Proceedings of the ASME 2026
International Design Engineering Technical Conferences
IDETC2026
August 23–26, 2026, Houston, TX

IDETC2026-193320

REFLECTIVE RESISTANCE: PROBING BEHAVIORAL RESPONSES TO THE ENVIRONMENTAL IMPACTS OF GENERATIVE AI THROUGH PHYSICAL AND DIGITAL INTERVENTIONS

Nicole Goridkov*, Allison Yuh, Diana Bolanos, Kosa Goucher-Lambert

Department of Mechanical Engineering, University of California, Berkeley, Berkeley, CA

ABSTRACT

Design for sustainable behavior (DfSB) encompasses a range of intervention strategies aimed at promoting pro-environmental behaviors by making otherwise invisible environmental costs salient and actionable. The widespread integration of generative AI (GenAI) into daily workflows creates a relevant application to study how intervention modality influences behavior and attitude change, especially as AI use increasingly contributes to carbon emissions and resource consumption. This paper explores how two DfSB strategies (eco-feedback and eco-steering) can be embedded into GenAI interactions to shape user awareness of AI use and invoke reflective behavior, specifically around environmental impact. To this end, we adopt a research through design approach, treating the development of novel interactive prototypes as a form of inquiry into how augmenting sustainability information shapes user behavior. We first design a browser-based eco-feedback plugin that visualizes real-time energy consumption associated with GenAI usage, establishing information as the foundation of environmental awareness. We then introduce a dynamically adaptive keyboard that modulates typing resistance in real time based on token usage during a research task – augmenting that informational baseline with a physical layer of engagement. Through a user study (n=30), we examine how these interventions influence patterns of AI use, perceived value of AI assistance, moments of reflection, and shifts in attitudes toward the sustainability implications of GenAI. Our findings demonstrate that embedding DfSB strategies into GenAI workflows can meaningfully shift user awareness, with participants weighing environmental cost against output quality and engaging more deliberately with AI, a process shaped by slow technology principles. The varied ways participants responded to physically-augmented versus purely digital feedback highlights the importance of modality diversity in reaching a wide range of users, while a consistent gap between awareness and available tools points to open questions for future sustainable AI interface design.

Keywords: design for sustainable behavior, AI usage, research through design, design intervention

1. INTRODUCTION

User behavior can drive a significant portion of a product's environmental impact, and intentional design choices made during early-stage design have the potential to support more pro-environmental behaviors. This stresses the importance of effectively designing products and interventions to drive purposeful behavior change toward more sustainable habits. The field of design for sustainable behavior (DfSB) tackles this challenge by studying and proposing design strategies for effectively driving sustainable behaviors from the early stages of the design process [1, 2]. DfSB encompasses a range of intervention strategies including eco-information, eco-feedback, and eco-steering, that seek to influence users by introducing information, feedback, choices, or prompts at critical moments in a product or service's life cycle [1, 3]. These intervention methods aim to nudge, support, or provoke reflection, allowing users to make more sustainable choices within their everyday lives.

As digital products and services have become increasingly relevant in our daily lives and workflows, they introduce new affordances and opportunities for applications of DfSB interventions in both physical and digital forms [4]. A recent systematic review of DfSB interventions across physical and digital product-service systems found that digital and physical modalities exhibit meaningfully distinct strategy profiles: digital interventions more frequently employ eco-information and eco-feedback, while physical interventions tend to cluster around functional constraints, changes in operational complexity, and influences on user emotion [5]. This work also found that emerging technologies are enabling broader and more diverse applications of DfSB than previously possible. Still, the field lacks an empirical understanding of how varying modalities of an intervention shape user perception, reflection, and long-term behavior change in practice.

The rapid adoption of generative artificial intelligence (GenAI) systems has introduced new forms of everyday inter-

*Corresponding author: goridkov@berkeley.edu
Documentation for `asmeconf.cls`: Version 1.34, August 6, 2026.

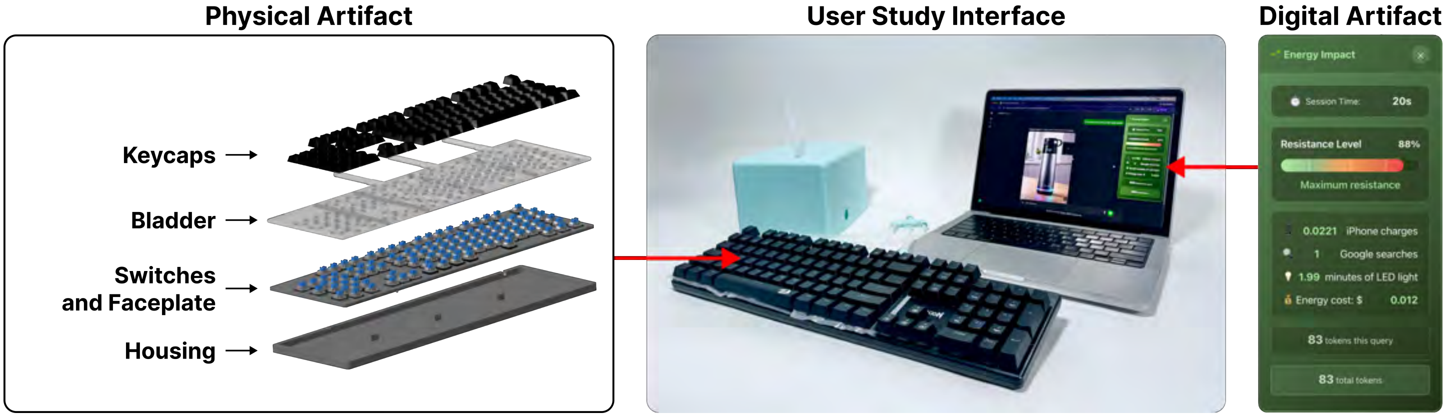

FIGURE 1: OVERVIEW OF THE ARTIFACTS PRESENTED IN THIS PAPER. (LEFT) EXPLODED VIEW OF THE ECO-STEERING KEYBOARD ARTIFACT, SHOWING THE FABRICATED BLADDER POSITIONED BETWEEN THE KEYCAPS AND THE SWITCH/FACEPLATE LAYERS. (CENTER) COMPLETE SYSTEM SETUP: THE PHYSICAL KEYBOARD PROTOTYPE SERVES AS INPUT TO A GENAI INTERFACE HOUSING THE DIGITAL ECO-FEEDBACK ARTIFACT. (RIGHT) CHROME EXTENSION OVERLAY DISPLAYING REAL-TIME ENERGY METRICS DURING CHATGPT USE, INCLUDING SESSION TIME, A RESISTANCE-LEVEL STATUS BAR, ENERGY USE EQUIVALENCIES, AND TOKEN COUNTS.

action that are fast, frictionless, and optimized, making GenAI a relevant application area to apply these design strategies. Many large language models (LLMs), including ChatGPT, Gemini, and Claude, offer free, on-demand access, making them highly accessible to the general public. Designers, students, and professionals alike rely on these tools, leading to widespread and sustained demand for computational resources. Yet, the environmental impacts are largely invisible to users. While training large AI models is energy-intensive, recent studies suggest that a significant portion of a model's total energy consumption occurs during inference (the running of AI models in real-time via queries), as many LLMs serve trillions of predictions to users every day [6, 7]. As generative AI becomes integrated into everyday workflows, the cumulative energy and resource costs of frequent use can quickly outweigh the emissions from training alone.

Thus, GenAI is an ideal application context for studying how different DfSB modalities can shape sustainable behavior and raises a central motivating question for this work: ***In what ways does augmenting digital eco-feedback with physical, embodied friction shape user awareness of and reflection on the environmental impacts of GenAI?*** Our project investigates this in a research through design [8, 9] approach, framing the design, prototyping, and deployment of two original DfSB interventions as a design exploration to explore how physically-augmenting eco-feedback can encourage reflection and support sustainable behaviors when engaging with GenAI tools. We contrast the rapid, frictionless nature of generative AI with two artifacts grounded in DfSB principles (Figure 1).

This work draws on counter-functional design as a framework for critically examining how designers and users interact with AI as an emerging technology. A counter-functional device is one that "figuratively counters some of its own essential functionality" [10]. Prior literature shows that intentionally limiting or opposing functionality can make users more aware of their actions and encourage reflection [10], and that increasing physical effort in digital interactions can disrupt habitual use and shift how users perceive value [11]. Research also suggests that many users are willing to accept slower systems in return for reducing environmental impact [12].

This paper presents two DfSB artifacts. The first, a browser-based eco-feedback plugin, visualizes real-time energy consumption associated with GenAI usage, establishing information as the foundation of environmental awareness. The second augments that informational baseline with a physical, embodied layer: a dynamically adaptive keyboard that modulates typing resistance in real time based on token usage. In support of the artifacts, we contribute the following:

- The design and implementation of a computing input device capable of dynamic, variable-resistance haptic feedback
- An open-source method for using 3D printers as precision heat sealers

We deploy these artifacts in a user study (n=30) that is generative rather than evaluative in intent, aimed at surfacing how these interventions may influence GenAI usage and invoke reflection among users. In support of studying DfSB interventions in practice, we contribute the following:

- An empirical user study comparing user behavior across physically-augmented and purely digital modalities.
- Insights into how augmenting GenAI interactions with embodied, physical feedback influences user reflection and AI perception.

