From Learned Helplessness to Digital Agency: Evaluating Seamful Design Interventions in Consent Management Platforms

BENCE SZABO*, Aalborg University in Copenhagen, Denmark
LOUISE FOLDØY STEFFENS*, Aalborg University in Copenhagen, Denmark
SARA SELMAN*, Aalborg University in Copenhagen, Denmark

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; • **Security and privacy** → *Usability in security and privacy*; *Privacy protections*.

Additional Key Words and Phrases: HCI, Seamfulness, Online Consent, Learned Helplessness, Privacy, Digital Agency, CMP, GDPR, Consent Fatigue, Dark Patterns

ACM Reference Format:
Bence Szabo, Louise Foldøy Steffens, and Sara Selman. 2025. From Learned Helplessness to Digital Agency: Evaluating Seamful Design Interventions in Consent Management Platforms. 1, 1 (December 2025), 79 pages. https://doi.org/XXXXXXX.XXXXXXX

Abstract

Current consent management platforms (CMPs) frequently fail to meet the GDPR's mandate for informed and unambiguous consent, often inducing consent fatigue and learned helplessness through manipulative dark patterns. This paper investigates how digital agency can be restored by balancing seamless and seamful automation, employing data abstraction to clarify the consequences of privacy choices, and providing traceable, revisitable consent mechanisms - all while attempting to minimise consent fatigue. Utilising a mixed-methods approach, including a participatory design workshop and a controlled user study ($N = 197$), we evaluated five distinct CMP variants: *Baseline*, *Informative*, *Seamful Automation*, *Seamless Automation*, and *Revisitability*. Our results indicate that while automation reduces interaction time, black-box seamlessness can increase perceived temporal demand and user stress. Conversely, seamful designs, specifically those offering plain-language explanations and retroactive controls successfully reduced perceived effort and strengthened some aspects of perceived control, despite requiring longer objective interaction times. We conclude that HCI researchers should move beyond the "all-or-nothing" paradigm of consent, instead designing "beautiful seams" that balance automated efficiency with strategic, meaningful transparency to empower the user.

*All authors contributed equally to this research.

Authors' Contact Information: Bence Szabo, bszabo21@student.aau.dk, Aalborg University in Copenhagen, Copenhagen, Denmark; Louise Foldøy Steffens, Aalborg University in Copenhagen, Copenhagen, Denmark, lfst21@student.aau.dk; Sara Selman, Aalborg University in Copenhagen, Copenhagen, Denmark, sselma24@student.aau.dk.

ACM XXXX-XXXX/2025/12-ART
https://doi.org/XXXXXXX.XXXXXXX

The functional prototypes, source code, and research data are provided as supplementary materials.

1 Introduction and motivation

The statement "*the internet is the Wild West* [16]" fittingly captures its early, chaotic, and largely unregulated phase, a period marked by both innovation and danger. While legal landmarks and the dominance of mega-corporations have since tempered this "Wild West" era, its hazards persist, having morphed into systemic challenges like privacy breaches, misinformation, and critically, unethical data collection. Today's websites are businesses driven by an imperative to profit and grow, often translating directly into the collection and sale of user data [21]. This practice, some argue, exploits the user's willingness to sacrifice personal privacy for the promise of a personalised experience - or simply in exchange for the digital product itself.

This monetisation of user behaviour creates stark ethical problems, notably in relation to the erosion of user agency, the fundamental issue of online identity articulation, and the development of learned helplessness in managing one's personal data [10]. Public awareness of these dangers was significantly amplified by the Cambridge Analytica case, which exposed the real-life consequences of online data tracking, where highly targeted political ads sparked controversy over the fairness of elections [22]. Despite the raised awareness, the disconnect between expressed privacy concerns and actual online behavior persists, not necessarily because of a privacy paradox, but because manipulative designs and an overload of complex choices prevent users from making informed decisions. Without a correct preconceived notion of what one would consent to, there is no true consent [3].

In response to data security concerns, regulations such as the European Union's General Data Protection Regulation (GDPR [7]) and Apple's App Tracking Transparency (ATT [2]) were introduced. A core tenet of these initiatives is that online consent must be freely given, informed, specific, and unambiguous [7]. However, research shows that the resulting consent management platforms (CMPs), while designed to secure user control, are often engineered to achieve the opposite [29]. Mega-corporations have an overwhelming tendency to prioritise maximising "accept all" clicks over fulfilling the spirit of the law, creating a system where regulatory compliance often functions as a facade for continued data harvesting. However, the challenge is arguably much more elaborate for those attempting to adhere to regulations. The issue cannot be boiled down to a trivial choice between withholding or providing sufficient

information to achieve informed consent, because elaborate forms and popups can swiftly result in potential cognitive overload for end users [8].

The primary goals of this research are to address the persistent issues of user fatigue, loss of agency and the resulting learned helplessness epidemic created by current CMPs. The project aims to take a step beyond the oversimplified dichotomy of informed and uninformed decision making, approaching the issue of online consent from the perspective of user-centered consent mechanisms. The research seeks to determine what genuinely empowers users, restores their sense of autonomy and agency, and simultaneously reduces consent fatigue by exploring a spectrum of design alternatives that strategically balance automation with transparency, informativeness with non-disruptiveness, and fatigue with agency. The core objective is to identify promising design approaches that can provide key information about consent quality and decision-making in a way that is clear and manageable. Ultimately, this research seeks to transform online consent from a painstaking and confusing obligation into a clearer, more intentional user experience.

2 Related work

To properly contextualise the paper's objectives and inspire the subsequent design approaches, this section reviews the existing landscape of related work. In this literature review, we trace the path and development of online consent research, establish a fundamental understanding of what consent currently entails in an online context and how it is critically shaped by prevailing design trends and provider-side decisions.

First, we examine the core problem that has driven much of this research: the widespread misuse of consent mechanisms. Driven by persistent controversies and documented misconduct, numerous studies have documented how CMPs frequently exploit the conflict of interest to maximise data collection instead of guaranteeing user autonomy. These studies have advanced the general understanding of the problem by providing metrics that prove the presence of phenomena such as the privacy paradox and learned helplessness within the context of CMPs. Furthermore, they describe the nature of how users perceive and experience online consent.

The phenomena of privacy paradox and learned helplessness are explored in Shklovski et al.'s paper, *Leakiness and creepiness in app space: perceptions of privacy and mobile app use* [35], defining privacy paradox as situations *"where intentions and behaviors around information disclosure often radically differ"*, and learned helplessness as *"when people come to believe that a situation is unchangeable or inescapable and will often construct reasons for why this is so even if solutions become available later on"*. Shklovski et al. draw a direct analogy: a breach of online privacy is experienced similarly to a physical invasion of personal space, viewing the mobile device as an extension of the body. Building on this idea, the expansion of computerisation has resulted in a growing power disparity between individuals and large institutions. This gap actively diminishes users' sense of agency and control over their personal data and their decision-making process concerning the companies' digital goods.

These perspectives on online power imbalances and privacy breaches are highly applicable in the context of CMPs. In this environment, the power imbalance manifests directly through dark patterns and manipulative design, forcing users into decisions that contravene their intentions. The privacy paradox is thus fueled by the practical friction of the banner, where the momentary reward of content access outweighs the long-term privacy intention. Crucially, the repetitive and frustrating nature of these interactions fosters learned helplessness, where users abandon informed decision-making, viewing the banner as an inevitable and unchangeable obstacle that must simply be dismissed. This positions the CMP not as a empowering tool of personal rights, but as the final point of friction that erodes user agency.

Nouwens et al.'s paper, *Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence* [29], provides crucial evidence of these ethically questionable tactics. CMP audits reveal that common manipulative practices include nudging users toward acceptance, hiding rejection options, and presenting overwhelming vendor lists that are too large to process.

The study demonstrates that these designs actively create a privacy paradox: although CMPs are intended to grant users control, they are deliberately engineered to push acceptance. This engineering is effective: the median number of vendors presented is 315, rendering informed consent nearly impossible. Furthermore, specific deliberate tactics like removing the "opt-out" button on the first page were found to boost consent by 22–23 percentage points, while providing more detailed controls lowers consent by 8–20 percentage points.

Overall compliance is low, as the research found that only a small minority, specifically 11.8% of the scraped CMPs, met the minimal requirements based on European law [7]. Compounding this, 32.5% of sites used implied consent, counting actions like navigating, closing the pop-up or even scrolling as agreement. User feedback confirms that most accept cookies simply to proceed, reinforcing the idea that no consent equals no service.

Further results confirm that the primary driver for accepting consent is not genuine agreement but annoyance and the desire to access content. Habib et al. found that 72.7% of users accepted all cookies, with half doing so *"just to get rid of the banner"*, a behavior often described as the "Okay, whatever" effect [8]. This high rate of acceptance is part of the broader phenomenon of privacy fatigue, *"the tendency of consumers to disclose greater information over time when using more complex and less-usable privacy controls"* [23].

This behavior is reinforced by misconceptions regarding the consequence of declining consent. Utz et al. (2019) reported that 30% of users wrongly believed rejecting cookies would block access to a website, while an iOS ATT study found 43% based their choice on incorrect beliefs (e.g., assuming tracking meant location sharing)

[39]. The pervasive failure of CMPs to meet GDPR requirements suggests that structural alternatives are necessary, as an enforcement of the "privacy by default" principle would likely result in less than 0.1% of users actively consenting to third-party tracking.

To clarify, some cookies, most commonly referred to as essential cookies, can in fact be vital for a website's correct functionality [31]. The issue arises when the separation between these, and non-essential cookies is unclear or buried beneath layers of user interfaces in a CMP. In theory, cookies that should be non-essential can be rendered essential due to the way a website's code is written. This practice creates monetarily advantageous, tight, and unnecessary coupling within the system. Furthermore, there is no standard consequence for rejecting cookies or, viewed from a more abstract perspective, not providing or withdrawing online consent. Users may sometimes be rerouted to a restricted version of the website, while in other cases, access may be completely restricted. The situation is generally clouded in uncertainty, and has not been clearly settled at the regulatory level [20].

These manipulative designs, characterised by nudging toward acceptance and concealing rejection options, ultimately result in learned helplessness. This failure is significant because it starkly contrasts with the GDPR's core notion of a transparent and unambiguous decision-making process. Enforcement remains problematic, ensuring that for the end-user, online consent is still a painstaking and confusing experience.

Given this structural environment of confusion and coercion, a key emergent research direction involves moving beyond the traditional perspective on the privacy paradox. While vast research proves the discrepancy between attitude and behaviour through a risk-benefit calculus, a less-explored area focuses on cognitive biases that disrupt rational decision-making, such as users making a decision based on prevalent benefits resulting in negligible risk assessment.

"Individuals constantly try to rationally maximise benefits but decision-making can only be rational within the limits of cognitive ability and available time [3]."

The stream of literature points toward the user's psychological state as a critical, unaddressed variable, noting that repeated invasion of privacy boundaries can lead to a state of resignation, where users feel little power to change the situation anyway. This suggests a research gap in exploring not just why users disclose, but how to overcome this lack of perceived control by developing solutions that reduce mental load and combat the privacy fatigue. This aligns with the necessity to move away from systems that demand total acceptance of a privacy policy as an all-or-nothing principle inherent to app use, towards designs that restore user autonomy and reduce the necessity of non-rational decision-making.

The persistent issue of CMP fatigue, caused by the burden of site-by-site consent, has driven research into more efficient, user-centric alternatives. Zimmeck et al. proposed the principle of Generalisable Active Privacy Choice (GAPC), where users make one active, intentional choice that is applied broadly (e.g., across all sites or data categories), thus shifting the choice mechanism from manipulative website settings to the browser level [42].

The paper, *Generalisable Active Privacy Choice: Designing a Graphical User Interface for Global Privacy Control [42]*, further illustrates that GAPC can cut CMP interactions from a median of 76 per week to as few as 1 using browser-level privacy control, while still allowing users to manage their preferences. Crucially, participants stressed the need for clear feedback that their choice was applied and highly valued the ability to revisit and correct decisions. This highlights that efficiency must go hand in hand with transparency and reversibility for consent to feel trustworthy.

From a more abstract perspective, these findings highlight the need for consent mechanisms that reduce fatigue, counter misconceptions, and restore user agency. The goal is to move beyond manipulative, repetitive consent forms toward designs that balance automation with transparency and traceability, ensuring privacy choices are both intentional and informed.

Beyond the measurable non-compliance, research has explored the emotional and psychological toll of manipulative designs. Gray et al.'s paper, *End User Accounts of Dark Patterns as Felt Manipulation* [14], highlights the personal accounts and feelings provoked by these perceived dark patterns. The study utilises the concept of "felt manipulation" as a proxy, revealing that users are often aware that something is "off" or "not correct" even if they lack the specific vocabulary of dark patterns. Common reported feelings include annoyance, frustration, betrayal, distrust, and sometimes even shame or guilt if users feel they have been fooled. This general awareness of a problem, combined with the frustration and annoyance, is interpreted as a sign of the loss of agency or autonomy users experience, which eventually devolves into ignorance and helplessness. The most common recurring themes related to these perceptions of manipulation were threats of big data (e.g., use of interaction data for business goals like targeted ads), barriers to security (e.g., threats of fraud or viruses), and manipulative design in freemium products that require unexpected payment to proceed. This evolution of awareness, where something might seem acceptable at first but becomes clearly manipulative after repeated interactions, highlights the crucial need to address these design flaws by focusing on real-world user experience and emotional consequences, rather than relying solely on abstract, academic definitions of dark patterns.

