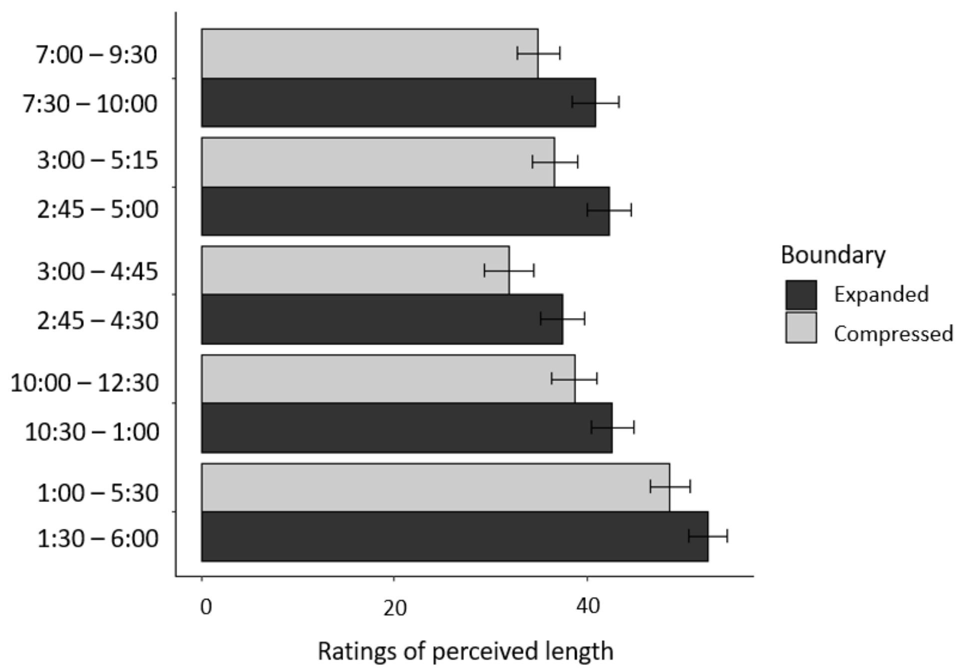
Study A2

In study A1, 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 S1. 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 < .0001).

Figure S1. Boundary-expanded periods were rated as “feeling longer” than boundary-compressed. Error bars represent standard error of the mean.



Discussion. The results of study A1 and A2 revealed that time periods felt longer when they spanned more hours. In study A1, participants selected boundary-expanded periods to feel longer more often than boundary-compressed. This finding was echoed by study A2: 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 A3 and A4 investigate these possibilities.

Study A3

Study A3 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 A3 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 A3, 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

[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.

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 significant intercept: Participants selected the boundary-expanded periods to “feel longer” more often than the boundary-compressed (75% vs. 25%, respectively; $z = 6.93$, $p < .0001$).

Study A4

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 S9 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 A4, 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.

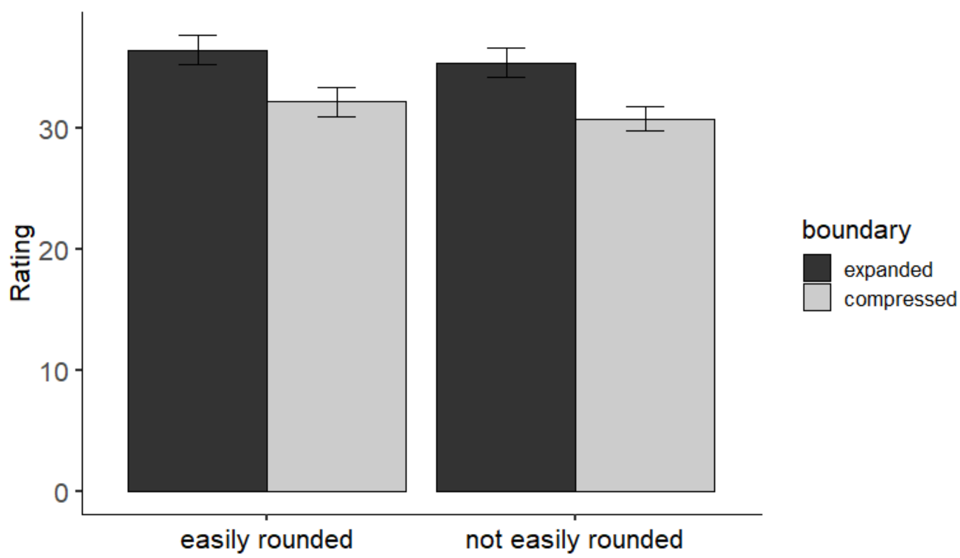
Participants and Procedure. One hundred and thirty-four MTurk workers participated in this study. Just as in study A2, boundary-expanded and boundary-compressed time periods were rated one at a time for "how long they feel" on a sliding scale from 0-100. However, in this study, we also included trials 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 S2).

Table S2. Time periods rated in study A4.

	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; $F(1, 2308.1) = 33.14$, $p < .0001$). 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 S3 displays these results (see also table S10).