2. RELATED WORK

In this section, we present related work on design for sustainable behavior (DfSB) research, counterfunctionality and critical making, and the environmental impact of generative AI systems. This work builds upon this existing array of research by using DfSB strategies to design two artifacts aimed to provoke reflection among GenAI users.

2.1 Design for Sustainable Behaviors

User behavior during the use phase of a product can have considerable environmental consequences, which has led to the development of various frameworks aimed at schematizing design strategies that can guide intentional design [2, 3, 13]. While these frameworks have collectively advanced our understanding of how design can promote sustainable behavior, the role of interaction modality in shaping the effectiveness of these strategies remains largely unexplored. This paper uses the DfSB strategies presented by [1] to explore this modality difference, focusing specifically on *eco-feedback* and *eco-steering*. Eco-feedback is the practice of showing users data about their resource consumption with the goal of promoting more sustainable behavior, and is one of the most common DfSB strategies found in the literature [5]. It has been applied across a wide range of domains [14–16], yet how its presentation ultimately affects decision-making is not yet well understood. Research has shown that surfacing environmental impact information at the point of choice can shift consumer priorities even when users lack full comprehension of the metric [17], and that the emotional valence of feedback matters: negative emotions tend to drive immediate behavior change, while positive emotions support longer-term engagement [18, 19].

Where eco-feedback informs, eco-steering goes a step further by structuring the environment so that unsustainable behavior becomes harder to perform, while still preserving user agency. It is defined as "to facilitate users to adopt more environmentally or socially desirable use habits through the prescriptions and/or constraints of use embedded in the product design" [1]. Empirical work on eco-steering has demonstrated that embedding physical constraints into product interactions can produce meaningful reductions in resource consumption, independent of users' prior environmental attitudes [20], and that active, system-controlled interventions can drive behavioral change that persists even after the intervention is removed [21]. Together, eco-feedback and eco-steering represent a spectrum of how design can shape sustainable behavior, and understanding how their effectiveness varies across physical and digital modalities remains an open question, which we seek to explore here.

2.2 Slow Technology Towards Reflection

The concepts of slow technology and counterfunctional design have been explored toward encouraging reflection and mental rest as a technological design philosophy [10, 22]. Where conventional interaction design aims to make technology fast, invisible, and efficient, slow technology argues for an alternative agenda: technology that claims time rather than eliminates it, inviting users into a slower, more conscious relationship with the tools they use. Counterfunctional design extends this idea, referring to objects that figuratively counter some of their own essential functionality, introducing friction or effort into otherwise effortless interactions so that users are forced to engage more intentionally. For example, [11] designed a hand-crank-powered Twitter interface to provoke reflection in an otherwise mindless and habitual behavior. From a cognitive standpoint, this friction may serve to shift users from fast, instinctive System 1 thinking toward more deliberate System 2 engagement [23]. In the context of environmental sustainability specifically, [12] demonstrates that users are often willing to trade system responsiveness for reduced carbon emissions, suggesting that deliberately slowing down an interaction can itself become a meaningful sustainability signal. We explore this concept within the context of GenAI, which has gained popularity in large part due to its low friction, high output, and highly iterative nature. As such, we employ the design philosophies seen in slow technology and counterfunctionality as a critical lens on this interaction, asking whether intentionally adding limitations can make AI use less automatic and more reflective, shifting what Hallnäs and Redström describe as time *disappearing* into time *appearing* [22].

2.3 Environmental Impact of GenAI

Researchers have been interested in quantifying the environmental impact of AI for many years [24, 25], and as GenAI continues to increase in popularity and use, it is more critical than ever to understand the environmental implications of the growing technology. In 2024, the International Energy Agency and US Department of Energy estimated that energy consumption of data centers could double by 2026-2028, indicating AI and data centers as a partial driver of growth [26, 27]. Many studies have sought to quantify the effects of AI development and training, largely attempting to measure the energy and water consumption gone into training the largest models [28, 29]. Recent life cycle assessments (LCAs) have sought to broaden the lens of analysis toward AI's environmental impact, and a recent paper suggests that inference, or the running of AI models in real-time via queries, dominates the energy impacts of GenAI services [30]. Tasks involving image generation, image processing, or heavy computation are especially energy-intensive, with some generation tasks consuming energy equivalent to multiple full smartphone charges [6]. At scale, the cumulative energy costs of deployment that are driven by millions of users generating billions of interactions annually can surpass the emissions generated during training within a relatively short time frame [6]. Despite the significance of these growing energy costs, the environmental footprint of everyday GenAI use remains largely invisible to users, motivating design interventions that introduce friction into these interactions to surface that impact and encourage more mindful habits of use.

3. DESIGN ARTIFACTS

This section discusses the development of two DfSB-embedded artifacts: a digital browser extension and a physically-resistive keyboard, seen in Figure 1. We highlight the fabrication process of the keyboard prototype and present an open-source method for replicating the heat-sealing step. Throughout this section, we refer to the browser extension as the *digital prototype*. We refer to the physical keyboard device as the *physical prototype*.

3.1 Digital Prototype

We designed and developed a digital browser extension that implements *eco-feedback* as a design strategy. Given the goal of "inform[ing] users clearly about what they are doing and to facilitate consumers to make environmentally and socially responsible decisions through offering real-time feedback" [1], we provided clarity and insights into the energy usage of interacting with a GenAI tool. We drew inspiration from [31], which delineates the major dimensions that are critical when designing an eco-feedback display: behavioral, temporal, data, metrics, valence, and contextual information. Specifically, we displayed individual behavior over an entire single session (*behavioral*), updated the display upon every prompt (*temporal*), used financial cost and energy equivalencies as *metrics*, and used a negative *valence* indicating resources consumed (versus saved). Future work could explore if incorporating *contextual information* regarding a proposed goal or comparing to peers would be effective.

We implemented this eco-feedback design strategy as a Chrome extension on the ChatGPT webpage, as seen in Figure 1. The extension provides an overlay that allows the user to start and end a session and toggle the Arduino communication to the air supply on or off. Once a prompt has been entered, a popup is displayed on the users' right side of the screen. The overlay is titled "Energy Impact" and provides data on the session time, a color graded status bar indicating level of resistance, energy conversions of queries in terms of full iPhone battery capacity charges, Google searches, minutes of LED light, and average energy cost, and finally token count per query and per session. The tokens were calculated using the ChatGPT API approximate conversion of 1 token for every 4 characters entered in a prompt[1]. To account for the substantially higher computational cost of image outputs, prompts that produced an image were scaled by a factor of 60x, as derived in [6]. A full table of the energy conversion factors and their respective sources can be found in the appendix (Section 8.3). This chrome extension can be used as a standalone tool and is available for future researchers upon request.

3.2 Physical Prototype

In addition, we designed and fabricated a physical keyboard device that implements *eco-steering* through counter-functionality as a distinct design strategy. We used the schematized design process from [10] to break down the counter-functional affordances we sought to achieve in GenAI interactions. This schema discretizes counter-functionality as a design tool into three distinct steps: 1) the typical affordance of a device, 2) the interaction that is now constrained, and 3) the new interactions that may emerge as a result. For this work, we frame this schema in the following way:

1. Normally one *can* engage in high-frequency, low-deliberation prompt interactions with a GenAI system.
2. Now one *can not* sustain high-frequency prompt interactions without expending significant physical effort.
3. But now, one might engage more deliberately and reflectively with a GenAI system, treating each prompt as a considered act rather than a reflexive one.

We developed a variable-resistance keyboard that adaptively adjusts keystroke resistance as a function of GenAI interaction patterns, penalizing high-frequency usage through increased

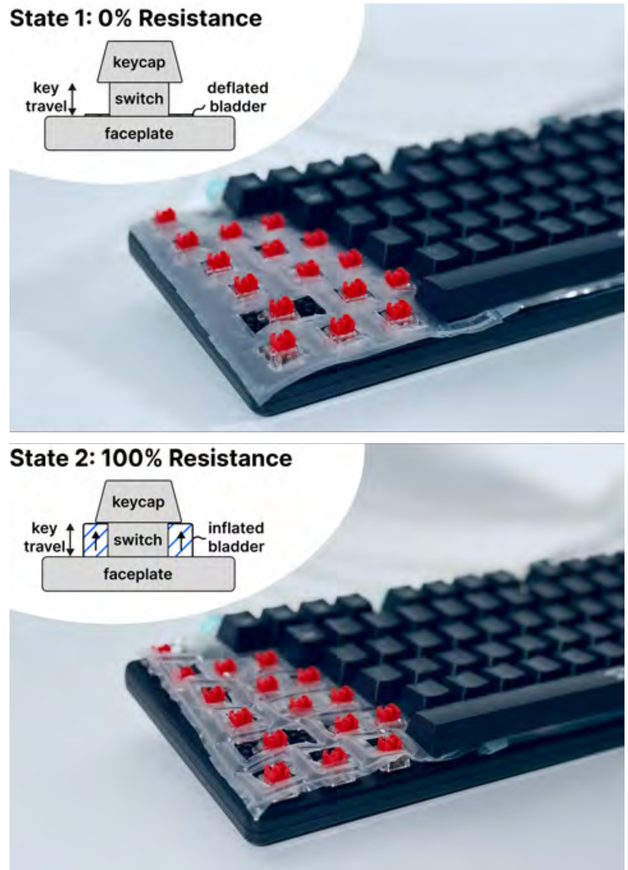

FIGURE 2: DIAGRAM AND PROTOTYPE SHOWING HOW THE KEYBOARD INHIBITS TYPING BY INTRODUCING AN INFLATABLE BLADDER IN THE PATH OF KEY TRAVEL (KEYCAPS REMOVED FOR DEMO).

physical demand and embedding friction directly into the interaction.