To analyse the root cause of these psychological and emotional failures from a design perspective, research must examine the seamlessness/seamfulness debate in human-computer interaction (HCI). This foundational design principle, which determines the extent to which a user interface strategically hides or reveals its internal workings, is critical to understanding the current trade-offs between system efficiency and user agency.

The deceitful absence of choice and lack of transparency in online data policies is not just a result of data-mining; rather, it is a

concurrent trend in the online space known as seamlessness. Inman and Ribes' paper, *Beautiful Seams: Strategic Revelations and Concealments* [18], discusses the pitfalls of such designs, arguing that showing the "seams" of systems can empower users and aid in their decision-making processes. Revealing a "seam", such as explicitly showing where or how data is being sent or processed, can increase user agency by providing information that allows them to make better decisions. Following this paper's framework, the design goal should be to strike a balance between providing enough transparency to be empowering without overwhelming the user with unnecessary technical details, contrasting the manipulative concealment found in most CMPs with a design strategy that prioritises user agency and informed choice.

Inman and Ribes quotes Mark Weiser, further emphasising the question of traceability and agency over choices in the digital world; *"If something has been taken over for me, is there a presentation of what has been taken over that I can bring to the fore whenever I like, including retroactively?"*

The current failure of online consent mechanisms arguably stems from a design philosophy that champions seamlessness to the point of deception, prioritising a friction-free experience for the business over user understanding and autonomy for the user. The paper provides a crucial theoretical counterpoint, arguing that "seams", the visible boundaries, uncertainties, and errors in a system are not flaws but a strategic design tool.

The design framework proposed by this paper is highly relevant for addressing learned helplessness and lack of agency in consent management platforms (CMPs) [18]:

- **Context-aware evaluation of user agency**: The framework encourages designers to ask: What does the user actually want managed for them?. For privacy, this means the system's extent of automatically managing non-essential tracking (seamless automation) but strategically revealing the results of that management and the ethical quality of the site's design ("beautiful seam").
- **Ethical considerations and trust**: Pushing designers to think ethically about what information to reveal or conceal is crucial for building trust and accountability in technological systems. Revealing ethical red flags (e.g., pre-ticked boxes) as a "seam" or signal can empower users to make intentional decisions, contrasting the deceptive seamlessness that currently underpins dark patterns.
- **Empowering users through traceability**: By making the "seams" visible, designers can create technologies that are more transparent where it makes sense in a specific context. In the context of CMPs, this could mean revealing:
  - **The history or state of a system**: Making the history of consent decisions visible over time.
  - **Adaptation, reuse, and appropriation**: Revealing how the system works (e.g., showing where data is being sent) to allow users to adapt or re-purpose it.

In summary, the review of related work demonstrates a systemic breakdown in online consent, rooted in design and enforcement failures. This situation is characterised by a pervasive lack of transparency and clarity around consent forms. As Nouwens et al.'s work documents, the majority of websites use manipulative dark patterns that undermine user autonomy by nudging acceptance and overwhelming users with hundreds of vendors.

These tactics are a direct consequence of a design philosophy that promotes seamlessness to the point of deception, prioritising a friction-free experience for the business over user understanding and autonomy. This manipulative concealment ensures that for the end-user, decision-making remains a painful and confusing experience. This results in the "Okay, whatever" effect, where users accept terms out of annoyance, and fosters learned helplessness, where users feel they lack real control. Furthermore, the work by Gray et al. on "felt manipulation" confirms the psychological toll of these designs, highlighting that users are often consciously aware of being deceived, leading to feelings of distrust, frustration, and a clear loss of agency.

The solution, therefore, lies not just in stricter enforcement, but in adopting a design strategy of "beautiful seams." This approach, where designers strategically reveal system workings—such as showing the history of consent decisions or ethical red flags—can empower users through traceability and counter the deceptive lack of choice prevalent today.

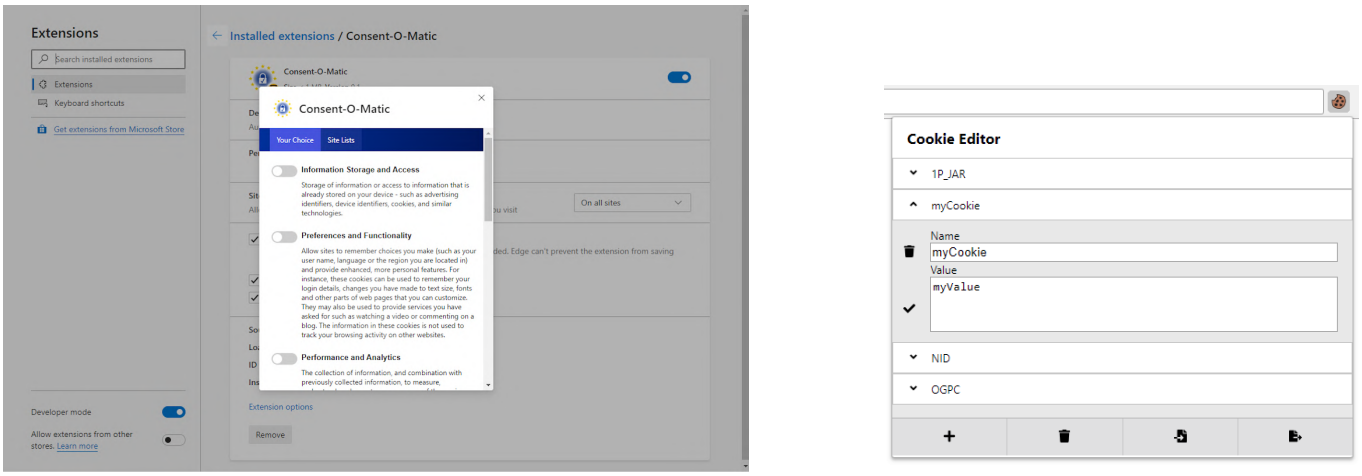
These chronic, structural failures, driven by deceptive seamlessness, highlight an urgent need to move beyond simply measuring non-compliance toward designing practical alternatives that restore user agency and reduce this significant mental burden. We now turn to reflect on the existing solutions designed to combat these documented failures.

2.1 State of the art

Existing solutions designed to combat the lack of GDPR compliance and manipulative design in online consent forms generally fall into two polarised categories, creating a trade-off between efficiency and transparency.

Automated, black-box solutions: These tools prioritise temporal efficiency by largely automating the consent process based on preset user preferences, but they sacrifice transparency. An example is Consent-O-Matic, which acts as a black-box, making case-by-case decisions for the user without providing feedback on the specific choices made or the underlying system operations [28]. This approach alleviates consent fatigue but perpetuates the user's lack of understanding and loss of agency, simply replacing manual annoyance with automated ignorance.

Informative, manual solutions: Conversely, tools like Cookie Editor offer granular, complete control over cookies [13]. These are primarily aimed at technically proficient users who desire fine-grained configuration. While highly informative and empowering



(a) Consent-O-Matic browser extension

(b) Cookie Editor browser extension

Fig. 1. Browser-based tools for managing online consent used as reference systems in this study.

for their target audience, these solutions are temporally inefficient for the average user, contributing to the very problem of information overload and disengagement that fuels the "Okay, whatever" effect.

This polarisation demonstrates a gap in the market for a solution that balances automation with genuine transparency, effectively addressing the user's learned helplessness without imposing a heavy mental load.

Despite the acknowledged flaws of current CMPs, empirical evidence is lacking on whether redesigned, user-centered consent mechanisms can simultaneously boost users' sense of autonomy and agency and reduce consent fatigue. Therefore, the aim of this study is not merely to measure efficiency, but to determine what truly empowers users, helping them regain agency against the feeling of learned helplessness and numbness towards personal privacy breaches.

2.2 Overview of our findings

The literature on online consent documents a systemic failure, where the regulatory intention of transparency and informed choice has been consistently undermined by a design philosophy prioritising business-friction reduction, often referred to as deceptive seamlessness. This approach manifests through manipulative dark patterns and information overload, leading directly to high rates of non-intentional consent and a deep-seated psychological toll characterised by privacy fatigue, learned helplessness, and conscious "felt manipulation". Existing counter-solutions are polarised: highly efficient tools sacrifice transparency for automation, while highly informative tools impose excessive cognitive burden. This gap highlights a critical need to transition away from the current paradigm toward a strategic application of "beautiful seams". The design challenge, therefore, is to empirically determine if a consent mechanism can effectively balance automation, to reduce fatigue, with traceability and strategic transparency, to restore agency and counter misconceptions. By developing and testing distinct CMP design variants, our study directly addresses this empirical gap, aiming to provide evidence-based guidelines on how to restore user autonomy against the structural coercion prevalent in the digital consent landscape.

3 Project goals and research questions

After reviewing the extensive literature on HCI design principles and the failures of current CMPs, we have identified a critical research gap. This focus directly informs our problem statement, which outlines the central issue this project seeks to address.

3.1 Problem statement

Despite GDPR's mandate for freely given, informed, and unambiguous consent, most CMPs employ dark patterns that induce user fatigue, learned helplessness, and a forced sacrifice of agency for convenience. While automation addresses the fatigue and cognitive burden, purely seamless solutions risk becoming opaque black-boxes, thereby sacrificing the traceability and understanding required for informed choice and agency. Therefore, the central problem this project addresses is the empirical failure to design a CMP solution that achieves a productive balance between these two extremes.

This study asks: *How can we build CMPs that strategically explore automation, informativeness, and revisitability to reduce cognitive load, restore informed agency, and enable user to easily verify and retroactively adjust how their consent is applied?*

3.2 Research questions

RQ1: Striking a fine balance between transparency and fatigue-reduction through automation

How can strategically seamful GDPR consent mechanisms provide automation without becoming black-boxes, ensuring that users remain informed and feel in control of their choices without experiencing fatigue?

The literature highlights a critical tension in digital privacy: achieving efficiency in consent management without sacrificing transparency and user agency. Automation has emerged as the primary solution to address consent fatigue, a phenomenon closely tied to learned helplessness, by eliminating the need for site-by-site consent - documented by Habib et al. amongst others [8]. While successful implementations, such as Nouwens et al.'s Consent-O-Matic and Zimmeck et al.'s GAPC, have significantly reduced fatigue and boosted explicit consent rates, they frequently neglect the fundamental requirement of informed consent, a key guiding principle of the GDPR [28] [42]. This sole focus on automation is proving too radical. Hutton et al.'s work on the iOS ATT feature demonstrates that users' flawed mental models lead to a critical loss of agency, while Kunze et al.'s research introduces the risks of automation-induced complacency and inappropriate trust in these systems [17] [25]. Given that these findings illustrate the inherent danger of a purely automated solution, we propose to investigate an alternative approach. Our research will explore the results of intertwining automation with context-aware, strategic revelations at points where user agency is most likely to be compromised. This method, inspired by Inman and Ribes' concept of "beautiful seams," aims to blend the efficiency of automation with targeted, timely transparency to effectively preserve informed consent and user control [18].

RQ2: Combating misconceptions and increasing comprehension by data abstraction
How does providing clear, plain-language explanations of intricate consent decisions influence users' high-level understanding and reduce misconceptions about GDPR consent, while being aware of its implications of potential consent-fatigue?

A significant body of work establishes that online consent choices, whether users choose to accept or reject data processing, are frequently rooted in fundamental misconceptions rather than informed intent [35]. This issue is compounded by the superficial nature of current consent interactions and the various reasons users cite for their compliance or defiance [28] [8]. A core driver of this lack of comprehension is the way information is presented. Acquisti et al. and Schaub et al. demonstrates that user comprehension can be significantly improved through optimised information design, specifically by employing data abstraction via short summaries, offering details on demand, and utilising clear visual cues [1] [34]. While data abstraction and plain language are proven effective for increasing comprehension, they must be integrated with the necessity of reducing user burden. Research by Zimmeck et al. confirms that users strongly desire fewer, less repetitive consent choices to combat fatigue [42]. Therefore, the central challenge for our research is determining how to optimally balance the demands of effective data abstraction techniques for increasing comprehension with necessary fatigue reduction techniques.

RQ3: Increasing agency through traceability
How does enabling users to easily revisit and adjust their consent decisions affect their sense of agency regarding CMPs and their willingness to exercise privacy-protective choices?

The perception of lost agency through lack of traceability in online consent is a critical concern, stemming from the premise that learned helplessness is a direct product of design, and can thus be reversed through intentional design choices [35]. The current online landscape is plagued by a systematic failure in consent platforms, characterised by the pervasive use of dark patterns [29]. This failure ultimately results in the "Okay, Whatever" effect, where user compliance is driven by apathy and surrender [8]. Deepening this psychological analysis, Gray et al. demonstrates that felt manipulation is central to the loss of agency, thereby underscoring the vital role of emotion in the consent experience [14]. Consequently, our approach, guided by Inman and Ribes' "beautiful seams" framework, seeks to reframe traceability as a visible design feature [18]. By deliberately designing traceability into the user experience, we can create a "seam" that allows users to clearly see which automated or previous choices were made on their behalf, enabling them to easily adjust them later. This strategy aims to actively mitigate design-induced helplessness and restore meaningful user agency.

