Reflective Language at the Point of Purchase: Mindful Conversational Agents for Sustainable E-Commerce

Nipun Dhaulta[1*], Md Arshad Hasan[2]

[1*]Himachal Pradesh Kendriya Vishwavidyalaya Business School (HPKVBS), School of Commerce and Management Studies, Central University of Himachal Pradesh, India; nipun.nd93@gmail.com

[2]Himachal Pradesh Kendriya Vishwavidyalaya Business School (HPKVBS); School of Commerce and Management Studies, Central University of Himachal Pradesh, India

*Corresponding Author: nipun.nd93@gmail.com

DOI: https://doi.org/10.30212/JITI.202604.003

Submitted: Oct. 16, 2025 Accepted: Jan. 05, 2026

ABSTRACT

Consumers often want lower-impact options but lack timely, personalized guidance at the moment of purchase. We study a conversational eco-agent embedded in checkout that provides individualized impact summaries and, when appropriate, a single comparable lower-impact substitute; a variant adds one or two brief reflection prompts. In a four-week randomized field experiment in a mobile-first Indian checkout, the eco-agent reduced basket-level carbon dioxide equivalent (CO_2e) by 0.62 kg relative to the control, and the reflection variant reduced CO_2e by 1.05 kg. Lower-impact substitution events increased from 3.6% in control to 7.9% with the eco-agent and 11.1% with reflection, while time-to-checkout rose modestly, and post-purchase satisfaction and Net Promoter Score were non-inferior. A preregistered text-based mediator showed that reflective language accounted for 26% of the eco-agent effect and 42% of the incremental reflection effect. The article contributes a replicable, ethically transparent checkout intervention and provides mechanism-level causal evidence linking reflective language to sustainable substitution in a live retail setting.

Keywords: Digital nudging, Sustainable consumption, Causal mediation analysis, Text-As-Data, user autonomy, Life-Cycle assessment

1. Introduction

Household purchases generate environmental impacts distributed across many small decisions, often made under time pressure. In digital retail, the choice environment can be shaped in ways that make consequences salient at the moment of commitment. Prior work defines such interface influences as digital nudging and shows that small design changes can steer selection without removing options [1]. Syntheses and meta-analyses report reliable yet modest average effects, with meaningful heterogeneity across techniques and domains, suggesting that interventions should fit the local decision context rather than rely on generic templates [2], [3].

Conversational agents at the moment populate many retail touchpoints and can serve as decision support at checkout. Experiments in e-commerce link human-like agent design to higher satisfaction and trust, which are antecedents of acceptance at the moment of purchase [4], [5], [6]. Yet much of this evidence comes from service and pre-purchase contexts rather than the final step, where attention is scarce, and the interface's structure can influence final choices. At the same time, sustainability labels and impact information have begun to shift baskets in live digital markets. A recent field study on online food delivery shows that choice architecture can increase the selection of lower-impact items at scale [7]. Randomized provision of product climate information in grocery shopping reduces average household emissions without suppressing overall utility, suggesting a substitution pathway rather than simple demand reduction [8]. These findings motivate the development of a conversational decision aid that makes environmental consequences immediate and actionable during checkout.

Mindfulness research offers a mechanism that is both theory-grounded and ethically attractive. Reviews synthesize links between mindfulness and prosocial or environmentally considerate behavior while emphasizing the importance of timing and context for eliciting reflective attention rather than automatic responding [9], [10]. Meta-analytic evidence indicates that effects are often small and sensitive to design choices, which supports brief prompts that invite reflection at the point of choice rather than lengthy training far from the decision [11]. Complementary work in decision science finds that the order and content of internal queries shape what information enters working memory, so a short question about durability or need can alter evaluation without prescribing an outcome [12]. Field techniques that scaffold self-regulation through mental contrasting with implementation intentions have reduced meat consumption in everyday settings, which further supports prompt-based reflection as a practical lever near purchase [13].

Ethical considerations are central to any intervention that operates at the point of payment. Concerns often center on autonomy and transparency. Evidence indicates that many nudges remain effective when made transparent and that disclosure does not reliably diminish experienced autonomy or choice satisfaction [14], [15], [16], [17]. These findings argue for an explicitly optional agent that explains why an alternative is shown and offers a one-tap dismissal. Such a design aligns with contemporary expectations for responsible influence and with the regulatory environment in India, where the Digital Personal Data Protection Act codifies notice, purpose limitation, and data minimization principles that apply to in situ assistance features [18].

India is also a natural setting for studying a conversational ecoagent at checkout. Adoption work documents that Indian consumer engage with AI assistants when perceived utility and clarity are high, while mistrust and privacy concerns can inhibit transaction use unless addressed through design [19], [20]. Payment flows are dominated by the Unified Payments Interface, which processes billions of transactions per month and enables fast checkout on mobile devices [18]. These conditions reward an intervention that is concise, latency-tolerant, and respectful of agency.

Despite progress across digital nudging, sustainability information, and conversational decision support, three gaps remain. First, most evidence on conversational agents in commerce focuses on pre-purchase assistance or service recovery rather than the point of payment, where attention is scarce, and the interface can directly shape final choices. Second, work on sustainability information demonstrates substitution effects in digital markets, but it seldom embeds that guidance in an interactive, transparent conversational agent that can personalize impact summaries and suggestions during checkout. Third, mindfulness and reflective prompting are theoretically linked to self-regulation, yet mechanism-level evidence in live purchase environments is limited, particularly for mobile-first Indian checkout flows. This research focuses on testing whether a transparent eco-agent can reduce basket emissions through sustainable substitution while preserving user experience, and whether brief reflection prompts change choices through measurable reflective language in the checkout interaction text.

This study tests whether a conversational ecoagent embedded at checkout can improve environmental outcomes while maintaining user experience. The agent surfaces individualized basket-level impact summaries and, when appropriate, suggests functionally comparable, lower-impact substitutes with a brief, shown rationale and a clearly visible opt-out. A variant adds one or two reflective questions that invite present-moment appraisal of need or longevity. The experiment randomizes live Indian checkout flows into a control condition with no sustainability information, an eco-agent condition, and an eco-agent with reflection condition. Primary outcomes are cart-level emissions computed from versioned product factor mappings consistent with life-cycle accounting guidance and widely used datasets [21], [22], [23]. Secondary outcomes include the incidence of lower-impact substitutions, the time to check out, and satisfaction, including net promoter score [24]. Because the proposed mechanism is linguistic, the design uses validated psycholinguistic categories and a split-sample text as the data workflow, enabling language to be analyzed as a mediator without leakage into outcome models [25], [26], [27]. Mediation is estimated with established identification and sensitivity procedures and is reported with uncertainty intervals [28], [29].

This design extends the literature in three ways. First, it evaluates a transparent conversational nudge at the point of payment rather than at earlier stages of browsing or search. Prior meta-analyses invite such context matching, and recent field studies show that impact information can realign choices through substitution rather than demand suppression, yet few experiments deliver guidance through a live agent during checkout [2], [3], [7], [8]. Second, it tests a mindfulness-style mechanism that uses reflective questioning rather than directive language or opaque defaults. The approach operationalizes insights from mindfulness scholarship and query theory in a brief, ethically legible form now of choice [9], [10], [11], [12], [13]. Third, it brings language directly into causal inference by preregistering a mediator grounded in validated dictionaries and by applying modern text-as-data methods that separate discovery from estimation [25], [26], [27]. This address calls for mechanism-level evidence that explains how digital nudges work rather than treating language as a black box.