Figure S3. Boundary-expansiveness versus susceptibility to rounding. Error bars represent the standard error of the mean.



Discussion. Studies A3 and A4 ruled out two possible explanations for the effect of boundary-expansiveness on perceived duration. The results of study A3 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 A4 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 B

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 A1, 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, $z = 6.91$, $p < .0001$. Expanded periods were selected approximately 73% of the time. Table S11 displays the results broken down by each period.

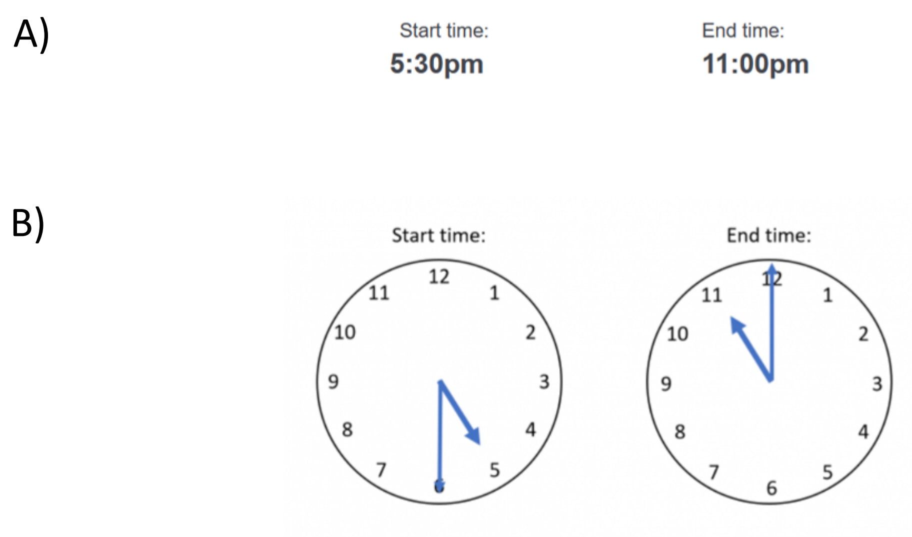
Study C

Time periods were presented numerically in all of our studies, although study 2b used a non-numeric category. Can the effect emerge when the presentation is fully visual, such as when one reads a clock? Study C 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.

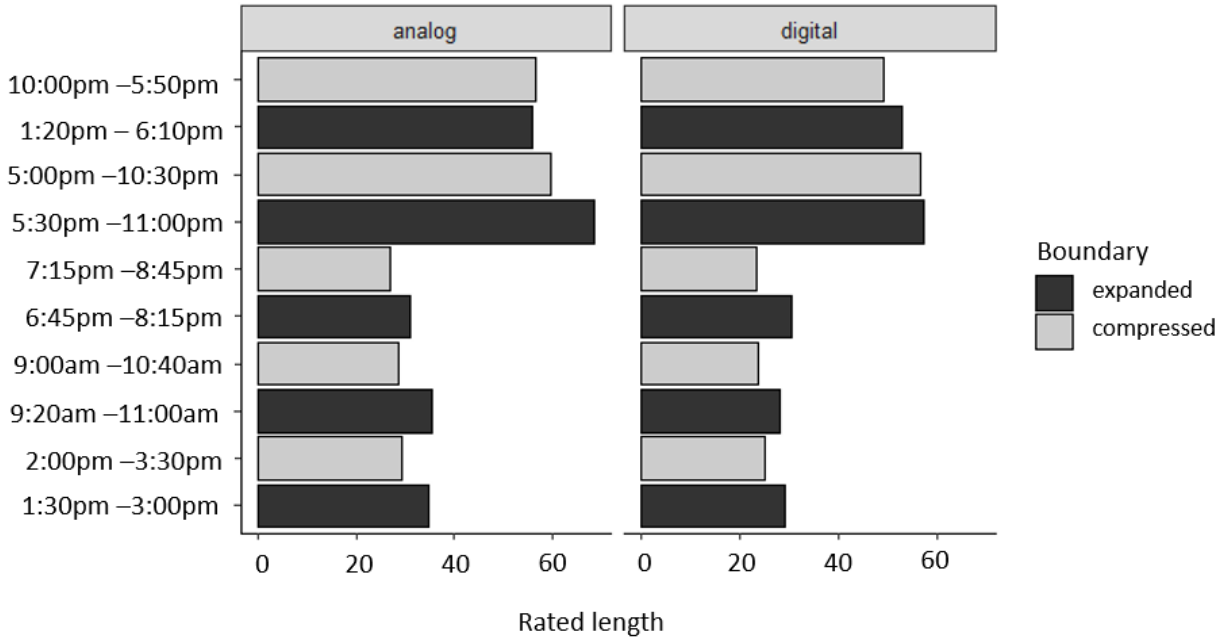
Just as in study A2 (and others), participants completed a series of trials in which they rated their perceived length of various time periods on a sliding scale. Participants were randomly assigned to view time periods that were expressed in either an analog or digital format. That is, half of the participants 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 S4 displays example stimuli.