3.3 Hypotheses

We derived a set of six testable hypotheses, two for each research question.

RQ1: Striking a fine balance between transparency and fatigue-reduction through automation

- **H1a**: CMPs utilising seamful automation (i.e., automation with explicit, persistent, context-appropriate visibility of system status) will yield significantly higher scores for perceived control among users compared to those utilising seamless (black-box) automation.
- **H1b**: Users will experience a stronger sense of agency in case of automated solutions by seeing what actions were taken on their behalf and by having options to override them.

RQ2: Combating misconceptions and increasing comprehension by data abstraction

- **H2a**: Pivoting away from highly accurate technical explanations and instead using everyday language explanations of consent categories will improve users' comprehension compared to standard CMP vendor lists.
- **H2b**: A simplified, visual CMP interface providing contextual explanations will result in a lower incidence of the privacy paradox (i.e., a smaller difference between stated privacy intentions and actual behavior) than a standard CMP design.

RQ3: Increasing agency through traceability

- **H3a**: Providing easily accessible, retro-active consent controls (traceability) will lead to a significant decrease in self-reported learned helplessness scores compared to the use of a standard, single-instance CMP banner.

- **H3b**: Users with revisitable consent options will be more willing to reject non-essential tracking than those without such options.

The core challenge remaining is designing a study that can yield valuable insights for potential design approaches, taking a step towards a balanced CMP solution, that achieves the ideal mix of automation, transparency, revisitability and minimal user burden.

The research should seek to identify promising design approaches for finding design approach that point towards a promising a middle-ground that effectively combines the temporal efficiency of automation, to counteract the "Okay, whatever" effect, with the clarity and trust-building aspects of informative design, thereby reducing the mental burden while maintaining user agency.

4 Methodology

Having established the theoretical framework, the next step is to empirically answer our research questions and validate the corresponding hypotheses. This section documents this transition from theory to experiment through a two-phase approach. We begin by outlining a qualitative, user-centered participatory design (PD) workshop. This initial step is crucial for grounding our solutions in genuine user needs. It ensures the CMP design variants are appropriate for addressing our research questions and reflect real-world user appropriations. Ultimately, the process is directly motivated by end-user pain points regarding fatigue, misconceptions, and loss of agency. The subsequent steps detail the controlled, quantitative user study used to test our hypotheses.

Disclaimer: for each one of the websites used in our studies, the first occurrence receives a citation, subsequent ones do not.

4.1 Participatory design workshop

As a preliminary step to the subsequent larger-scale user study, a PD workshop is facilitated. The main motivating points are to promote a user-focused-design approach before any design element is set in stone, broaden the plethora of design ideas for potential solutions, and facilitate meaningful discussion with a small group of potential end-users. Bødker et al. describes PD, as a design process that *"includes activities where users, designers, and researchers collaborate toward shared goals"*, and as *"the basis of a democratic concern to empower people through design and development processes where alternatives are developed to illustrate and counterbalance mainstream solutions or technologies"* [4].

The focus of this PD workshop is not to find a solution. This would not be a realistic prospect in the first place. What we are looking for is a widened spectrum of possibilities and alternative approaches - a result of meaningful discussions followed by iterative, hands-on, quick and dirty prototyping. Participants and facilitators are both encouraged to explore diverse avenues, enriching the diversity of options we had to begin with.

4.1.1 Motivation and goals. We primarily follow Clay Spinuzzi's article, *The Methodology of Participatory Design* [36] to establish the fundamental motivational factors and key principles of our participatory design workshop - aided by Hansen et al.'s paper, *How Participatory Design Works: Mechanisms and Effects* [15], and crucially Bødker et al.'s book, *Participatory Design* [4]. Although Spinuzzi's article primarily focuses on implementing participatory design within larger organisations, we have adapted its core principles to fit the practical requirements and relatively small-scale ambitions of this project.

Spinuzzi's more tangible description of an applicable methodology is deeply rooted in the Scandinavian model largely credited to Professor Susanne Bødker and her collaborators' research going back to the 1980s. As Bødker et al. describes, *"PD was developed in order to help white- and blue-collar workers in traditional industries"* [4]. This is likely the root of Spinuzzi's choice of organisational context, and the consequent need for the methods' recontextualisation for this project.

Although it requires recontextualisation from its original organisational focus, participatory design's core principles [36] make it a valuable approach for guiding contemporary product development:

Democratisation of design: Including users in design can provide valuable new, user-centric perspectives and out-of-the-box ideas, thus facilitating user-centric development. Similarly to how Inman and Ribes' "Beautiful Seams" paper [18] presents, users will appropriate products to serve their own needs, so why not have their appropriation act as a design guide?

Design that is grounded in actual practices: The way that users recontextualise an idea of a product, and their perceived usage will likely stem from personal experiences. Observing this first-hand is super valuable for understanding individual users.

A new perspective: Designers are already constrained by their preconceptions and perceived limitations for a potential product. Users do not have to consider the burden of development, and can therefore represent broader perspectives for approaching the solution.

Inspiration goes both ways: In an ideal setting, participants meet domain experts with richer understanding of the issues at hand and possibilities, inspiring for new ideas and discussion. Meanwhile, facilitators meet people with unique individual perspectives and opinions, providing a semi-representative picture of how users would envision an adept product. Bødker et al. further emphasises the importance of emancipatory practices to support mutual learning [4].

Collaboration with equal power relation: Although nearly impossible to impose, it is an important quality to strive for. Prototyping should be a process where users and designers have equal power relations. Everyone should be incentivised to participate actively and creatively, drawing on their different qualifications despite a lack of technical expertise. As Bødker et al. emphasises, participants should feel that they can make meaningful choices and possess decision-making power in the design as facilitators, in the context

of the problem and its constraints [4].

To translate these foundational principles - mutual learning, collaboration and agency - into a practical methodology, the participatory design process is typically structured around three core phases. These phases ensure a clear progression from understanding the problem to co-creating the solution:

Key objectives in a participatory design workshop [36]

- Exploration and discovery: Deep contextual inquiry, including ethnographic observation and interviews, to understand current practices.
- Design and envisioning: Collaborative brainstorming and prototyping sessions where users and designers envision future tools.
- Implementation and deployment: Testing low-commitment/fidelity prototypes in a quick mock environment and iterating based on observed use and first-hand experience.

Adapting these high-level objectives into a practical, actionable plan, the PD methodology is executed through a three-stage structure, where each phase is designed with unique mechanisms and goals to ensure a clear and productive progression from contextual understanding to artifact creation. The key activities can be seen in appendix Appendix J.

Another key step in moving from defining the PD approach to executing the workshop is acknowledging its inherent practical and logistical limitations. Successfully managing participant expectations and data representativeness requires a clear understanding of these constraints from the outset.

Pre-established challenges and limitations

- The flexibility of and level of commitment to the goals and a preconceived design from the facilitators' side is a factor that must be clearly established before the workshop. Highlighting the looseness or tightness of some design elements can be crucial factors in getting the results the team is looking for.
- While feedback, discussions, and user-side design initiatives can be open up interesting avenues for design work, it can be difficult to justify representativeness of the results due to the limited sample-size. This is an inherent limitation of how PD is designed - such a workshop is to be facilitated with a small number of future users [4].
- A participatory design workshop is a time-intensive event, requiring designers to spend significant time in the field to be confident in providing proper assistance to users who might not be domain experts.

4.1.2 Practical information. The workshop involves 8 participants aged 20-30, recruited through university and personal networks. All participants are EU residents, and are regular internet users familiar with online consent banners. The session takes place at Aalborg university in a highly controlled environment and lasting approximately 1.5 hour(s).

4.1.3 Procedure. The PD workshop follows the three-part setup based on the framework described in detail above Appendix J. However, the team will take a unique approach considering the iterations of prototyping.

After the collective discussion of personal anecdotes and common goals concludes, and crucially before prototyping begins, pairs are formed. Said tuples are either constructed based on contextual relevancy (e.g. common perspectives), self-chosen, or random. Any inner-conflicts are not treated as something to be resolved by compromise, but opportunities for sparking new design ideas.

Initially, users are not nudged in any particular direction, and their interpretation of the provided materials (e.g. UI elements) are respected. If an icon is interpreted vastly differently from its intended use, participants were encouraged to embrace their novel interpretation. Facilitator intervention only occur in subsequent iterations, or if a pair appear complacent in their design. At these points, facilitators step in to challenge and discuss the existing design, or introduce specific, unexplored avenues. These avenues (focused on automation, global settings, history/traceability, or system feedback) are tailored to the pair after a brief discussion and inspection of their current mock-up. It is important to emphasise that groups are not prompted to explore features, but goals - thus brainstorming their own feature to achieve the given goal. Although interpretations of the tools and approaches might differ, the overall, established product goals remain constant.

With each iteration, the A3 print of the website acting as the shared design space is changed out to spark new concerns. First being Zalando [41] representing commercial interactions, then Skat [9] for finances, and lastly Sundhed.dk [37] for health.

Zalando was chosen as the primary study context because Hutton and Ellis' research indicates that cookie consent and tracking permission decisions are particularly frequent and impactful in shopping and e-commerce apps [17]. This ensures the scenarios presented have high ecological validity. The other websites were intentionally selected to represent vastly different, yet sensitive contexts (finance and health), ensuring that the final design variants were tested across diverse settings where users' consent decisions are both common and consequential.

The workshop is conducted in-person primarily relying on physical materials to facilitate hands-on, quick-and-dirty prototyping and meaningful discussion.

For this study's prototyping phase, we implemented a dual-material strategy. Common UI elements, which users are expected to be familiar with, were pre-printed to facilitate rapid prototyping. Conversely, advanced or novel UI elements are not necessarily meant to make sense to participants to begin with - yet they are available from the start. This deliberate vagueness inspired participants to determine for themselves how these advanced elements should fit in and how they would practically appropriate them into their respective design. If they are overlooked, nudging pairs of users towards specific goals

to diversify the resulting designs should prompt them to make use of these materials.

The complete list of materials and talking points with corresponding justifications can be seen in appendix E.

4.1.4 Data collection. The design workshop generated qualitative data, including photographs of participant sketches, sticky notes, and audio recordings. Observations and transcripts are analysed in a thematic analysis, focusing on users' expressions of agency, control, and transparency.

5 Results of the participatory design workshop

The paragraphs below summarize our key takeaways from the data analysis. To see the raw analysis, mock-ups, along with the pre-established discussion points and questions, refer to appendices B and E.

Fig. 2. PD group work in action.

The analysis of Q1 revealed a consistent, negative user perception: participants overwhelmingly described CMPs as "annoying", "manipulative", and "effortful". They frequently cited the use of common dark patterns, specifically mentioning the concealment of "hard-to-find" "reject all" options, visually understated "accept all" buttons, and the presence of pay-or-accept walls. Furthermore, users criticised the high obstructiveness caused by overly complex flows that require granular, categorical decision-making for partial consent. Crucially, participants noted that opting for privacy-protective choices, such as selecting "only necessary" options, often resulted in perceived technical failures (e.g., videos or articles being blocked on media sites). These recurring experiences directly reinforced user fatigue and the feeling that rejecting consent "breaks" the site, thereby establishing the key challenge addressed by RQ1.

The answers to Q2 focused on the lack of clarity, revealing that participants felt critical information was routinely missing. They expressed a need to know: what specific data is being collected, the concrete purpose of collection, which third-party vendors receive it, how long it will be stored, and a clear definition of what "necessary cookies" actually means. Furthermore, they criticised the use of overly vague labels, such as the blanket term "marketing", and legalistic, opaque text. This unanimous feedback on the deficiency of information presentation speaks directly to the core of RQ2, which aims to combat user misconceptions and increase comprehension through data abstraction.

Responses to Q3 provided crucial insight into the emotional and psychological dimensions of consent decisions. The widespread act of clicking "accept all" was frequently described as a last resort to "get on with it," often accompanied by cynical humor about "selling their soul." Crucially, participants' willingness to grant consent was highly context-dependent: they were more permissive on sites they used frequently (e.g., YouTube and Twitch) but significantly more skeptical of unfamiliar ones. This indicated that decisions were often driven by a subjective judgment of a site's "vibe" or established trust, rather than an objective evaluation of its actual data practices.

Q4 focused on the issue of long-term control, showing that although participants cared about the consequences of their consent decisions, most were unable to articulate how to review or change their choices once made, often relying solely on the imprecise method of clearing browser cookies. This fundamental lack of visibility and control over past decisions strongly underscores the critical need for traceability and revisiting mechanisms, which is the central focus of RQ3.

During the prototyping phases, we noted four overarching goals that participants repeatedly returned to and referred to when discussing ideas on the whiteboard: Increase agency, combat misconceptions, reduce fatigue, and improve transparency. These goals implicitly guided how they developed and refined concepts across the three iterations. They acted as a very briefly and simply formulated assessment of the research goals of this project, prioritising easy understandability over providing a completely accurate representation of what we are going for.