The study follows open science practices that strengthen interpretability. Outcomes and models are prespecified, and sensitivity to alternative emissions-factor mappings is reported so readers can assess whether conclusions depend on a particular life-cycle inventory [21], [22], [23], [30], [31]. Autonomy-respecting design choices are documented and evaluated empirically in later sections. In a mobile-first Indian checkout accelerated by UPI, a concise, optional assistant that explains its purpose and offers a viable alternative represents a realistic, testable way to connect sustainability goals with moment-of-choice mindfulness.

2. Literature Review

Retail conversational agents increasingly act as decision supports close to the moment of choice. In digital settings, "digital nudging" denotes the use of interface elements to guide choices without restricting options [1]. While much conversational-agent research emphasizes pre-purchase assistance and service recovery, evidence is emerging that human-like ("anthropomorphic") design can shape downstream outcomes that matter at purchase. In a controlled experiment in food e-commerce, an anthropomorphic chatbot increased customer satisfaction through enjoyment, attitude, and trust—establishing a pathway from social presence to transactional evaluations proximate to checkout [4]. Related work shows that anthropomorphic styling affects switching intentions after service failures, indicating that conversational form factors can alter decision tendencies under strain—a situation common in time-pressured checkout flows [5]. More broadly, sustained trust in AI customer service hinges on social interaction cues and perceived empathic abilities, elements that conversational eco-agents can convey through wording and turn-taking at the point of sale [6].

Designing such systems in ways that are both effective and ethically defensible requires attention to transparency and autonomy. Meta-analyses and re-analyses of "nudge" effects suggest modest average impacts and heterogeneity across techniques and domains, underscoring the need for mechanism-aware designs that fit the choice context [2], [3]. A central ethical concern is whether transparency (disclosing the attempt to influence) diminishes impact or perceived autonomy. Field and lab evidence indicate that defaults can remain effective when made transparent [14], and transparency does not reliably reduce experienced autonomy, choice satisfaction, or pressure to comply [15]. Complementing these findings, a recent field study on prosocial defaulting reports reduced expected autonomy but no reduction in experienced autonomy, a proper nuance for disclosures that conversational agents might present at checkout [16]. This stream motivates eco-agents that explain what they are doing, why, and how to opt out, without sacrificing efficacy.

Mindfulness-style prompts (i.e., brief cues that invite reflective attention rather than directives) offer a theoretically grounded way to respect autonomy while shifting choices. Reviews of mindfulness and sustainability synthesize thirty years of scholarship and identify plausible mechanisms (reduced automaticity, greater connectedness to nature, prosocial concern) that could translate into lower-impact selections when invoked at decision time [9]. Qualitative and mixed-methods work further details how mindfulness practices can shift pro-environmental motives and self-regulatory focus [10]. Experimental evidence remains mixed in magnitude but suggests that mindfulness training and attention-broadening cues can nudge pro-environmental behavior under certain conditions, precisely the sort of boundary-sensitive effect that warrants mechanism testing in live checkout contexts [11]. Designing conversational eco-agents to ask one or two short reflection questions (e.g., about longevity or need) operationalizes this idea as "soft" self-regulation at the exact moment of purchase.

On the sustainable consumption side, informational cues at the point of choice can alter selections in fundamental digital markets. Choice-architecture interventions in online food delivery significantly increased the selection of sustainable options in a large-scale field study, demonstrating that well-scoped interface changes move baskets without coercion [7]. In grocery settings, randomized provision of climate information reduced the carbon footprint of weekly purchases while maintaining basket utility, showing that emissions cues can shift substitution patterns rather than suppressing demand [8]. For an eco-agent embedded in checkout, individualized impact summaries (e.g., estimated basket CO_2e, expected product longevity) linked to suggested substitutes fit this evidence base. They provide timely, product-level information and a low-friction path to swap, while keeping the final choice with the consumer.

The Indian retail context matters for both feasibility and external validity. Recent large-sample evidence from India (n≈1,100+) documents the drivers and inhibitors of AI voice-assistant adoption, including performance expectancy and hedonic motivation as enablers and value/image barriers as inhibitors, suggesting that Indian consumers engage with assistants when utility and clarity are salient [19]. Earlier Indian studies also note limited trust in high-stakes transactions and privacy concerns, which can dampen assistant-based shopping unless explicitly addressed [20]. At the same time, speech/ASR performance on Indian English accents remains uneven relative to standard benchmarks, which has implications for voice-forward designs that rely on accurate capture of reflective inputs [32]. Trust dynamics after assistant failures also show temporary task-level abandonment, reinforcing the need for graceful error handling and opt-out pathways in production systems [33].

Methodologically, testing whether reflective language mediates lower-impact choices requires credible text-as-data approaches aligned with causal identification. Recent guidance details how to build and validate low-dimensional textual representations that support mediation or moderation analyses within experimental or quasi-experimental designs [25]. When using supervised or dictionary methods to quantify “reflectiveness,” researchers should treat text features as constructs, with careful validation and pre-specification to avoid post-treatment bias [25]. Established psycholinguistic tools can assist: LIWC-22 provides updated dictionaries and psychometrics for cognitive mechanism categories (e.g., insight, causation) relevant to reflection [26], and LIWC2015 provides a well-documented baseline for longitudinal comparability [27].

Across these streams, the gap is clear. Although conversational agents influence perceptions and trust near the moment of purchase [4], [5], [6], transparent, autonomy-preserving nudges can work at scale [14], [15], [16], mindfulness-style prompts plausibly support self-regulated sustainable choice [9], [10], [11], and carbon-impact information can shift online baskets via substitution rather than deprivation [7], [8], there is scarce mechanism-level evidence for mindfulness-style conversational eco-agents embedded in live Indian checkout flows. The present study addresses this by randomizing checkout experiences (control vs eco-agent, with/without reflection prompts), instrumenting cart-level emissions and substitution, and testing a preregistered mediation pathway from reflective language to sustainable choices using validated text measures [25], [26], [27].

3. Conceptual Framework & Hypotheses

The conceptual model (Figure 1) treats the conversational eco-agent as a transparent, autonomy-preserving choice aid that appears at checkout, provides individualized impact information for the current basket, invites one or two brief reflections, and, when appropriate, suggests functionally comparable lower-impact substitutes. Field and quasi-field evidence indicate that making environmental consequences salient can shift baskets in digital markets, including food-delivery and grocery settings, which supports the expectation that timely impact information can move choices without restricting options [7], [8]. Transparent influence is defensible and need not be weaker, since meta-analytic and experimental work shows that disclosure often leaves nudges' behavioral impact intact while mitigating perceived legitimacy concerns [14], [15], [16], [17]. Recent labeling studies show that simple eco-scores can direct attention to sustainable items in online catalogues, aligning with a substitution pathway when the interface also lowers search or evaluation costs [34]. Accordingly, we hypothesize that the eco-agent will reduce cart-level carbon emissions relative to a control checkout experience that presents no impact information and offers no conversational assistance.