The physical prototype consisted of a mechanical tactile keyboard augmented with an inflatable Thermoplastic Polyurethane (TPU) bladder positioned beneath the key switches, capable of exerting variable opposing resistance against users' keystrokes (Figure 2). Resistance was dynamically scaled to the cumulative token count generated by the user during the design task, which is described in depth in Section 3.2.2. An overview of the complete system interaction is shown in Figure 3.

3.2.1 Bladder Fabrication. The bladder was fabricated from double-layered 0.1mm TPU film, heat-sealed to conform to the negative geometry of the keyboard's underside. Heat sealing was achieved by repurposing a fused deposition modeling (FDM) 3D printer: operated without filament, the heated nozzle served as a controlled thermal bonding tool guided along predefined seam paths. To generate these toolpaths, we developed an open-source, in-house G-code generation system that accepts an arbitrary geometry in DXF format and outputs a finalized G-code file executable directly by the printer. The instructions for following this process have been made available at tinyurl.com/RRheatsealing for use in future precision heat-sealing applications. The sealing geometry was designed in CAD, exported to DXF, and processed

[1] https://platform.openai.com/tokenizer

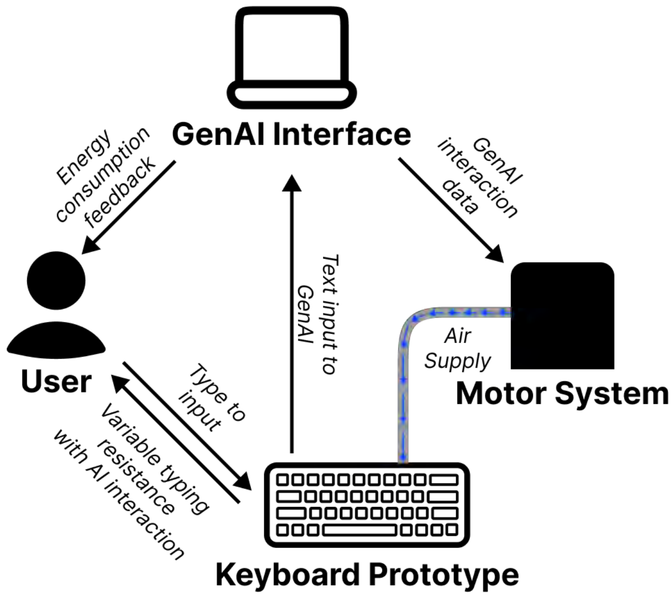

FIGURE 3: SYSTEM OVERVIEW OF THE FEEDBACK LOOP BETWEEN USER TYPING, A BROWSER-BASED ENERGY TRACKER, GENAI TOKEN USAGE, AND REAL-TIME MOTOR-CONTROLLED KEYBOARD RESISTANCE.

through this pipeline before fabrication. This method provided greater repeatability than manual heat sealing, yielding consistent seam patterns across samples, and the programmable workflow enabled systematic variation of parameters such as temperature, travel speed, and pressure depth. Due to bed size constraints, the bladder was fabricated in three parts and subsequently joined.

Assembly and Enclosure The fabricated bladder was trimmed to accommodate the keycaps and switches, then connected to an air supply via flexible silicone tubing and 3D-printed valve splitters, sealed with rubber bands and secured with 3D-printed retaining rings. The air supply, motor driver, and Arduino microcontroller are housed within a 3D-printed enclosure and powered via a wall-connected power supply. Both the Arduino and keyboard interface with the host laptop via USB.

3.2.2 Token to Physical Feedback Translation. In order to map GenAI usage to keyboard resistance, a three-stage communication pipeline translates the participant token usage to physical feedback via pulse width modulation (PWM) motor control. The three stages are token calculation, resistance mapping, and serial command transmission.

Stage 1: Token Calculation When the user submits a prompt, the content script of the Chrome extension (described in Section 3.1) estimates token count using a character-based heuristic, which approximates an equivalency of 1 token for every 4 characters. Tokens accumulate across a session and are stored in a running *totalTokens* counter. Special multipliers are applied for image generation, reflecting the disproportionately higher computational cost of image based inferences.

Stage 2: Resistance Mapping Accumulated tokens are mapped to a resistance percentage (0-100%) via a linear function with a max of 90 tokens.

$$R = \min\left(100,\ \frac{T}{T_{\max}} \times 100\right) \tag{1}$$

where R is the resistance percentage, T is the token count, and $T_{\max}$ is the token threshold for maximum resistance. This resistance value is not static, as it decreases by 1.25% every 5 seconds during idle periods, mimicking a cooldown effect for computational systems. As an example, a full resistance bar value of 90 tokens would translate to ~60-70 words or ~1 paragraph. Values for maximum tokens and decay rate were chosen empirically through iterative pilot testing.

Stage 3: Serial Command Transmission The resistance percentage is relayed from the content script through the Chrome extension's background service worker to a local Node.js/Express bridge server via HTTP POST. The bridge writes serial commands to the Arduino over USB. On the Arduino side, incoming resistance values are parsed and mapped to PWM duty cycles: resistance below 25% produces zero output (motor off) but increases the status bar to signal to the user on increased token count. Values from 25-100% are linearly scaled to PWM 153-255 (60-100% duty cycle). A ramp-step of 15PWM units per 10ms loop iteration provides smooth motor acceleration and deceleration, avoiding abrupt transitions. The 31 kHz PWM frequency (achieved via Timer1 prescaler configuration) minimizes audible motor whine and vibration.

4. USER STUDY METHODS

We conducted a within-subjects user study with both prototypes to observe how people interact with and respond to each design. Our goal was to understand how these interventions shape patterns of AI use, perceived value of AI, and moments of reflection. The study was approved by the authors' Institutional Review Board (IRB), conducted in-person, and participants were compensated $20 for their time.

4.1 Participant Information

Participants were recruited through university mailing lists and needed to have prior experience interacting with generative AI tools. Data was collected from 30 participants. Participants ranged from 18 to 34 years old. 19 participants were women, 10 were men, and 1 preferred not to share. There were 14 undergraduate students, 15 graduate students (PhD and masters), and 1 working professional. 12 participants indicated that they had 3-5 years of engineering or design experience, 11 participants indicated that they had 0–2 years of engineering or design experience, while 4 indicated that they had 6-9 years of engineering or design experience and 3 indicated they had 10+ years of engineering or design experience. 5 participants indicated they used GenAI tools a few times a month, 14 participants indicated they used GenAI tools a few times a week, and 11 participants indicated they used GenAI tools daily.

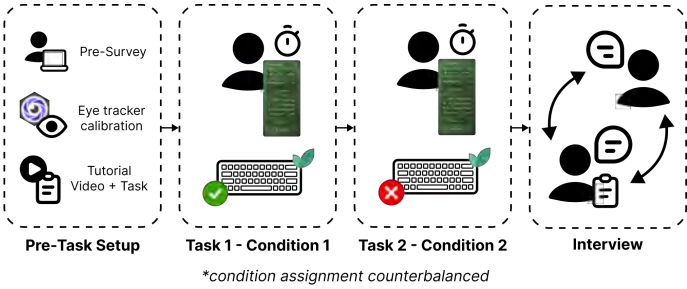

FIGURE 4: OVERVIEW OF USER STUDY FLOW USED TO CAPTURE INTERACTIONS AND RESPONSES TO ARTIFACTS.

4.2 Study Design

As shown in Figure 4, participants first completed a pre-survey regarding their demographics and GenAI habits, along with questionnaires on reflection and environmental attitudes. Then, participants completed eye-tracking calibration and engaged with a tutorial where they gained familiarity with the system. Next, they completed a design task where they used ChatGPT to generate images for a storyboard of a new product. In the first condition, they completed the task with the browser plugin providing feedback on their interactions. In the second condition, they completed the task with both the browser plugin and the keyboard prototype. Conditions were counterbalanced across subjects. Participants were instructed to think aloud during both conditions. After each condition, participants completed a post-task survey including questions on their reflection of the task as well as NASA TLX cognitive load questionnaire. After both tasks were completed, a post-task interview was conducted to better understand the participants' experience. Details on the study task can be found in the appendix, Subsection 8.2.

