Journal of Experimental Psychology:
Learning, Memory, and Cognition
2009, Vol. 35, No. 1, 81–93

0278-7393/09/$12.00 DOI: 10.1037/a0013665

Anchoring on the "Here" and "Now" in Time and Distance Judgments

Robyn A. LeBoeuf
University of Florida

Eldar Shafir
Princeton University

Time and distance estimates were elicited with either unit-based (e.g., "How many days until . . .") or end-based (e.g., "On what date . . .") questions. For intervals of uncertain extent, unit-based estimates were consistently lower than were the corresponding end-based estimates. The observed patterns are consistent with an anchoring and adjustment process: When people generate unit-based estimates of uncertain dates or distances, they may anchor on the "here" or "now" and adjust incrementally by the unit; such adjustment, however, is often insufficient and yields systematic underestimation. Although this anchoring and adjustment cannot be directly observed, consistent with the hypothesized process, reliance on larger units yielded higher estimates and warning about insufficient adjustment reduced the effect. Implications for research on anchoring, the planning fallacy, and everyday judgment are discussed.

Keywords: anchoring and adjustment, judgment, time perception, planning fallacy, distance estimation

A future event can be demarcated by its anticipated date or by the estimated amount of time until its occurrence. Although logically comparable, these two methods can prompt different senses of event proximity. For example, toward the end of February one may contemplate a March engagement, only to realize that it will happen in 2 days. Similarly, during the last days of 1999, many noted that "the next millennium" seemed much farther off than did "days away." Even a pending project may be perceived differently when one contemplates the number of days until its completion instead of the corresponding completion date.

The studies reported in this paper investigated these two ways of gauging time. We found that temporal estimates were influenced by whether they were unit based (e.g., number of days) or end based (e.g., final date), and we documented similar effects in distance estimation. These effects arise, we suggest, from a process of anchoring and (insufficient) adjustment in the generation of unit-based estimates. Although our methods precluded direct observation of the posited adjustment processes, this article presents evidence consistent with this account. This article also considers related effects of question phrasing and tasks to which the findings may be relevant. We open with a brief review of research on time estimation and on elicitation effects.

Time Estimates

People's estimates of past and future time intervals have been studied in various contexts and have been found to be biased in several ways. For example, for past time intervals, there seems to be systematic "forward telescoping," or a tendency for past events to be perceived as having occurred more recently than they actually did (Neter & Waksberg, 1964). Whereas initial telescoping research examined perceptions of mundane events, such as the timing of household expenditures (Neter & Waksberg, 1964), telescoping tendencies have even been observed in the context of more dramatic events, such as being the victim of a crime (Loftus & Marburger, 1983).

Future time estimates can be distorted as well. Multiple studies have documented systematic overoptimism in people's estimates of the time required to complete various tasks (Buehler, Griffin, & MacDonald, 1997; Buehler, Griffin, & Ross, 1994; Griffin & Buehler, 1999). For example, participants in one study thought they would require an average of 34 days to finish their academic theses, whereas, on average, the theses were completed in 56 days (Buehler et al., 1994). This "planning fallacy" has been documented for all manner of projects, both academic (writing essays and lab reports) and nonacademic (cleaning apartments and submitting tax returns), and is most pronounced when people are motivated to complete a project quickly (Buehler et al., 1994, 1997).

In most studies of the planning fallacy (e.g., Buehler et al., 1994, 1997; Griffin & Buehler, 1999), people are asked for estimated completion dates; although the fallacy has proved robust, it is not clear to what extent the effect might depend on this specific elicitation method. For example, compared with providing a date, estimating the number of days to completion might prompt more grounded and accurate responses; alternatively, a day-by-day simulation of progress might lead respondents to disregard competing demands and to further underestimate completion times (cf. the simulation heuristic, Kahneman & Tversky, 1982). Along similar lines, even though researchers investigating telescoping have obtained both date and time-length estimates (Huttenlocher, Hedges, & Bradburn, 1990; see Friedman, 1993, for a discussion), they have not examined whether these methods yield estimates of comparable accuracy. An open question thus remains as to whether patterns such as telescoping and the planning fallacy, as well as

Robyn A. LeBoeuf, Marketing Department, University of Florida; Eldar Shafir, Department of Psychology and Woodrow Wilson School of Public Affairs, Princeton University.

We thank Susanne Klockner, Molly Maxfield, and Whitney Meyers for their help with data collection. Portions of this research were awarded the Student Poster Award at the Annual Meeting of the Society for Judgment and Decision Making.

Correspondence concerning this article should be addressed to Robyn A. LeBoeuf, Marketing Department, University of Florida, P.O. Box 117155, Gainesville, FL 32611-7155. E-mail: LeBoeuf@ufl.edu

time estimation in general, are sensitive to variations in how the estimates are obtained.

Elicitation Effects

Outside of the domain of time estimation, a rich research tradition has documented numerous ways in which judgments and decisions can be influenced by subtle and normatively immaterial changes in response-elicitation procedures. Preferences, for example, can systematically change when outcomes are described in terms of losses rather than gains (Tversky & Kahneman, 1981), when people are asked to reject rather than to choose (Shafir, 1993), when products are priced rather than chosen (Lichtenstein & Slovic, 1971, 1973), and when options are evaluated separately rather than directly compared (Hsee, 1996). In fact, dozens of studies have shown preferences to be malleable; preferences are often constructed on the fly, and are thus susceptible to influence by the method of elicitation (see Kahneman & Tversky, 2000). Similar patterns have been observed in judgment and estimation studies and in survey research, wherein the framing and ordering of questions, as well as the provided response scales, can affect estimates, reported attitudes, and behaviors (Schwarz, Groves, & Schuman, 1998; Tanur, 1992; see Gilovich, Griffin, & Kahneman, 2002, for a review).

Of particular relevance to the present research is the documented impact of irrelevant "anchors" on judgment. In one classic experiment, participants first decided whether an apparently arbitrary number was greater or smaller than the percentage of African nations in the United Nations; they then estimated the actual percentage. Estimates were significantly higher among those initially exposed to a high number than among those exposed to a low number (Tversky & Kahneman, 1974; see also Chapman & Bornstein, 1996; Jacowitz & Kahneman, 1995; Wilson, Houston, Etling, & Brekke, 1996; Wong & Kwong, 2000). One possible explanation is that people use the anchor as a starting point and adjust but that they fail to adjust sufficiently and thus provide a final estimate too close to the initial anchor (see, however, Chapman & Johnson, 2002, and Mussweiler & Strack, 1999, for alternative explanations of anchoring to which we later return).

Anchoring effects are likely quite prevalent in everyday judgment. In fact, anchors need not be explicitly provided but can instead be naturally triggered (Epley & Gilovich, 2001). For most Americans, for example, contemplating the boiling point of water on Mount Everest naturally triggers an anchor of 212°F, from which they then adjust. Epley and Gilovich (2001, 2004) have demonstrated that estimates further from or closer to such anchors can be produced when (typically insufficient) adjustment is implicitly encouraged or discouraged.

Predictably, opinions and attitudes that are held with some conviction tend to remain stable across elicitation procedures and are less affected by problem frames or suggested anchors. However, people often evaluate objects and estimate quantities about which they have only a vague impression. In those cases, elicitation methods in general, and anchors in particular, can prove consequential. In what follows, we integrate research on elicitation effects, and especially anchoring, with research on estimates of the temporal proximity of events. We then extend our research to estimates of physical distance.