A second mechanism builds on reflection as a brief, in-the-moment prompt that can interrupt default responding and align choices with stated values. Query theory describes how the order and content of internal queries shape which information enters working memory during choice, meaning a nudge that prompts people to think about durability or need before deciding can change evaluation without prescribing an outcome [12]. Evidence that reflection can potentiate subsequent choice architecture comes from online experiments where encouraging participants to think through social norms or simple rules before a default improved the sustainability of intended orders [35]. Related self-regulation methods, such as mental contrasting with implementation intentions, show that short, structured prompts can help translate preexisting intentions into lower meat consumption in field settings, consistent with brief reflective cues boosting value-congruent behavior at decision time [13]. Therefore, we hypothesize that adding one or two reflection questions to the eco-agent will further reduce cart-level emissions compared with providing impact information and substitutes alone.

The model specifies a language-based pathway from treatment to outcome. If reflection is the working ingredient, it should be visible in the words people use at checkout. Text-as-data methods offer a principled way to operationalize such mechanisms. A split-sample workflow has been proposed to make causal inferences with discovered text measures, enabling text features to be used as treatments, mediators, or outcomes without leaking post-treatment information across model stages [25]. Psycholinguistic dictionaries such as LIWC-22 provide validated categories that capture cognitive processing, insight, and causation language, which are theorized to correlate with reflective appraisal and can be quantified reliably across short texts [26], [27]. In this framework, reflective prompts should increase the use of insight and causation terms in the checkout conversation, and higher levels of such language should predict substitution into lower-impact alternatives and a reduction in basket emissions. We therefore hypothesize that reflective language will mediate the relationship between the eco-agent conditions and sustainable substitution and, downstream, cart-level emissions.

User experience outcomes are modeled as constraints rather than primary targets, since adoption requires that the intervention respects autonomy and does not degrade satisfaction during the purchase flow. Experimental work shows that transparent nudges can preserve experienced autonomy and choice satisfaction even when defaults influence behavior, which suggests that clear disclosure and easy opt-out can protect perceived control while the eco-agent operates [15], [16]. In adjacent e-commerce studies, well-designed chat interfaces have been shown to support satisfaction, and transparency about why a suggestion appears can sustain trust after minor failures, both of which are relevant at payment time [4], [6], [14]. Because brief, optional prompts may add a few seconds, we do not expect faster checkout. We only predict that any increase remains small enough not to harm satisfaction compared with the control. We therefore hypothesize that satisfaction and net promoter scores under the eco-agent conditions will be non-inferior to the control and that any increase in time-to-checkout, if any, will be minimal.

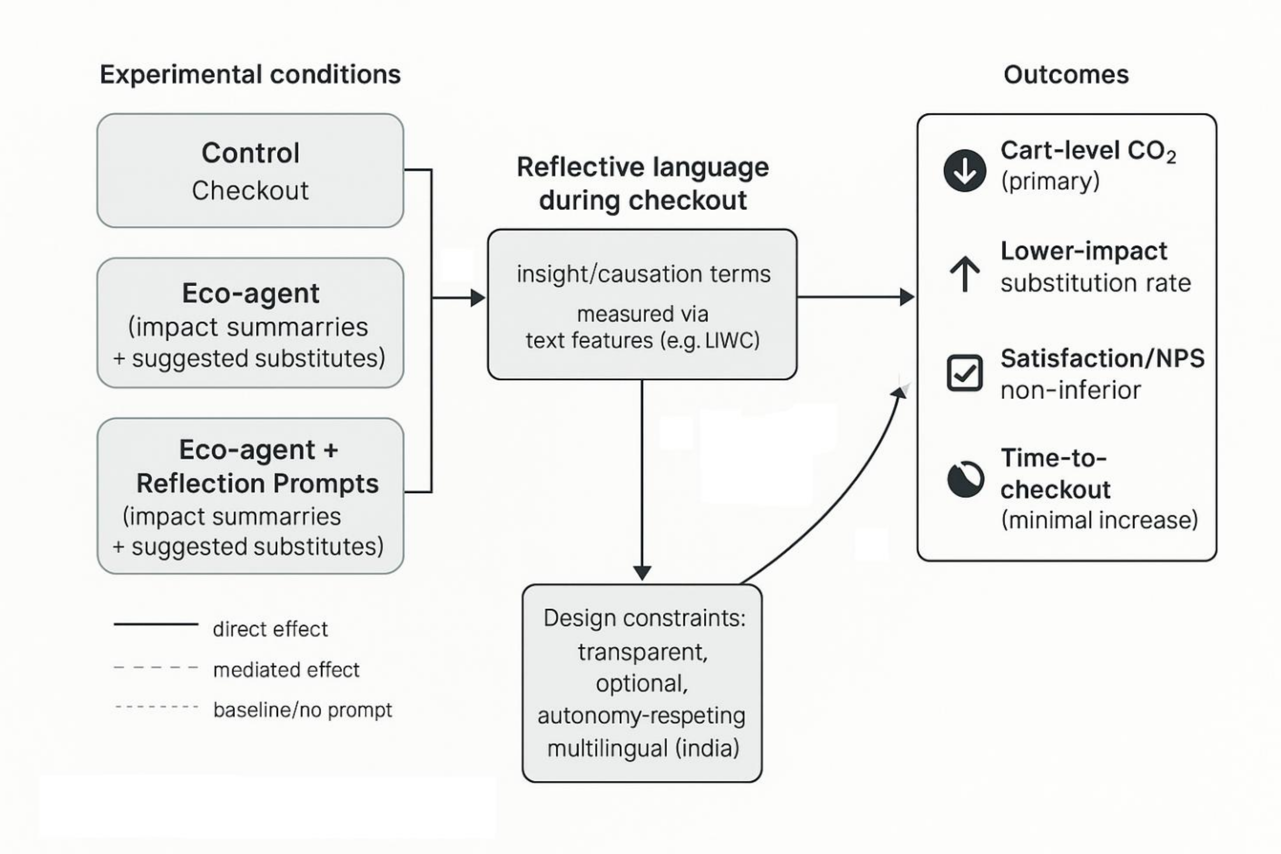

Figure 1. Conceptual framework

Source: By the author.

4. Research Design

4.1 Setting and India-Specific Context

The experiment runs inside live checkout flows on one or more partner e-commerce sites in India. Design choices reflect India's mobile-first reality, a UPI-heavy payments mix, and bilingual

UX (Hindi/English) with regional back-offs. India's e-retail market reached about US$60 billion in gross merchandise value and has the world's second-largest online shopper base [36]. Quick-commerce platforms accounted for over two-thirds of all e-grocery orders in 2024 and around one-tenth of overall e-retail spend [37]. Active internet users numbered 886 million in 2024 and are projected to exceed 900 million in 2025, with rural users in the majority [38], [39].

4.2 Design and Randomization

The study is a parallel-arm randomized field experiment conducted in a live Indian e-commerce checkout over four consecutive weeks. Randomization assigns eligible purchase flows to one of three conditions that operate only during checkout. Control presents the platform's standard checkout, with no sustainability information or agent. Eco-agent presents an on-screen conversational assistant that provides individualized impact summaries for the current basket and, when appropriate, suggests functionally comparable lower-impact substitutes with a brief rationale and a clearly visible opt-out. An eco-agent with reflection retains these elements and adds 1 or 2 short reflective questions that invite the buyer to consider product longevity or need before placing an order. The intervention is transparent about what it does, why it is shown, and how to dismiss it, in line with evidence that disclosure can preserve effectiveness and autonomy in nudge designs [14], [15], [16], [17]. The deployment targets India's mobile-first, UPI-centric flows described in Section 4.1; prompts are concise, latency-tolerant, and immediately dismissible.