4.3 Data Collection

During the study, participants were asked to think aloud while completing tasks, with sessions captured via audio recording and screen capture – the latter also serving as a basis for gaze tracking analysis (see Subsubsec. 4.3.1). Think-aloud recordings were transcribed alongside post-task interview data, then analyzed collaboratively by two of the authors using thematic analysis [32]. Qualitative codes were developed iteratively: grouped, discussed, and refined until consensus was reached.

When the browser extension was active, each prompt interaction was automatically logged for that session, capturing the prompt text, length, and token count, as well as whether an image was generated. Image generation was a considerable variable because of the significant increased energy cost of generating an image as opposed to text [6].

4.3.1 Gaze Tracking. To measure and track the attention of the user throughout the study, we used the commercially-available Beam Eye Tracker software. Commonly used for academic studies and in the gaming industry, the software uses a face camera to create an overlay showing a bubble of approximately where the user is looking on the screen. Throughout the study, the bubble was hidden, but recorded for data collection purposes. Users were asked to do a short calibration task before the study to ensure accurate tracking. Post study, we ran a computer vision algorithm to track the bubble at all times and detect instances where the user was looking at the feedback display, as well as what metrics they were looking at [33].

We used the following Locus-of-Attention Index (LAI) metric from [34] to track attention over time, defining our area of interest (AOI) as the chrome extension itself:

$$LAI = \frac{f_{non-extension} - f_{extension}}{f_{non-extension} + f_{extension}} \tag{2}$$

where $f_{non-extension}$ represents # of frames where attention was on the screen but not on the browser extension, while $f_{extension}$ represents # of frames where attention was on the browser extension. Thus, $LAI = 1$ when attention is fully outside the extension and $LAI = -1$ when attention is fully inside the extension for a given window of time.

5. RESULTS

Here, we present results from the user study where participants interacted with a GenAI tool and our prototype(s). For brevity, we refer to the purely digital condition as the *digital condition*, where feedback was presented only through the browser extension. Similarly, we refer to the physically-augmented condition as the *physical condition*, where feedback was presented through both the browser extension and prototyped keyboard.

5.1 GenAI Prompting Patterns

Prompt inputs were logged during the task to observe querying behaviors under both conditions. Across all participants, a total of 306 prompts were logged (digital: n=150, physical: n=156), with prompt lengths ranging from 1 to 1,495 characters (Mdn = 103.5, M = 179.9). Specifically, three metrics were recorded per prompt: prompt length (characters), prompt token count, and total number of prompts submitted per session. Figure 5 presents an overview of these measures aggregated across all participants. Prompt length showed no significant difference between conditions ($Mdn_{\text{digital}} = 143.54$, $Mdn_{\text{physical}} = 146.72$, $W = 160$, $p = .140$), though variance was higher in the digital condition (SD = 275.31) than physical (SD = 174.70), suggesting greater consistency in prompting behavior in the physical condition. Prompt tokens showed a slight decline from the digital (Mdn = 1790.00, SD = 4178.24) to physical condition (Mdn = 1646.42, SD = 2680.61), which was statistically significant ($W = 136$, $p = .047$). The number of prompts per condition showed no significant difference ($Mdn_{\text{digital}} = 5$, $Mdn_{\text{physical}} = 5$, $W = 152$, $p = .541$), with variance again lower in the physical condition (SD = 2.12 vs. SD = 2.80). Although variance was consistently higher in the digital condition across all three metrics, Levene's tests indicated these differences were not statistically significant.

5.2 Attention Over Time

Figure 6 shows the mean Locus-of-Attention Index (LAI) across trial quartiles for both the digital and physical condition. In the digital condition, participant attention remained stable throughout the task ($F(3, 81) = 0.71$, $p = .550$, $\eta^2 p = .026$), ranging from 0.882 to 0.835, indicating gaze was consistently directed *outside* the browser extension with no significant change over time. The physical condition showed similar stability during

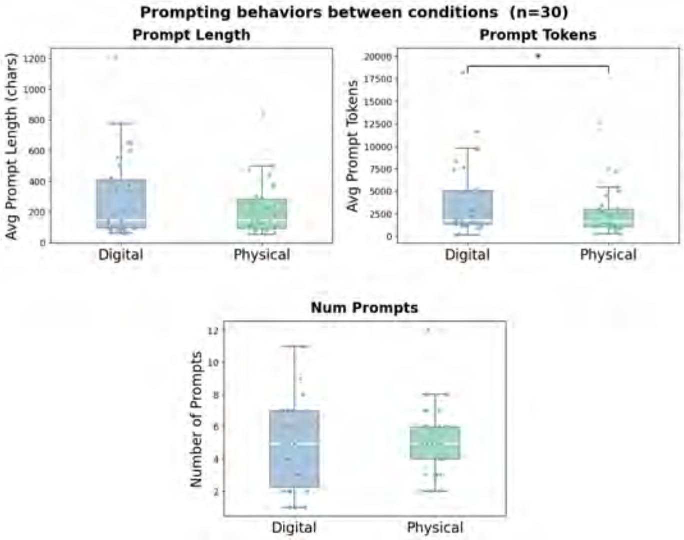

FIGURE 5: DISTRIBUTION OF PROMPTING BEHAVIOR ACROSS BOTH CONDITIONS, SHOWING PROMPT LENGTH, TOKEN COUNT, AND NUMBER OF PROMPTS PER SESSION.

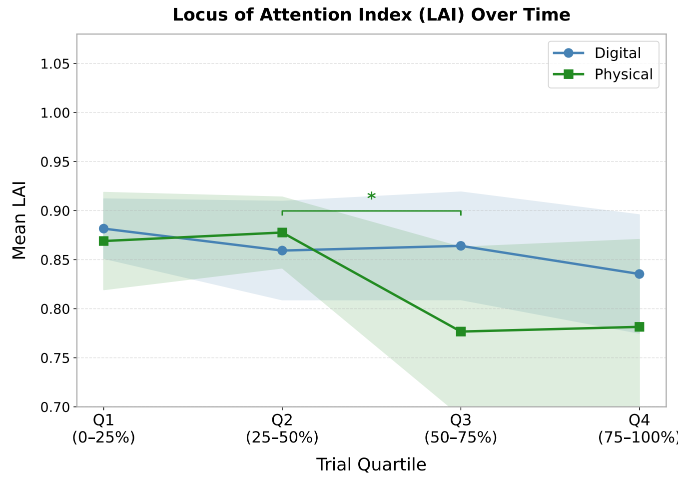

FIGURE 6: MEAN LOCUS-OF-ATTENTION INDEX ACROSS TRIAL QUARTILES FOR BOTH CONDITIONS, WITH 95% CONFIDENCE INTERVALS SHADED. THE PHYSICAL CONDITION MAINTAINED STABLE LAI THROUGH Q2 BEFORE DROPPING SHARPLY AT Q3 AND REMAINING SUPPRESSED, WHILE THE DIGITAL CONDITION REMAINED STABLE THROUGHOUT.

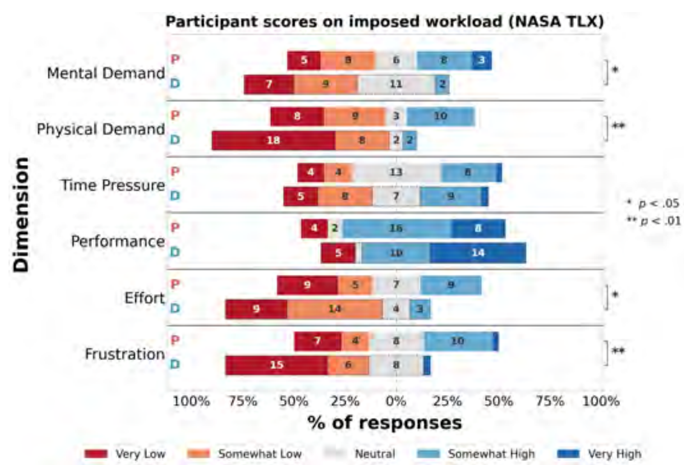

FIGURE 7: POST-TASK SURVEYS CAPTURED PARTICIPANTS' PERCEPTIONS OF IMPOSED WORKLOAD PER CONDITION.

the first half of the task, but exhibited a significant decline in the second half (F(3, 81) = 2.81, p = .045, η^2p = .094). Post-hoc paired comparisons revealed this effect was driven by a significant drop between Q2 and Q3 (t(27) = 2.18, p = .039), after which attention remained at this lower level through Q4. This mid-task shift may reflect a growing awareness of the physical keyboard as resistance increases, leading participants to check the energy tracker as they enter their prompts and feel the effects directly.