In Iteration 1 (Zalando), most proposals stayed at site/CMP level. Participants focused mainly on making the existing flow less opaque and less frustrating. They suggested "more info" panels with concrete vendors and data types, icons distinguishing internal vs external cookies, and a clearer button to change consent later on the site. One group also proposed a browser extension that could explain cookies and take decisions on the user's behalf, hinting at moving some work away from repeated site-by-site interactions towards more automated handling.

In Iteration 2 (Skat), the discussion shifted towards agency and consistency. Participants suggested a common CMP layout across sites to reduce fatigue, a "safe to accept" badge for compliant sites, and moving detailed explanations into a browser-side view to keep banners simpler while still offering transparency. A browser-level "consent history" was proposed so users could revise earlier decisions without hunting through each site's settings, introducing a first concrete idea of a central place to revisit consent over time.

In Iteration 3 (Sundhed.dk), some groups moved beyond CMP tweaks to more explicit browser-level control, particularly because of the

sensitive health context. They proposed global cookie category preferences (optionally supported by AI that learns from previous choices), small real-time pop-ups when new data is collected with a short explanation, a coloured gauge or icon showing how well a site respects the user's settings, and a central consent history that lets users review and change past decisions. Here, the goals of increasing agency, reducing fatigue, combating misconceptions, and improving transparency are all visible in small fragments scattered between each one of the proposed design mock-ups.

Taken together, the three iterations highlight four recurrent design features that characterise what participants wanted from a "better" consent system and that align with our research questions:

Automation of decisions (RQ1 / Reduce fatigue): Reducing repetitive, granular clicking by letting a tool handle routine consent choices on the user's behalf, instead of "starting from zero" on every site.

System feedback (RQ2 / Combat misconceptions & Improve transparency): Providing clear, visible system status indicators about which cookies and vendors were accepted or rejected, and whether the site follows the rules.

Global settings (RQ1 & RQ3 / Reduce fatigue & Increase agency): Allowing users to express preferences once (e.g., per category) and apply them across sites, rather than only through per-site banners.

History and traceability (RQ3 / Increase agency & Improve transparency) Offering an accessible overview of past consent decisions and an easy way to revisit and adjust them later.

Overall, the workshop confirmed that current CMPs create fatigue, confusion, and learned helplessness, and that users want tools that simplify decisions, explain clearly, and make their consent traceable and adjustable.

The workshop provided a crucial foundation, confirming that current CMPs create a cycle of fatigue, confusion, and learned helplessness, the exact issues this project addresses. The findings translated directly into three key user demands: tools that simplify decisions (addressing RQ1 by minimising effort and combating resignation), tools that explain clearly (addressing RQ2 by combating misconceptions through data abstraction), and tools that make consent traceable and adjustable (addressing RQ3 by restoring long-term agency and control). These pain points directly informed the four recurrent design features, namely automation, seamfulness, informativeness, and history/traceability, which now serve as the core variables for the subsequent quantitative study. Also, we now have some promising design approaches ready for refinement for our upcoming user study.

6 Design process

For our design process we followed a iterative and user-centered approach that translated our findings from our participatory design (PD) workshop to into concrete interface alternatives. During our work, the research questions and hypotheses written on a blackboard beside us to keep the bigger picture in focus. Our course of action entailed analysing each iteration of the PD workshop design alternatives creating paper prototypes. In this phase, we were attentive of any interesting ideas that might have been lost through the participants' iterations. Next, we created medium-fidelity mock-ups on Canva, discussed and simplified these in light of the research questions, then tested them in a small focus group probing for minor usability improvements. Finally, we moved on to creating the final prototypes.

6.1 Initial mock-ups

Our prototyping began with quick paper prototypes, where we combined key findings from the PD workshop into tangible interface concepts. The paper-prototype phase was crucial for identifying which ideas were promising before investing time in crafting higher-fidelity prototypes.

We then translated the paper prototypes into medium-fidelity digital mock-ups using Canva. This stage preserved the conceptual structure of the sketches while enabling clearer visual organisation, more consistent spacing, and exploration of layout and color.

We produced four distinct digital Canva prototype variants: Seamful Automation, Seamless Automation, Informative, and Revisitability, together with the Baseline CMP mirroring empirically "average" GDPR popups based on Nouwens et al.'s paper [29]. These mock-ups can be found in appendix F. Rather than carrying all features forward, each prototype was boiled down so it only included the functionality directly relevant to one of the core themes. This deliberate simplification ensured that each experimental condition cleanly isolated a single design mechanism for empirical comparison.

Before finalising them for the user study, each variant went through a small-scale validation tests with 5 participants. All participants were fellow AAU master's students enrolled in the Software program. These ad-hoc sessions were crucial for refining accessibility, visual clarity, and interaction flow.

A recurring theme in the validation tests concerned visibility and clarity. Participants requested that buttons, text, and icons must be significantly larger and higher-contrast than typical CMPs. To address this, we increased the size of all key interface elements and adopted high-contrast colors to ensure recognisability. Also, a vivid color (neon green) was selected as the color of the "Go to the next page" button used to navigate between pages on our website containing the CMPs, because it remained the most visually distinguishable across the static page contexts used in testing. This adjustment should ensure that participants do not have any issues proceeding with our test.

Another important insight from the validation tests was the need to include fake loading screens between site transitions. Without

them, several participants did not notice that they had entered a new webpage, leading to confusion about why they were seeing a new CMP.

Together, these refinements directly shaped the visual identity and interaction logic of all four prototypes. Ensuring that colors, text, and buttons were highly visible and legible meant that comprehension did not depend on close reading or technical familiarity,

6.2 Prototypes

Based on the medium-fidelity Canva mock-ups and the insights gained from the validation sessions, we produced five final prototype variants, each designed to isolate a single design mechanism along with the Baseline CMP. Across all variants, we implemented a brief loading screen between websites to ensure that participants clearly perceived each context change, an adjustment motivated by feedback from the validation group. Every CMP variant was designed with the intent to not contain any dark patterns or nudging, using the same blue color and size for every button, and neutral wording.

Across all variants, based on the most common online consent types, we chose the following grouping of consent categories:

- Essential cookies
- Statistics and internal development
- Tracking across devices
- Ads and third party consent

6.2.1 Baseline variant. The Baseline is a standard GDPR CMP with a first layer offering "accept all", "reject all", and "more options" leading to a category/vendor view. It reflects the current standard CMP previously selected from Nouwens et al.'s paper [29], slightly simplified to enhance the contrast in the amount of information provided to the user. Its characteristics match the expectations participants described in the workshop; no automation, no global control, extremely limited explanation(s), and no obvious way to revisit or modify previous decisions. It includes no explicit automation, enhanced feedback, global settings, or history mechanisms, and is used as the reference condition.

6.2.2 Seamful Automation variant. In the Seamful Automation variant, we chose to present quantitative feedback (the number of cookies accepted or rejected) rather than detailed explanations of internal system processes. The numbers of accepted/rejected cookies are dynamically generated by an algorithm adding a semi-random number in a realistic range to the total sum. In this condition, the user is not forced through a blocking banner on any visit. Instead, cookie decisions are applied automatically in the background according to a pre-configured privacy profile. Although earlier design iterations aimed to describe system behaviour in a more elaborate manner, focus group feedback indicated that concise numerical summaries provided sufficient transparency while keeping cognitive load low. Additionally, this design decision better distinguishes this variant from the Informative variant. The CMP is primarily inspired by Groups 1 and 3 B, where participants proposed a browser extension that "manages the choice for you" at browser level, either by automation or defaults set by global settings. The design of the one-time, although revisitable, pop-up where this configuration is done was also aided by Zimmeck et al.'s design ideas [42], but strived to be a lot more straightforward and easily digestible compared to Nouwens et al.'s Consent-O-Matic solution [28]. This variant isolates fully automated, low-friction consent handling aimed at reducing fatigue and interruption (RQ1), while still providing basic feedback on system status in line with Nielsen's visibility heuristic [27].

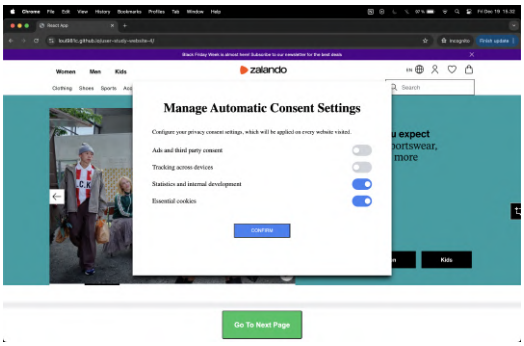


Fig. 4. Seamful automation variant (1/2)

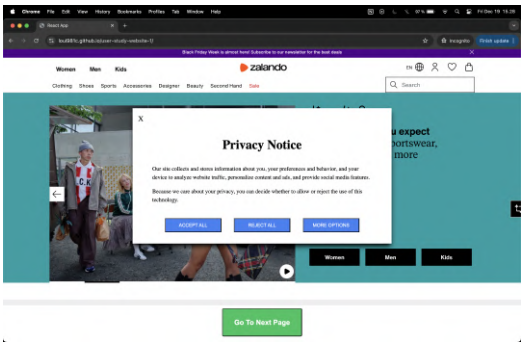


Fig. 3. Baseline CMP

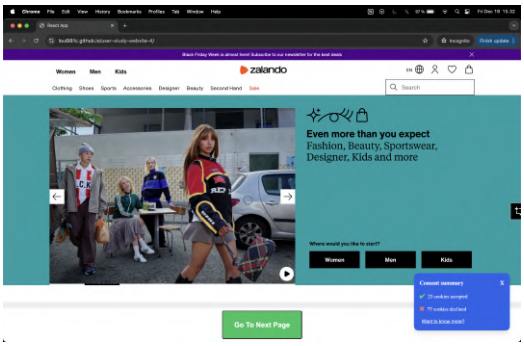


Fig. 5. Seamful automation variant (2/2)

6.2.3 Seamless Automation variant. The Seamless Automation variant presents the user with a slightly modified version of the Baseline popup once on the first website visited, emphasising the permanence of choices. For each subsequent website, no system functionality is communicated to the user.

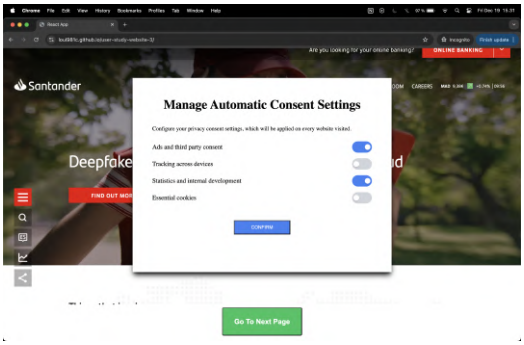


Fig. 6. Seamless automation variant

6.2.4 Informative variant. In the Informative variant, the interface was made more compact by introducing dynamic information panels that changes based based on the user's selections. This ensures that users are only presented with information relevant to the choice they are actively making, thereby reducing clutter on the main screen and maintaining a manageable information load. The presented information is site-specific, and primarily focuses on the consequences of one's choices. To further support comprehension, the explanatory text is color-coded, with warnings and potential risks highlighted in red, making it easier for users to distinguish between neutral information and content that may require closer attention. Bold text is also used to highlight the key takeaways in the brief text presented to the users. The feedback variant draws on prototypes that focused on explanation and combating misconceptions, especially Group 2, and some elements from Group 1 Iteration 2 (Fig. 2) and Group 3 Iteration 2 (Fig. 8) B.

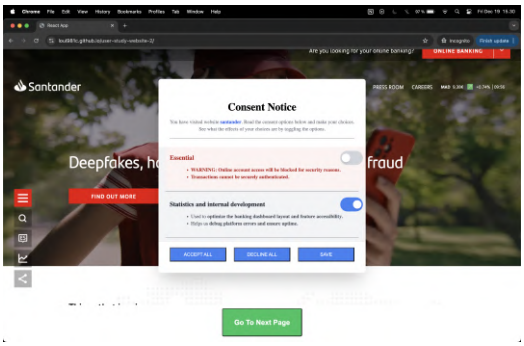


Fig. 7. Informative CMP variant

6.2.5 Revisitability variant. The Revisitability variant emphasises user agency by allowing users to easily revisit and adjust their previous consent decisions in a pop-up positioned on the right-bottom corner of the page. The design features a visually prominent red retract button below the name of each previous website, making revisitation both immediate and unmistakable. This design choice ensures that the most important functionality is visible from the outset of the survey, increasing the likelihood that participants notice and make use of the revisitation feature. It is also worth noting that the Revisitability variant still presents the user with the Baseline variant before displaying the overview of previous consent choices. This variant builds on prototypes that introduced a history-like overview and revisit mechanisms, especially Group 4, Group 2, Iteration 2 (Fig. 5) and the warning/indicator concepts from Group 3, Iteration 3 (Fig. 9 B). [27].