The Current Research

In the studies reported below, we first manipulated the temporal estimates that participants were induced to provide: Some estimated the date of occurrence of past or future events, whereas others estimated the number of days remaining until (or that had elapsed since) those events. Consider first those estimating in days: We predicted that, much as those who estimate the quantity $1 \times 2 \times 3 \ldots 7 \times 8$ appear to begin at the implied, low starting point (Tversky & Kahneman, 1974) and those who estimate annual fatalities per 100 citizens appear to adjust upward from 1 (Yamagishi, 1994), those who estimate an uncertain temporal distance in terms of days from today would begin at a low anchor of "now" and adjust upward in time.

Specifically, we hypothesized that respondents who are asked to estimate in days from the present might begin at the (self-generated anchor of the) present and adjust by the proposed day-based units. We further hypothesized (following the view of adjustment outlined by Epley and Gilovich, 2006) that such adjustment would proceed serially, with respondents adjusting from the anchor, testing whether the resulting estimate is plausible, and adjusting and testing until a plausible estimate is reached. Because such adjustment processes are effortful, people often "satisfice" and report the first plausible estimate reached; as a result, adjustment tends to be insufficient (Epley & Gilovich, 2004, 2006). Consequently, we predicted that estimates requested in terms of days might be relatively low.

Estimation of a calendar date, on the other hand, was expected to yield a more global, bird's-eye view (perhaps monthly or yearly but largely independent of the moment) and thus to prove less conducive to stepwise adjustment from the present. Therefore, we predicted that time intervals would be shorter when estimated in days (and based on adjustment from the present) than when estimated in dates (and likely based on a calendar). Furthermore, we predicted that, if day-based underestimation results from insufficient unit-based adjustment, increased unit size (e.g., weeks instead of days) might yield somewhat greater estimates. Finally, because estimates of events whose dates are known (e.g., last Christmas or next Valentine's Day) likely are less malleable, elicitation effects were predicted to be most pronounced for uncertain future events (i.e., for events that have not yet occurred and whose date of occurrence is not known in advance).

Study 1a: Time Estimation, Initial Demonstration

Method

Materials. Participants estimated the time of occurrence of various past and future events, which were further divided into "private" and "calendar" events. Calendar events, such as Mother's Day, have a fixed, publicly available date; private events, such as getting a haircut, are largely determined by the individual and have no predetermined calendar date. Whenever possible, future and past versions of the same events were used. For example, some estimated the time remaining until next Mother's Day, whereas others estimated the time elapsed since last Mother's Day.

The main manipulation was in response elicitation. Approximately half of the participants estimated the number of days elapsed since, or remaining until, the target event (e.g., "How many days has it been since . . ." or "How many days will it be until . . .”), hereafter called the day-format condition. The remaining participants estimated the target event’s date (e.g., “On what date will . . .” or “On what date did . . .”), hereafter called the date-format condition. The 58 target events ranged in distance from 1 to over 30 days and are listed in the Appendix.

Procedure and participants. Participants were undergraduate volunteers. Because of a limited formal participant pool and the relatively simple nature of the task, data were collected on several occasions. On five occasions, participants were recruited to spend up to an hour completing the current task, as well as other, unrelated tasks. On another three occasions, participants were approached on campus and presented with a brief questionnaire. The number of participants per occasion ranged from 58 to 475. Target events were arbitrarily assigned to occasion, and more events were presented during longer sessions. Each participant thus estimated between 2 and 10 events; the number of participants who responded to each event is summarized in the Appendix.

On any given data-collection occasion, response format (day/date) and event locus (past/future) were manipulated between participants, who were arbitrarily assigned to condition. Some participants estimated both private and calendar events, whereas others estimated only calendar or only private events; this had no systematic effect on the results. The data are collapsed across the different data-collection occasions in the analyses that follow.

Results and Discussion

Date estimates were converted into the corresponding number of days. Thus, for responses given on April 9, “April 10” was coded as 1 day and “April 29” was coded as 20 days. For each event, outlier responses (those beyond the 75th/25th percentile plus/minus 1.5 times the interquartile range of the data set; Hinkle, Wiersma, & Jurs, 1994, p. 68) were excluded from further analysis.

Each item’s mean date- and day-format estimates appear in the Appendix, along with an indication of the reliability of the format effect for each item. Because of the somewhat unorthodox data collection, some participants estimated more events than did others, thereby precluding standard 2 (response format) $\times$ 2 (event locus) $\times$ 2 (event type) analyses of variance (ANOVAs). Instead, we summarize the effects of response format separately for each Event Locus $\times$ Event Type combination.

Consider first the 15 future-private events (such as reading a book or getting a haircut). The day-format estimate was less than the date-format estimate for all but 1 event; the one exception (speaking to family on the telephone) showed a null effect and yielded estimates so low as to suggest a floor effect. The difference between day and date estimates was reliable ($p < .05$) for 8 events, and four such comparisons were even reliable when a far stricter criterion was used ($p < .003$). This criterion was derived from a Bonferroni correction that divides $\alpha = .05$ by 15, the number of pairwise comparisons; this fixes the overall Type 1 error rate for future-private events at $\alpha = .05$ (Myers & Well, 1995). For another way to explore the effect, we computed the day–date discrepancy per item (as a percentage). Averaged across items, day-format estimates amounted to only 67% of corresponding date-format estimates.

Thus, for future events with no predetermined date, estimating the number of days until the event yields consistent underestimation compared with estimating the date of occurrence directly. This underestimation perhaps reflects insufficient adjustment from the present when people adjust by the number of days.

Another possibility, however, is that people simply err in converting between dates and days. When asked for day estimates, people may initially envision a date similar to that reported in the date condition but may then convert improperly. Such error could yield systematic underestimation. To test for this possibility, we examined estimates for future-calendar events (i.e., future events whose dates are predetermined). We hypothesized that, when asked for the number of days until an event with a known, easily retrieved date (e.g., Valentine’s Day), people would retrieve the date and then convert it to the number of days.

Indeed, estimates of future-calendar events were hardly affected by elicitation procedure. Across 16 future-calendar events, format effects were rarely reliable. Day estimates proved reliably higher than date estimates once and reliably lower twice; only one of the latter comparisons survived a more stringent Bonferroni correction. Furthermore, averaged across items, day estimates were 98% of corresponding date estimates. Evidently, response-format manipulations mainly affect judgment when a date is not readily available to be converted to days. When a date is predetermined, day-format responses are consistent with it. Such consistency argues against the error-of-conversion hypothesis.

We have proposed that the important difference between future-private and future-calendar events is the predetermination of their dates. We contend that such predetermination allows conversion when day-format estimates of future-calendar events are required but that the lack of predetermination prompts anchoring and adjustment when such estimates of future-private events are needed. There is, however, a potential confound: Some other intrinsic private-calendar difference could underlie the differential susceptibility to response-format effects. To explore this possibility, we shifted our attention to past events, private and calendar, all of which (by virtue of having already occurred) had predetermined dates that respondents either knew or could attempt to directly recall.

Response format had no reliable effect on any of the past-private events, and only 2 of the 14 past-calendar events showed reliably lower day- than date-format estimates. (Only one of these comparisons was reliable after a Bonferroni correction.) Day estimates averaged 96% of date estimates for both event types. Thus, it does not appear that private events are inherently more subject to response-format effects than are calendar events.

Instead, it appears that response format has a negligible effect on estimates of events with a known date (such as past or calendar events) but has a systematic impact on estimates of events whose time of occurrence is not fixed. Estimates of calendar dates likely derive from top-down contemplation of a monthly or yearly calendar. Estimates of days seem to be simply deduced when the date is known. However, when the date is not known, individuals may estimate days by anchoring on the present and progressively adjusting; the recruited adjustment process, being typically insufficient, may produce the observed relative underestimation.

