The Cost of Progress: How Ad Creative Features Drive User-Ad Interaction in Blind Lead-In Advertising

Taiwan Summer Workshop on Information Management, TSWIM 2026

Jaewon ("Jay-one") Yoo

Institute of Service Science, College of Technology Management
National Tsing Hua University

co-authored with
Seokjoon Yoon (Ridi Corp.)
Minki Kim (KAIST)
Wonjoon Kim (KAIST)
Min Sok Lee (Univ. of Chicago)

The BLI Format

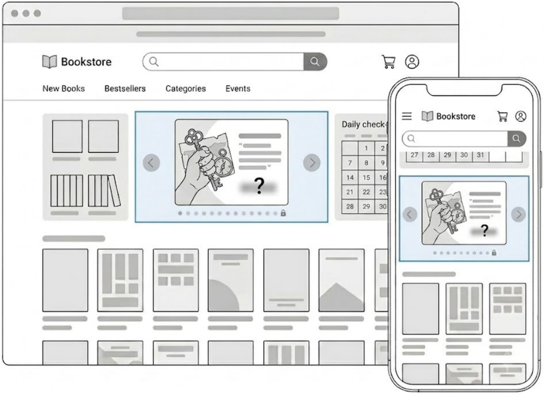
Users swipe through story panels; the book is revealed only on the last panel, with a click-through button.

story panels (swipe) ⟶ reveal + click-through

each ad: 10–20 story panels

- Product identity appears only after progression; completion (reaching the reveal) is the behavioral gate.
- Only 3.3% of treated users complete (0.7–7.7% across ads) $\Longrightarrow$ assignment to BLI $\neq$ exposure to the product.

Ad Placement



- **Paid** media (Instagram, TikTok) vs. **owned** media (the firm's own platform).
- Ours: the retailer's **owned** e-commerce site (daily check-in); no incremental media cost.
- Users meet it amid routine browsing; BLI hides the product until the reveal.

Schematic illustration informed by the retailer's app/web environment; the field experiment ran on the daily check-in page.

Motivation

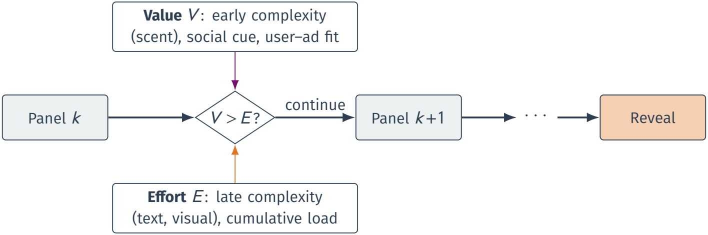
- **Big market, no evidence.**
  - Interactive ads (carousel, story, slider) are everywhere; the market is growing fast ($40B to $120B by 2030).
  - Firms run them already, but no one has tested whether they work or how to design them.
- **Most users never reach the product.**
  - It appears only at the end, and most stop before they get there.
  - So the ad works only if users keep going.

Research Questions

- **RQ1. Does interactive BLI outperform a static banner on purchase conversion?**
  - Practitioners already use sequential/card-based ad formats; a close analogy is carousel advertising, but here the ad runs on an owned retail platform.
  - The field experiment compares the BLI slider against the retailer's static-banner control, not against a no-ad condition.
- **RQ2. Which creative features make users complete the reveal sequence?**
  - Completion is the behavioral gate: non-completers do not learn the product identity.
  - Panel-level creative features give a way to study user–ad interaction inside the format.

Cost of Progress (CoP) Framework

Each panel's value and effort are functions of its features, $V_k = v(x_k)$ and $E_k = e(x_k)$; the user advances while $V_k > E_k$.



- **Stage asymmetry:** the same complexity is *value early, effort late*: information scent at the entry panel, processing load once it accumulates.
- **Completion is the causal channel to purchase:**
  - Assignment to BLI (**H1a**) vs. actually completing it (**H1b**).

Theory & Hypotheses

- **Context-specific theorizing** (Hong et al. 2014, *ISR*): established theories are adapted to the interactive-BLI setting to generate stage-specific predictions.
- Our contribution is explanation-and-prediction (Gregor 2006, *MISQ*), not a new grand theory.