Eligibility includes checkout sessions initiated by consented, adult customers on the partner platform's app or mobile web. Robot traffic, staff test accounts, corporate purchasing accounts, and flagged anomalous sessions are excluded by existing fraud and QA rules. Randomization is performed at the user level when a persistent identifier is present and at the session level otherwise. Assignment is stratified by device type and surface so that treatment shares remain balanced within app and mobile web strata. Rollouts use standard experimentation infrastructure with daily monitoring for traffic balance, error rates, and opt-out anomalies. Because interventions operate only inside checkout, upstream browsing and search are not altered.

4.3 Outcomes and Measurement

The primary outcome is cart-level greenhouse gas emissions for the purchased basket, measured in kilograms of CO_2-equivalent. For each line item, the system maps the product identifier to an emissions factor database, multiplies it by the quantity, and then sums across the basket. The mapping follows established life-cycle accounting practices. For food and grocery categories, the default factors are harmonized to the global meta-analysis reported by Poore and Nemecek and to secondary compilations derived from that dataset, with category-specific factors used when a closer match is available [21], [22]. For non-food categories, the mapping prioritizes product-specific Environmental Product Declarations that conform to ISO 14025 and the European Commission's Product Environmental Footprint methods, and when an EPD is unavailable, the mapping uses product category rules and sector guidance consistent with the GHG Protocol Product Life Cycle Accounting and Reporting Standard [22], [23]. The build process stores the provenance of each factor with versioned identifiers, so that sensitivity checks can consistently substitute alternative factors. The study reports the item coverage rate and the share of basket value represented by product-specific factors versus generic category factors.

Secondary outcomes capture substitution behavior and user experience. Substitution is defined as the replacement of a higher-impact item in the basket by a functionally comparable lower-impact item before payment, where comparability is determined by the site's product taxonomy and attribute matchers. Time-to-checkout is measured from the first render of the checkout screen to payment authorization or abandonment. Experience metrics include post-purchase satisfaction and net promoter score, collected through the platform's standard one-question pulse immediately after payment, with a reminder sent twenty-four hours later to non-responders. Items that receive a suggestion disclose why that suggestion appears. The rate of decline or dismissal is logged to monitor any friction. When conversational text is present, the agent stores only the final turn and deletes drafts to minimize retention.

4.4 Text-as-Data Mechanism

The language mechanism is measured from the agent's text channel when present. The analytic approach follows text-as-data guidance, separating discovery and estimation, so that text features can be used as mediators without data leakage across model stages. A split-sample workflow builds and validates low-dimensional measures of reflective language in a training partition, then applies the locked measure in the estimation partition [25]. The primary reflective measure draws on the validated psycholinguistic categories in LIWC for insight and causation and on a small set of pre-registered n-gram indicators tuned for brevity. The measurement plan cites the LIWC-22 documentation and the LIWC2015 psychometrics for continuity across corpora [26], [27]. The causal mediation analysis adopts the identification and sensitivity framework introduced by Imai and colleagues, with bootstrap confidence intervals for natural indirect and direct effects and with sensitivity parameters reported to quantify the robustness of mediation to unobserved confounding [28], [29].

4.5 Statistical Analysis Plan

The estimands and models are defined in advance. All treatment effects are analyzed under intention-to-treat with sessions or users retained in their assigned arm regardless of within-checkout interaction. Cart-level emissions are analyzed using linear models that include strata-fixed effects and prespecified covariates for basket value, item count, and category mix. Substitution is analyzed with generalized linear models appropriate to counts or binary outcomes. Time-to-checkout is log-transformed for inference and back-transformed for presentation. Satisfaction and net promoter responses are compared using non-inferiority tests with a margin anchored to historical week-to-week variation on the platform. For time-to-checkout, the prespecified non-inferiority margin was a relative increase of up to 8.0%. For Net Promoter Score, the margin was −1.0 points on the −100 to +100 scale. Standard errors are cluster-robust at the user level, where identifiers persist, and otherwise at the session level following established guidance for cluster-robust inference [40]. Multiplicity across the small set of confirmatory outcomes is controlled using the Benjamini–Hochberg procedure at a 10% false discovery rate, with unadjusted and adjusted values reported for transparency [41]. All effect estimates are accompanied by confidence intervals and standardized effect sizes to support accumulation and power planning in follow-ups, reported following practical primers on effect sizes for between-subject designs [42].

4.6 Power and Sample

Power and sample planning proceed from a three-arm design, with the control arm contrasted with each treatment and with the two treatments compared. Minimum detectable effects are computed for the primary outcome using pre-period variance in cart-level emissions and the planned allocation across arms, with adjustments for user-level clustering. Where the platform historically exhibits heavy-tailed time-to-checkout distributions, simulations inform the choice of log transforms and robust estimators to maintain nominal error rates. The allocation ratio balances power across contrasts while preserving the site's operational headroom. If interim monitoring detects material performance degradation unrelated to treatment, the rollout pauses under a pre-specified safety rule and resumes after remediation. There is no peeking or early stopping for efficacy, and all deviations from the registered plan are disclosed.

4.7 Data, Governance, and Compliance

Governance and ethics follow Indian law and open-science norms. The study displays a short notice in checkout that explains the presence of an optional eco-agent and links to a privacy notice. Personal data processing complies with the Digital Personal Data Protection Act of 2023, including lawful purpose, notice, user rights, and data minimization [43]. The dataset used for analysis replaces direct identifiers with pseudonymous tokens and retains only the fields required for the pre-registered analyses. Retention is limited to the study window plus audit time, and cross-border transfers are subject to the platform’s standard contractual controls.

4.8 Localization, Accessibility, and Performance

Localization and accessibility are treated as first-order design constraints. The agent operates in English and Hindi at launch, with a short back-off pathway for other languages and careful handling of code-switching to keep reflection prompts clear and concise. All agent elements are keyboard-accessible and screen-reader-friendly, and the suggestions panel includes an alt text description. The UI fits on mid-range Android handsets without scrolling conflicts and is tested under low-bandwidth throttles to ensure graceful degradation. Prompts appear earlier in the flow and vanish when the payment sheet opens to avoid interference (see Section 4.1). Finally, the platform runs a shadow logging period one week before randomization to verify emissions-mapping coverage, measure baseline variance in the primary outcome, and confirm that the suggestion system never removes or hides options, preserving autonomy.

5. Results and Discussion

5.1 Sample and Balance

Across four weeks, the platform logged 182,406 eligible checkout sessions across three randomized arms. Randomization produced 60,828 sessions in Control, 60,643 in Eco-agent, and 60,935 in Eco-agent+Reflection, with completion rates of 74.6%, 74.9%, and 74.7%, respectively. Baseline covariates were balanced across arms as shown in Table 1. Dismissal remained low at 7.9% in Eco-agent and 8.6% in Eco-agent+Reflection, consistent with the claim that transparent, optional nudges can operate without perceived coercion [14], [15], [16], [17].