Study 1b: Time Estimation, Conceptual Replication

Study 1a suggests the existence of response-format effects in temporal estimates: Participants who estimated the number of days until an event (with no publicly known date) foresaw lesser temporal extension than did those who estimated the event’s date. Data collection, however, was somewhat unorthodox, as it was spread over a number of occasions. Although the patterns obtained on separate occasions did not differ (and it was heartening that they were sufficiently robust to arise in "noisy" nonlaboratory conditions), we replicated the findings with the more conventional procedure described here.

Method

Undergraduate students ($N = 137$) completed this study among other short tasks in the lab for course credit. They were randomly assigned to one cell of a 2 (event locus, future or past) × 2 (response format, day or date) design. Each participant estimated the occurrence of four private events (i.e., events with no publicly available date): finishing a book, shopping at the mall, completing a project, and shopping at a campus bookstore. (The particular book and project to be estimated were specified by each participant prior to estimation.) As noted, response format and event locus were manipulated between subjects, so that, for each event, participants were asked, "How many days do you think it will be until you . . ." (future day), "On what date do you think you will . . . (future date), "How many days do you think it has been since you . . ." (past day), or "On what date do you think you . . ." (past date).

Results and Discussion

As before, date-format responses were converted to the corresponding number of days for analysis. As shown in Table 1, for all future events, day-format estimates were lower than date-format estimates, but the effect was less consistent for past events. Averaged across events, future day-format estimates ($M = 27.2$, $SD = 23.8$) were reliably lower than the corresponding date estimates ($M = 43.5$, $SD = 37.7$), $t(66) = 2.17$, $p = .03$, but this difference was not observed for past events, $t(67) = .004$, $p > .99$. Supporting these conclusions, a 2 (format) × 4 (event) mixed ANOVA for future events revealed only a main effect of response format, $F(1, 51) = 7.33$, $p = .009$ (all other ps $> .25$; degrees of freedom in the ANOVA and the t test differed because those who gave unusable responses to any one item were excluded entirely from the ANOVA). A similar ANOVA on past events revealed only a main effect of event, $F(3, 168) = 4.43$, $p = .005$ (all other ps $> .15$).

The results of this study support Study 1a's findings: Future events without a publicly known date were predicted to occur sooner when estimated in days rather than in dates. This effect did not emerge for comparable past events, presumably because the date of a recent past event can be deduced through recollection of the event and contemplation of the recent calendar; this date can then be converted to days. When the date in question is not fixed, on the other hand, those who estimate days seem to use a different procedure than that used by those who envision a date. We suggest that those estimating days may anchor on the present and incrementally adjust. This adjustment is typically insufficient (Epley & Gilovich, 2004) and yields relative underestimation.

Study 1c: A Consideration of Conversational Norms

We posit that a process of anchoring and insufficient adjustment underlies the observed response-format effects, but it is prudent to consider one salient alternative account. Specifically, a frequent critique of demonstrations of judgmental bias appeals to conversational norms (e.g., Grice, 1975; see Shafir & LeBoeuf, 2002, for a review of such critiques). Perhaps a request for an estimate in terms of days conveys an expectation of greater temporal proximity than does a request in terms of dates (cf. Golding, Magliano, & Baggett, 1995; Schwarz, 1996). For example, asking about doing laundry in terms of days may convey the expectation that laundry will be done soon. Note, however, that such suggestions about the expected proximity of events should be equally applicable to past and future events: To the extent that the requested response format is informative or suggestive, it should be equally suggestive for future and for past events. However, in Studies 1a and 1b, response format influenced only future events.

This systematic asymmetry between past and future judgments casts doubt on conversational norms as a main explanation for the current findings; nevertheless, it is possible that the uncertainty surrounding predictions of future events makes them more susceptible to conversational influences than are past events. That is, although the observed asymmetry suggests that any conversational norms are not powerful enough to influence responding in general, the asymmetry alone may not rule out some influence of norms for future events. In Study 1c, we further evaluated the conversational norms account by examining whether the mere suggestion of specific units, absent their use in generating a response, affects estimates.

Method

Undergraduate students ($N = 227$) completed this study, among other tasks, for course credit. All participants were presented with a brief description of the general procedure of Studies 1a and 1b (e.g., "We are asking . . . students to think of a book that they recently resolved to read"); by random assignment, participants read either the day-format version (e.g., "We are asking them how

Table 1
Mean Estimates, in Days, by Event Locus and Response Format, Study 1b

	Past		Future	
Event	Day format	Date format	Day format	Date format
Read a book	72.1 (91.9)	66.4 (78.8)	30.0 (26.5)	36.0 (33.5)
Shop at a mall	32.7 (66.4)	19.5 (35.9)	15.7 (29.6)	53.6 (125.6)
Complete a project	29.0 (37.9)	17.5 (21.7)	31.0 (69.5)	37.2 (37.0)
Shop at a bookstore	21.9 (58.9)	49.6 (124.0)	30.7 (40.7)	58.8 (68.5)
Average estimate	38.1 (38.8)	38.1 (44.7)	27.2 (23.8)	43.5 (37.7)

All data are *M* (*SD*).

many days it will be until they finish reading that book") or the date-format version (e.g., "We are asking them on what date they will finish reading that book"). All respondents were then asked, "On average, by what [calendar] date do you think these students think they will have read the book?" Thus, half of the respondents were encouraged to think about completion in enumerable days, but all of them had to generate a date for a response. Participants made estimates for three events (reading a book, shopping at a bookstore, and doing laundry).

Results and Discussion

If a day format engenders inferences that events should be proximal, one would expect participants reading the day format to generate lower estimates than participants reading the date format. Instead, for each item, the two groups' estimates were statistically indistinguishable (ps $> .15$), and the trends were the opposite of what a conversational norms account might suggest (book: $M_{day} = 62.0$, $SD = 111.9$; $M_{date} = 45.4$, $SD = 48.4$; bookstore: $M_{day} = 45.3$, $SD = 49.1$; $M_{date} = 41.5$, $SD = 50.3$; laundry: $M_{day} = 8.3$, $SD = 18.3$; $M_{date} = 8.0$, $SD = 15.4$). Overall, estimates averaged 38.4 days ($SD = 42.2$) in the days version and 31.4 days ($SD = 26.3$) in the date version, $t(225) = -1.50$, $p = .14$.[1]

Although null results are destined to remain inconclusive, there is no indication that the mere suggestion of smaller units alters perceived time frames and conveys suggestions about event proximity. We suggest, instead, that small units exert their influence when they are used in response generation and that underestimation results from insufficient unit-based adjustment.

Study 2: Unit Change

If insufficient adjustment from the present underlies the lower estimates observed with a day response format, an increase in unit size (e.g., from days to weeks) should increase estimates. That is, the same number of adjustment steps should yield greater absolute adjustment, because a "week" step would be larger than a "day" step (cf. Janiszewski & Uy, 2008). Study 2 examined this possibility.

Method

Study 2 was run concurrently with Study 1a. Eight future-private and 5 past-calendar events (presented in Study 1a) were presented in a week-and-month format (e.g., "How long, in weeks and months, will it be until . . .") and were run along with Study 1a's other two conditions, with participants randomly assigned to provide date, day, or week-and-month estimates for each of these 13 events (listed in Table 2). Because a week-and-month format was predicted to alter response, by increasing the units of adjustment because an anchoring-and-adjustment process was expected primarily for events with an unknown date of occurrence, we expected the week-and-month format to differ from the day format mainly for future-private events.

Results and Discussion

Week-and-month estimates were converted to the corresponding number of days by multiplying the number of weeks and months by 7 and 30, respectively.[2] As predicted, for future-private events, the change in unit size had a substantial impact, with week-and-month estimates reliably exceeding day estimates for each item (see Table 2). In stark contrast, day estimates were statistically indistinguishable from week-and-month estimates for all but one of the past-calendar events. Thus, consistent with the proposed adjustment account, larger units prompted greater estimates in the context of future-private, but not past-calendar, events.