Established account	Mapping to interactive BLI	Prediction
Information foraging / scent (Pirolli & Card 1999)	early panels signal the value of continuing	entry text ↑ completion (**H2**)
Cognitive load / ad complexity (Pieters et al. 2010)	later panels accumulate processing effort	late text & visual complexity ↓ completion (**H3**)
Diagnosticity & narrative transportation (Jiang & Benbasat 2004; Green & Brock 2000)	faces add interpretability and engagement	faces ↑ completion (**H4**: role of social cues does not change in early vs. late stages, “stage-agnostic”)
Personal relevance / fit (Petty et al. 1983)	genre match raises the expected payoff	user–ad fit ↑ completion (**H5**)
Affect-as-information (Schwarz & Clore 1983)	post-completion feelings inform judgment	positive experience ↑ persuasion (**H6**)

Related Literature

- **User interaction and continuation with digital artifacts**
    - How users decide to keep exploring digital content: information scent / foraging (Pirolli & Card 1999, *Psychol. Rev.*); interactivity and diagnosticity (Jiang & Benbasat 2004, *JMIS*; Jiang & Benbasat 2007, *ISR*; Yi, Jiang & Benbasat 2017, *ISR*; Animesh et al. 2011, *MISQ*).
- **Delayed reveal, interactivity, and ad measurement**
    - Delayed branding in *passive* media (Romaniuk 2009, *JAR*; Guitart et al. 2018, *IJRM*); interactive ads and field-experiment measurement (Goldfarb & Tucker 2011, *Mktg. Sci.*; Gordon et al. 2019, *Mktg. Sci.*; Gordon et al. 2023, *Mktg. Sci.*).
- **Contextual IS theorizing:** adapt established theory to a new context (Gregor 2006, *MISQ*; Hong et al. 2014, *ISR*).

Field Experiment

Customers (54,901 treatment / 54,255 control)	109,156
BLI ads, one per book title	76
Panels per ad (swipe-through slides)	10–20
Check-in impressions (unit of observation)	1,443,059
Experiment length	80 days
Treatment vs. control	BLI slider vs. static banner
Platform	retailer's daily check-in page

- **Design:**
  - Assignment is randomized at the individual level; completion (reaching the reveal) is endogenous.
- **Creative-feature coding:**
  - Coded panel-by-panel with the Google Cloud Vision API (faces, color diversity) and text parsing (character counts); panels grouped into entry / middle / late stages.
  - Features: text richness, visual complexity, social cue (faces), user–ad fit.

Empirical Strategy: Conversion

- **Encouragement design with one-sided noncompliance.**
  - Z = assignment to BLI (randomized); D = completion (reach the reveal); Y = purchase (same-day, 30-day).
  - Control users cannot complete by construction $\Rightarrow$ Z instruments D (Angrist et al. 1996).
- **ITT** (effect of assignment; LPM; additional check with fully interacted model, Lin 2013):

$$Y_{ij} = \alpha + \theta Z_i + \Gamma \mathbf{X}_i + \rho_j + \varepsilon_{ij}$$

- **LATE** (effect of completion; 2SLS):

$$\text{first stage:}\ \ D_{ij} = \beta + \gamma Z_i + \Gamma \mathbf{X}_i + \rho_j + \nu_{ij}$$

$$\text{second stage:}\ \ Y_{ij} = \alpha + \theta \widehat{D}_{ij} + \Gamma \mathbf{X}_i + \rho_j + \varepsilon_{ij}$$

  - $\mathbf{X}_i$: pretreatment covariates (spending, books, orders, recency, tenure, gender, age, B2B); ρ_j: product FE; SE clustered at the level of treatment assignment (user/individual; Abadie et al. 2023).
  - θ_{LATE} = causal effect of a completed interaction (vs. static banner), for compliers (**H1a**, **H1b**).