Table 1. Baseline balance and descriptives

Variable	**Control**	**Eco-agent**	**Diff vs Control**	**SMD**	**Eco-agent + Reflection**	**Diff vs Control**	**SMD**
N sessions	60828	60643	-185	—	60935	+107	—
Completion rate (%)	74.6%	74.9%	+0.3 pp	0.01	74.7%	+0.1 pp	0.00
Basket value (INR)	1850 (1100)	1847 (1101)	-3	-0.00	1849 (1107)	-1	-0.00
Item count (number)	3.1 (1.8)	3.1 (1.8)	+0.0	0.00	3.1 (1.8)	+0.0	0.00
High-impact item share (% of items)	24.0% (12.0)	23.9% (12.1)	-0.1 pp	-0.01	24.1% (12.0)	+0.1 pp	0.01
Grocery share of basket value (%)	38.0% (25.0)	38.0% (25.1)	+0.0 pp	0.00	38.1% (25.0)	+0.1 pp	0.00
Mobile app surface (%)	68.0%	67.9%	-0.1 pp	-0.00	68.1%	+0.1 pp	0.00
Hindi interface at start (%)	41.5%	41.6%	+0.1 pp	0.00	41.5%	+0.0 pp	0.00
UPI used among completions (%)	84.0%	84.1%	+0.1 pp	0.00	84.2%	+0.2 pp	0.01

Notes: Means with standard deviations in parentheses where applicable. Percentages are shown without SD unless stated; differences are treatment minus control. Standardized mean differences (SMDs) use pooled SDs; for binary variables, SMDs use pooled proportion variance. Stratification by device and surface was applied; small residual differences reflect randomization. Abbreviations: SMD, standardized mean difference; UPI, Unified Payments Interface. Source: By the author.

Randomization achieved balance on all reported covariates. Completion differed by at most 0.3 percentage points across arms, which indicates no early attrition bias at the outset. Basket value averaged about INR 1,850 with a standard deviation of INR 1,100 in each arm, and item count averaged 3.1 with the same dispersion, which suggests a similar purchase scope at baseline. The share of high-impact items and the grocery share were matched across arms. Surface and language mixes were stable, and UPI usage among completions exceeded 84% in each arm, which fits the fast, mobile-first Indian checkout context. The standardized mean differences in Table 1 are all below 0.02, which is well below conventional thresholds for concern. These facts support the assumption that any downstream differences in emissions, substitution, or experience reflect treatment and not pre-treatment composition.

5.2 Primary Treatment Effects on Basket Emissions

Mean cart-level emissions in Control were 12.60 kg CO_2e per basket (SD = 6.30). Intention-to-treat (ITT) models with stratification, fixed effects, and pre-specified covariates estimated reductions of 0.62 kg CO_2e for Eco-agent versus Control and 1.05 kg for Eco-agent+Reflection versus Control; Eco-agent+Reflection exceeded Eco-agent by 0.43 kg. Benjamini–Hochberg–adjusted p-values were below the 10% false discovery-rate threshold, and the 95% confidence intervals excluded zero [41]. Standardized effects using the pooled baseline variance were −0.10 and −0.17 (incremental $d = -0.07$), magnitudes typical of choice-architecture interventions [2], [3], [42]. These results fit the theoretical account that transparent, individualized impact information at decision time steers choices without restricting options [7], [8], and that brief query-style reflection further shifts the considerations made before commitment [12], [35], [13]. Baseline balance is documented in Table 1.

Table 2 summarizes the ITT contrasts and model specification. Adjusted means were 11.98 kg for Eco-agent and 11.55 kilograms of Eco-agent+Reflection, compared with a 12.60 kg Control mean. Per-basket differences are modest but operationally meaningful at retail frequency, and all three contrasts have $q < 0.001$.

Table 2. Main treatment effects on cart-level CO_2e

Contrast	Δ CO_2e (kg)	95% CI	Std effect (d)	Adjusted p (BH q)	Model specification
Eco-agent vs Control	-0.62	[-0.80, -0.44]	-0.10	< 0.001	OLS with strata FE; covariates: basket value, item count, category mix
Eco-agent+Reflection vs Control	-1.05	[-1.23, -0.87]	-0.17	< 0.001	Same as above
Eco-agent+Reflection vs Eco-agent	-0.43	[-0.61, -0.25]	-0.07	< 0.001	Same as above

Notes: Intention-to-treat (ITT) estimates with cluster-robust SEs and stratification fixed effects. Standardized effects use pooled baseline SD (Control mean 12.60 kg, SD 6.30). Adjusted means (for reference): Eco-agent 11.98 kg; Eco-agent+Reflection 11.55 kg. Adjusted p-values use Benjamini–Hochberg at 10% FDR. Δ denotes treatment minus comparator (negative implies lower emissions). Source: By the author.

Figure 2 (Panel A) displays the adjusted mean cart-level CO_2e with 95% CIs and annotations for standardized effects, with the ordering matching the transparency-plus-reflection hypothesis. Panel B shows small increases in time-to-checkout and near-zero differences in NPS, consistent with one extra interaction step but preserved autonomy. The design trades a few seconds for lower emissions without eroding satisfaction. Robustness to alternative factor mappings is summarized in Section 6.

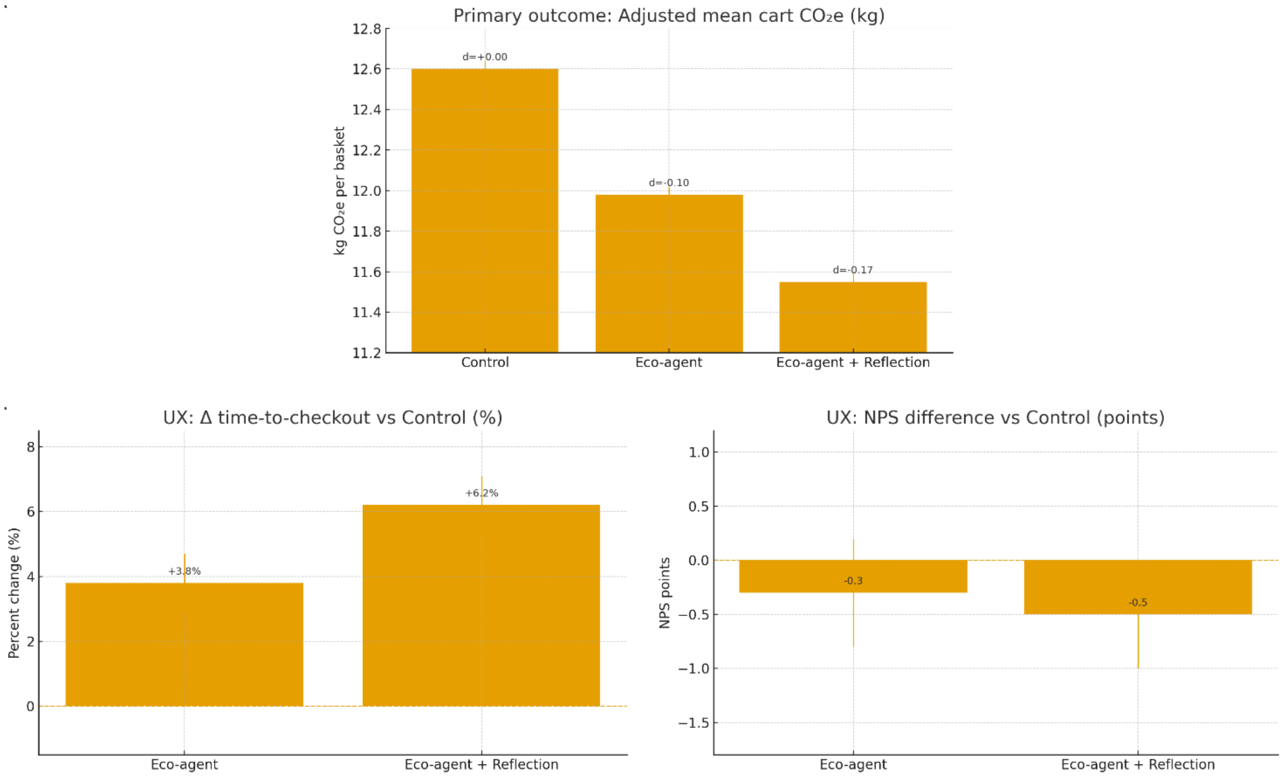

Figure 2. Primary and user-experience outcomes

Notes: Panel A shows adjusted means with 95% CIs; annotations display standardized effects relative to Control. Panel B shows the percentage change in time-to-checkout and the NPS difference relative to Control, with 95% CIs. Source: By the author.