Table 2
Mean Estimates, in Days, by Response Format, Study 2

	Response format	
Event	Day	Week and month
Future-private events		
Finish writing letter	5.0 (5.2)	15.9** (8.9)
Finish reading book	10.7 (10.6)	26.7** (20.3)
Buy semester's books	5.0 (3.6)	16.4** (9.4)
Find a job	4.8 (5.3)	16.7** (10.6)
Visit museum	17.6 (20.9)	27.4** (22.8)
Reach weight-loss goal	51.1 (38.6)	91.7* (55.7)
Attain fitness goal	40.0 (53.4)	68.5* (54.3)
Set up dorm room	5.4 (4.7)	9.8** (4.8)
Past-calendar events		
Lewinsky scandal	426 (175)	503* (195)
Columbine shootings	129 (25.7)	138 (28.0)
Turkey earthquake	23.7 (8.4)	24.4 (8.3)
Kennedy death	48.7 (14.9)	50.1 (16.4)
World Cup win	58.5 (20.9)	61.7 (19.3)

Note. All data are M (SD). More details about these items appear in the Appendix. p values correspond to Mann–Whitney comparisons between the estimates. We used this analysis because of skew in the data, but standard t tests yield highly similar results.
* $p < .05$. ** $p < .005$.

One might wonder whether the use of larger units essentially forced higher estimates among the future-private events due to a larger "graininess" in acuity: Respondents may have attempted to provide neat responses by rounding their estimates to, say, the nearest week. Note, however, that a rounding explanation would predict a similar discrepancy for past-calendar events, but there was no systematic influence of response format for those events. Note also that, in the context of the future events, the size of the format effect was substantially larger than rounding would warrant. For example, the estimated time to find a job increased from 4.8 days with a day format to 16.7 days with a week-and-month format, much beyond the 7-day estimate expected from merely rounding up to the nearest week.

This pattern is thus consistent with the notion that increased estimates under the week-and-month format are attributable not to rounding error but to differential absolute adjustment with larger versus smaller units when no date is available for conversion. We consider these effects again in the General Discussion, but first we replicate the findings in the estimation of physical distances.

[1] For approximately half of the participants in this study, the key manipulation (e.g., "on what date" vs. "how many days") was italicized to allay concerns that participants might not attend to the suggested units. Italicization had no effect.

[2] Because some months have 31 days, this formula is, if anything, conservative in testing the hypothesis that the week-and-month format increases estimates.

Study 3: Distance Estimation

When people estimate unknown temporal distances with provided units, people appear to adjust insufficiently from the present, particularly when the units of adjustment are small. The following studies replicate this pattern with physical distances, which also can be estimated by using various units. As with calendar dates, which can be estimated without anchoring on the present, distances can be estimated without adjusting from a starting point. When asked to estimate the distance between a restaurant and the theater, for example, people are likely to engage in some form of memory retrieval, mental comparison, or inferential reasoning; they likely will not simulate walking the distance. On the other hand, for those asked to provide an estimate in specific units, such as "footsteps," no simple inference or comparison is available. Although these individuals could do arithmetic (e.g., divide the estimated distance by the estimated length of a step), we predicted that this is not how people tend to make such estimates. Like people gauging temporal proximity, who may anchor on the present and adjust (rather than choose a date and compute the implied number of days), people estimating physical proximity in terms of specific units would begin at the suggested starting location and then adjust, typically insufficiently.

Method

Materials. Participants estimated distances between familiar campus landmarks. In the specified-unit condition, participants provided estimates in specific units, such as footsteps. The five specified-unit questions were (a) "How many steps would it take to cover the distance between the Garden Theater and Palmer Square?" (b) "How many cars, placed end to end, would it take to cover the distance between the Garden Theater and Palmer Square?" (c) "How many steps would it take to cover the distance between McCosh lecture hall and Firestone Library?" (d) "How many cans of soda, laid end to end, would it take to cover the distance between McCosh lecture hall and the Garden Theater?" and (e) "How many steps would it take to cover the distance between McCosh lecture hall and Forbes College?" Participants later estimated the relevant unit lengths (e.g., "How much distance do you cover in an average step?"; "How long is an average can of soda?").

Participants in the unspecified-unit condition were not given suggested units. Instead, they were asked to estimate the corresponding distances directly (e.g., "What is the distance from the Garden Theater to Palmer Square?").

Procedure and participants. Respondents were undergraduate volunteers. Some completed this task embedded among unrelated tasks in a 1-hr session; others were approached on campus and presented with a brief questionnaire. The number of participants recruited on these occasions ranged from 74 to 221, and each participant gave estimates for one or three distances, depending on the occasion. Questions were arbitrarily assigned and response format was randomly manipulated between participants. Results did not differ among occasions, and so the data were collapsed for analysis.

Results

All responses were converted to feet. In the unspecified-unit condition, this involved mere translation (e.g., from miles to feet). In the specified-unit condition, each participant's estimated number of units was multiplied by his or her estimated unit size to generate the estimated distance in feet. As before, we excluded outliers (Hinkle et al., 1994, p. 68).

Specified-unit estimates were smaller than unspecified-unit estimates for every item. Participants who estimated the distance from the Garden Theater to Palmer Square in number of steps produced significantly lower estimates (M = 500.6 ft [152.6 m], SD = 360.1 [109.8 m]) than did those in the unspecified-unit condition (M = 1,130 ft [344.4 m], SD = 629.8 [192.0 m]), $t(72) = 5.3$, $p < .001$. The actual distance was estimated from a map to be 1,030 ft (313.9 m); this is significantly greater than the specified-unit estimate, $t(37) = 9.08$, $p < .0001$, but is not different from the unspecified estimate, $t(35) = -0.95$, $p = .35$.

Another group of participants estimated the same distance in terms of car lengths or of unspecified units. Again, estimates were lower when units were specified (M = 557.9 ft [170.0 m], SD = 470.1 [143.3 m]) than when they were not specified (M = 1,103 ft [336.2 m], SD = 716.5 [218.4 m]), $t(220) = 6.81$, $p < .0001$. The former estimate was reliably shorter than the actual distance, $t(122) = -11.2$, $p < .0001$, whereas the latter was not reliably different, $t(98) = 1.00$, $p = .32$.

Estimates of the distance from McCosh to Firestone in footsteps were marginally lower than were those in unspecified units (M = 461.1 ft [140.5 m], SD = 356.4 [108.6 m], vs. M = 532.7 ft [162.4 m], SD = 356.1 [108.5 m]), $t(321) = 1.80$, $p = .07$. Specified-unit estimates were reliably lower than the actual distance of 578 ft (176.2 m), $t(171) = -4.30$, $p < .0001$, whereas unspecified-unit estimates were not, $t(150) = -1.56$, $p = .12$.

When the distance from McCosh to the Garden Theater was estimated in number of soda cans, specified-unit estimates diverged as predicted from unspecified-unit estimates (M = 603.1 ft [183.8 m], SD = 638.6 [194.6 m], vs. M = 998.7 ft [304.4 m], SD = 705.9 [215.2 m], respectively), $t(147) = 3.59$, $p < .001$. Use of unspecified units yielded slight overestimation of the actual distance of 859 ft (261.8 m), $t(71) = 1.68$, $p = .10$, whereas use of specified units again generated clear underestimation, $t(76) = -3.52$, $p < .001$.