Reporting Ad Effects on Conversion

For cumulative purchase window h ($h = 0$: same-day; $h = 30$: 30-day):

- **Point estimate**: the conversion effect $\widehat{\theta}_h$, in percentage points (pp).
- **Economic magnitude**: lift and NNE re-express $\widehat{\theta}_h$ for interpretation.

$$\widehat{\text{Lift}}_h = 100 \times \frac{\widehat{\theta}_h}{\widehat{\mu}_{0h}}, \qquad \text{NNE}_h = \frac{1}{\widehat{\theta}_h}.$$

  - **Lift**: % increase over the counterfactual baseline $\widehat{\mu}_{0h} = \widehat{\Pr}(Y_h(0) = 1)$, the completers' mean minus the effect (Gordon et al. 2019, 2023, *Mktg. Sci.*).
  - **NNE**: impressions or completions per incremental purchase.

- Each horizon uses its own static-banner baseline $\overline{Y}_{0,h}$; the headline is the 30-day effect, where $\overline{Y}_{0,30} = 0.051\%$.

Empirical Strategy: Completion

- **Model completion** (treated users) on panel-level creative features:

$$\Pr(\text{Completed}_{ij} = 1) = g\left(\alpha + C_j'\phi + \text{Fit}_{ij}\,\psi + \rho_j\right)$$

  - Logistic regression; report average marginal effects (AMEs) in pp (treatment group only).

- **Creative features C_j, by stage:**
  - **Entry** (the hook) **/ middle / late** character count (text richness) and color diversity (visual complexity).
  - **Any-face** indicator (social cue).

- **User–ad genre fit (Fit_{ij}):**
  - The advertised book's genre matches the customer's past purchases.

- Maps onto CoP predictions **H2–H5**: entry value vs. late effort, social cue, relevance.

Result 1a: Conversion

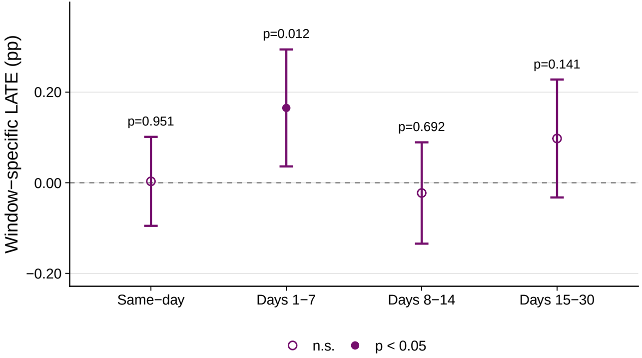
Horizon	Estimand	Coef. (prob.)	Effect (pp)	p	Static-banner baseline	Economic scale
Same-day	ITT	0.00001 (0.00002)	0.001 (0.002)	.678	0.009%	–
Same-day	LATE	0.00018 (0.00048)	0.018 (0.048)	.708	0.009%	–
30-day	ITT	0.00008 (0.00004)	0.008 (0.004)	.055	0.051%	+16%; NNE 12,500
30-day	LATE	0.00229 (0.00113)	0.229 (0.113)	.043	0.051%	+449%; NNE 437

Notes: Coef. are probability-unit estimates as in the paper; pp = $100\times$ coef. ITT/LATE compare BLI to the static-banner control, not no-ad. Lift = $\widehat{\theta}/\overline{Y}_{0,h}$; NNE = $1/\widehat{\theta}$.

- Main read: 30-day LATE = 0.229 pp over a 0.051% baseline; same-day LATE is small and n.s.
- Purchases often lag exposure (Chapelle 2014), so same-day and 30-day conversion are tracked; the next slide splits the window to show when the effect arrives and fades.