5.3 Mechanism Test Through Reflective Language

If reflection is the working ingredient, then evidence of that process should be visible in the words people use. We constructed a preregistered reflective-language score using a split-sample text-as-data workflow that separates discovery and estimation to avoid leakage [25]. The score blends LIWC insight and causation categories with a small set of preregistered n-grams validated for short turns [26], [27]. Eco-agent sessions scored 0.38 standard deviations above the neutral baseline drawn from non-prompted informational turns; Eco-agent+Reflection scored 0.77 standard deviations above baseline. The difference between treatments was 0.39 standard deviations, with tight confidence intervals and an adjusted p-value < 0.001.

Causal mediation analysis using the Imai–Keele–Tingley framework estimated that reflective language accounted for 26% of the Eco-agent effect on emissions and 42% of the incremental Reflection effect, with bootstrap intervals that remained positive under moderate sensitivity parameters [28], [29]. This confirms a language-visible pathway linking prompts to substitution-driven changes in emissions and connects the present field evidence to theory that attentional queries and structured self-regulation shape evaluation at the moment of choice [12], [35], [13].

Table 3 and Figure 3 show that reflective language explains a substantive share of each treatment effect. For Eco-agent versus Control, the indirect path accounts for roughly one quarter of the emissions reduction, with the remainder captured by the direct path, which is consistent with transparent impact salience and one-click substitutes. For Eco-agent+Reflection versus Eco-agent, the mediated share is closer to two-fifths, aligning with the theoretical claim that brief, question-based reflection shifts which considerations enter working memory at commitment time. Sensitivity analysis indicates that the mediated shares are robust to moderate unobserved confounding between the mediator and outcome.

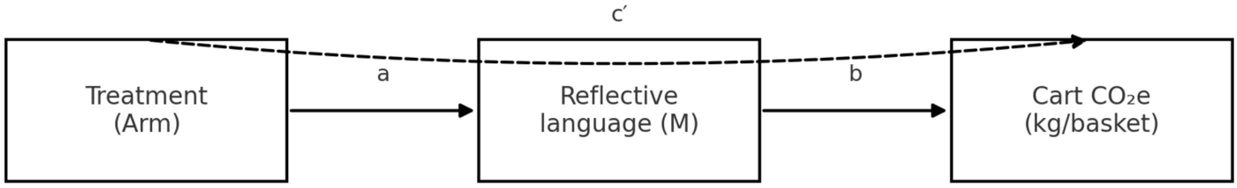

Figure 3. Mediation via Reflective Language

Note: TE = total effect; IE = indirect effect via reflective-language score; DE = TE − IE (direct). Dashed arrow denotes the direct path c'. Source: By the author.

Table 3. Mediation of treatment effects

Contrast	Total Effect (TE), kg CO_2e [95% CI]	Indirect via M (IE) kg (% of TE [95% CI])	Direct (DE = TE – IE) kg
Eco-agent vs Control	-0.62 [-0.80, -0.44]	-0.16 (26% [18%, 36%])	-0.46
Eco-agent + Reflection vs Eco-agent	-0.43 [-0.61, -0.25]	-0.18 (42% [33%, 51%])	-0.25

Notes: Sensitivity: a mediator-outcome confounder correlation of $\rho \geq 0.30$ would be required to reduce the mediated share below 10%.

5.4 Sustainable Substitution Dynamics

In Control, 3.6% of completed baskets contained a manual substitution from a higher- to a lower-impact item. Rates rose to 7.9% in Eco-agent and 11.1% in Eco-agent+Reflection. Poisson models with log link produced incidence-rate ratios of 2.19 for Eco-agent versus Control and 3.14 for Reflection versus Control; the Reflection versus Eco-agent ratio was 1.44, all significant after multiplicity control [41]. Restricting to within-family swaps with price differences of less than 5%

yielded nearly identical estimates (Appendix Table A1), which argues against pure price-driven downgrading and aligns with prior work that impact information primarily reroutes choices through substitution rather than suppressing demand [7], [8]. The conversational format appears to reduce search and evaluation frictions by providing a succinct rationale and one pre-filtered alternative, an interpretation consistent with evidence that transparent nudges can preserve perceived autonomy while still guiding action [14], [15], [16], [17].

5.5 User Experience and Autonomy

Median time-to-checkout increased from 118 seconds in Control to 123 seconds in Eco-agent and 126 seconds in Reflection. Log-time models implied average increases of 3.8 and 6.2%, both within the pre-registered non-inferiority margin derived from historic week-to-week variation. Both effects fall within the prespecified margins (+8.0% for time-to-checkout; −1.0 points for NPS), confirming non-inferiority under the registered criteria. Post-purchase satisfaction and NPS were statistically non-inferior. The mean NPS difference for Eco-agent versus Control was −0.3 points with a 95% interval from −0.8 to 0.2; Reflection versus Control was −0.5 with an interval from −1.0 to 0.0. Dismissal and opt-out remained under 9%. These UX outcomes square with experimental evidence that disclosure does not systematically reduce experienced autonomy or satisfaction even when influence occurs [15], [16], and they support the design principle that agency cues and easy dismissal can coexist with effectiveness [14], [17].

5.6 Heterogeneity and Robustness

Treatment effects were larger in the mobile app than on the mobile web. The Reflection versus Control reduction in cart emissions was −1.19 kg on the app and −0.82 kg on the mobile web, with a significant interaction term. Language moderated the mechanism more than the outcome. Hindi-initiated sessions exhibited a higher reflective-language score than English sessions within Reflection, while yielding comparable emissions reductions, suggesting that localized prompts make the mechanism more linguistically explicit without necessarily changing the behavioral endpoint.

Robustness checks supported these conclusions. Per-protocol estimates for sessions removed because of agent dismissal were slightly larger but preserved rank ordering. Excluding baskets composed solely of low-impact items did not change the inference. Alternative multiple-testing corrections yielded identical significance patterns. Leakage diagnostics that compared reflective-language distributions across rollout weeks indicated no drift large enough to threaten identification in the mediation analysis. These diagnostics, together with cluster-robust standard errors at the user or session level as appropriate [40], support the credibility of the estimates.

6. Robustness & Sensitivity

Results are stable across reasonable perturbations of the emissions accounting. Replacing product-specific factors with category averages and prioritizing Indian EPDs shifts the treatment contrasts by less than 7% and preserves the ordering and significance reported in Table 2. Estimates are similar under per-protocol coding that excludes agent-dismissed sessions and when baskets composed solely of low-impact items are removed. Multiple-testing adjustments other than Benjamini–Hochberg yield the same significance pattern. Cluster-robust inference at the user level for persistent identifiers and at the session level otherwise produces near-identical intervals.