Finally, the distance from McCosh to Forbes estimated in number of steps was significantly shorter than the distance estimated in unspecified units (M = 2,119 ft [645.9 m], SD = 1,505 [458.7 m], vs. M = 3,008 ft [916.8 m], SD = 1,631 [497.1 m]), $t(148) = 3.49$, $p < .001$. Both groups underestimated the actual distance of 3,750 ft (1,143 m; ps $< .001$).

Discussion

As with time estimates, distance estimates were significantly lower when gauged in specified small units than in unspecified units. In fact, responses obtained in the specified-unit condition reflected substantially greater error in the estimation of actual, and otherwise well-gauged, distances. We propose that such a bias arises in the specified-unit condition because participants begin at the suggested starting location (i.e., anchor on "here") and incrementally adjust by the suggested units, simulating progress to the endpoint. However, such adjustment is effortful and is consequently often insufficient, as participants stop near the lower end of the range of plausible responses (Epley & Gilovich 2004, 2006).

Recall that the specified-unit estimates were a function of participants' estimated number of units and the estimated unit size. Although we have attributed the persistently low estimates to an insufficient number of units, it is possible that the problem resides in an underestimation of unit size. Instead of misjudging the number of soda cans or steps required, participants perhaps underestimated the actual length of a can or a step. However, this does not appear to have been the case. For example, the average estimated length of a can of soda was 5.78 in. (14.7 cm; $SD = 1.27$ [3.2 cm]), reliably longer than a standard can's actual length of 5.0 in. (12.7 cm), $t(76) = 5.37$, $p < .001$. In general, there was not a reliable tendency to underestimate unit length; if anything, slight overestimation—as in the case of the soda cans—might suggest that the unit-based adjustments posited above were even more insufficient than is apparent.

Finally, it is worth noting that for known distances, as for time estimates, relatively more accurate conversion was observed. In a follow-up study, we asked 30 undergraduates, "How many cans of soda, laid end to end, do you think it would take to cover the length of a standard American football field, from one goal line to the other?" We also asked them to estimate the length of a can of soda. The average estimated length of the field (computed by multiplying each person's estimated number of cans by his or her estimated can length) was 302.2 ft (92.1 m; $SD = 57.6$ ft [17.6 m]), which was not different from the correct length of 300 ft (91.4 m), $t(29) = 0.21$, $p = .84$. Thus, participants are able to convert accurately from known lengths to suggested units, but they seem not to engage in such conversion when they estimate unknown lengths.

Study 4: Evidence for Adjustment

We have suggested that low specified-units estimates arise from a process of anchoring on the "here" or "now" followed by insufficient unit-based adjustment. Our case for insufficient adjustment, however, is admittedly indirect. Another mechanism with potentially similar manifestations would be for the suggested anchor to act as a prime: Recent evidence suggests that some anchors are briefly evaluated as potential responses to estimation questions and that confirmatory hypothesis testing ensues (Mussweiler & Strack, 1999, 2000; see Chapman & Johnson, 2002, for a review). Although the anchor is ultimately rejected as the correct answer, such a process activates information that is biased toward the anchor and that influences the final response. For example, when asked whether the mean temperature in Germany is higher or lower than 5°C, respondents may retrieve information consistent with Germany being rather cold; this accessible information may, in turn, promote low estimates of Germany's temperature without involving adjustment from the anchor (Mussweiler & Strack, 1999, 2000). It is possible that the unit-based questions in the previous studies prompted the testing of a hypothesis that the proposed unit (one day, one can, one footstep) was the correct answer. Such a hypothesis would have been quickly rejected, but the process, quite different from adjustment, may have rendered accessible thoughts that supported a low final response.

Although the results of Studies 1 through 3 do not allow us to tease apart these competing explanations conclusively, Epley and Gilovich (2005, 2006) have suggested a technique with which to elucidate the process underlying a given anchoring effect. They argue that the proclivity to engage in systematic, effortful thought ought to moderate the extent to which one adjusts from an anchor but is unlikely to moderate the influence of information inadvertently primed by the anchor. That is, if effects arise because participants adjust insufficiently from an anchor, one ought to be able to lead participants to produce less (or more) anchored responses by inducing them to adjust more (or less; Epley & Gilovich, 2001, 2005). If, on the other hand, effects arise because anchor-consistent information is primed, anchoring effects should operate more like standard knowledge-accessibility or priming effects, which tend to occur automatically (i.e., with little control and with little effect of ability to adjust). Indeed, Epley and Gilovich (2005, 2006) have found that increasing the proclivity to adjust reduces the effects of self-generated anchors (which have been argued to recruit adjustment processes; Epley & Gilovich, 2001) but does not alter the effects of experimenter-provided anchors (which have been attributed to knowledge accessibility, Mussweiler & Strack, 1999).

Thus, in Study 4a, participants estimated distances either in specified or unspecified units. We manipulated whether or not they were informed about insufficient adjustment (cf. Epley & Gilovich, 2005). We predicted that, if insufficient adjustment underlies low specified-unit estimates, participants made aware of adjustment should exhibit less bias. If, however, low estimates result from an automatic priming process, those estimates should remain relatively unaffected by such a manipulation. Note, incidentally, that we predicted no effect of learning about adjustment on unspecified-unit estimates, which were not assumed to involve adjustment.

Study 4a

Method

Materials. Three distances between campus landmarks were estimated. Respondents were randomly assigned to the specified- or unspecified-unit condition. The three specified-unit questions were: "How many steps do you think you would take if you walked from this room to the Reitz Union?" "How many steps do you think you would take if you walked from this room to the football stadium?" and "How many cans of soda, laid end to end, would it take to cover the distance from this room to the Century Tower?" The unspecified-unit questions requested estimates of the same distances but made no mention of units (e.g., "What do you think the distance is from this room to . . ."). Participants later also estimated the lengths of the units (footsteps and soda cans) that they had used.

Whereas roughly half of the participants responded without prior warning, the rest were randomly assigned to a "warning" condition, in which they read about the potential for insufficient adjustment (modeled after Epley & Gilovich, 2005, p. 207):

> Previous research has demonstrated that people's judgments are often biased by the first pieces of information that come to mind. For example, one study examined real estate agents who were setting prices for houses. Agents tended to be biased in the direction of the price of the *last* house they visited. This bias seems to occur because people start with the first number that comes to mind, and then they *insufficiently adjust* away from that value. In the following questions, you will need to generate estimates; some information will probably come to mind immediately. Please try not to be overly influenced by the first thing that comes to mind, and try not to adjust too little.

Participants and procedure. A sample of 484 undergraduate students completed this task, which was embedded in a packet of unrelated questionnaires, for course credit. As in Study 3, all responses were converted to feet and extreme outliers were excluded from analyses.

Results and Discussion

Responses were collapsed across items to derive each participant's average estimate. The averages were subjected to a 2 (unit, specified or unspecified) × 2 (adjustment warning, present or absent) ANOVA, which revealed a main effect of unit type, $F(1, 390) = 10.6$, $p < .001$. More important, this main effect was qualified by a reliable Unit × Warning interaction, $F(1, 390) = 4.02$, $p < .05$. The interaction is best understood by considering the impact of unit type in the two warning conditions (see Figure 1). In the absence of warning, participants in the specified-unit condition produced reliably lower estimates than did those in the unspecified-unit condition (M =1,800 ft [548.6 m], SD = 1,389 [423.4 m], and M = 2,569 ft [783.0 m], SD = 1,421 [433.1 m], respectively), $t(197) = 3.85$, $p < .001$. However, when participants were warned about insufficient adjustment, estimates in the specified- and unspecified-unit conditions proved statistically indistinguishable (M = 2,154 ft [656.5 m], SD = 1,689 [514.8 m], and M = 2,337 ft [712.3 m], SD = 1,266 [385.9 m], respectively), $t(193) = 0.86$, $p = .39$.