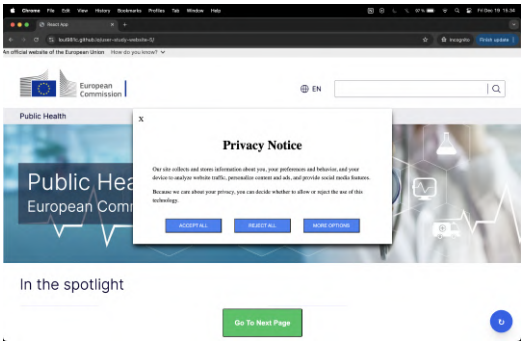


Fig. 8. Revisitability variant (1/2)

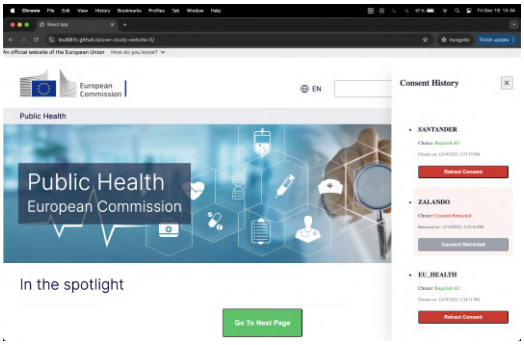


Fig. 9. Revisitability variant (2/2)

The universal UI elements present across all prototypes, namely the welcome page, artificial loading page, and the thank you page with secret UUID code can be found in appendix G.

7 System architecture

With the designs refined, verified, and ensured that they are distinct enough to answer our research questions, we are now ready to implement each variant.

7.1 Overview

To support the empirical evaluation of multiple CMP variants under controlled yet ecologically acceptable conditions, we implemented a modular web-based system composed of a React frontend and

a Supabase backend. The architecture is intentionally designed to decouple interface logic, experimental conditions, and data persistence, enabling clean isolation of the five design variables. This separation is essential for maintaining internal validity across CMP variants while allowing semi-realistic interaction flows resembling contemporary web environments of different contexts.

The core requirements for the system are the following:

- The system must be a fully deployed and stable website that, to account for the worst case, is accessible by up to 200 users concurrently. This is crucial, since performance issues could have major effects on the validity of the results. Furthermore, no expensive server-side processing should occur to ensure that none of our our free hosting and database solutions (GitHub Pages, Supabase) are producing any bottlenecks under the heavy workload.
- The front end system should by responsive to most desktop environment and window sizes.
- The system must contain an independently accessible version of the webpage for each of the CMP variant (Baseline, Informative, Seamful Automation, Seamless Automation, Revisitability).
- The websites must consist of a welcome page with a brief introduction to the task, a random series of non-interactive versions of Zalando, Santander and EU Health respectively, with one of the CMP variants correctly functioning on and in cases continuously across each one (e.g. the Revisitability variant).
- Between each page, an artificial loading animation must play for a brief period of time. This was found to be very valuable in communicating to the user that they are redirected.
- All UI elements letting the user progress with their task (e.g. I am done, take me to the next page) must be clearly visible.
- Every relevant interaction on the pages must be logged and saved in a database.

The logged data consists of the following:

- Decision time (time between CMP shown and CMP closed).
- Time on webpage (time between page render and redirect).
- Button presses (all interactions with UI elements).
- The webpage the above-mentioned data is recorded on.
- The CMP the above-mentioned data is recorded for.
- Trial indices (the order the websites are shown in).
- Session IDs (all webpages are just redirects in the same session).

See complete log list in appendix M.

7.2 Component responsibilities

7.2.1 Frontend architecture (React). The frontend is implemented using React due to its component-based architecture and support for state-driven UI rendering [32]. React allows each CMP variant to be implemented as a self-contained component with clearly defined responsibilities, while sharing a common interaction framework and visual baseline. The state-driven UI rendering also makes it easier to implement logging. Every interaction (e.g. opening a menu, clicking buttons or hovering over an info icon) updates the state, so it is much easier to hook into those updates to generate logging data.

To ensure functional consistency across the study, each CMP variant is built as a distinct component following a shared interface contract. This framework standardises the consent categories and primary actions ("accept", "reject", and "customise") for every version. Interaction logic is managed through predefined state transitions; for instance, the system tracks a user's journey from an initial "undecided" state, through the "customisation" state, and finally to a "saved" state once a choice is confirmed. CMP Variant-specific logic, such as dynamic explanations, system-status feedback, or history access, is encapsulated within each component rather than distributed across the application.

This design results in a plug-and-play CMP layer, where consent mechanisms can be swapped without altering page structure, navigation flow, or surrounding content. The architecture therefore provides the technical foundation required to isolate design differences while keeping all other aspects of the user experience constant.

7.2.2 Experimental variable: CMP Variant and Switching Mechanism. Building on this modular architecture, the CMP variant serves as the primary independent variable in the study. At session initialisation, participants are assigned to a single CMP condition, which remains fixed throughout their interaction with the simulated websites.

Variant switching is implemented at the component level, ensuring that all participants encounter the same type of websites, transitions, and contextual cues. Only the consent mechanism and sequences of websites varies between conditions to ensure contextual validity. This approach attempts to prevent unintended carry-over effects and to some extent ensures that observed differences in behavior or perception can be attributed to the CMP design.

7.2.3 Backend and data management (Supabase). Supabase is selected as the backend solution due to its lightweight integration and database capabilities, all of which aligns with the project's scope and data collection requirements [38]. Rather than functioning as a complex business logic layer, the backend primarily serves as a centralised data store for consent decisions, interaction logs, and session metadata.

Each participant interaction, such as accepting or rejecting categories, opening informational panels, or revisiting consent history, is logged as a discrete event. This event-based logging strategy supports efficient analysis of behavioral patterns while remaining neutral to the specific CMP variant in use. By maintaining a uniform logging schema across all conditions, the system ensures that quantitative comparisons between variants are methodologically sound.

Supabase also enables persistence of consent state across the simulated sequence of websites. This persistence is crucial for variants emphasising automation and revisitability, where previous decisions influence subsequent system behavior. At the same time, storing

consent state externally reinforces the conceptual separation between frontend presentation and backend traceability, mirroring real-world CMP infrastructures while remaining fully controllable in an experimental context.

7.3 Summary

Overall, the system architecture is designed to support controlled experimentation with multiple CMP designs while preserving a realistic browsing experience and ensuring that no technical difficulties occur that might plague the quality of our study. By combining a modular React frontend with a lightweight Supabase backend, the system cleanly separates interface logic, experimental conditions, and data persistence. This structure enables CMP variants to function as interchangeable components with identical capabilities but differing interaction designs, ensuring that observed differences stem from design choices rather than implementation artifacts. At the same time, centralised logging and local persistent of consent state allow reliable behavioral analysis and support features such as automation and revisitability. Together, this architecture provides a stable and flexible foundation for evaluating how different consent mechanisms affect user agency, comprehension, and fatigue.

The entire codebase can be found in the supplementary materials.

8 Methodolody (cont.)

8.1 User study

While the PD workshop provided rich qualitative insights into user needs and design preferences, a large-scale, controlled user study is required to empirically validate our hypotheses and measure the impact of specific design approaches. Addressing abstract research questions requires an approach that separates specific design features and measures how they affect both user behavior and thoughts.

To ensure legitimacy for the selected format of the user study in the context of UX surveys, we were primarily inspired by a paper and a literature review, namely Perrig et. al.'s *Measurement practices in user experience (UX) research: a systematic quantitative literature review* [30], and took a deeper look into one of its foundational papers, Law et al.'s *Understanding, Scoping and Defining User eXperience: A Survey Approach* [26].

Building on our PD workshop, Law et al. characterise workshops as a valuable qualitative method for understanding user experience, particularly when later expanded through quantitative surveys. So, continuing with our study, we strive to strike a fine balance in approaching the tension, described by Perrig et al., that entails capturing the holistic view of experiences contradicting the reductionist measurement process.

As previous research clearly suggests, standardised/validated scales can offer objectivity, reliability, and the possibility of scientific generalisation. Furthermore, seasoned Prolific users are likely familiar with some of a the standardised measures. However, blindly applying them will not yield valuable results. Therefore, we will carefully expand on these generally useful scales to better fit them to the context of our research questions. Although performing any modifications to standardised scales is frowned upon, it is a common practice. The quality of a scale is determined by *how accurately a test measures the thing which it does measure* [24]. Every project has a unique aim, so we argue that making some adjustments can aid in specialising a scale to serve our goals. To enhance validity, we exclusively extend upon established, validated scales rather than developing entirely new ones. This strategy is advantageous, as extending a validated scale is less detrimental to its validity than modifying its core content, and significantly more advisable than employing self-developed methods. When expanding on the scales, any potential ambiguity introduced by the new questions is aided with clarifications.

For the Prolific user study, our methodology must be structured to capture the holistic nature of user experience (UX) as defined by Law et al., who assert that UX is dynamic, context-dependent, and highly subjective. Critically, as highlighted by Perrig et al.'s review, most current research over-measures pragmatic facets like usability and aesthetics, leaving a significant gap in capturing the emotional and experiential facets. To address this, the study will apply a multi-method approach that triangulates objective measures capturing behavioral data (such as task success and time-on-task) with subjective self-reports.

When translating research questions into the survey format, it Perrig et al.'s recommendations (1-4) [30] acted imperative for the group. They are recontextualised for this project and are documented in appendix H.

To follow these guidelines, we implement the following specific measures and focus areas, taking major inspiration from Law et al.'s paper of UX as *Experience before, during, and after interaction* [26]:

Pre-use: Measure initial expectations, perceived quality, and sense of agency before interaction. For this step, the group decided to deviate from the recommended best practice by not having any questions. We are not looking for a within-subjects-style comparison of a baseline CMP variant against a new design, and demography data is directly available through Prolific. Furthermore, giving a randomly selected variant to each participant without a preliminary baseline ensures a clean-sheet and no order-effects.

In-use: Measure behavioral performance and real-time subjective experience. To this effect, we can record task success rate and time-on-task for primary tasks (e.g., completing checkout). Logging will be super useful when comparing behavioral and attitudinal data across design variations.

Post-use: To capture the subjective aspect of the user actions, we use extended versions of the EUQ-S and NASA-TLX to gather metrics to evaluate our research questions and corresponding hypotheses on. Originally, we looked into expanding on this to gather more interesting information, like using the NPS-inspired approach or questions based on technology acceptance models. These were rejected to minimise the time spent on the survey.

8.1.1 Study type. This study employs an online between-subjects design via Prolific, where each participant is assigned to a single experimental condition. This approach allows for a direct comparison of results across different groups without the risk of learning or order effects.

A between-subjects design was chosen to ensure that each participant's responses reflect their genuine first encounter with a CMP, rather than a learned comparison across multiple versions. This approach prevents transfer and order effects, providing more reliable data on participants' immediate understanding, perceived control, and agency [6]. It also allows for shorter study sessions, which are essential for remote, unmoderated testing on Prolific, where longer or repetitive tasks can lead to reduced attention and survey fatigue. While this design requires a larger sample and may introduce greater variability than a within-subjects design, random assignment helps maintain internal validity and minimise confounding variables that could bias group comparisons. For further justification, see appendix A.

The study aims to capture how users behave when faced with online consent popup variants, and how each variant influences their perceptions of control, clarity, and awareness, allowing for actionable insights into design factors that can restore autonomy in digital privacy choices.

This study uses behavioral logging and subjective scales to measure how specific CMP features - automation, seamfulness, data abstraction, and traceability - impact user interaction time, comprehension, and perceived agency. The results identify which design interventions most effectively restore user autonomy while minimisaing the cognitive burden of online consent.

8.1.2 Practical information. We facilitate a 10 minute long between-subjects survey with 200 participants on Prolific, meaning 40 participants per experimental condition. Eligibility criteria require participants to be 18 years or older, fluent in English, regular internet users, and EU residents only (every country must fall under the coverage of GDPR). Participants are compensated £0.5 per session (£7.20 / hr). The survey should only be accessible from a desktop computer.

Using the demography data available on Prolific and some simple python code, we have also found out the following about our sample-size:

- Nationality: The majority of our participants are from Poland (51), Portugal (24), Italy (22), Spain (20) and Greece (13).
- Age: The average age of our participants is 33,42.
- Gender: 57 of our participants were women, and 143 were men.
- Experience: Average total approvals (number of previously approved survey completions) of our participants is 946.62.

8.1.3 Procedure. In this user study, participants are tasked to interact with a randomly selected CMP variant in a random sequence of three carefully selected websites.

To evaluate our hypotheses, we manipulated the independent variable (IV), the CMP variant, across five distinct levels: the Baseline, Seamful Automation, Seamless Automation, Informative, and Revisitability. We selected three distinct contexts as our repeated measure - Zalando [41] (e-commerce), Santander [33] (finance), and EU Health [11] (healthcare) - maintaining the same rationale used in the PD workshopC. To improve international recognition for this study, we replaced the specific Danish sites used previously with these globally familiar counterparts.

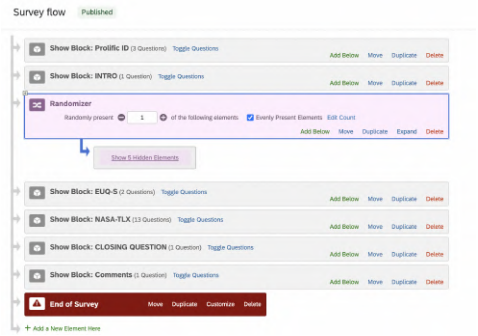
To increase ecological validity, participants interact with the system across these varied environments while behavioral data is captured via background interaction logs. To triangulate these logs with subjective experience, participants then complete an extended User Experience Questionnaire (UEQ-S) and an extended NASA Task Load Index (NASA-TLX) on Qualtrics. By building specifically upon these validated scales rather than developing entirely new ones, we maintain high psychometric validity while ensuring the metrics are sufficiently specialised to address our research questions.