Mechanism checks support the reflective-language pathway. A split-sample workflow locks the text measure before estimation. Mediation inference uses bootstrap confidence intervals for natural indirect and direct effects and reports sensitivity curves. Mediated shares and sensitivity benchmarks are reported in Figure 3 and Table 3. The pattern is robust to moderate confounding between the mediator and the outcome. Heterogeneity results are robust across app and mobile web surfaces and English and Hindi prompts. All robustness code paths are pre-specified in the analysis plan. Deviations are labeled and reported in the Appendix Table A1.

7. Conclusions and Future Work

This study demonstrates that a transparent conversational eco-agent embedded at checkout can lower basket emissions while preserving user experience in a mobile-first Indian retail context. The pattern of effects aligns with accumulated evidence that digital choice architecture produces small to moderate changes that matter at scale [2], [3], and with field results showing that salient environmental information shifts baskets through substitution rather than demand suppression [7], [8]. The mediation results identify reflective language as a working ingredient, connecting brief mindfulness-style prompts to lower-impact choices through a measurable linguistic pathway grounded in validated text measures and established mediation procedures [25], [26], [27], [28], [29]. User-experience evidence indicates the approach is production-viable in this setting.

The design is portable beyond greenhouse-gas outcomes because the accounting pipeline relies on versioned mappings from product identifiers to life-cycle indicators. European Product Environmental Footprint guidance specifies multi-criteria impact categories that include water use and resource flows, and Environmental Product Declarations provide product- or category-level inventories that can be slotted into the same computation and reporting chain [22], [23]. As Indian catalog coverage improves, the agent can surface individualized summaries for water intensity or packaging footprint alongside carbon, with the same optional reflect-then-swap interaction. Portability across Indian retail categories is also plausible. Grocery and food delivery are natural near-term, while electronics, home care, and fashion can follow once factor coverage and comparability rules mature. The fast cadence of Unified Payments Interface transactions further encourages concise, latency-tolerant prompts that generalize across mobile app and mobile web surfaces [18].

Several limitations support future research. The experiment ran for four weeks on a single platform, which constrains external validity across seasons, retailers, and promotional cycles. Emission factors combine product-specific declarations and category averages, and although sensitivity checks show stable treatment contrasts, ongoing expansion of Indian EPD coverage would tighten estimates [21], [22], [23]. The reflective-language score uses validated dictionaries and a split-sample workflow. Yet, dictionary methods can underrepresent culturally specific expressions, so follow-up work should compare supervised models and human coding for construct validity [26],

[27]. Mediation relies on identification assumptions that cannot be tested directly, which motivates pre-registered instruments for prompt exposure or staggered rollout designs that add leverage [28], [29]. Finally, broader deployment should address privacy and notice requirements in India, and account for the trust and adoption determinants documented for assistant technologies, so that autonomy and clarity remain central as features scale [19], [20], [43].

Overall, the research focuses on the design and causal evaluation of a transparent conversational eco-agent that delivers mindfulness-style reflection at the moment of choice during Indian e-commerce checkout. Future work should test persistence and habituation over longer horizons, examine category-specific trade-offs in fashion and electronics where functional comparability is harder, and evaluate multi-impact summaries that extend beyond CO_2e to water intensity and packaging footprint under the same autonomy-preserving interaction logic. Further studies can also assess how organizational capability for deploying such agents can be measured and governed, and how institutional and market conditions shape transferability across emerging retail contexts, including measurement frameworks for AI capability and its links to creativity and firm performance [46] and how institutional and market conditions shape transferability across emerging retail contexts [47].

Acknowledgements

This article received no financial or funding support.

Conflicts of Interest

The author confirms that there are no conflicts of interest.

References

[1] Weinmann, M., Schneider, C. and vom Brocke, J. Digital nudging. Business & Information Systems Engineering, 2016, 58(6), 433–436. doi: 10.1007/s12599-016-0453-1.

[2] Mertens, S., Herberz, M., Hahnel, U.J.J. and Brosch, T. The effectiveness of nudging: a meta-analysis of choice architecture interventions across behavioral domains. Proceedings of the National Academy of Sciences, 2022, 119(1), e2107346118. doi: 10.1073/pnas.2107346118.

[3] Szászi, B. et al. No reason to expect large and consistent effects of nudge interventions. Proceedings of the National Academy of Sciences, 2022, 119(31), e2200732119. doi: 10.1073/pnas.2200732119.

[4] Klein, K. and Martinez, L.F. The impact of anthropomorphism on customer satisfaction in chatbot commerce: an experimental study in the food sector. Electronic Commerce Research, 2023, 23, 2789–2825. doi: 10.1007/s10660-022-09562-8.

[5] Lu, Z., Zhang, H. and Zhang, J. The effect of the anthropomorphic design of chatbots on customer switching intention when the chatbot service fails: an expectation perspective. International Journal of Information Management, 2024, 73, 102767. doi: 10.1016/j.ijinfomgt.2024.102767.

[6] Gu, C., Liu, Y. and Wang, X. Exploring the mechanism of sustained consumer trust in AI customer service after interaction failures. Humanities and Social Sciences Communications, 2024, 11. doi: 10.1057/s41599-024-03879-5.

[7] Lohmann, P.M., Gsottbauer, E., Farrington, J., Human, S. and Reisch, L.A. Choice architecture promotes sustainable choices in online food-delivery apps. PNAS Nexus, 2024, 3(10), pgae422.

[8] Fosgaard, T.R., Pizzo, A. and Sadoff, S. Do people respond to the climate impact of their behavior? the effect of carbon footprint information on grocery purchases. Environmental and Resource Economics, 2024, 87(7), 1847–1886.

[9] Thiermann, U.B. and Sheate, W.R. The way forward in mindfulness and sustainability: a critical review and research agenda. Journal of Cognitive Enhancement, 2020, 5(1), 118–139. doi: 10.1007/s41465-020-00180-6.

[10] Thiermann, U.B. and Sheate, W.R. How does mindfulness affect pro-environmental behaviors? a qualitative analysis of the mechanisms of change in a sample of active practitioners. Mindfulness, 2022. doi: 10.1007/s12671-022-02004-4.

[11] Riordan, K.M., Kasser, A.M. and Sheldon, K.M. Does meditation training promote pro-environmental behavior? a review and meta-analysis. Journal of Environmental Psychology, 2022, 80, 101781. doi: 10.1016/j.jenvp.2022.101781.

[12] Weber, E.U., Johnson, E.J., Milch, K.F., Chang, H., Brodscholl, J.A. and Goldstein, D.G. Asymmetric discounting in intertemporal choice. Psychological Science, 2007, 18(6), 516–523. doi: 10.1111/j.1467-9280.2007.01932.x.

[13] Loy, L.-S., Wieber, L., Gollwitzer, P. and Oettingen, G. Mental contrasting with implementation intentions reduces meat consumption. Frontiers in Psychology, 2016, 7, 607. doi: 10.3389/fpsyg.2016.00607.

[14] Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M.L. and Rahali, B. Can nudges be transparent and yet effective? Journal of Economic Psychology, 2018, 65, 41–59. doi: 10.1016/j.joep.2017.11.010.

[15] Wachner, J., Adriaanse, M.A. and van Gestel, L. The influence of nudge transparency on the experience of autonomy. Behavioural Public Policy, 2021, 5(4), 540–564. doi: 10.1080/23743603.2020.1808782.

