

Imperfect Attention in Discrete Choice Experiments

Response to Jing Li

Sen Zeng

Cornell University

Annual Health Econometrics Workshop

Hunter College, CUNY

October 2, 2026

Recap: What this paper asks about DCEs

How reliable are the insights we draw from stated choices when respondents do not fully process the attributes?

- A weak response may reflect a weak preference or incomplete processing *in the survey*.
- I compare counterfactual predictions relative to each model's own status quo.
- I currently do not assume survey inattention persists in actual decisions.

1. Tastes and attention: two restrictions

Comment: Can tastes learned in attentive states describe non-attenders?

The intended exclusion: W shifts survey attention, not the taste distribution.

The current likelihood also factorizes:

$$p(A_i, \beta_i \mid W_i) = p(A_i \mid W_i) f(\beta_i).$$

This additionally imposes population independence between attention and tastes.

Respondents contribute to states probabilistically. But when legality is ignored, their fully processed legality taste is not directly revealed by choices.

1. Do attention groups have different tastes?

Comment: Allow constants and price tastes to vary with legality attention.

All four constants and price means vary by latent group. Selected means:

Legality attended to	Menthol-cigarette constant	Price mean
Neither legality attribute	5.451	−0.466
Cigarette legality only	5.960	−0.438
E-cigarette legality only	5.803	−0.497
Both legality attributes	5.875	−0.424

This exercise lets W affect legality tastes, removing the original exclusion.

The point estimates differ, but equality has not been tested.

2. Why exclude survey-process variables?

Comment: Completion time, phone use and repeat visits may also reflect tastes.

Here, W contains log DCE completion time, phone use and multiple survey visits.

A committed menthol smoker may answer quickly because the choice is easy, not because the respondent skipped the attributes.

- The baseline uses W to distinguish processing from preferences; the exclusion needs substantive justification.
- Letting *all* of W enter tastes removes that identifying restriction. Identification would then need a separate argument.

2. What the expanded taste specification shows

Comment: Let W enter taste means and compare the estimates.

All three variables enter nine taste locations.

	Restricted model	W also in tastes
Menthol-cigarette legality attendance	48.2%	47.8%
Menthol-cigarette street-illegality mean	-5.123	-5.035
Change in quit share (percentage points)	12.01	12.09

Quit changes compare joint street-illegality with each model's status quo.

These quantities change little, but the expanded model no longer retains the exclusion.

3. What recall adds to choices

Comment: What information does recall contribute beyond choices?

Recall adds a noisy signal of processing; it is not needed to run EM.

Attribute	Fitted probability of correct recall	
	$A = 0$	$A = 1$
Menthol-cigarette legality	42.6%	64.9%
Menthol e-cigarette legality	31.1%	44.1%
Non-menthol-cigarette price	31.6%	43.2%
Menthol-cigarette price	45.5%	62.0%
Tobacco e-cigarette price	26.8%	34.3%
Menthol e-cigarette price	34.3%	43.7%

All gaps are not negligible, recall is informative but far from a perfect attention label. A richer measurement model is future work.

3. Sensitivity to the recall weight

Comment: Does choosing $\omega = 0.1$ drive the findings?

Parametric model. Quit changes are relative to each model's status quo under joint street-illegality.

Recall weight	Menthol-cigarette street-illegality mean	Change in quit share (percentage points)
0 (recall-free start)	-4.411	10.94
0.05	-4.197	11.29
0.10	-4.205	11.35
0.50	-4.050	11.62
1.00	-3.289	8.31

Low weights give similar contrasts; full joint-likelihood weight ($\omega = 1$) does not.
Cross-validation deferred to next steps.

4. DCE comparisons, not forecasts of cessation

Comment: How should we interpret the predictions and their baseline gaps?

These are DCE-based comparisons, not an assumption that survey inattention persists in actual decisions.

Model	Status quo quit share	Both menthol products street-illegal	Change (pp)
Benchmark	15.66%	26.38%	10.71
Parametric attention	10.41%	21.76%	11.35
ML attention	10.29%	21.61%	11.32

Changes are similar even though levels differ. Subtracting each model's status quo focuses on the scenario contrast (it may not circumvent hypothetical bias)

4. What if legality were fully processed?

Comment: Set legality attention to one.

Change in quit share: both menthol products street-illegal minus own status quo.

Model	Survey processing (percentage points)	Full legality processing (percentage points)
Parametric attention	11.35	19.15
ML attention	11.32	19.05

Full legality processing raises the DCE-implied change by about 8 points.

This compares processing assumptions, not actual cessation or a general correction for hypothetical bias.

5. Can attention change across tasks?

Comment: Use task-level timing to capture fatigue or learning.

A respondent could read carefully at first and skim later. One post-DCE recall answer cannot locate that change.

- The current model uses one attention state across 12 tasks as a simplification, not evidence that processing is constant.
- Will check whether usable timestamps can be linked to individual choices, then consider task-varying attention.
- Timing also reflects difficulty and interruptions. It would remain a noisy measurement, not an attention label.

5. Building better attention shifters

Comment: Randomize attribute order or screen position.

A future survey could present the same attributes in randomly assigned positions, alongside recall and task-level timing.

- Random assignment separates the display from pre-existing preferences, unlike naturally occurring completion time.
- Exclusion still requires the display to affect choices through the modeled processing channel.
- This could strengthen the identification strategy.

6. When is an attention model useful?

Comment: Can divergence from full attention diagnose DCE quality?

Great question! The intended use is to ask whether the substantive answer depends on how survey processing is modeled.

- Nesting full attention allows the models to agree; it does not guarantee agreement in finite samples.
- Differences may also reflect measurement restrictions, taste distributions or numerical fitting.
- In the original validation, later-task probability-based prediction improves, but new-respondent and aggregate-share prediction do not.

Disagreement is a good reason to investigate.

6. From this application to design guidance

Comment: Simulate when inattention and task complexity change the answer.

Great idea! This application studies one threat to DCE interpretation; a simulation could show when the extra modeling and recall questions help.

Vary	Inattention, task complexity and recall noise.
------	--

Compare	Benchmark, choice-only attention and recall-augmented attention.
---------	--

Assess	Taste recovery, scenario contrasts and false alarms under full attention.
--------	---

The goal is to understand which insights from a DCE are reliable, and under which assumptions.