Figure S4. Example stimuli, digital (A) and analog (B) formats.



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 ($M = 38.08$, $SD = 24.21$), $F(1,2586) = 79.13$, $p < .0001$. 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 S5.

Figure S5. Ratings of “how long does it feel” for time periods shown in analog versus digital format.


analog
digital
10:00pm –5:50pm
1:20pm – 6:10pm
5:00pm –10:30pm
5:30pm –11:00pm
7:15pm –8:45pm
6:45pm –8:15pm
9:00am –10:40am
9:20am –11:00am
2:00pm –3:30pm
1:30pm –3:00pm
0
20
40
60
Rated length
Boundary
expanded
compressed

TRIAL-LEVEL INFORMATION

Study 1

Table S1. 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 S2. 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 2b

Table S3. Ratings broken down by number of classes and hours spanned.

	Spans 2 classes	*Spans 3 classes*
Spans 2 hours	*M* = 64.46 *SD* = 23.74	*M* = 66.92 *SD* = 23.90
Spans 3 hours	*M* = 64.60 *SD* = 25.35	*M* = 68.98 *SD* = 22.94

Study 3

Table S4. 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 12:30pm – 2:00pm	196 80	12:00pm – 1:30pm	324
	Getting blood drawn	10:45am – 11:00am 9:45am – 10:00am	86 183	10:00am – 10:15am	331

	Dental cleaning	1:30pm – 3:00pm 12:30pm – 2:00pm	90 171	1:00pm – 2:30pm	339
	DMV appointment	1:30pm – 2:00pm 12:30pm – 1:00pm	102 157	1:00pm – 1:30pm	341
Time-maximizing	Free time	2:30pm – 4:00pm	327	2:00pm – 3:30pm 3:00pm – 4:30pm	133 140
	Exploring on a work trip	12:30pm – 2:00pm	316	1:00pm – 2:30pm 12:00pm – 1:30pm	191 93
	Watching favorite show	7:30pm – 9:00pm	285	7:00pm – 8:30pm 8:00pm – 9:30pm	122 193
	Taking a nap	2:15pm – 3:00pm	328	2:00pm – 2:45pm 3:00pm – 3:45pm	170 102

Study 4

Table S5. Flight Scenario: Compensation required to endure boundary-expanded and compressed waiting periods broken down by period.

Expanded period	*Compressed period*	*Boundary-exp compensation*	*Boundary-comp compensation*
9:47am – 2:12pm	9:24am – 1:49pm	*M* = 273.48 *SD* = 168.93	*M* = 258.75 *SD* = 171.05
10:47am – 3:12pm	11:24am – 3:49pm	*M* = 269.80 *SD* = 188.55	*M* = 250.18 *SD* = 184.85

Table S6. Bus Scenario: Willingness-to-pay to avoid boundary-expanded and compressed waiting periods broken down by period.

Expanded period	*Compressed period*	*Boundary-exp WTP*	*Boundary-comp WTP*
10:40am – 5:05pm	11:15am – 5:40pm	*M* = 64.22 *SD* = 56.26	*M* = 55.26 *SD* = 46.73

12:40pm – 7:05pm	12:15pm – 6:40pm	*M* = 72.02 *SD* = 61.74	*M* = 63.79 *SD* = 52.20

Supplemental Studies

Study A1

Table S7. 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 A2

Table S8. 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 A3

Table S9. 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 A4

Table S10. 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 B

Table S11. 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

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: $z = 3.06$, $p = .002$.

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. We observe the anticipated interaction between boundary expansiveness and boundary salience ($F(1,6951) = 10.88$, $p = .001$). When boundaries were salient, boundary-expanded periods ($M = 45.62$, $SD = 22.20$) were rated as feeling longer than compressed ($M = 44.09$, $SD = 21.83$; $t = 3.73$, $p = .0002$, $d_z = .24$), and we observe no difference in ratings when boundaries were not salient ($t = -.968$, $p = .333$, $d_z = -.07$).

Study 2b. This study had no exclusions.

Study 3. The predicted interaction emerges such that choice of period (boundary-expanded vs. compressed) differed between the two types of activities; $z = -5.35$, $p < .0001$. For the time-maximizing activities, participants disproportionately selected the boundary-expanded periods over the compressed (52% vs. 48%; $z = 2.23$, $p = .026$). For the time-minimizing activities, participants selected boundary-compressed periods more often than the expanded (56% vs. 44%; $z = -5.35$, $p < .0001$).

Study 4. 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.

Study 5. 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* excluding trips that 1) had fewer than 100 trips in the dataset with the same route, start hour, and ride type and 2) were two SDs above or below the average duration of their similar trips. Results are presented in table S12.

Table S12. Alternate analyses of rideshare data in Study 5.

		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 without exclusions (logistic regression)	Estimate	-0.468***	0.006**	-0.163***
	Standard Error	0.3482	0.0016	0.0015

****p* < 0.0001

***p* < 0.001

Does the Effect Differ Based on Duration?

In this section, 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 S13. 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	$z = -5.48, p < .0001$	The difference between boundary-expanded and compressed decreases as duration increases.