5.3 Perceptions of Workload

Participants self-reported the imposed workload they felt during each condition, providing multidimensional insights into how each modality was experienced from a cognitive perspective, as seen in Figure 7. The most striking difference appears in Physical Demand from the prototypes, where 67% of digital condition participants rated it "Very Low" or "Somewhat Low" while 37% of physical condition participants rated it "Somewhat High" or "Very High". This contrast is expected given that the physical condition required additional manipulation of a tangible input. A similar pattern holds for Effort and Frustration: digital participants clustered heavily toward the low end (77% rating Effort either "Very Low" or "Somewhat Low"; 50% rating Frustration "Very Low"), whereas physical participants reported more mixed and elevated responses, likely reflecting the added demands of working with friction-embedded physical objects. On Performance, satisfaction skewed positively for both conditions, but the digital condition yielded a notably higher proportion of "Very High" scores (47% vs. 27%), suggesting participants felt more successful in the digital environment, though they were able to effectively complete the task in both conditions.

Wilcoxon signed-rank tests confirmed these patterns statistically. Significant differences emerged for Mental Demand ($W = 28.5$, $p = .02$), Physical Demand ($W = 9.0$, $p = .001$), Effort ($W = 38.0$, $p = .035$), and Frustration ($W = 35.0$, $p = .005$), with the physical condition yielding higher ratings across all four dimensions (Mental Demand: $Mdn = 3.0$ vs. 2.0; Physical Demand: $Mdn = 2.0$ vs. 1.0; Effort: $Mdn = 3.0$ vs. 2.0; Frustration: $Mdn = 3.0$ vs. 1.5). Time Pressure and Performance showed no reliable differences between conditions ($p = .434$ and $p = .604$). These findings indicate that the physical condition imposed a significantly greater burden in terms of mental and physical exertion, and emotional frustration, while temporal demands and perceived performance remained largely equivalent across modalities.

In addition to questions about cognitive workload, participants answered post-task questions about the perceived value and added awareness that both forms of feedback provided. Across both conditions, participants broadly reported that the feedback was useful and prompted reflection on their GenAI use. As shown in Figure 8 of the Appendix, the majority of respondents agreed or strongly agreed that the feedback helped them maintain an overview of their GenAI use (P: 87%; D: 67%), made it easy to

Theme	Subtheme	Description
Prompting reflection	Haptic	Keyboard resistance physically interrupting or slowing the act of typing
	Digital	The visual energy tracker drawing attention to consumption during prompting
	Audio	Motor noise creating an auditory signal of computational effort
	Mental	The cognitive load of overcoming the hindrance itself prompting pause
Behavior change strategies	Tradeoffs	Weighing the cost of further prompts against another factor
	Prompt Modifications and Intentionality	Changes to how prompt composition was approached
Ownership	Guilt and anxiety	Emotional discomfort arising from seeing one's own specific consumption quantified
	Consequences	Environmental and resource costs that once felt distant becoming personal
Digital feedback and intangibility	Unobtrusive Nature	The extension's peripheral placement limiting how much it registered during tasks
	Intangible Metrics	Difficulty making sense of the raw numbers without a meaningful frame of reference

TABLE 1: OVERVIEW OF QUALITATIVE THEMES THAT AROSE DURING USER STUDY WITH BOTH ARTIFACTS.

reflect as the task progressed (P: 83%; D: 80%), and made them more deliberate about how often they used the AI (P: 67%; D: 53%). Agreement was similarly high for the feedback prompting new ways of thinking about GenAI use and increasing attentiveness when writing prompts. The physical condition tended to show slightly stronger positive responses across most statements, with the largest divergence on active engagement with the feedback (P: 60% agree or strongly agree vs. D: 38%). Full response distributions for all items are reported in the appendix 8.4. These findings are expanded upon with think-aloud and interview data, which probed at participants' thought process as they engaged with these tools, but generally show that participants can be accepting of this feedback and perceive it as useful.

5.4 Qualitative Insights

Thematic analysis of think-aloud and interview transcripts yielded 79 initial codes iteratively refined into five themes, four of which are discussed below and summarized in Table 1.

5.4.1 Reflective Triggers Vary Across Participants. Artifacts generally prompted moments of reflection during the study, however, the mechanism or source that triggered the reflection varied among participants. Participants were most often interrupted by one or more of the following: 1) the haptic feedback of keyboard resistance (15 participants), 2) the energy tracker in the browser extension (18 participants), 3) the noise generated by the motor powering the keyboard prototype (9 participants), or 4) the mental load of overcoming the hindrance itself (4 participants). For some, the combination of modalities proved effective (13 participants), as P33 noted:

> *"It's both the resistance from the keyboard as well as those numbers that made me rethink the way I use AI." (P33)*

However, for others, attending to the presented feedback introduced additional demand that competed with task focus, as P57 described:

> *"Because I'm constantly thinking about my task, I don't really have the cognitive space to also think about those numbers [the energy equivalencies] and then think what that means in terms of my task." (P57)*

These varied responses suggest that no single mechanism of interruption consistently reached all participants, with individual differences in cognitive style and task focus shaping whether and *how* the feedback registered.

5.4.2 Behavior Change and Tradeoffs. In response to both physical and digital feedback, participants demonstrated a range of adaptive strategies that reflected a shift toward more resource-conscious interaction with AI, including more concise prompting, greater intentionality, and reduced iteration. A common strategy was prompt refinement (19 participants), where participants became more deliberate with their inputs and moved from exploratory to efficiency-driven interaction. As P77 described,

> *"How can I more effectively express the prompt to achieve the same effect that I want with less input?" (P77)*

Similarly, P52 noted that the feedback introduced a natural brake to habitual prompting:

> *"When you don't have the [modified] keyboard, you just type away, but when you start getting the feedback you're like okay, I have to slow down." (P52)*

In this way, the artifacts successfully embodied the concept of "slow technology" [22], disrupting habitual System 1 prompting behaviors in favor of the slower, more deliberate System 2 thinking described by [23].

Beyond efficiency strategies, participants also began to consciously navigate the tradeoff between output quality and the environmental cost of each additional prompt (10 participants). The most common source of this tension was participants' desire to iterate further until they were satisfied with the output, which often required additional prompts, as described by P43:

> *"So when I see the resistance level, it made me think, maybe I should not ask it for more edits. But at the same time, I wanted a level of quality in my task, so I was not going to shy away from making those edits, because at the end of the day, the quality of the image mattered a lot to me." (P43)*

This internal negotiation was echoed by P89, who reflected that the digital feedback tool could help gauge whether a task was worth its environmental cost:

> *"I really liked that, if there was something like [the digital feedback tool available], I would use it for myself in my daily life. I would be like, okay, can I figure out this bug in two minutes? Or should I use, I don't know, eight bottles of water?" (P89)*

Together, these responses suggest that the artifacts make environmental cost a salient factor in participants' decision-making, even if they are eventually out-prioritized.

5.4.3 Ownership Versus Guilt.
A few users found that the feedback system fostered a sense of ownership over their actions. By physically feeling the resource consumption of their GenAI use and viewing the numerical impacts, some participants felt that the consequences of their usage were brought closer to the individual. For example, P23 described how the eco-feedback made the impacts of their usage more difficult to ignore:

> *“I do feel more of a sense of ownership over what’s happening, instead of just saying the processing is happening somewhere else." (P23)*

It was also noted that these reminders encouraged them to reflect on whether their actions aligned with their stated environmental values. Relatedly, P23 commented:

> *“Having a reminder that takes you back and actually makes you take accountability for saying you’re somebody who cares about those things [environmental impact] was really useful." (P23)*

This participant described how the feedback created a moment of accountability, prompting them to reconsider the disconnect between identifying as environmentally conscious and the impact of their interactions with AI. Participants also described how the feedback system made them more conscious of their behavior and forced them to acknowledge the impacts that their actions have. In fact, five participants characterized the metrics as creating negative emotions, like feeling a “guilt trip" or anxiety. In addition, a few participants questioned where responsibility for the resource consumption should lie, wondering whether the costs were being borne by the companies running the models or other external sources. Together, these responses suggest that showing resource consumption through feedback can bring otherwise invisible impacts closer to the user, prompting reflection on personal responsibility and the consequences of AI use.

5.4.4 The Intangibility and Invisibility of Digital Feedback.
An interesting commonality between multiple participants was a shared feeling of ambiguity and intangibility of the digital feedback. Metrics such as energy equivalencies, costs, or tokens were often hard for participants to interpret and relate back to their interactions with the system. For example, P38 explained that the numerical nature of the feedback made it difficult to understand its implications:

> *“So I felt that it [the digital readout] didn’t really influence me much on how I was typing or what I was asking it to do, because I just like generating numbers, and I kind of didn’t really understand.” (P38)*

These responses suggest that while digital metrics provided information about resource consumption, their abstract nature made it hard for participants to fully form the connection between their actions and the feedback and incorporate that into their decision-making. The digital feedback interface was also often perceived as unobtrusive to the point that it did not meaningfully interrupt participants’ workflows. Because the plugin appeared to the side of the primary work interface rather than directly within the interaction itself, several participants reported that it was easy to ignore it while focusing on completing the task. These responses suggest that the unobtrusive nature of the popup wasn’t invasive enough to invoke change or sustained reflection, especially when users are focused on completing time-constrained tasks.

