How Should Time Estimates be Structured to Increase Consumer Satisfaction?

Beidi Hu[1], Celia Gaertig[2], and Berkeley Dietvorst[3]

[1]The Wharton School, University of Pennsylvania, [2]Haas School of Business, University of California, Berkeley, and [3]Booth School of Business, University of Chicago

Poster session: Friday Feb. 11, 3:45pm-4:45pm ET
https://upenn.zoom.us/j/99001005367

Research Question

Consumers often encounter time estimates in inherently uncertain contexts (e.g., driving, food delivery). How should these estimates be structured to increase consumers' satisfaction?

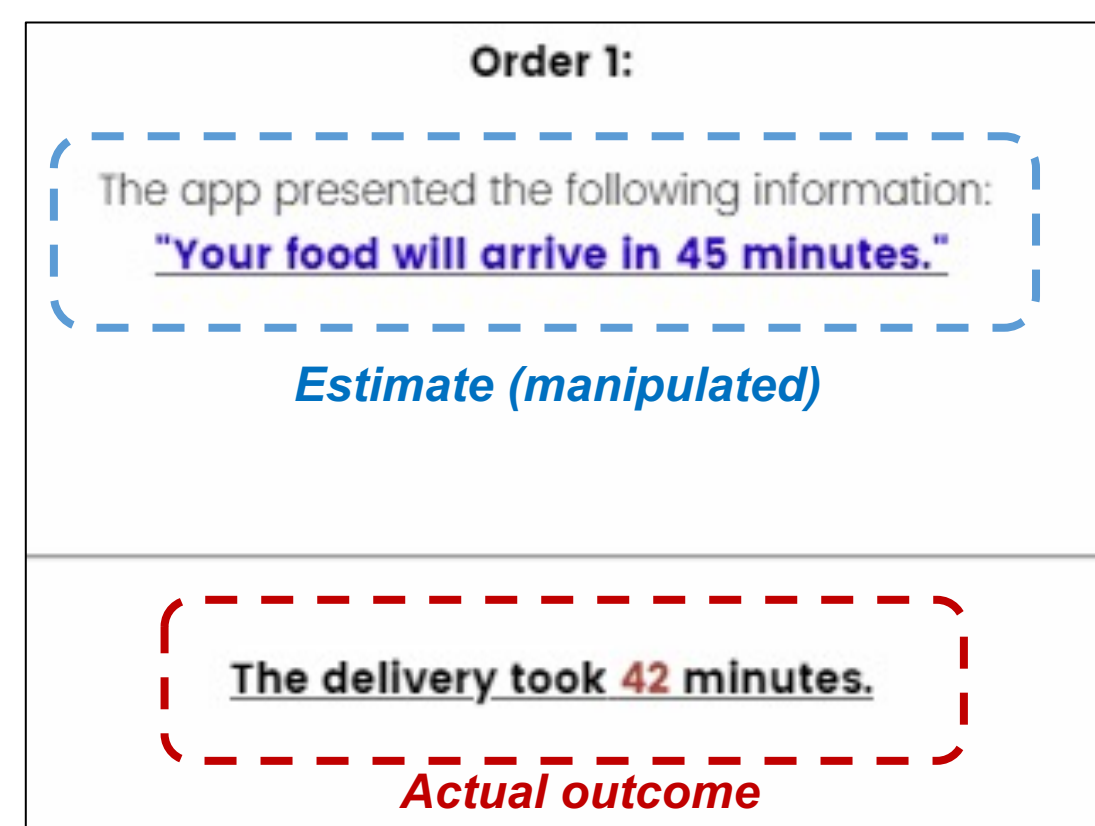The current research examines two aspects of this question:

(1) Do consumers prefer point estimates or ranges?

(2) Where in the distribution of outcomes should companies put the estimates?

Study Design

Nine pre-registered studies (N = 6,301)

Procedure: 20 simulated trials with a hypothetical food delivery app / GPS app

Sample stimulus (Study 1):

> Order 1:
> The app presented the following information:
> **"Your food will arrive in 45 minutes."**
> *Estimate (manipulated)*
>
> **The delivery took 42 minutes.**
> *Actual outcome*

Dependent measures:

Subjective Liking of the App (α ≥ 0.92)
1. How much do you like the delivery (arrival) estimates provided by this app? (7-point)
2. How informative do you find the delivery (arrival) estimates provided by this app? (7-point)
3. How useful do you find the delivery (arrival) estimates provided by this app? (7-point)
4. How accurate do you think this app is? (7-point)
5. How much do you trust this app? (7-point)