Warning about insufficient adjustment thus reduced the anchoring effect induced by the reliance on specified units. Because such warning has been shown to moderate anchoring effects that stem from insufficient adjustment but not those that are due to increased accessibility (Epley & Gilovich, 2005), the observed pattern lends support to the anchoring and adjustment interpretation. It is notable that, compared with no warning, warning tended to decrease estimates in the unspecified-unit condition, $t(202) = -1.23$, $p = .22$, albeit nonsignificantly. This result suggests that the impact of warning was not simply creation of a demand for larger estimates, and it is consistent with the notion that estimates in the unspecified-unit condition are not produced by adjusting upward from an anchor.

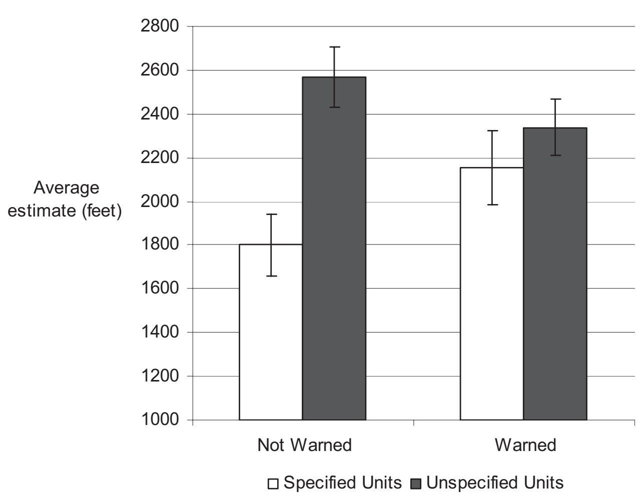

Figure 1. Average distance estimates, as a function of unit type and warning about insufficient adjustment, Study 4a. Error bars represent the standard error of the mean.

Study 4b

An alternative interpretation of Study 4a is that the warning manipulation did not affect adjustment per se; rather, it may have led participants to adopt a different, more successful strategy (perhaps generating an unanchored estimate and converting to the requested units), in a quest to be more accurate.[3] Although prior research suggests that such warnings about adjustment are effective only when adjustment is already operating (Epley & Gilovich, 2005), we conducted a follow-up study to lend further support to the adjustment interpretation. This study replicated Study 4a's procedure but included a condition in which the insufficient-adjustment warning was replaced with a paragraph of identical length that merely warned participants to be "accurate." We predicted that, if warning about adjustment prompts participants to adopt a wholly new strategy in a quest to increase their accuracy, the new accuracy instructions certainly should have a similar effect. If, on the other hand, the adjustment warning has specific effects on adjustment, the new instructions (which gave no insight into how accuracy could be achieved) should not be as effective (as respondents likely lack insight into the optimal strategy to deploy to be accurate; Arkes, 1991).

Method

Participants (N = 78 undergraduates, who participated for course credit) were randomly assigned to one cell of the 2 (unit, specified or unspecified) × 3 (warning, absent, adjustment, or accuracy) between-subjects design. After reading the warning (if any, depending on the condition), participants estimated the distance to the same three campus landmarks as in Study 4a. Participants completed this task along with other short tasks in the lab. As before, all responses were converted to feet and extreme outliers were excluded from analyses.

Results and Discussion

A 2 × 3 ANOVA on participants' average estimates revealed the predicted Unit × Warning interaction, $F(2, 64) = 2.92$, $p = .06$, as well as a main effect of unit type, $F(1, 64) = 21.3$, $p < .001$. Although the anchoring effect was replicated in the no-warning condition ($M_{\text{specified units}}$ = 1,672 ft, [509.6 m], SD = 792.1 [241.4 m]; $M_{\text{unspecified units}}$ = 2,841 ft [865.9 m], SD = 1,037 [316.1 m]), $t(22) = 3.10$, $p = .005$, this effect was attenuated (and was no longer reliable) when participants were warned about insufficient adjustment ($M_{\text{specified}}$ = 1,966 ft [599.2 m], SD = 1,361 [414.8 m]; $M_{\text{unspecified}}$ = 2,395 ft, [730.0 m], SD = 1,160 [353.6 m]), $t(19) = 0.78$, $p = .45$. Finally, when participants received an accuracy warning, with no mention of adjustment, the anchoring effect persisted ($M_{\text{specified}}$ = 1,004 ft [306.0 m], SD = 507.7 [154.7 m]; $M_{\text{unspecified}}$ = 2,964 ft [903.4 m], SD = 1,351 [411.8 m]), $t(23) = 4.72$, $p < .001$, and specified-unit estimates were somewhat diminished relative to when no warning was given at all, $t(22) = 2.46$, $p = .02$.

[3] We thank an anonymous reviewer for suggesting this possibility.

Note that, as predicted, one-way ANOVAs revealed that the warnings had no effect on unspecified-unit estimates, $F(2, 33) = 0.73$, $p = .49$, but reliably affected specified-unit estimates, $F(2, 31) = 3.21$, $p = .05$. This result is consistent with the suggestion that, when generating specified-unit estimates, people rely on a strategy (i.e., anchoring and adjustment) that can be accessed and improved upon with appropriate instructions. Warnings about general accuracy or inaccuracy, however, did not improve specified-unit estimates and, if anything, deteriorated them further, perhaps because respondents, not knowing specifically how to improve (Arkes, 1991), grew conservative and wary of overadjusting (Arkes, Dawes, & Christensen, 1986; Simmons, LeBoeuf, & Nelson, 2008).

Taken together, the results of Studies 4a and 4b are consistent with the hypothesis that specified-unit (in contrast to unspecified-unit) estimates are produced via a process of anchoring and adjustment. Providing information about insufficient adjustment increased specified-unit estimates, which were products of adjustment, but not unspecified-unit estimates, which were not products of the same, serial adjustment.

General Discussion

Time and distance estimates were influenced by the suggestion of particular measurement units. Prompting participants to respond in terms of small units produced lower estimates and, consequently, greater underestimation than did prompting larger, or no, units.

These response-format effects, robust and significant in the estimation of uncertain quantities, were not observed for quantities that were known or readily inferred. Naturally, estimation biases are most likely to arise when a readily computable response is lacking. For example, the number of African nations in the United Nations is not readily available and is thus susceptible to biased estimation, whereas the number of states in the United States is known by many and is thus unlikely to show anchoring effects. Of course, even generally familiar entities sometimes are not precisely known and must be estimated. Often, they can be gauged in the context of what is known; this leads to inferences such as, "It is 1 kilometer from home to the mall, and this distance is somewhat shorter, so I'll estimate it at 800 m." Such inferential reasoning seems less likely to occur when responses are given in unconventional units, such as footsteps, for which general knowledge is less directly applicable.

These effects are consistent with patterns observed in frequency estimates, for which higher estimates are observed when people estimate, for example, annual fatalities out of 100 rather than 10,000 individuals (Yamagishi, 1994, 1996, 1997). Yamagishi has suggested that anchoring and adjustment may underlie such effects in the estimation of rare events: Respondents might start from an anchor of "1" and adjust upward, with successive increments having a larger impact on estimated frequency when the denominator is 100 instead of 10,000. These and the present findings are consistent with much research showing that responses are often malleable and formed on the fly. People typically "accept" the perspective imposed by a problem; they produce responses in line with the proposed frame or the suggested units rather than convert between logically equivalent frames in an attempt to generate an unbiased response (LeBoeuf & Shafir, 2003; Tversky & Kahneman, 1981).