6. DISCUSSION

This work presents a research through design approach using two artifacts to probe behavioral responses to GenAI’s environmental impacts, surfacing several considerations for DfSB in this emerging context. We discuss how limited aggregate differences in prompting behavior across conditions can reveal individual variation in behavior change triggers, the challenge of motivating change without inducing guilt, and the lens of slow technology as a new DfSB approach for uncovering mechanisms of reflection as AI tools reshape design practice.

6.1 Interpreting Prompting Behavior Across Conditions

Quantitative prompting data revealed limited behavioral differences between conditions in aggregate, which may reflect several factors: constraints inherent to the design task itself, the wide range of individual responses to the interventions, and the fact that both conditions included eco-feedback, making it difficult to isolate the effect of either intervention alone. Aggregate behavioral metrics may also understate the true impact of the interventions, as interview responses indicate that many participants modified their interactions in ways that departed meaningfully from their day-to-day GenAI use, even when this did not manifest as statistically significant differences in prompt length or count. This pattern of individual variation points to a broader challenge for the field of designing strategies that may successfully resonate with a diverse set of users.

Prior work has identified a persistent gap in AI eco-feedback research: tools that visualize energy consumption are rarely accompanied by empirical evaluation [35]. This study begins to address that gap. Our findings suggest that engagement with eco-feedback interventions spans a wide spectrum: the physical condition imposed significantly higher workload across mental demand, effort, and frustration, yet produced stronger agreement that feedback prompted deliberate reflection, while digital energy equivalencies were frequently described as abstract and difficult to act on. Fogg’s Behavior Model holds that motivation, ability, and a well-timed trigger must align simultaneously [36], but these factors vary considerably across individuals and contexts. Future work should more distinctly characterize this range of responses in order to build tools that are better calibrated to the diversity of users they aim to reach.

6.2 Calibrating Emotional Valence and Behavior Change

While counterfunctional and slow technology approaches proved effective at interrupting habitual prompting behaviors in our study [10, 22], they also risk crossing an emotional line between productive reflection and feelings of guilt or blame. Negative valence messaging, which frames use as consumption as opposed to a positive alternative, is often used in eco-feedback technologies and can provide a sense of urgency that motivates behavior change [36]. However, invoking *too* much negative emotion can backfire and lead to outright avoidance [18]. In our study, five participants described the metrics as inducing guilt or anxiety, and several questioned whether responsibility for consumption should fall on individual users at all rather than the companies running the models. Yet participants were often encountering quantified information about their personal AI use for the first time, and overall curiosity and expressed desire to change behavior remained high despite this discomfort – P46 described seeing their own consumption data as novel and eye-opening, and P43 expressed frustration that existing interfaces offer no support for acting on this awareness. This coexistence of guilt and curiosity suggests that users can hold discomfort and genuine interest simultaneously, and that negative affect does not necessarily bar motivated engagement. This is a broader challenge for DfSB interventions that rely on friction or provocation as a mechanism for behavior change [1, 2]. Friction-based interventions must be carefully calibrated to invoke the deliberate, reflective reasoning that behavior change requires [23] while framing that reflection in ways that confer a sense of agency rather than culpability.

6.3 Slow Technology as a DfSB Design Approach

This work introduces slow technology as a design approach in the context of DfSB, a framing that has seen limited application in engineering design contexts. We view this lens as complementary to, and a tool for investigating, phenomena around dual-process theories in design work [23, 37]. Particularly, prior research has shown deep interplay between Type 1 and Type 2 thinking during early stages of the design process, including ideation [37], and slow technology offers a mechanism for deliberately shifting designers toward the reflective, deliberate reasoning that Type 2 thinking requires. AI tools are enabling phases from concept generation to prototyping to become increasingly automated [38, 39], and understanding how these tools change the way designers think and approach the design process is an ongoing and prominent research area. However, prior work has shown that designers tend to favor concrete, solution-oriented thinking over abstract analysis [23], and GenAI tools may amplify this tendency. While rapid ideation is encouraged across much of the design process to promote volume and diversity of ideas, some evidence suggests that ideating over longer time horizons improves quality and novelty [40]. As such, we propose the use of slow technology as a design tool for investigating our understanding of dual-process theory and how it may shift with the rise of GenAI tools.

7. FUTURE WORK AND LIMITATIONS

It is worth noting that this work does not seek to identify an optimal intervention modality. Instead, the aim is to surface how different modalities shape the nature of user reflection and engagement with environmental cost. The presented study was limited by the laboratory setting, which may not completely capture real-life or long-term GenAI usage. Future work should explore deploying these interventions in a longer-term, ethnographic setting to explore how effects shift with familiarity and learning effects. The physical constraints of the bladder geometry also capped the maximum resistance participants could feel, and refining the design to achieve full typing inhibition would enable exploration of a wider range of constrained behaviors and their effects. Additionally, the token conversion and energy calculations relied on estimates from prior work, and life cycle assessments conducted on the specific model in use could yield more precise figures. Beyond these refinements, future work might also vary the framing of sustainability information presented to users, such as manipulating valence (positive vs. negative) or impact granularity (individual vs. collective) [16], to examine whether different informational framings elicit novel behaviors in the context of these prototypes. Finally, the variable resistance keyboard artifact presented in this paper could provide a useful tool for studying *slow technology* as a philosophy in other computer interactions, as the mechanism of inhibiting typing resistance may prove useful in a wide range of tasks.

8. CONCLUSION

Design for sustainable behavior (DfSB) strategies explore interventions for supporting pro-environmental behaviors, and we draw upon two of these strategies (*eco-feedback* and *eco-steering*) in this work. Following a research through design approach, we design and prototype a browser-based extension that establishes a digital informational baseline by displaying real-time feedback on energy usage, and a dynamically adaptive keyboard that augments that baseline with physical, embodied friction by modulating typing resistance in real-time. We situate these prototypes within GenAI interactions, a technology whose environmental impact is still being fully characterized, though emerging evidence suggests it is substantial. In a user study with 30 participants, interaction behaviors do not change significantly across conditions, yet many participants report meaningful shifts in environmental awareness and describe actively negotiating tradeoffs between output quality and environmental cost – suggesting that slow technology principles can be productively applied to AI interactions, disrupting habitual prompting behaviors and creating space for more deliberate, reflective engagement. Notably, augmenting the digital eco-feedback baseline with physical friction surfaced distinct patterns of reflection not observed with purely digital feedback alone, underscoring that the two modalities are not interchangeable but complementary. However, participants also noted that current AI interfaces offer little scaffolding to act on this awareness, pointing to a gap between motivation and available tools. Findings underscore that modality diversity is key to reaching a broad range of users, while the persistent gap between environmental awareness and actionable tools points to open questions for future DfSB research in the context of GenAI.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 2146752. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This work was also supported by the UC Berkeley Jacobs Institute for Design Innovation's Innovation Catalyst grant.

REFERENCES

[1] Bhamra, Tracy, Lilley, Debra and Tang, Tang. "Design for Sustainable Behaviour: Using Products to Change Consumer Behaviour." *The Design Journal* Vol. 14 No. 4 (2011): pp. 427–445. DOI 10.2752/175630611X13091688930453.

[2] Lilley, Debra. "Design for sustainable behaviour: strategies and perceptions." *Design Studies* Vol. 30 No. 6 (2009): pp. 704–720. DOI 10.1016/j.destud.2009.05.001.

[3] Lockton, Dan, Harrison, David and Stanton, Neville. "Making the user more efficient: design for sustainable behaviour." *International Journal of Sustainable Engineering* Vol. 1 No. 1 (2008): pp. 3–8. DOI 10.1080/19397030802131068.

[4] Colombo, S., Montagna, F., Cascini, G. and Palazzolo, V. F. "Digital Artefacts and The Role of Digital Affordance." *Proceedings of the Design Society* Vol. 2 (2022): pp. 11–20. DOI 10.1017/pds.2022.2.

[5] Goridkov, Nicole and Goucher-Lambert, Kosa. "Harnessing digital vs physical design for sustainable behavior strategies: A review." *Proceedings of the Design Society* Vol. 5 (2025): pp. 1973–1982. DOI 10.1017/pds.2025.10211.

[6] Luccioni, Sasha, Jernite, Yacine and Strubell, Emma. "Power Hungry Processing: Watts Driving the Cost of AI Deployment?" *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*: pp. 85–99. 2024. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/3630106.3658542.

[7] Wu, Carole-Jean, Raghavendra, Ramya, Gupta, Udit, Acun, Bilge, Ardalani, Newsha, Maeng, Kiwan, Chang, Gloria, Aga, Fiona, Huang, Jinshi, Bai, Charles, Gschwind, Michael, Gupta, Anurag, Ott, Myle, Melnikov, Anastasia, Candido, Salvatore, Brooks, David, Chauhan, Geeta, Lee, Benjamin, Lee, Hsien-Hsin, Akyildiz, Bugra, Balandat, Maximilian, Spisak, Joe, Jain, Ravi, Rabbat, Mike and Hazelwood, Kim. "Sustainable AI: Environmental Implications, Challenges and Opportunities." *Proceedings of Machine Learning and Systems* Vol. 4 (2022): pp. 795–813.