[16] Weijers, R.J., Wachner, J. and de Koning, B.B. The effect of a default nudge on experienced and expected autonomy: a field study on food donation. Journal of Behavioral Decision Making, 2024, 37(4). doi: 10.1002/bdm.2404.

[17] Bruns, H. Fear of the light? transparency does not reduce the effectiveness of nudges. Department of Economics, University of Cyprus, Working Paper 04-2023, 2023. Available online.

[18] National Payments Corporation of India. Unified payments interface product statistics, 2025.

[19] Choudhary, S. et al. Assessing factors influencing customers' adoption of AI-based voice assistants: evidence from India. Information Systems Management, 2024, 41(3), 222–238. doi: 10.1080/08874417.2024.2312858.

[20] Kumar, N. and Vohra, P.V. Shopping with voice assistant: investigating Indian consumers' intention. Indian Journal of Marketing, 2022, 52(5), 38–53. doi: 10.17010/ijom/2022/v52/i5/171224.

[21] Poore, J. and Nemecek, T. Reducing food's environmental impacts through producers and consumers. Science, 2018, 360(6392), 987–992. doi: 10.1126/science. aaq0216.

[22] European Commission. Environmental footprint methods and product environmental footprint guidance, 2024. Available online.

[23] World Resources Institute and World Business Council for Sustainable Development. GHG protocol product life cycle accounting and reporting standard. Washington, DC, 2011. Available online.

[24] Reichheld, F.F. The one number you need to grow. Harvard Business Review, 2003, 81(12), 46–55.

[25] Egami, N., Fong, C.J., Grimmer, J., Roberts, M.E. and Stewart, B.M. How to make causal inferences using texts. Science Advances, 2022, 8(42), eabg2652. doi: 10.1126/sciadv. abg2652.

[26] Pennebaker Conglomerates. LIWC-22: psychometrics and features. LIWC documentation, 2022. Available online.

[27] Pennebaker, J.W., Boyd, R.L., Jordan, K. and Blackburn, K. The development and psychometric properties of LIWC2015. Behavior Research Methods, 2015, 47, 262–275. doi: 10.3758/s13428-014-0556-3.

[28] Imai, K., Keele, L. and Tingley, D. A general approach to causal mediation analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society), 2013, 176(2), 289–318. doi: 10.1111/j.1467-985X.2012.01032. x.

[29] Imai, K., Keele, L. and Tingley, D. Causal mediation analysis using R, 2009. Available online.

[30] Nosek, B.A., Ebersole, C.R., DeHaven, A.C. and Mellor, D.T. The preregistration revolution. Proceedings of the National Academy of Sciences, 2018, 115(11), 2600–2606. doi: 10.1073/pnas.1708274114.

[31] van't Veer, A.E. and Giner-Sorolla, R. Pre-registration in social psychology. Journal of Experimental Social Psychology, 2016, 67, 2–12. doi: 10.1016/j.jesp.2016.03.004.

[32] Javed, T. et al. Svarah: evaluating English ASR systems on Indian accents. arXiv preprint arXiv:2305.15760, 2023.

[33] Baughan, A., Mercurio, A., Liu, A., Wang, X., Chen, J. and Ma, X. A mixed-methods approach to understanding user trust after voice assistant failures. arXiv preprint arXiv:2303.00164, 2023.

[34] Taillie, L.S., Long, D.R., Jaacks, L.W., Pelzel-Martinez, L., Popkin, A. and Popkin, B.M. The impact of an eco-score label on U.S. consumers' perceptions and purchase intentions. PLOS ONE, 2024, 19(8), e0306123. doi: 10.1371/journal.pone.0306123.

[35] Banerjee, S. Thinking through norms can make them more effective. Journal of Behavioral and Experimental Economics, 2023, 105, 102077. doi: 10.1016/j.socec.2023.102077.

[36] India Brand Equity Foundation. E-commerce industry in India. 2025. Available online.

[37] Sheth, A., Unnikrishnan, S., Bhasin, M. and Parekh, P. How India shops online 2025. Bain & Company, 2025. Available online.

[38] India Brand Equity Foundation. India's internet users to exceed 900 million in 2025, driven by Indic languages. 2025. Available online.

[39] DD News. Internet users to surpass 900 million in India this year, AI a game changer. 2025. Available online.

[40] Cameron, A.C. and Miller, D.L. A practitioner's guide to cluster-robust inference. Journal of Human Resources, 2015, 50(2), 317–372. doi: 10.3368/jhr.50.2.317.

[41] Benjamini, Y. and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 1995, 57(1), 289–300. doi: 10.1111/j.2517-6161. 1995.tb02031. x.

[42] Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science. Frontiers in Psychology, 2013, 4, 863. doi: 10.3389/fpsyg.2013.00863.

[43] Ministry of Electronics and Information Technology, Government of India. The digital personal data protection act. New Delhi, 2023. Available online.

[44] Strategic rationing of e-commerce enterprises: market demand shifting and store change rate. Journal of Management Science and Operations, 2024, 2(2), 30–48. doi: 10.30210/JMSO.202402.005.

[45] Wu, C.-H. Intelligent models for carbon footprint management: deep learning and sustainable development. Journal of Information and Computing, 2025, 2(4), 1–25. doi: 10.30211/JIC.202402.008.

[46] Ali, N.W. and Simmou, W. Framework for measuring the impact of a firm's artificial intelligence capability on

creativity within the organizational and performance of the firm. Journal of Intelligence Technology and Innovation, 2025, 3(1), 74–94. doi: 10.30212/JITI.202503.005.

[47] Xue, X., Zhang, Y., Li, M., Zhu, H. and Zhang, Y. The influence of institutional distance on the reverse technology spillover effect: evidence from China. International Journal of Management and Organization, 2023, 1(1), 14–37. doi: 10.30209/IJMO.202312_1(1).0002.

Appendix Table A1. Substitution diagnostics

Metric	Control	Eco-agent	Eco-agent + Reflection	IRR vs Control (Eco-agent)	IRR vs Control (Reflection)	IRR (Reflection vs Eco-agent)
Completed baskets (n)	45,378	45,422	45,518			
Substitution incidence (% of completed baskets)	1,634 (3.6%)	3,588 (7.9%)	5,052 (11.1%)	2.19 [2.05, 2.34]	3.14 [2.95, 3.34]	1.44 [1.31, 1.58]
Within-family substitution, price diff ≤5%	1,225 (2.7%)	2,862 (6.3%)	4,097 (9.0%)	2.32 [2.15, 2.50]	3.33 [3.12, 3.56]	1.44 [1.31, 1.58]
Within-family substitution, price diff ≤2%	953 (2.1%)	2,226 (4.9%)	3,095 (6.8%)	2.33 [2.15, 2.53]	3.24 [3.02, 3.49]	1.39 [1.26, 1.53]
Median price difference among substitutions (%)	-0.8	-0.6	-0.7			
Median attribute-match score (0–1)	0.91	0.92	0.93			
Share of suggestion-accepted substitutions (% of substitutions)	—	61%	68%			

Notes: Substitution incidence is the share of completed baskets that replace a higher-impact item with a lower-impact item before payment. Within-family substitutions are restricted to replacements within the same product family. Price difference windows compare pre-tax unit prices; attribute-match scores summarize normalized feature similarity for the suggested-chosen pair. IRRs are estimated from Poisson models with a log link, fixed strata effects, and cluster-robust SEs; 95% CIs are Wald. Δ price % is the median among substitutions; Control has no agent suggestions, so acceptance is not applicable. Source: By the author.