Perceived Accuracy of the App's Predictions
1. What percentage of the app's time estimates are very accurate? (0% - 100%)
2. What percentage of the app's time estimates are very inaccurate? (0% - 100%)
3. When this app gives an inaccurate estimate, how many minutes is the app off on average? (0 minute to 20 minutes)
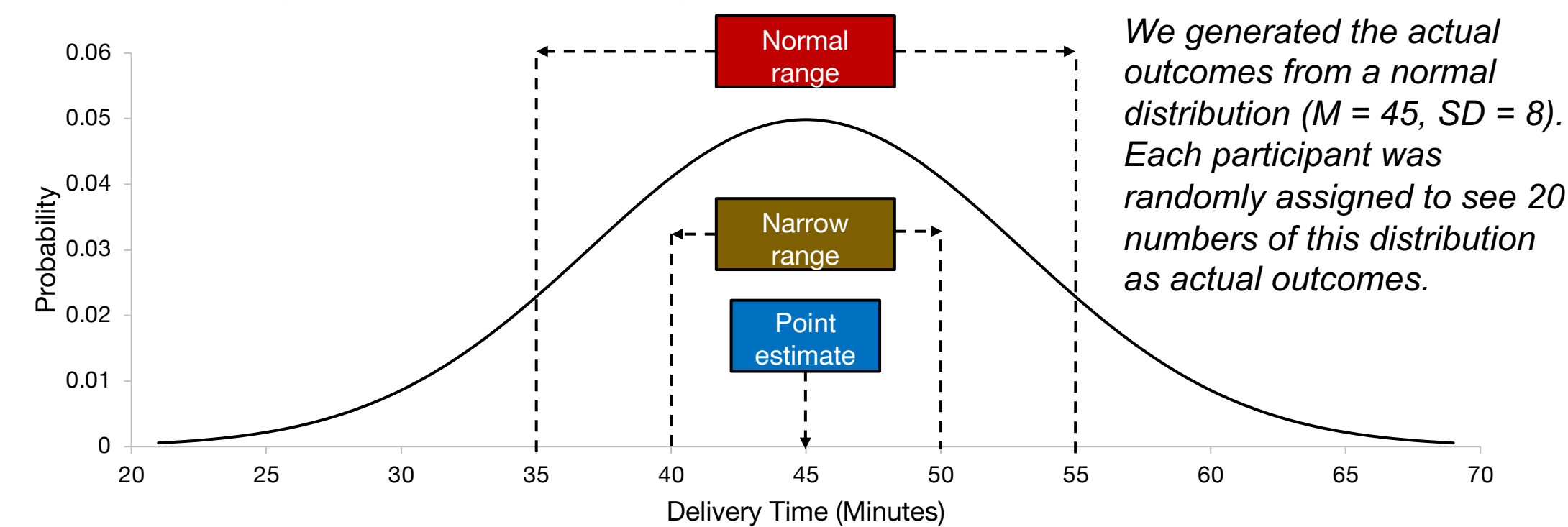
Pre-registrations, data, and materials available at:
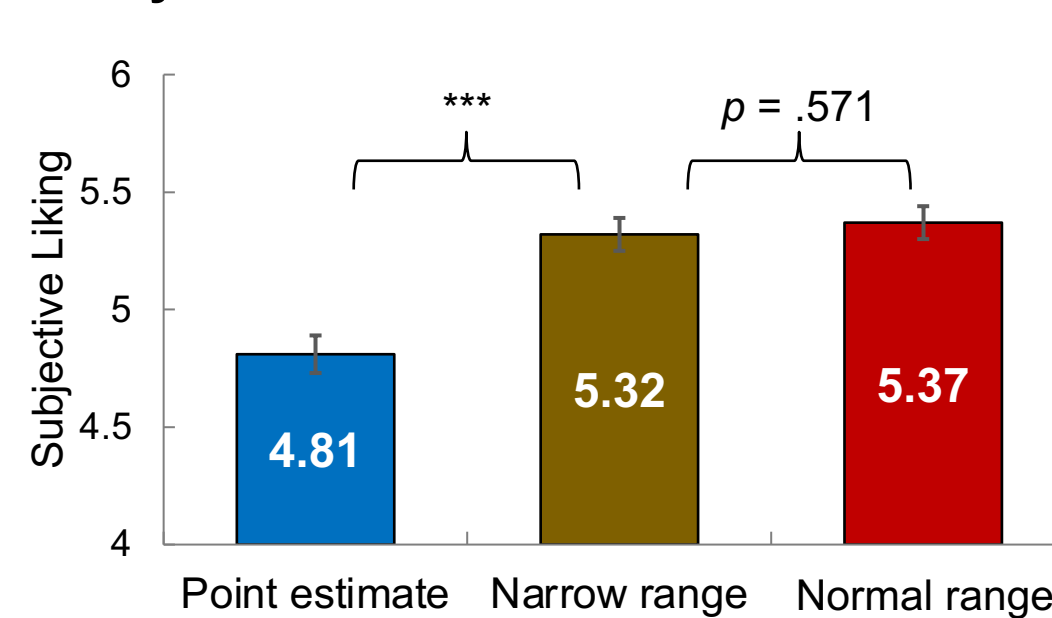https://researchbox.org/482&PEER_REVIEW_passcode=KNDQRR