To see the complete Qualtrics survey, see appendix N. To view all responses, see the supplementary materials. See the entire study flow and configuration in figure 10a.

To get an overview of how the CMP variants differ in each focus area, we have drafted a color-coded table. This table depicts the presence of automation, explanations, and revisitability in each variant shown in 10b. When designing the variants, it was also crucial to append and take away details, so that our results would prove helpful in comparing design variations and addressing our research questions and hypotheses.

8.1.4 Study contexts. The study uses a random sequence of three websites, described in the paragraph above, as the context for the online consent interface. All other website elements are static and non-interactive, ensuring that participants focus solely on the CMP. This context reflects a common online experience where consent decisions are embedded in commercial websites. The design emphasises ecological validity while allowing controlled measurement of interactions with the consent interface and feedback mechanisms.

We selected a CMP rather than a simple banner, as Nouwens et al. found that banners were ignored 3.6 times more often than CMP overlays, effectively resulting in implicit consent [29]. In accordance with GDPR requirements, stating that rejecting must be as easy as accepting, the first layer of the CMP includes both "accept All" and "reject All" options. This design choice is further supported by the finding that 93.1% of participants interacted only with the first layer, while removing the "reject all" button increased consent rates by 22–23 percentage points, demonstrating a strong nudge effect. Although only 6.9% of participants selected "more options" to customise their choices, and rarely scrolled or changed settings—the authors note that "anything requiring interaction to access might as well not exist." Nevertheless, providing that option remains important to preserve a sense of agency. Based on these findings, the standard CMP used in this study corresponds to Option B from



(a) Qualtrics study flow

↓ Design | Feature →	Automation	Explanations	Revisitability
Baseline	None	Minimal	None
Seamful Automation	Full	Limited	Minimal
Seamless Automation	Full	No	None
Informativeness	None	Full	None
Revisitability	None	Minimal	Full

(b) Feature presence across CMP variants

Fig. 10. Overview of the study structure and the distribution of features across CMP variants.

Nouwens et al.'s paper [29], as it represents the most suitable baseline for comparison with the traceable CMP prototype developed in this research.

8.1.5 Data collection methods. To provide a complete view of the user experience, we employ a mixed-methods approach that triangulates objective behavioral data with subjective self-reports.

Interaction logs: We utilise background logging to capture granular behavioural data, including click patterns, toggle interactions, visits to the consent history, and precise time-on-task and decision-time for each screen. This allows us to measure actual user performance and friction points across the different CMP variants.

Subjective surveys: Following the interaction phase, participants complete specialised versions of the NASA-TLX and UEQ-S. We adapted these validated scales to specifically target our research questions regarding cognitive load, perceived agency, and transparency.

To maximise our sample size within the established budget and minimise participant fatigue, we intentionally omitted open-ended qualitative measures other than an optional comment at the very end of the survey. This streamlined survey design reduces the total mental load, ensuring higher completion rates and more reliable quantitative data across all experimental conditions.

9 Results of the user study

Having established the experimental design and data collection protocol, this section presents the empirical findings derived from our mixed-methods analysis of $N = 197$ participants. Our objective is to evaluate how different CMP design variants impact user behaviour and subjective experience. The following subsections detail our findings as they relate to each of our core research questions: the critical tension between automation-induced fatigue and necessary transparency (RQ1); the effectiveness of data abstraction techniques in enhancing user comprehension (RQ2); and the restorative effect of explicit traceability controls on user agency (RQ3).

9.1 Quantitative data analysis

In this section, we go through the key takeaways of our analysis. To see the complete results, graphs, and the full analysis, refer to appendices K and I and the supplementary materials.

The quantitative phase of this study adopted a mixed-methods approach, utilising both subjective, self-reported survey data and objective interaction log data to evaluate the impact of the five experimental CMP designs, the independent variable, on key dimensions of user experience. This methodology was selected to provide a systematic statistical comparison of the various design principles, allowing us to address our three core research questions. By employing statistical methods on a final sample of $N = 197$ participants, this analysis moves beyond anecdotal evidence to provide empirically grounded findings on the effectiveness of various consent mechanism designs.

While 200 participants initially completed the survey via Prolific, a systematic manual review process resulted in the omission of three cases due to failed sanity checks. This data verification protocol involved verifying Prolific IDs, completion codes, and internal response consistency across the start, middle, and end of the survey. Most critically, participants were required to successfully complete a hidden code challenge, manually copying a secret code only accessible upon reaching the final website in the experimental flow to ensure complete engagement with the interactive task and the validity of the associated interaction logs.

Using the Qualtrics survey flow logic, we implemented a design supporting our between-subjects approach, ensuring that the CMP variants were randomly and equally distributed across the participant pool (≈ 40 participants per variant). This randomisation was used to mitigate selection bias and ensure a balanced sample size for each experimental condition. Furthermore, using additional randomisation on our webpage(s), we ensured that participants would each receive the sites in a random sequence, seen in figure 16, supporting internal validity by mitigating order and context related biases.

9.1.1 Methodology and metrics. The core analytical framework employed analysis of variance (ANOVA) with Tukey's HSD post-hoc testing for dependent variables (DVs) and the Chi-Square test of independence for the categorical behavioral outcome. These metrics, gathered from the post-task survey, addressed psychological states and perceived effort.

A metric commonly referred to in the following section is the *p-value*. The *p-value* tells us: If there was actually no difference between my CMP variants (the Null Hypothesis), how likely is it that I would see results at least as extreme as the ones I just measured? [5] A below 0.05 p-value suggests that the difference in user behavior (e.g., interaction time or perceived agency) is statistically significant. It is unlikely to have happened by chance, so the CMP design likely caused the effect. *P-values* above this value suggest that the results are not statistically significant. Any difference one observe might just be noise in the data from the 197 participants. We also highlight p-values that are close to significant because of our limited sample size.

In addition to the *p-value*, we report the *mean difference* (ΔM) for our pairwise comparisons. While the *p-value* indicates the likelihood that a result is not due to chance, the *mean difference* quantifies the magnitude and direction of the effect. It is calculated by subtracting the average score of one group from another ($\Delta M = \bar{x}_1 - \bar{x}_2$) [12]. In this study, a positive ΔM indicates that a specific design variant (e.g., Seamful Automation) outperformed the comparison group (e.g., Baseline) on a given scale. Reporting the *mean difference* is essential because it allows us to evaluate the practical significance of our results, showing not just if a design changed user perception, but by how much.

Ordinal dependent variables (DVs)

User burden and fatigue (RQ1): This was operationalised using the NASA-TLX subscales, specifically temporal demand and effort, which directly quantify the cognitive load imposed by the consent mechanism. The objective time-on-task (captured from log data) served, although not exclusively, as a direct behavioral measure of this burden. We observed that decision time in the seamless designs, seen in figure 18, tended to be higher than the seamful designs, seen in figure 19, as these conditions required an initial setup step. However, overall session duration, seen in figure 20, systematically increased across the designs that introduced more seamful elements (e.g., the Seamful Automation, Revisitability, and Informative variants) compared to the Baseline and Seamless Automation designs. This highlights the inherent trade-off: greater transparency costs more time. The overall ANOVA found marginal significance for both temporal demand ($P = 0.0585$) and effort ($P = 0.0502$). Interestingly, the Seamless Automation showed a trend in increased perceived temporal demand compared to Baseline ($\Delta M = 1.41, p = 0.0687$) and the Revisitability design ($\Delta M = 1.31, p = 0.0947$), suggesting that the lack of transparency in the automated process led to a higher perception of time pressure or "hurriedness". This may suggest that the ambiguity of the Seamless Automation design, which offers no context or results, can be stressful for users, leading to a perceived rush or time pressure.

Analysis of perceived effort revealed a slight trend suggesting that the Informative design required less effort from users compared to both the Seamless Automation ($\Delta M = -1.06, P_{\text{adj}} = 0.1038$) and the Baseline ($\Delta M = -1.05, P_{\text{adj}} = 0.1206$) designs. While these comparisons sit just outside the traditional 0.10 threshold for marginal significance, the consistent reduction of over 10 point on the NASA-TLX scale suggests that providing plain-language translations may lower the perceived work required to navigate consent choices. This indicates that increasing the quality of information and decresing the responsibilities managed for a user, even if it lengthens the objective time-on-task, may successfully reduce the subjective effort required to process the choice, providing a key insight into mitigating fatigue.

Agency and Control (RQ1, RQ3): Assessed through self-reported perceived agency (Q19), perceived control (Q20), and the specific traceability metric, perceived revisitability (Q24). These metrics were selected to directly operationalise the psychological state of felt manipulation and learned helplessness detailed in the literature, which RQ1 and RQ3 seek to counteract. Notably, while the general metrics of perceived agency and perceived control did not show significant differences, the specific design element targeting revisitability proved effective. The Revisitability design, which explicitly offered retro-active consent controls, resulted in a marginally significant increase in perceived revisitability ($\Delta M = 1.77, P_{\text{adj}} = 0.054$) compared to the Baseline, and had the highest overall score for revisitability among all five variants. This, to a certain degree of doubt, suggests that the inclusion of a revisit mechanism successfully increased the users' sense of future autonomy over their data. This directly supports the goal of RQ3 by confirming that intentional design can restore the perception of choice flexibility. Beyond the survey metrics, the log data for the Revisitability design provided strong behavioral evidence of exercised agency: this variant showed a clear shift in user interaction patterns, specifically exhibiting lower implicit consent rates. Revisit had about 12% of all sessions with implicit consent, displayed in figure 17, and thus higher rates of explicit actions compared to other designs. Furthermore, approximately 18% of participants revisited their choices, and 67% of those revisits resulted in the retraction of previously accepted consent, demonstrating that the design enables users to actively overturn permissive decisions.

Comprehension (RQ2): Measured by survey responses for understandability (Q21), confidence in reasons (Q22), and understanding of consequences (Q23). These metrics directly addressed RQ2's focus on whether data abstraction and plain-language explanations could effectively enhance the requirements of informed consent. These were chosen over a performance-based quiz to capture the user's subjective confidence and cognitive grasp, which are crucial components of their mental model of the system. The consistent non-significance across all three comprehension DVs ($P > 0.30$) indicates that simply simplifying the language and providing details-on-demand within the CMP's first layer was insufficient to yield a statistically significant improvement in self-reported understanding when compared against the other designs. This highlights the persistent challenge of addressing the fundamental misconceptions

that plague consent decisions, regardless of the clarity of the presentation.

Categorical dependent variables (DV)
The user's self-reported final consent choice (Q25) was used to test for shifts in privacy-protective behavior (H2b, H3b). To ensure the validity of the Chi-Square test and maintain adequate expected cell counts, the raw responses were consolidated into three primary groups: "accept all", "accept necessary", and "reject/custom" (merging "reject all," "manually choose," and "close the banner").

Main findings and hypothesis testing
The statistical analysis revealed significant and marginal effects in specific areas related to user effort and the perception of agency, while yielding largely non-significant results for comprehension and final consent behavior.

RQ1: Striking a fine balance between transparency and fatigue-reduction through automation: Our findings offer nuanced, yet incomplete support for the hypothesis that seamful automation can mitigate fatigue. Our findings offer nuanced support for the hypothesis that seamful automation can mitigate fatigue. One-way ANOVA tests revealed marginally significant effects of design variant on temporal demand ($F(4, 192) = 2.32, P = 0.059$) and Perceived Effort ($F(4, 192) = 2.42, P = 0.050$). Notably, post-hoc analysis showed that the Informative design trended toward lower effort scores compared to both the Seamless Automation ($\Delta M = -1.06, P_{\text{adj}} = 0.104$) and the Baseline ($\Delta M = -1.05, P_{\text{adj}} = 0.121$). This suggests that while automation is intended to reduce burden, providing plain-language information may be more effective at lowering the actual cognitive work required for consent management than full automation alone. The post-hoc analysis for temporal demand revealed a trend where the Seamless Automation design was perceived as more hurried than the standard Baseline (mean difference = 1.41, $P_{\text{adj}} = 0.0687$).

However, the core hypotheses about restoring user agency were not supported. Neither perceived control ($F(4, 192) = 0.34, P = 0.851$) nor perceived agency ($F(4, 192) = 0.70, P = 0.595$) showed significant differences across the designs. The very low F-statistics (both < 1.0) indicate that the variance between the design variants was minimal, suggesting that these interventions were unsuccessful in significantly enhancing the general feeling of control or agency compared to a standard Baseline.

While the core hypotheses H1a (perceived control) and H1b (perceived agency) were not supported by the general metrics, two major trends emerged regarding user burden.

Objective time cost vs. subjective effort: The trend on NASA_TLX_Effort suggests the Informative variant was perceived as requiring less effort than the Seamless Automation and Baseline designs. This indicates a potential path to achieving transparency without increasing psychological burden, even if the objective time-on-task is longer.