[8] Zimmerman, John, Forlizzi, Jodi and Evenson, Shelley. "Research through design as a method for interaction design research in HCI." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: pp. 493–502. 2007. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/1240624.1240704.

[9] Godin, Danny and Zahedi, Mithra. "Aspects of Research through Design: A Literature Review." *DRS Biennial Conference Series* (2014).

[10] Pierce, James and Paulos, Eric. "Counterfunctional things: exploring possibilities in designing digital limitations." *Proceedings of the 2014 conference on Designing interactive systems*: pp. 375–384. 2014. ACM, Vancouver BC Canada. DOI 10.1145/2598510.2598522.

[11] Song, Katherine W, Vivrekar, Janaki, Yeom, Lynn, Paulos, Eric and Salehi, Niloufar. "Crank That Feed: A Physical Intervention for Active Twitter Users." *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*: pp. 1–6. 2021. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/3411763.3451817.

[12] Kim, Hyeonwook, Young, Sydney, Chen, Xuesi, Gupta, Udit and Hester, Josiah. "Slower is Greener: Acceptance of Eco-feedback Interventions on Carbon Heavy Internet Services." *ACM Journal on Computing and Sustainable Societies* Vol. 3 No. 2 (2025): pp. 1–21. DOI 10.1145/3723038.

[13] DiSalvo, Carl, Sengers, Phoebe and Brynjarsdóttir, Hrönn. "Mapping the landscape of sustainable HCI." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: pp. 1975–1984. 2010. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/1753326.1753625.

[14] Kappel, Karin and Grechenig, Thomas. ""show-me": water consumption at a glance to promote water conservation in the shower." *Proceedings of the 4th International Conference on Persuasive Technology*: pp. 1–6. 2009. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/1541948.1541984.

[15] Azzi, S., Reymond, G., Mérienne, F. and Kemeny, A. "Eco-Driving Performance Assessment With in-Car Visual and Haptic Feedback Assistance." *Journal of Computing and Information Science in Engineering* Vol. 11 No. 041005 (2011). DOI 10.1115/1.3622753.

[16] Mohanty, Vikram, Filipowicz, Alexandre L. S., Bravo, Nayeli Suseth, Carter, Scott and Shamma, David A. "Save A Tree or 6 kg of CO2? Understanding Effective Carbon Footprint Interventions for Eco-Friendly Vehicular Choices." *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*: pp. 1–24. 2023. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/3544548.3580675.

[17] Goucher-Lambert, Kosa and Cagan, Jonathan. "The Impact of Sustainability on Consumer Preference Judgments of Product Attributes." *Journal of Mechanical Design* Vol. 137 No. 8 (2015): p. 081401. DOI 10.1115/1.4030271.

[18] Bao, Qifang, Burnell, Edward, Hughes, Ann M. and Yang, Maria C. "Investigating User Emotional Responses to Eco-Feedback Designs." *Journal of Mechanical Design* Vol. 141 No. 2 (2019): p. 021103. DOI 10.1115/1.4042007.

[19] Saadi, Jana I. and Yang, Maria C. "Motivating Sustainable Behavior Using Cognitive Interventions in Product Design." 2020. American Society of Mechanical Engineers Digital Collection. DOI 10.1115/DETC2020-22464.

[20] Halabieh, S. and Shu, L. H. "Reducing Waste Outflow to Motivate Water Conservation." *Journal of Mechanical Design* Vol. 146 No. 2 (2024): p. 020905. DOI 10.1115/1.4064042.

[21] Jou, William, Beaulieu, Samantha M., Lim, Adrienne K. and MacDonald, Erin F. "A Wizard-of-Oz Experiment to Demonstrate Water Reduction and User Training With an "Autonomous" Faucet." *Volume 7: 31st International Conference on Design Theory and Methodology*: p. V007T06A052. 2019. American Society of Me-

chanical Engineers, Anaheim, California, USA. DOI 10.1115/DETC2019-98468.

[22] Hallnäs, Lars and Redström, Johan. "Slow Technology – Designing for Reflection." *Personal and Ubiquitous Computing* Vol. 5 No. 3 (2001): pp. 201–212. DOI 10.1007/PL00000019.

[23] Kannengiesser, Udo and Gero, John S. "Design thinking, fast and slow: A framework for Kahneman's dual-system theory in design." *Design Science* Vol. 5 (2019): p. e10. DOI 10.1017/dsj.2019.9.

[24] Strubell, Emma, Ganesh, Ananya and McCallum, Andrew. "Energy and Policy Considerations for Deep Learning in NLP." Korhonen, Anna, Traum, David and Màrquez, Lluís (eds.). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*: pp. 3645–3650. 2019. Association for Computational Linguistics, Florence, Italy. DOI 10.18653/v1/P19-1355.

[25] Lacoste, Alexandre, Luccioni, Alexandra, Schmidt, Victor and Dandres, Thomas. "Quantifying the Carbon Emissions of Machine Learning." (2019). DOI 10.48550/arXiv.1910.09700. ArXiv:1910.09700 [cs].

[26] Çam, Eren, Casanovas, Marc and Moloney, John. "Electricity 2025: Analysis and Forecast to 2027." (2025).

[27] energy.gov. "DOE Releases New Report Evaluating Increase in Electricity Demand from Data Centers." (2024).

[28] Verdecchia, Roberto, Sallou, June and Cruz, Luís. "A systematic review of Green AI." *WIREs Data Mining and Knowledge Discovery* Vol. 13 No. 4 (2023): p. e1507. DOI 10.1002/widm.1507.

[29] Luccioni, Alexandra Sasha, Viguier, Sylvain and Ligozat, Anne-Laure. "Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model." .

[30] Berthelot, Adrien, Caron, Eddy, Jay, Mathilde and Lefèvre, Laurent. "Understanding the Environmental Impact of Generative AI Services." *Commun. ACM* Vol. 68 No. 7 (2025): pp. 46–53. DOI 10.1145/3725984.

[31] Sanguinetti, Angela, Dombrovski, Kelsea and Sikand, Suhaila. "Information, timing, and display: A design-behavior framework for improving the effectiveness of eco-feedback." *Energy Research & Social Science* Vol. 39 (2018): pp. 55–68. DOI 10.1016/j.erss.2017.10.001.

[32] Braun, Virginia and Clarke, Victoria. "Using thematic analysis in psychology." *Qualitative Research in Psychology* Vol. 3 No. 2 (2006): pp. 77–101. DOI 10.1191/1478088706qp063oa.

[33] Holmqvist, Kenneth, Nyström, Marcus, Andersson, Richard, Dewhurst, Richard, Jarodzka, Halszka and Weijer, Joost van de. *Eye Tracking: A comprehensive guide to methods and measures*. OUP Oxford (2011).

[34] Alzayat, Ayman, Hancock, Mark and Nacenta, Miguel A. "Quantitative Measurement of Tool Embodiment for Virtual Reality Input Alternatives." *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*: pp. 1–11. 2019. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/3290605.3300673.

[35] Görücü, Sinem, Morais, Luiz A. and Panagiotidou, Georgia. "A Critical Analysis of Machine Learning Eco-feedback Tools through the Lens of Sustainable HCI." *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*: pp. 1–18. 2025. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/3706598.3713198.

[36] Fogg, BJ. "A behavior model for persuasive design." *Proceedings of the 4th International Conference on Persuasive Technology*: pp. 1–7. 2009. Association for Computing Machinery, New York, NY, USA. DOI 10.1145/1541948.1541999.

[37] Gonçalves, Milene and Cash, Philip. "The life cycle of creative ideas: Towards a dual-process theory of ideation." *Design Studies* Vol. 72 (2021): p. 100988. DOI 10.1016/j.destud.2020.100988.

[38] Ma, Kevin, Grandi, Daniele, McComb, Christopher and Goucher-Lambert, Kosa. "Do large language models produce diverse design concepts? A comparative study with human-crowdsourced solutions." *Journal of Computing and Information Science in Engineering* Vol. 25 No. 2 (2025): p. 024501.

[39] Edwards, Kristen M., Man, Brandon and Ahmed, Faez. "Sketch2Prototype: rapid conceptual design exploration and prototyping with generative AI." *Proceedings of the Design Society* Vol. 4 (2024): pp. 1989–1998. DOI 10.1017/pds.2024.201.

[40] Tsenn, Joanna, Atilola, Olufunmilola, McAdams, Daniel A. and Linsey, Julie S. "The effects of time and incubation on design concept generation." *Design Studies* Vol. 35 No. 5 (2014): pp. 500–526. DOI 10.1016/j.destud.2014.02.003.

[41] You, Josh. "How much energy does ChatGPT use?" https://epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use (2025).