Mechanisms and Alternative Explanations

Studies 2 and 4 presented evidence generally supportive of the insufficient adjustment hypothesis. However, even in those studies the evidence for insufficient adjustment remained indirect; our methodology did not allow direct observation of adjustment processes, nor were we able to estimate, for example, separate anchoring and adjustment parameters to elucidate further the underlying processes. Thus, although insufficient adjustment provides a parsimonious explanation, other accounts are possible.

One possibility is that intervals described by days (or other temporal extents) seem longer than equivalent intervals described by dates, precisely because such extent descriptions highlight the expanse of time involved (LeBoeuf, 2006). Thus, "7 days" may seem longer than an interval that spans "May 1 to May 7"; this would lead participants to feel that a task is more likely to be completed during the former interval than the latter. This explanation, however, seems less able to account for the increased estimates observed in the week-and-month format (a format that highlights temporal extent, just like the day format) and for the fact that similar patterns occur in distance estimation.

As discussed in the context of Study 1c, another possibility is that a request for an estimate in terms of days implies greater temporal proximity than does one in terms of months or dates (cf. Grice, 1975). However, the fact that response format alters future but not past estimates casts doubt on a pure conversational norms account, inasmuch as one might expect such norms to apply equally to the reporting of past and future events. Beyond pure conversational norms, terms such as *day* might play a larger inferential role for potentially more uncertain future events (compared with past events); that said, in Study 1c we explored a closely related idea and found no evidence that the mere suggestion of smaller units alters respondents' perceived time frames.

On the whole, the data are consistent with an anchoring and adjustment interpretation, but the evidence remains indirect and other possible mechanisms cannot be conclusively ruled out. Further research focusing on a more direct estimation of the amount of adjustment or the precise nature of the anchors may provide more direct evidence of the underlying process.

Implications and Concluding Remarks

If anchoring and adjustment underlie the current findings, one should consider when people might naturally engage in such processes when they estimate times and distances. One boundary condition suggested thus far is that anchoring and adjustment are more likely when quantities are uncertain than when an answer can be readily retrieved and converted into the desired units. Even with uncertain quantities, however, unit-based estimates are more likely to arise spontaneously in some contexts than others. For example, the finding that estimates in days are more natural for near-term than for distant events (Golding et al., 1995) suggests that day-based adjustment, and unit-based adjustment more generally, may more naturally arise for events anticipated to occur in the near future. Similarly, in the case of distances, specific contexts are likely to promote estimation in apparently relevant units (e.g., footsteps when mentally gauging the dimensions of a room but city blocks when gauging downtown locations).[4] Such contexts are likely to trigger a process of adjustment from a low anchor,[5] and this adjustment will often prove insufficient, in part because it is effortful and people often stop when a plausible value is reached (Epley & Gilovich, 2004, 2006).

The current findings vindicate the tendency not to dwell on response format when studying the perceived temporal distance of well-defined past events (e.g., Huttenlocher et al., 1990; Kemp, 1999; Rubin & Baddeley, 1989). In making such estimates, people seem able to convert specific dates into the requested units. That said, response format is not immaterial to retrospective judgment. People generally prefer to estimate the time elapsed since an event than to provide a precise date (Huttenlocher et al., 1990). Also, assessments of the time elapsed since ill-defined historical events (e.g., "the beginning of modern psychology"), as opposed to concrete events ("the introduction of the telephone"), are typically lower when estimated in years than in calendar dates (Teigen, 1987), perhaps as a result of insufficient adjustment from the present.

As to future events, recall that studies of planning routinely observe an optimistic bias in predicted task-completion dates (e.g., Buehler et al., 1994). According to the present findings, such studies may have conservatively estimated the extent of this optimism. Had participants in planning fallacy studies estimated the number of days to task completion, a more pronounced planning fallacy might have been observed. Indeed, because estimation in days rather than dates is more natural for close-range events (Golding et al., 1995), one might expect the naturally occurring planning fallacy to be greater than previously anticipated, at least in short-range contexts. That said, if setting a target completion time provides motivation to be consistent with one's declared deadline (e.g., Festinger, 1957; Swann, Pelham, & Krull, 1989), estimation in days might lead to quicker completion. Of course, to the extent that no amount of effort will enable a person to meet an overly optimistic target, people may consistently fall short of day-based estimates. Such shortfalls would lead to lowered perceptions of self-efficacy and perhaps eventually to diminished motivation.

Indeed, the present findings may bear on person perception as well as self-perception. Many of us repeatedly make commitments and plans to which we are later unable to adhere. An important fact about intuitive heuristics that lead to biased judgment is that their causes and consequences are typically imperceptible. People may be prone systematically to misjudge the resources required for attaining goals and subsequently to misunderstand the causes of failures to achieve them. It seems that relying on small units when gauging uncertain temporal or physical distances can lead people to anchor on the "here" or "now" and to adjust insufficiently. The resulting underestimation, in turn, may have implications for planning and motivation.

[4] Also relevant is work suggesting that distance estimates are influenced by factors (such as the donning of a heavy backpack) that increase the perceived effort required to walk those distances (see Proffitt, 2006, for a review). Such findings suggest that distances that one anticipates walking may be gauged by simulating step-by-step progression.

[5] Our results do not shed light on whether the initial anchor is likely to be numeric (e.g., 1 day, one step) or whether it may initially be somewhat more abstract (e.g., today, this place). We remain agnostic on this point, inasmuch as our data lead us to conclude only that the anchor is likely quite low.

References

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110,* 486–498.

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes, 37,* 93–110.

Buehler, R., Griffin, D., & MacDonald, H. (1997). The role of motivated reasoning in optimistic time predictions. *Personality and Social Psychology Bulletin, 23,* 238–247.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67,* 366–381.

Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology, 10,* 519–540.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 120–138). Cambridge, England: Cambridge University Press.

Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science, 12,* 391–396.

Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin, 30,* 447–460.

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making, 18,* 199–212.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17,* 311–318.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin, 113,* 44–66.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment.* Cambridge, England: Cambridge University Press.

Golding, J. M., Magliano, J. P., & Baggett, W. (1995). Answering when questions about future events in the context of a calendar system. *Discourse Processes, 20,* 249–271.

Grice, H. P. (1975). Logic and conversation. In D. Davidson & G. Harman (Eds.), *The logic of grammar* (pp. 64–75). Encino, CA: Dickenson.

Griffin, D., & Buehler, R. (1999). Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology, 38,* 48–78.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1994). *Applied statistics for the behavioral sciences.* Boston: Houghton Mifflin.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation of preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67,* 247–257.

Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 196–213.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin, 21,* 1161–1166.

Janiszewski, C., & Uy, D. (2008). Precision of the anchor influences the amount of adjustment. *Psychological Science, 19,* 121–127.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames.* Cambridge, England: Cambridge University Press.

Kemp, S. (1999). An associative theory of estimating past dates and past prices. *Psychonomic Bulletin and Review, 6,* 41–56.

LeBoeuf, R. A. (2006). Discount rates for time versus dates: The sensitivity of discounting to time-interval description. *Journal of Marketing Research, 43,* 59–72.

LeBoeuf, R. A., & Shafir, E. (2003). Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making, 16,* 77–92.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89,* 46–55.

Lichtenstein, S., & Slovic, P. (1973). Response-induced reversals of preferences in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology, 101,* 16–20.

Loftus, E. F., & Marburger, W. (1983). Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory & Cognition, 11,* 114–120.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35,* 136–164.

Mussweiler, T., & Strack, F. (2000). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology, 78,* 1038–1052.

Myers, J. L., & Well, A. D. (1995). *Research design and statistical analysis.* Hillsdale, NJ: Erlbaum.

Neter, J., & Waksberg, J. (1964). A study of response errors in expenditures data from household interviews. *American Statistical Association Journal, 59,* 18–55.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1,* 110–122.