Stress of seamlessness: The trend in temporal demand highlights that the Automation designs induced a higher perceived hurry/rush compared to the Baseline and Revisitability designs. This may suggest that users find the black-box nature of the Seamful Automation design stressful, demonstrating a failure of black-box automation to fully address fatigue.

Finally, the overall ANOVA for NASA_TLX_Performance ($F(4, 192) = 3.19, P = 0.014$) showed a significant difference in how successful participants felt in completing the task. Post-hoc comparisons revealed a potential trend: the Seamful Automation design led to higher perceived performance than both the Baseline ($\Delta M = 1.07, P_{\text{adj}} = 0.077$) and Informative ($\Delta M = 1.07, P_{\text{adj}} = 0.0668$) designs. This suggests that providing explicit feedback on the system's actions contributes positively to user confidence. These results are further aided by multiple participants expressing their confusion surrounding the Seamless Automation design's correct functionality.

RQ2: Combating misconceptions and increasing comprehension by data abstraction: The data provided no statistical support for the effectiveness of the design interventions on user comprehension. H2a and H2b were not supported, as the design variant had no significant effect on understandability ($F(4, 192) = 0.96, P = 0.430$), confidence in reasons ($F(4, 192) = 0.95, P = 0.437$), or understanding of effects ($F(4, 192) = 1.17, P = 0.323$). These low F-statistics indicate that participants' perceived comprehension remained largely uniform across all conditions. This suggests that the implemented data abstraction techniques were insufficient to produce a measurable increase in user comprehension compared to the Baseline or Automation variants. Although, the consent choices for Informative variant were more scattered than the other designs, which proposes an interesting new discussion point.

RQ3: Increasing agency through traceability: This research question yielded the most compelling evidence for the potential of seamful design. H3a received partial support from a marginally significant finding for perceived revisitability ($F(4, 192) = 2.21, P = 0.070$). Specifically, the post-hoc comparison revealed that the Revisitability variant resulted in a notable increase in perceived revisitability compared to the Baseline ($\Delta M = 1.77, P_{\text{adj}} = 0.054$). This suggests that explicitly communicating the ability to change consent choices later significantly bolsters users' sense of long-term autonomy, even if their immediate agency remains unchanged. This result demonstrates that the deliberate introduction of a "revisitable seam" successfully enhanced the users' perception of their ability to retroactively adjust their consent decisions.

On the other hand, the behavioral hypothesis H3b was not supported. The Chi-Square test on final consent choice found no significant association between the design variant and the distribution of final consent decisions ($\chi^2(8, N = 197) = 10.28, P = 0.246$). While the Revisitability design successfully increased users' perceived empowerment (supporting H3a), this psychological perception did not translate into a statistically significant shift in their actual behavior regarding the rejection of non-essential tracking. This suggests a perception-behavior gap, where feeling empowered to change a choice later does not necessarily change the immediate choice itself.

9.2 Qualitative data analysis

After having completed the survey, users were met with a final freeform question where they had the freedom to share any comment they found relevant by answering the question; *"Would you like to share any other comments about cookie popups?"*.

Out of the 197 valid participants, 93 chose to answer this optional closing question. Out of these 93 answers, we removed those that essentially meant "no comment", resulting in 72 answers. We grouped the raw text into three categories through an inductive thematic analysis, meaning that we let themes and patterns emerge organically from the data itself, without relying on pre-existing theories or frameworks:

- Behaviour-related responses (17)
- General opinions on CMPS (34)
  - Positive (0)
  - Negative (27)
  - Neutral (7)
- Opinions on our CMPs (21)
  - Baseline (5)
  - Informative (5)
  - Seamful Automation (5)
  - Seamless Automation (3)
  - Revisitability (3)

Remark: We choose the one most fitting category for each answer, despite some being eligible for multiple. To see all answers with author comments above, see L.

The comments offer valuable insights into the participants' experiences, confirming issues targeted by our research questions and providing qualitative evidence for our hypotheses. We have chosen to group the comments to directly address the core themes of our research questions.

To see the complete inductive thematic analysis, refer to appendix L. To view the dataset in its entirety containing all answers, refer to the supplementary materials.

9.2.1 RQ1: Striking a fine balance between transparency and fatigue-reduction through automation. The users' general lack of trust in all CMPs, and the contrast in responses regarding seamful and seamless automation reinforces the necessity of the strategic, seamful design in H1a and H1b to restore faith and provide the necessary feedback/control to counter the black-box feeling.

Users explicitly expressed their appreciation for the concept and functionality of seamful automation: *"I wish there was a system (browser extension?) that would [do what the Seamful Automation CMP variant does]."*. The simple system status visibility was valued for its clarity: *"I liked the blue box explaining briefly what I had accepted. It was clear and concise."* The positive feedback regarding the simple, visible system status indicator, (the *"blue box"*), suggests explicit feedback successfully increased perceived control, validating the seamful approach (H1a).

The Seamless Automation variant led to confusion and doubt, some participants even doubting the system's correct functionality: *"[...] I am confused with that one in this study,"* and *"Maybe I did my part wrong [...]"*. One user reported the CMP seemingly not working: *"Only showed up on the first page..."*. The confusion, doubt, and lacking control when the system acted without visible indication supports the idea that transparency is essential for agency. The seamless approach failed to provide the necessary visibility (H1b).

9.2.2 RQ2: Combating misconceptions and increasing comprehension by data abstraction. The contrast in comments related to the Informative CMP variant against the Baseline's lack of clarity offer direct evidence regarding the importance of clear and concise explanations and their impact on user comprehension. Users generally distrust current systems: *"I don't 100% trust the website"* and are unsure of the consequences: *"I am not sure how my choices will affect the usability/functionality of the sites."*

The general negative comments often underline dark patterns like burying refusal and defaulting to "accept". While providing explanations can increase the textual load, the high user appreciation for the clarity and ease of use suggests the benefit to comprehension outweighs the potential for fatigue, provided the information is concise and dynamically presented, as in our Informative variant.

When confronted with the Informative variant, users highly appreciated the plain-language explanations being responsively updated based on the user's current choices: *"I liked the explanation of the obstructive/innovative etc. terms; it made it easier to understand what exactly am I rating."* and *"I liked that the popups in this survey switched between explanations depending on if the cookies are allowed or not, this made it easy to understand what each slider would do."* The brief and simple-language explanations stating the potential consequences of one's action on each platform specifically made it easier for some individuals to understand directly support the hypothesis that plain-language explanations improve user comprehension compared to standard vendor lists/terms (at least in self-reported data) (H2a).

In non-informative variants, users expressed clear frustration over the lack of information regarding the consequences of their choices: *"[...] I don't actually know what happened: were the cookies accepted or rejected?"*. This highlights the lack of context and feedback in standard designs. The other variants' (Baseline, Seamless Automation, and Revisitability) lack of information contributes to users' misconceptions and uncertainty about the outcome, reinforcing the need for the clarity provided by the Informative design (H2a, H2b).

9.2.3 RQ3: Increasing agency through traceability. This RQ is supported by comments on the Revisitability variant, general feelings of helplessness, and ingrained behaviour.

Neutral answers emphasise the necessity of choice: *"we (the users) must always have the choice to reject them if we want to."* The desire to reject loosely reinforces the idea that if the mechanism, revisitability, and the explanation (RQ2) are clear, users want to exercise

privacy-protective choices, validating H3b. However, nobody explicitly commented on the potential of making retroactive decisions.

Multiple comments express resignation and habitual acceptance: *"I am used to seeing cookie pop ups and I almost always accept all cookies. It has become a habitual action."* and *"My data will be stolen anyway."* This pervasive sense of learned helplessness underscores the potential motivation behind urgent a mechanism, such as revisitability, to restore agency and counteract the feeling of fatalism ("data will be stolen anyway") (H3a).

Although limited, the comments show a varied response, including some distrust: *"I do not trust [the Revisitability CMP variant] in your study"*. While one user showed distrust, the design goal is to provide a mechanism to counteract the feeling of being trapped by the initial choice. The quantitative data is more crucial here, but qualitatively, the prevalence of learned helplessness strongly suggests that easy revisitability is a necessary tool to empower users to re-engage and exercise protective choices (H3a, H3b).

9.3 Discussion of results

The quantitative and qualitative analyses reveal a pattern of selective success. The primary success of this study lies in the ability of explicit, seamful design features to restore specific psychological dimensions of control and reduce perceived workload. However, core efforts to boost general comprehension and shift overall self-reported behavioral intent remained largely unsuccessful.

The most compelling finding stems from RQ3. The hypothesis that enabling users to easily revisit their consent choices (H3a) would increase their sense of agency was partially supported by a marginally significant increase in perceived revisitability compared to the Baseline ($\Delta M = 1.77, P_{\text{adj}} = 0.054$). This result confirms that deliberately introducing a "revisitable seam" is a potent mechanism for combating the learned helplessness prevalent in the qualitative data, where users described habitual acceptance and a sense of resignation. This is reinforced by the behavioral log data, which shows a non-trivial rate of users actively revisiting their decisions, with a significant majority of those revisits resulting in the retraction of initially accepted consent as seen in figure 21. Conversely, the non-significant result for H3b (final consent choice, $P = 0.246$) indicates a perception-behavior gap: the design increased the feeling of choice flexibility without immediately altering the initial self-reported tendency to reject tracking. This suggests the value of revisitability lies in long-term autonomy rather than immediate behavioral shifts.

In addressing RQ1, the data highlights that automation is not a uniform solution for user burden. While all seamful designs increased objective time-on-task (e.g., Seamful Automation at $\sim 70s$ vs. Baseline at $\sim 58s$), the black-box Seamless Automation design was associated with a marginal trend of higher perceived temporal demand (hurry/rush) compared to the Baseline ($P_{\text{adj}} = 0.069$). This supports the notion that a lack of transparency can generate user stress. In contrast, the Informative design showed a trend toward reducing perceived effort compared to both Seamless Automation ($P_{\text{adj}} = 0.104$) and the Baseline ($P_{\text{adj}} = 0.121$). This suggests that strategic investment in high-quality information (a "meaningful seam") can reduce psychological burden even if the objective task duration is longer. Furthermore, the overall significant effect on NASA-TLX Performance ($F(4, 192) = 3.19, P = 0.014$) revealed that using the Seamful Automation variant led to higher perceived task success ($P_{\text{adj}} = 0.077$ vs. Baseline). This emphasises that system status visibility is vital for user confidence. The failure of H1a (perceived control) and H1b (perceived agency) across general metrics reinforces the specificity of these effects; design interventions appear to impact focused psychological constructs, like revisitability or performance confidence, rather than broad, general feelings of overall control.

Finally, the findings for RQ2 were the most limited, as all metrics for understandability, confidence, and effect-awareness were non-significant ($F < 1.17, P > 0.32$). This indicates that the implemented plain-language abstraction was insufficient to produce a measurable change in self-assessed comprehension. Additionally, while interaction logs showed that participants spent significantly more time making consent choices, both in initial decision-making and total session duration, this objective increase in engagement did not result in a corresponding improvement in self-reported understanding. This quantitative result stands in contrast to the qualitative feedback, where users explicitly appreciated the system status indicator "blue box" and responsive explanations. This discrepancy may suggest a ceiling effect in self-report metrics or that the inherent complexity of data privacy transcends simple text-based interventions. Consequently, the failure of H2b is expected, as a significant shift in behavior likely requires a more fundamental change in comprehension than what was achieved here.

10 Closing discussion

This discussion synthesises the findings from our mixed-methods study to evaluate how alternative CMP designs, moving away from the "all-or-nothing" paradigm of traditional banners, can reshape the digital consent landscape.

10.1 Reflecting on Research Questions and Hypotheses

Our results suggest that a seamful approach, which reveals the inner workings of automation, is a viable path toward reducing user fatigue (RQ1) without necessarily sacrificing agency. While the Seamless Automation variant resulted in the lowest interaction time, it was associated with a marginal trend toward higher perceived temporal demand, suggesting a black-box stress effect. In contrast, the Seamful Automation variant maintained higher perceived task success. This supports our hypothesis that providing feedback on automated choices mitigates the learned helplessness often associated with background privacy tools by providing necessary closure.

The findings extend the theory of "Beautiful Seams" into the domain of automated privacy. While Inman and Ribes' work [18] suggests that making infrastructure visible (seamfulness) can foster understanding, we specifically identified that in the context of CMPs, seamfulness acts as a psychological bridge. While previous work on

tools like Consent-O-Matic focused on the technical efficiency of seamless removal of banners, our study demonstrates that moving too far toward seamlessness creates a black-box stress that actually increases temporal demand. This suggests a new design principle for privacy tools: automation should not be invisible, but rather traceable to maintain user agency.

Regarding comprehension and misconceptions (RQ2), the quantitative metrics for understandability and confidence did not reach statistical significance. However, qualitative feedback suggests that the Informative Variant was highly valued for its plain-language explanations. This discrepancy highlights a comprehension-perception gap: while users felt more informed and appreciated the clarity, this did not translate into a measurable shift in their objective self-assessment of the system's effects. This suggests that while plain-language abstraction improves the experience of consenting, it may not be sufficient to overcome the deep-seated mental models users hold regarding data tracking.