(1) People prefer range estimates to point estimates…

Study 1 (N = 698, MTurk, pre-registered)
Conditions: Point estimate (45 minutes); Narrow range (40-50 minutes; 50% CI); Normal range (35-55 minutes; 80% CI)

We generated the actual outcomes from a normal distribution (M = 45, SD = 8). Each participant was randomly assigned to see 20 numbers of this distribution as actual outcomes.

Study 1 Results

	Point estimate	Narrow range	Normal range
Subjective Liking	4.81	5.32	5.37

*** (Point estimate vs Narrow range); p = .571 (Narrow range vs Normal range)

The result is robust to:
✓ Different locations of point estimates in the distribution (Studies 4-5)
✓ Different underlying distributions (normal vs. log-normal; Study 6)
✓ Different duration lengths (Study 6)
✓ Different contexts (food delivery vs. driving; Study 6)

…as long as the ranges are not excessively wide.

Study 2 (N = 599, MTurk, pre-registered)

New condition:
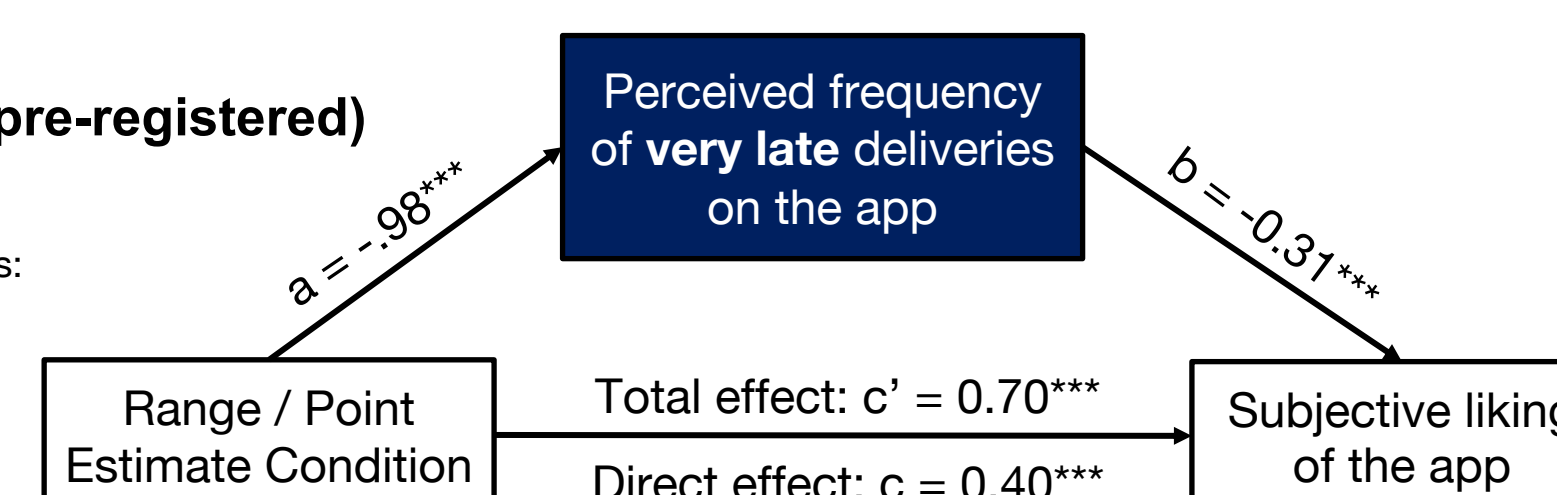Really wide range (15-75 minutes; 99% CI)