[42] Hölzle, Urs, President, Senior Vice and Operations. "Powering a Google search." https://googleblog.blogspot.com/2009/01/powering-google-search.html (2009).

APPENDIX

8.1 Pneumatic Model of the Keyboard Actuation System

The pneumatic system is modeled as a decaying pressure source feeding three independently valved inflatable bladders. This simplified formulation provides an analytical approach for predicting bladder pressures and guiding design trade-offs in soft, rapidly prototyped pneumatic interfaces. In this analysis, all pressures are absolute and the air is assumed ideal and isothermal.

Supply pressure dynamics. The one-way pump and upstream tubing are modeled as a lumped pressure source with exponential relaxation toward atmospheric pressure:

$$\dot{P}_s(t) = -\lambda_s\big(P_s(t) - P_{\text{atm}}\big) + u_p(t), \tag{3}$$

where P_s is the supply pressure, λ_s is the pump-side decay rate, and $u_p(t)$ represents user pumping input.

Bladder pressure dynamics. Each bladder $i \in \{1, 2, 3\}$ is modeled as a compliant pneumatic chamber. Using mass conservation and the ideal gas law, the pressure dynamics are

$$\dot{P}_i(t) = \frac{RT}{V_i(P_i) + P_i \dfrac{dV_i}{dP_i}} \left(\dot{m}_{\text{in},i}(t) - \dot{m}_{\text{leak},i}(t)\right), \tag{4}$$

where R is the specific gas constant, T is temperature, and $V_i(P_i)$ is the effective bladder volume.

Valve inflow. Flow into bladder i through its electrically triggered valve is approximated by an orifice model:

$$\dot{m}_{\text{in},i}(t) = s_i(t)\, K_{v,i} \sqrt{\max\big(P_s(t) - P_i(t),\, 0\big)}, \tag{5}$$

where $s_i(t) \in [0, 1]$ is the valve activation signal and $K_{v,i}$ is a lumped valve coefficient.

Leakage model. Because the tube–bladder interface is sealed using a compliant rubber constraint rather than a rigid fitting, each bladder exhibits pressure-dependent leakage to atmosphere:

$$\dot{m}_{\text{leak},i}(t) = K_{\ell,i}(t) \sqrt{\max\big(P_i(t) - P_{\text{atm}},\, 0\big)}. \tag{6}$$

To capture increased leakage during keypress events, the leak coefficient is allowed to vary with typing activity:

$$K_{\ell,i}(t) = K_{\ell 0,i}\big(1 + \alpha_i\, u_i(t)\big), \tag{7}$$

where $u_i(t)$ is a normalized keypress signal and α_i controls sensitivity to mechanical disturbance.

Linearized RC form. As a simplified first-order model, the system can be represented by an equivalent pneumatic RC network:

$$C_i\, \dot{P}_i = s_i(t) \frac{P_s - P_i}{R_{v,i}} - \frac{P_i - P_{\text{atm}}}{R_{\ell,i}(t)}, \tag{8}$$

where C_i is the effective pneumatic capacitance and $R_{v,i}$ and $R_{\ell,i}(t)$ are the valve and leak resistances, respectively.

8.2 User study protocol

In the user study, participants were tasked with creating a visual storyboard for a new smart water bottle or new smart alarm clock. Participants were given 3 minutes of independent ideation time before they were allowed access to ChatGPT. Then, participants were given 15 minutes to generate their visual storyboard using ChatGPT. Storyboards were required to be at least 5 panels in length, with each panel needing a caption. When time was up or the participant had finished the storyboard (whichever came first), participants completed a post-task survey on Qualtrics. After this, the same steps were repeated for the second condition, with the respective intervention.

8.3 Energy Equivalency Table

Table 2 presents the values used to convert energy usage per token in the browser extension, along with the corresponding sources.

8.4 Post-Task Survey Results

Figure 8.4 shows participants' perceptions of the feedback provided during each condition in the study, providing insight into aspects of reflection, AI approach, and engagement.

Equivalency	Energy Conversion (per token)	Source
Energy per token	0.004 Wh/token	Luccioni et al. [6]; Epoch AI [41]; Estimated for GPT-5 given values for GPT-4
Smartphone charge	$\frac{0.004 \text{ Wh/token}}{22 \text{ Wh/charge}} = 1.82 \times 10^{-4}$ charges/token	Luccioni et al. [6]
Google search	$\frac{0.004 \text{ Wh/token}}{0.3 \text{ Wh/search}} = 0.0133$ searches/token	Google [42]
LED light usage	$\frac{0.004 \text{ Wh/token}}{0.1667 \text{ Wh/min}} = 0.0240$ min/token	10 W LED assumption
Image generation multiplier	$\frac{2907 \text{ Wh/image generation inference}}{47 \text{ Wh/text generation inference}} = 61.85 \approx 60$	Luccioni et al. [6]; Estimated by comparing mean image generation energy to mean text generation energy, rounded for implementation.

TABLE 2: ENERGY CONVERSION FACTORS USED TO ESTIMATE THE ENERGY IMPACT OF CHATGPT USAGE.

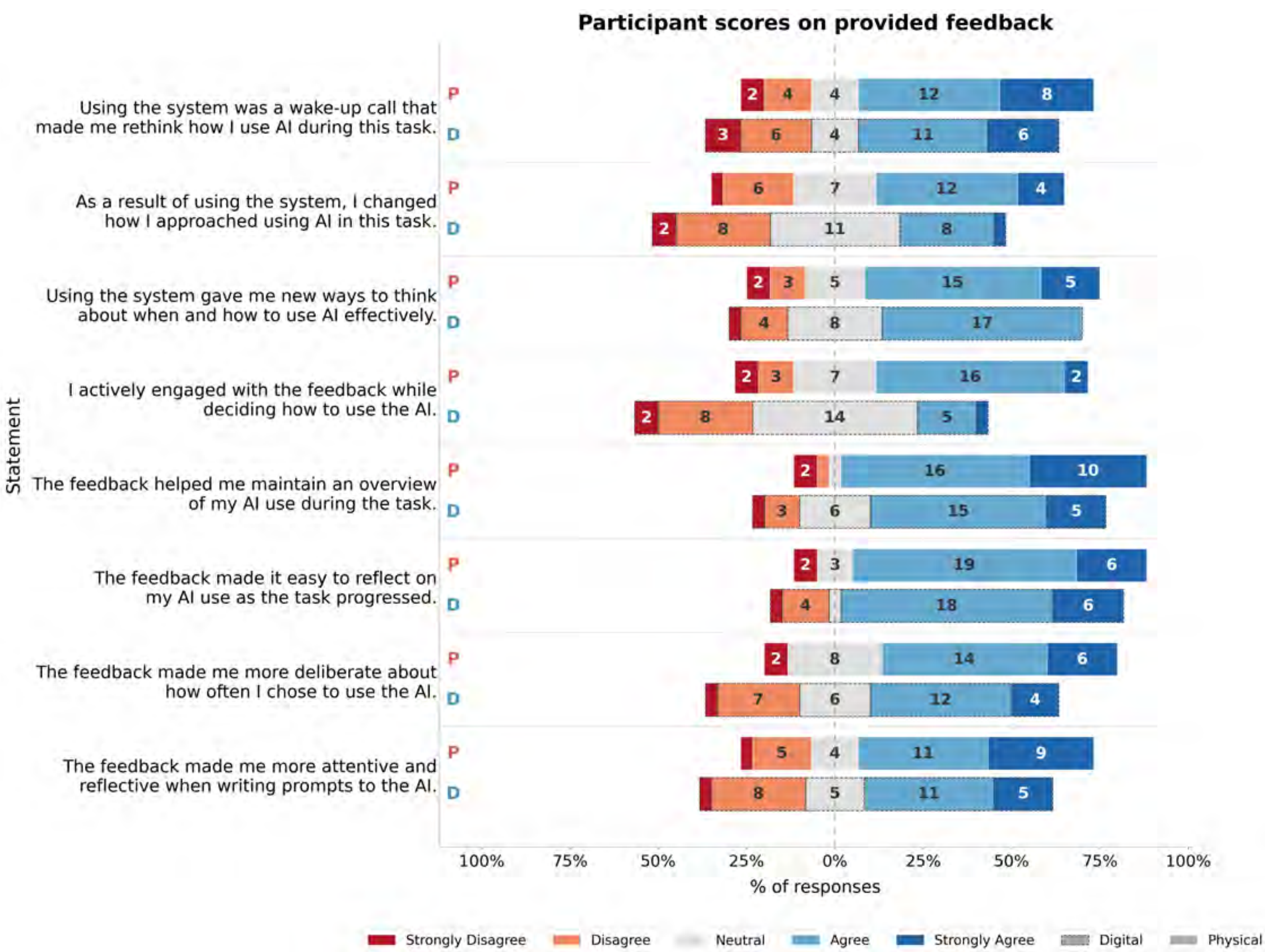

FIGURE 9: POST-TASK SURVEYS CAPTURED PARTICIPANTS' PERCEPTIONS OF THE FEEDBACK PROVIDED DURING EACH CONDITION.