Result 1b: Timing

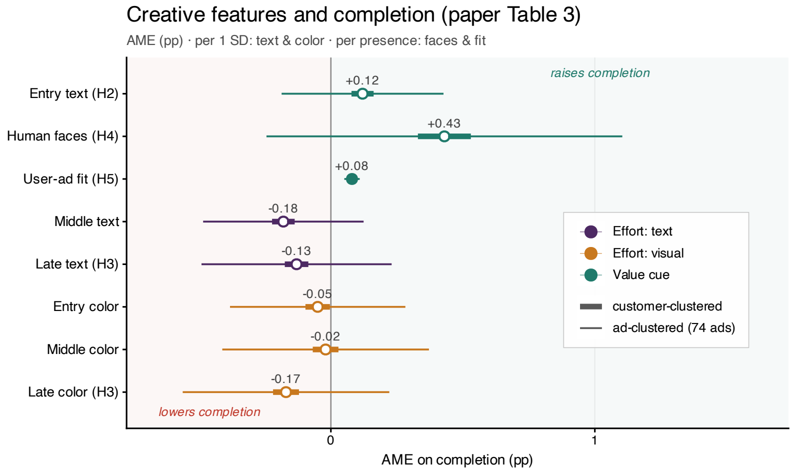
- The follow-up is split into non-overlapping post-exposure windows to see when the conversion effect appears and when it fades.



- The purchase response is not concentrated on the exposure day; the detectable window-specific effect appears in days 1–7.

Notes: Window-specific outcomes are non-overlapping and reconstructed from full follow-up purchase data.

Result 2: Completion



Entry text ↑, late complexity ↓, faces and user–ad fit ↑ completion (**H2–H5 supported**).

Supplementary Survey

- 242 participants recruited to match the RCT sample's gender and age distribution.
- Each respondent views 10 randomly selected BLI sliders.
- After each slider, four outcomes are measured on 5-point scales (1 = strongly disagree, 5 = strongly agree):
  - hedonic experience (Ruan, Hsee & Lu 2018):
    - "I am having a good time right now."
  - persuasion (attitudes and intentions; Morwitz, Steckel & Gupta 2007):
    - consideration: "Would you consider purchasing the product";
    - favorable opinion: "My opinion of this product is favorable";
    - purchase intention: "I plan to buy this brand in the future."
- Individual and product fixed effects isolate within-person, within-product variation.

Result 3: Hedonic Experience and Persuasion

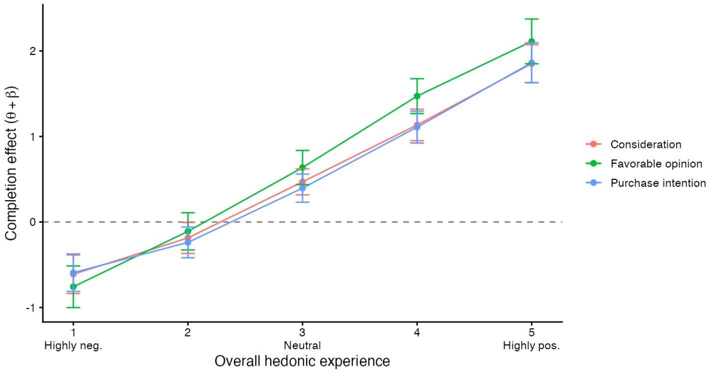
	Consideration	Favorable opinion	Purchase intention
Completion	-0.611***	-0.757***	-0.592***
	(0.113)	(0.122)	(0.110)
× Hedonic 2	0.425***	0.648***	0.354***
	(0.092)	(0.096)	(0.091)
× Hedonic 3	1.080***	1.394***	0.988***
	(0.098)	(0.102)	(0.099)
× Hedonic 4	1.746***	2.229***	1.702***
	(0.110)	(0.107)	(0.110)
× Hedonic 5	2.461***	2.869***	2.454***
	(0.128)	(0.117)	(0.133)

Notes: OLS with individual and product fixed effects. Participant-clustered s.e. in parentheses. Outcome scales 1–5. Hedonic level 1 is the omitted baseline. $^{***}p < .001$.

- The completion main effect is negative when the interaction feels bad (lowest hedonic level).
- The completion × hedonic interactions rise monotonically; completion helps once the experience is positive.

Result 3b: Hedonic Experience, Visualized

Figure 2. Completion persuades only when the experience is positive



Conclusions (one-page summary)

- **Effectiveness:** assignment to interactive BLI has little average effect, but completion raises 30-day conversion among compliers by 0.229 pp over a 0.051% static-banner baseline.
- **Timing:** the response is not an impulse purchase on exposure day; the incremental effect arrives mainly after exposure, concentrated in days 1–7.
- **Creative mechanism:** completion is governed by a stage-specific value–effort tradeoff: early information helps, late complexity hurts, and fit/social cues raise continuation.