Rubin, D. C., & Baddeley, A. D. (1989). Telescoping is not time compression: A model of the dating of autobiographical events. *Memory & Cognition, 17,* 653–661.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation.* Mahwah, NJ: Erlbaum.

Schwarz, N., Groves, R. M., & Schuman, H. (1998). Survey methods. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 143–179). Boston: McGraw Hill.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition, 21,* 546–556.

Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology, 53,* 491–517.

Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2008). *The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors?* Manuscript submitted for publication.

Swann, W. B., Pelham, B. W., & Krull, D. S. (1989). Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *Journal of Personality and Social Psychology, 57,* 782–791.

Tanur, J. M. (Ed.). (1992). *Questions about questions.* New York: Russell Sage Foundation.

Teigen, K. H. (1987). Fortiden i vare hender [The past in our hands]. In J. P. Myklebust and R. Ommundsen (Eds.), *Psykologiprofesjonen mot ar 2000* (pp. 378–383). Bergen, Norway: Universitesforlaget.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and psychology of choice. *Science, 211,* 453–458.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125,* 387–402.

Wong, K. F. E., & Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes, 82,* 314–333.

Yamagishi, K. (1994). Consistencies and biases in risk perception: I. Anchoring process and response-range effect. *Perceptual and Motor Skills, 79,* 651–656.

Yamagishi, K. (1996). Effects of response range in frequency judgments of common versus uncommon events. *Perceptual and Motor Skills, 82,* 1371–1376.

Yamagishi, K. (1997). Upward versus downward anchoring in frequency judgments of social facts. *Japanese Psychological Research, 39,* 124–129.

(*Appendix follows*)

Appendix

Stimuli and Results for Study 1a

Target event	Day-format mean estimate	Date-format mean estimate	$p <$	Cohen's d
Future-private events: Participants were asked to estimate the extent of time until they . . .				
finished a letter they had resolved to write (n = 67)	5.0 days	12.7 days	.001	−1.28
finished reading a book they had resolved to read (n = 64)	10.7 days	36.8 days	.001	−1.24
did their laundry (n = 38)	3.5 days	7.3 days	.02	−0.91
bought their books (n = 73)	5.0 days	10.1 days	.001	−0.90
found a job (if trying to find a job; n = 20)	4.8 days	10.3 days	.14	−0.84
visited a museum they had resolved to visit (n = 73)	17.6 days	35.0 days	.001	−0.83
finished a paper they were planning to write (n = 58)	13.8 days	22.5 days	.01	−0.74
reached a weight-loss goal (if losing weight; n = 71)	51.1 days	77.1 days	.04	−0.60
attained a non-weight-loss fitness goal (if applicable; n = 37)	40.0 days	54.7 days	.03	−0.34
had dinner off campus (n = 31)	4.8 days	7.1 days	.97	−0.32
checked their e-mail[a] (n = 43)	62.2 min	70.7 min	.68	−0.18
set up their dorm rooms for the academic year (n = 79)	5.4 days	6.2 days	.54	−0.16
got a haircut (n = 87)	38.6 days	39.6 days	.31	−0.03
saw a movie in a theater (n = 84)	19.0 days	19.4 days	.56	−0.03
spoke to their families on the telephone (n = 70)	2.6 days	2.5 days	.99	0.08
Future-calendar events: Participants were asked to estimate the extent of time until . . .				
classes started in the fall (n = 90)	164 days	175 days	.001	−1.05
Mother's Day (n = 119)	23.3 days	26.5 days	.03	−0.41
Dean's Date (the date papers are due; n = 192)	32.8 days	33.5 days	.11	−0.38
the university's next graduation (April respondents; n = 127)	40.1 days	41.8 days	.07	−0.35
they had a birthday (n = 133)	155 days	183 days	.13	−0.25
the university's next graduation (December respondents; n = 210)	168 days	170 days	.36	−0.23
the new president of the United States took office (n = 67)	289 days	294 days	.17	−0.22
Spring Break begins (n = 208)	96.7 days	98.0 days	.08	−0.14
Houseparties (an end-of-year tradition; n = 102)	10.2 days	10.3 days	.64	−0.13
the next presidential election (n = 171)	325 days	327 days	.39	−0.12
they graduated from college (n = 91)	813 days	830 days	.43	−0.04
Easter (n = 134)	5.0 days	5.0 days	.62	0.03
the last day of classes (n = 102)	10.2 days	10.1 days	.35	0.19
Valentine's Day (December respondents; n = 174)	66.7 days	66.2 days	.23	0.28
President's Day (n = 57)	12.9 days	10.8 days	.10	0.45
Valentine's Day (February respondents; n = 84)	6.1 days	6.0 days	.04	0.48
Past-private events: Participants were asked to estimate the extent of time since they . . .				
went out with their friends (n = 111)	3.4 days	3.8 days	.13	−0.30
last began checking their e-mail[a] (n = 54)	98.2 min	123 min	.16	−0.29
received a letter in the mail (n = 200)	3.8 days	4.9 days	.12	−0.24
resolved to lose weight (if trying to lose weight; n = 89)	24.1 days	31.1 days	.08	−0.23
resolved to complete a paper (n = 166)	5.8 days	7.0 days	.64	−0.19
resolved to read a book (n = 167)	22.6 days	26.2 days	.95	−0.11
did their laundry (n = 123)	6.3 days	6.7 days	.68	−0.09
saw a movie in a theater (n = 193)	16.6 days	16.8 days	.34	−0.01
got a haircut (n = 206)	42.9 days	40.8 days	.68	0.06
had dinner off campus (n = 195)	10.7 days	9.8 days	.23	0.10
received an e-mail from an instructor (n = 114)	2.9 days	2.6 days	.32	0.12
spoke to their parents on the phone (n = 202)	2.6 days	2.3 days	.54	0.13
last started eating dinner[a] (n = 36)	123 min	102 min	.53	0.37
Past-calendar events: Participants were asked to estimate the extent of time since . . .				
Valentine's Day (n = 32)	205 days	209 days	.26	−0.86
Spring Break began (n = 73)	189 days	205 days	.001	−0.73
Dean's Date of the previous semester (n = 85)	139 days	153 days	.005	−0.55
the university's most recent graduation (n = 73)	126 days	132 days	.13	−0.38
Easter (n = 42)	134 days	142 days	.37	−0.37
the Monica Lewinsky scandal first became news (n = 53)	426 days	487 days	.33	−0.33

(*Appendix continues*)

Appendix (*continued*)

Target event	Day-format mean estimate	Date-format mean estimate	$p <$	Cohen's d
the shootings at Columbine High School (n = 123)	129 days	136 days	.24	−0.29
Father's Day (n = 89)	105 days	111 days	.54	−0.24
the university's president announced his resignation (n = 57)	15.4 days	16.2 days	.29	−0.19
a devastating earthquake hit Turkey (n = 83)	23.7 days	24.5 days	.61	−0.09
classes started in the fall (n = 60)	30.3 days	30.3 days	.97	0
John F. Kennedy, Jr., was killed in a plane crash (n = 158)	48.7 days	48.0 days	.75	0.04
Mother's Day (n = 98)	140 days	138 days	.67	0.06
the United States won the Women's World Cup (n = 151)	58.5 days	56.0 days	.67	0.13

Note. p values refer to Mann–Whitney comparisons between the estimates. We used this analysis because of skew in the data, but standard t tests yield highly similar results. Numbers of participants exclude those who gave estimates classified as outliers.
[a] For these items, participants reported either minutes (corresponding to day format) or a specific time (e.g., 7:00, corresponding to date format).

Received April 20, 2006
Revision received July 15, 2008
Accepted July 21, 2008 ■