Study 2 Results

	Point estimate	Normal range	Really wide range
Subjective Liking	4.62	5.31	4.32

*** / *** / *
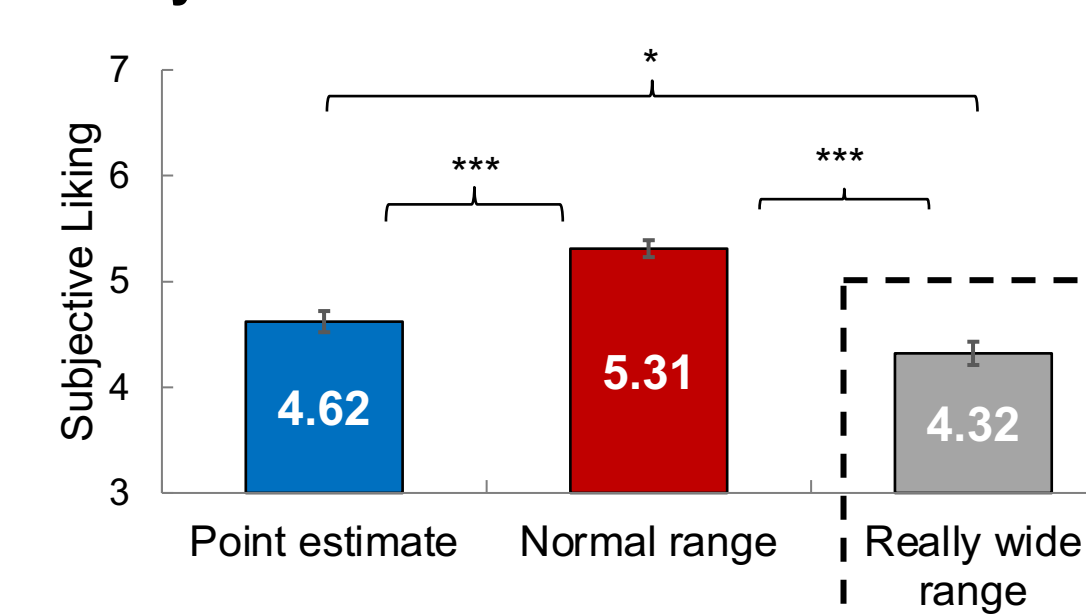
The preference for ranges is mediated by the perception of *extremely late* outcomes.

Study 3 Results (N = 485, MTurk, pre-registered)

Bootstrap with 1,000 samples:
Indirect effect = 0.31,
95% CI = [0.17, 0.44]

Range / Point Estimate Condition → Perceived frequency of **very late** deliveries on the app (a = -.98***) → Subjective liking of the app (b = -.31***)
Total effect: c' = 0.70***
Direct effect: c = 0.40***

(2) No simple answer regarding the preference of time estimates' location:

People's preference for time estimates' location differs across domains.

In Studies 4-6, we additionally manipulated the location of estimates.

Studies 4-6 Results (Ns = 801, 998, 888, pre-registered):

When evaluating a food delivery app: Participants liked conservative estimates (that overestimate durations) no less than accurate (i.e., "best guess") estimates (Studies 4-5).
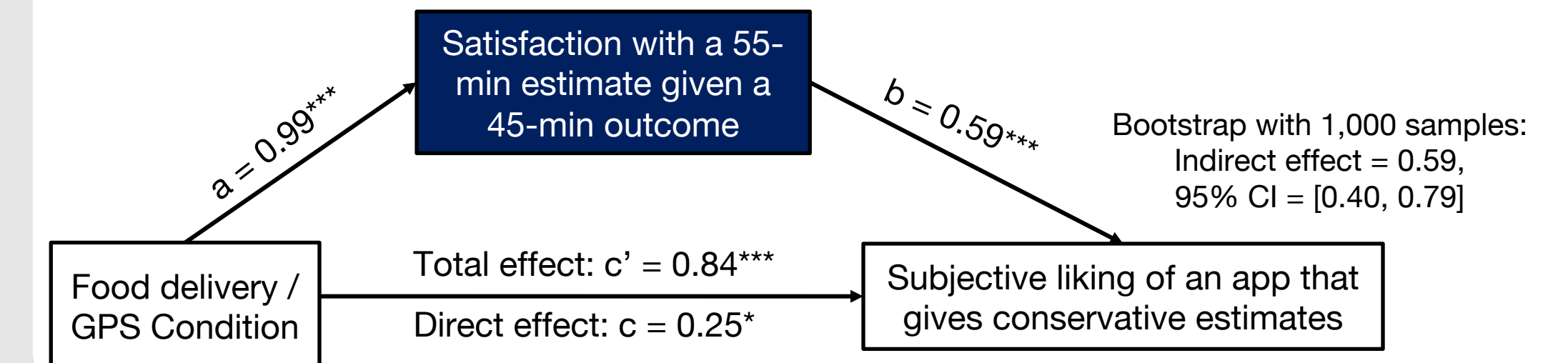
When evaluating a GPS app: Participants preferred accurate estimates (Study 6).