- Managerial implication: design creative sequences for completion before evaluating persuasion at the purchase margin.

Thank you for your attention!

Contact Information:
jaewon.yoo@iss.nthu.edu.tw
https://j1yoo.github.io/

Sources I

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics* 113(2), 231–263.

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics* 138(1), 1–35.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434), 444–455.

Animesh, A., Pinsonneault, A., Yang, S.-B., & Oh, W. (2011). An odyssey into virtual worlds: exploring the impacts of technological and spatial environments on intention to purchase virtual products. *MIS Quarterly* 35(3), 789–810.

Chapelle, O. (2014). Modeling delayed feedback in display advertising. In *Proc. 20th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 1097–1105.

Consegic Business Intelligence (2024). *Interactive Advertising Market: Size, Share & Forecast to 2030* (industry report).

Goldfarb, A., & Tucker, C. (2011). Online display advertising: targeting and obtrusiveness. *Marketing Science* 30(3), 389–404.

Gordon, B. R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. (2019). A comparison of approaches to advertising measurement: evidence from big field experiments at Facebook. *Marketing Science* 38(2), 193–225.

Gordon, B. R., Moakler, R., & Zettelmeyer, F. (2023). Close enough? A large-scale exploration of non-experimental approaches to advertising measurement. *Marketing Science* 42(4), 768–793.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology* 79(5), 701–721.

Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly* 30(3), 611–642.

Guitart, I. A., Gonzalez, J., & Stremersch, S. (2018). Advertising non-premium products as if they were premium: the impact of advertising up on advertising elasticity and brand equity. *International Journal of Research in Marketing* 35(3), 471–489.

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., & Dhillon, G. (2014). A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research* 25(1), 111–136.

Jiang, Z., & Benbasat, I. (2004). Virtual product experience: effects of visual and functional control of products on perceived diagnosticity and flow in electronic shopping. *Journal of Management Information Systems* 21(3), 111–147.

Jiang, Z., & Benbasat, I. (2007). Investigating the influence of the functional mechanisms of online product presentations. *Information Systems Research* 18(4), 454–470.

Sources II

Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: reexamining Freedman's critique. *The Annals of Applied Statistics* 7(1), 295–318.

Morwitz, V. G., Steckel, J. H., & Gupta, A. (2007). When do purchase intentions predict sales? *International Journal of Forecasting* 23(3), 347–364.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: the moderating role of involvement. *Journal of Consumer Research* 10(2), 135–146.

Pieters, R., Wedel, M., & Batra, R. (2010). The stopping power of advertising: measures and effects of visual complexity. *Journal of Marketing* 74(5), 48–60.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review* 106(4), 643–675.

Romaniuk, J. (2009). The efficacy of brand-execution tactics in TV advertising, brand placements, and internet advertising. *Journal of Advertising Research* 49(2), 143–150.

Ruan, B., Hsee, C. K., & Lu, Z. Y. (2018). The teasing effect: an underappreciated benefit of creating and resolving an uncertainty. *Journal of Marketing Research* 55(4), 556–570.

Sahni, N. S., Zou, D., & Chintagunta, P. K. (2017). Do targeted discount offers serve as advertising? Evidence from 70 field experiments. *Management Science* 63(8), 2688–2705.

Yi, C., Jiang, Z., & Benbasat, I. (2017). Designing for diagnosticity and serendipity: an investigation of social product-search mechanisms. *Information Systems Research* 28(2), 413–429.

Appendix: hypothesis-test summary

Hyp.	Prediction	Status
H1	Interactive BLI increases purchase conversion	Partially supported
H2	Entry-slide informational value increases completion	Supported
H3	Late-stage textual and visual complexity lowers completion	Supported
H4	Human faces increase completion	Supported
H5	User-ad genre fit increases completion	Supported
H6	Positive hedonic experience strengthens completion's persuasive effect	Supported

- Empirical pattern supports the CoP account: creative design shapes completion.
- Completion persuades when the experience justifies the user's effort.

Appendix: survey and panel mechanism checks

Term	Est.	s.e.	p	Interpretation
Question prompt	1.58	0.43	$< .001$	curiosity / value cue
Effort index	-2.48	0.82	.004	effort cost
User-ad fit	2.75	1.44	.061	relevance
Value index	1.17	1.17	.319	composite value cue
Position × text density	1.24	0.37	.001	late-stage text cost

Notes: First four rows are survey completion regressions, reported in percentage points. Last row is the panel-hazard interaction, scaled by 100 for readability.

- The survey does not replace the field RCT; it tests whether the CoP primitives behave as predicted in a higher-completion setting.
- The mechanism evidence is strongest for question prompts, effort cost, user-ad fit, and the later-panel text penalty.

Appendix: who are the compliers? (Abadie 2003)

- We cannot identify *which* units are compliers from observed data.
  - Compliance type depends on $(D_i(0), D_i(1))$, both never observed together.
- But we can learn about compliers' **characteristics**. For any pretreatment covariate X_i:

$$\mathbb{E}[X_i \mid \text{complier}] = \frac{\mathbb{E}[X_i D_i \mid Z_i = 1] - \mathbb{E}[X_i D_i \mid Z_i = 0]}{\mathbb{E}[D_i \mid Z_i = 1] - \mathbb{E}[D_i \mid Z_i = 0]}$$

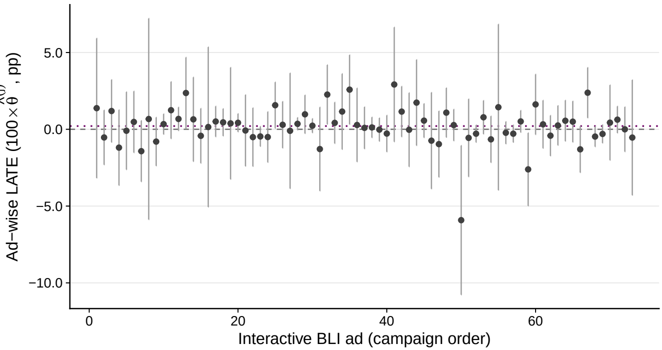
  - Same Wald-ratio structure, with $X_i D_i$ in place of Y_i.
- Why it matters:
  - **External validity**: are compliers representative of the population of interest?
  - **Managerial relevance**: if interactive BLI is deployed more broadly, will newly induced completers resemble current compliers?
  - Compare complier means to population means to assess.

Appendix: cumulative horizon estimates

Cumulative window	LATE (pp)	95% CI	*p*
Same-day	0.003	[-0.095, 0.101]	.951
7-day	0.168	[0.005, 0.332]	.044
14-day	0.146	[-0.058, 0.349]	.161
30-day	0.243	[-0.001, 0.488]	.051

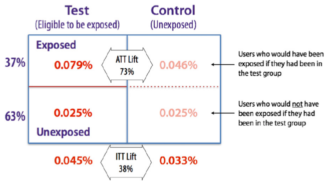
- Cumulative version of the horizon reconstruction.
- Main result remains Table 2's official 30-day LATE.
- Non-overlapping-window plot is used in the talk only to decompose timing.

Appendix: ad-wise LATE diagnostic



- Ad-wise 2SLS treats each BLI creative as a campaign-level execution, following Sahni et al. (2017, *Mgmt. Sci.*).

Appendix: rare-event ad benchmarks



- Gordon et al. (2019, *Mktg. Sci.*): Facebook RCT lift benchmark.
- Study 4 checkout: 0.033% control, 0.045% test; ATT lift 72.8%.
- Study 5 checkout: 0.009% control, 0.022% test; ATT lift 449.6%.
- Gordon et al. (2023, *Mktg. Sci.*): 663 Facebook experiments; DML/SPSM do not reliably recover RCT effects.
- Sahni et al. (2017, *Mgmt. Sci.*): 0.1% redemption; +37.2% spending.