These domain differences can be predicted by different preferences for early outcomes.

Study 8 (N = 371, Prolific, pre-registered)
- We manipulated whether participants evaluated a food delivery app or a GPS app.
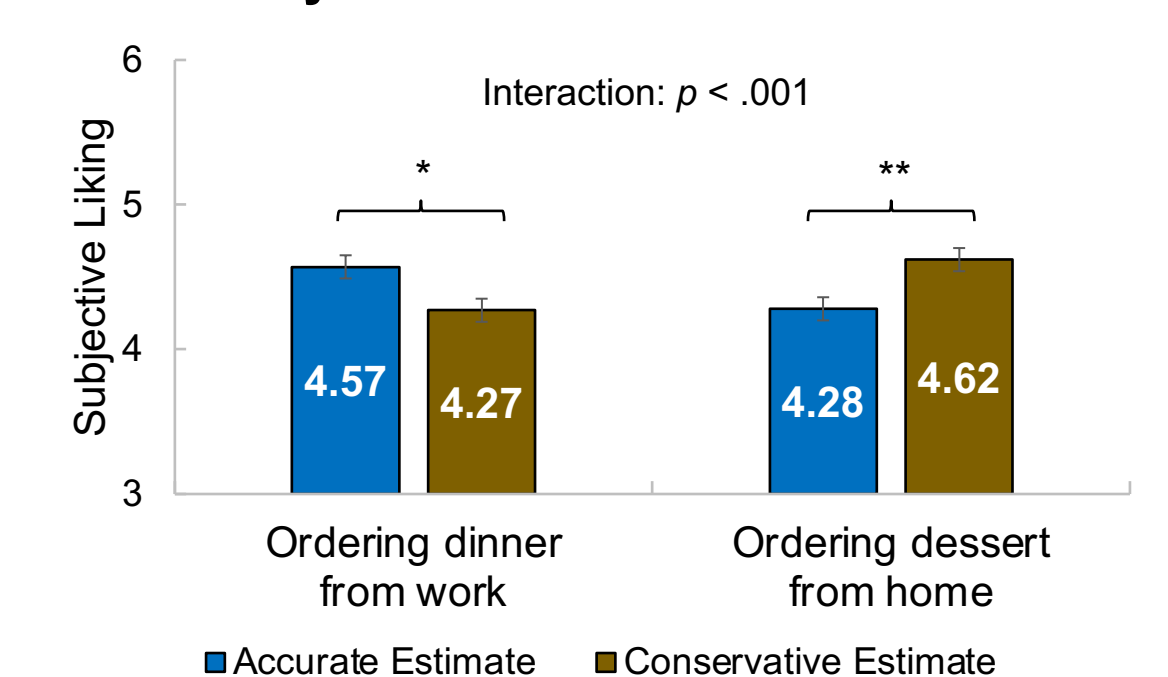- All participants saw conservative estimates.

Food delivery / GPS Condition → Satisfaction with a 55-min estimate given a 45-min outcome (a = 0.99***) → Subjective liking of an app that gives conservative estimates (b = 0.59***)
Total effect: c' = 0.84***
Direct effect: c = 0.25*
Bootstrap with 1,000 samples: Indirect effect = 0.59, 95% CI = [0.40, 0.79]

Even within the same domain, preference for early outcomes predicts whether people prefer an accurate or conservative estimate.

Study 9 (N = 1,088, Prolific, pre-registered)
- We manipulated within the food delivery domain people's preference for an early outcome.
- **Conditions:** 2 (location: accurate vs. conservative) x 2 (context: ordering dinner from work vs. ordering dessert from home)

Study 9 Results

Interaction: p < .001

	Ordering dinner from work	Ordering dessert from home
Accurate Estimate	4.57	4.28
Conservative Estimate	4.27	4.62

* (Ordering dinner from work); ** (Ordering dessert from home)

■ Accurate Estimate ■ Conservative Estimate

All error bars are +/-1 SE. * p < .05; ** p < .01; *** p < .001