This comprehension-perception gap offers a new perspective on the concept of felt manipulation described by Gray et al. [14]. Their work focuses on how dark patterns strip away agency; our results suggest that even when these patterns are removed and replaced with user-centered design (e.g. dynamic explanations), the psychological residue of previous manipulative experiences remains. We found that users feel more informed, yet their objective mental models of tracking remain unchanged. This implies that design interventions cannot exist in a vacuum; they are fighting against a baseline of learned cynicism created by years of exposure to deceptive CMPs.

Finally, for revisitability (RQ3), our findings indicate that users value the ability to revisit decisions. The Revisitability Variant acted as a psychological safety net that strengthened perceived long-term control, even when the history logs remained unaccessed. This confirms that the mere presence of a "revisitable seam" can restore a sense of autonomy that traditional, ephemeral banners destroy.

This psychological safety net effect nuances the findings of the Leakiness and Creepiness paper [35] regarding privacy resignation. Shklovski et al. argue that users give up because they feel they have no choice. Our study shows that providing a "revisitable seam", even if it is not necessarily clicked, can reverse this resignation to a certain extent. Unlike traditional all-or-nothing banners that force a permanent decision, the mere visibility of a history log signals to the user that their agency is persistent rather than ephemeral. This suggests that perceived control is often as important as actual use in restoring digital agency.

10.2 Discussion of the execution of our study procedures

10.2.1 Participatory design workshop. The selection of users for the PD workshop was grounded in the principle of democratisation of design. In practice, this process faced challenges; power relations did not always equalise as planned due to the inherent complexity of online privacy. While PD provided rich qualitative context, the difficulty for non-experts to imagine entirely new approaches without being mentally constrained by existing legal frameworks remains a limitation. Our choice to go with PD over co-design to prioritise personal narratives, which proved essential in identifying the learned helplessness that our Seamful and Revisitability variants eventually targeted.

10.2.2 User study design choices. We intentionally omitted a "pre-use" phase to avoid priming participants and to prevent the "testing effect", where answering general privacy questions might have influenced their subsequent behavior. While a partial within-subjects study comparing preconceptions with post-use opinions might have been interesting, we determined it would not significantly aid in answering our specific RQs regarding design interventions. Also, it would have been another factor that would have potentially influenced both the behaviour of our participants and their answers to the subsequent questions.

Our study adopted the ISO 9241-210 [19] definition of UX as a person's perceptions resulting from use. However, we acknowledge that our study does not fully capture the long-term, hedonic aspects of UX. Given that UX is temporally expanded, our choice of three distinct website environments was a necessary step toward a contextualised understanding, though the short-term nature of the tasks remains a limiting factor.

10.3 Reflections on nudging and dark patterns in our CMP variants

Our study attempted to exclusively compare dark-pattern-free CMPs to document how a GDPR approved baseline solution compares to our unique approaches.

Although variants were designed to exclude dark patterns, unintended nudging may have persisted. The left-to-right order of buttons and the rejection defaults stand in stark contrast to the deceptive CMPs users encounter daily. This lack of obstruction might skew results; users conditioned to "click to bypass frustration" in the real world may have found our "best-case" baseline surprisingly easy to navigate. Additionally, to sharpen the experimental contrast, we removed informative elements from all designs except the Informative variant. While this highlighted the impact of our interventions, it arguably borders on information withholding. This design choice underscores the fundamental HCI challenge: how to provide enough clarity for the user to stay engaged without sacrificing the comprehensive disclosure required by regulations.

10.4 Limitations of the results

Although the experimental design was structured to isolate the effects of the CMP variants, the following limitations should be considered when interpreting the results.

Internal validity: We cannot fully guarantee that participants read every explanation in the Informative Variant. Self-report bias may also lead participants to answer as "socially desirable" privacy-conscious users rather than reflecting their true internal state.

Ecological validity: Using static websites prevents us from seeing

how users react to "pay-or-accept" walls or dynamic content that might change their time-vs-privacy valuation. Additionally, as a remote study, we cannot observe the physical context of the user, potentially missing external distractions or stressors.

External Validity: The study was conducted within a simulated environment grounded in the GDPR legal framework. Consequently, the findings may not fully generalize to jurisdictions with different privacy standards (e.g., the US or China).

Automation perception: Since this was a short-term study, the Automation variants may have been perceived more positively than it would be in a longitudinal setting, where "automation surprise" or shifts in preference over time might occur.

Sample demographics: The Prolific sample was skewed toward "professional" survey-takers (average approvals > 900). These users are likely more digitally literate than the general population.

Statistical Test Assumptions: Although ANOVA is frequently used in HCI research with large sample sizes due to its robustness against non-normality, the violation of Levene's test for the NASA-TLX Performance metric suggests that a non-parametric alternative, such as the Kruskal-Wallis H Test, may be more appropriate [40]. ANOVA assumes that the sample is normally distributed, and ours is not, because the data comes from ordinal scales. A re-analysis using a Kruskal-Wallis test would be a valuable step to take. Kruskal-Wallis does not assume a normal distribution or homogeneity of variance, making it better suited for the ordinal nature of our UX metrics. A post-hoc sensitivity analysis using non-parametric tests is a necessary next step to confirm whether the significant differences observed in our ANOVA results remain stable or if the violation of variance assumptions has led to a Type I error (false positive).

Levene's test violation: The violation of homogeneity of variance in the NASA-TLX Performance metric suggests that the spread of success-perception was unequal across groups, requiring a cautious interpretation of the significant ANOVA result.

Omitted Global Settings variant: We originally planned a Global Settings variant that would apply pop-up defaults based on a one-time preference profile. This was omitted to ensure a larger sample size across fewer variants. However, its exclusion limits our ability to distinguish between the effects of context-aware automation (Seamful/Seamless variants) and preemptive automation (Global Settings). Understanding whether user fatigue is better mitigated by a one-time "set-and-forget" profile versus dynamic, on-page automation remains a critical gap for future comparative research.

10.5 Future work

A primary area for future investigation is the long-term habituation to "seams." While our findings indicate that users currently appreciate the "blue box" system status indicator and responsive explanations, it remains unclear if this strategic transparency eventually leads to notification fatigue. Furthermore, although our study highlighted a perception-behavior gap in the short term, a multi-week study could reveal if persistent exposure to explicit revisitability controls eventually shifts actual privacy behavior. Overall, a multi-week longitudinal study is required to determine the point at which a "meaningful seam" devolves into a background nuisance, potentially triggering a secondary level of learned helplessness.

Second, we propose the development of hybrid agency models that bridge the gap between the omitted "Global Settings" and the context-aware "Seamful Automation" evaluated here. Given that participants expressed a desire for browser-level controls during our participatory design workshop, future iterations should explore how global privacy profiles can be dynamically "surfaced" only when a website's specific data practices deviate significantly from a user's established preference baseline. This would address the comprehension-perception gap by focusing user attention solely on high-stakes or anomalous data requests.

Finally, we identify AI-driven "agency proxies" as a critical frontier. Building on our Seamful Automation results, which showed higher perceived task success when system status was visible, future work should evaluate whether Large Language Models (LLMs) can act as intermediaries that translate complex legal vendor lists into the plain-language abstractions our participants found valuable. The core research question remains: can we maintain the psychological "safety net" of revisitability in a mobile-first environment, where reduced screen real estate demands even higher levels of automated efficiency without sacrificing the user's ultimate veto power?

11 Conclusion

Our study demonstrates that the current erosion of digital agency is not an inevitable byproduct of complex data regulations, but rather a consequence of specific design choices that prioritise friction reduction over user autonomy. By evaluating a palette of CMP alternatives, we established that automation in privacy contexts fundamentally requires transparency. Strategic revelations of system operations are essential to prevent automated tools from becoming opaque black-boxes that induce stress and leave users feeling powerless.

Furthermore, while our quantitative data suggests that plain-language abstraction alone may not immediately shift objective comprehension scores, our qualitative findings indicate it is highly valued by users. Participants proved capable of engaging with nuanced privacy concepts when technical language was replaced with abstracted explanations of consequences. We also found that traceability significantly aids agency; the presence of an explicit revisitability mechanism acts as a psychological safety net that restores perceived long-term control, even when not actively accessed.

Ultimately, restoring agency in the digital landscape requires moving away from deceptive seamlessness toward a strategy of "beautiful seams", designing interactions that prompt reflection without causing privacy fatigue. As automated decision-making becomes more

integrated into the digital experience, these principles offer a design framework for ensuring that efficiency does not come at the cost of transparency, ultimately fostering a more intentional and respectful relationship between users and the consent management systems they navigate.

Acknowledgments

We would like to thank our supervisors Carla Florencia Griggio for her support and input throughout the project. We also extend our thanks to all participants in the user study and focus group for their time and valuable contributions to the evaluation of our work.

References

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang, and Shomir Wilson. 2017. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *ACM Comput. Surv.* 50, 3, Article 44 (aug 2017), 40 pages. doi:10.1145/3054926 Accessed: 2025-12-19.

[2] Apple Inc. 2024. *App Tracking Transparency.* Apple Developer Documentation. https://developer.apple.com/documentation/apptrackingtransparency Accessed: 2025-12-19.

[3] Susanne Barth and Menno D. T. de Jong. 2017. The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics* 34, 7 (2017), 1038–1058. doi:10.1016/j.tele.2017.04.013 Accessed: 2025-12-19.

[4] Susanne Bødker, Christian Dindler, Ole Sejer Iversen, and Rachel Charlotte Smith. 2022. *Participatory Design.* Springer International Publishing, Cham, Switzerland. doi:10.1007/978-3-031-02235-7 Accessed: 2025-12-19.

[5] Iain E. Buchan. 2024. StatsDirect Statistical Software: Analysis of Variance (ANOVA). https://www.statsdirect.com/help/analysis_of_variance/anova.htm. Accessed: 2024-12-19.

[6] Raluca Budiu. 2018. Between-Subjects vs. Within-Subjects Study Design. Nielsen Norman Group. https://www.nngroup.com/articles/between-within-subjects/ Accessed: 2024-12-19.

[7] EU/Intrasoft consulting. 2025. *GDPR Consent.* The European Union. https://gdpr-info.eu/issues/consent/ Accessed: 2025-11-04.

[8] Hana Habib Megan Li Ellie Young Lorrie Faith Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *CHI 2022.* Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3491102.3501985 Accessed: 2025-12-19.

[9] Danish Customs and Tax Administration. 2025. *Skat for borgere (Tax for Individuals).* Skatteforvaltningen. https://skat.dk/borger

[10] DataGuard. 2024. *Consent management: the complete guide for business leaders.* DataGuard. https://www.dataguard.com/blog/consent-management-platform/ Accessed: 2025-11-04.

[11] European Commission. 2025. *Public Health: Promoting health, preventing diseases and fostering healthy lifestyles in the EU.* Directorate-General for Health and Food Safety. https://health.ec.europa.eu/index_en

[12] Jim Frost. 2016. Understanding Analysis of Variance (ANOVA) and the F-test. Minitab Blog. https://blog.minitab.com/en/blog/adventures-in-statistics-2/understanding-analysis-of-variance-anova-and-the-f-test Accessed: 2024-12-19.

[13] Christophe Gagnier. 2024. Cookie-Editor (Version 1.13.0). https://cookie-editor.com/. https://cookie-editor.com/ Accessed: 2025-12-19.

[14] Colin M. Gray, Jingle Chen, Shruthi Sai Chivukula, and Liyang Qu. 2021. End User Accounts of Dark Patterns as Felt Manipulation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2, Article 372 (10 2021), 25 pages. doi:10.1145/3479516 Accessed: 2025-12-19.

[15] Nicolai Brodersen Hansen, Christian Dindler, Kim Halskov, Ole Sejer Iversen, Claus Bossen, Ditte Amund Basballe, and Ben Schouten. 2019. How Participatory Design Works: Mechanisms and Effects. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction (OZCHI '19).* Association for Computing Machinery, New York, NY, USA, 30–41. doi:10.1145/3369457.3369460 Accessed: 2025-12-19.

[16] Suvi Andrea Helminen. 2025. *The Internet is the Wild West.* MIT Open Documentary Lab. https://opendoclab.mit.edu/the-internet-is-the-wild-west-in-many-ways-suvi-andrea-helminen/ Accessed: 2025-11-04.

[17] Hannah J. Hutton and David A. Ellis. 2023. Exploring User Motivations Behind iOS App Tracking Transparency Decisions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23).* Association for Computing Machinery, New York, NY, USA, Article 165, 12 pages. doi:10.1145/3544548.3580654 Accessed: 2025-12-19.

[18] Sarah Inman and David Ribes. 2019. "Beautiful Seams": Strategic Revelations and Concealments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland, UK) *(CHI '19).* Association for Computing Machinery, New York, NY, USA, Article 285, 14 pages. doi:10.1145/3290605.3300508 Accessed: 2025-12-19.

[19] International Organization for Standardization. 2019. *ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems.* Standard. International Organization for Standardization, Geneva, CH. https://www.iso.org/standard/77520.html Accessed: 2025-12-19.

[20] iubenda. 2024. *Is the use of a cookie wall allowed in European Countries?* iubenda s.r.l. https://www.iubenda.com/en/help/24487-cookie-walls-gdpr Accessed: 2025-12-19.

[21] Tanya Kant. 2021. *A History of the Data-Tracked User.* The MIT Press. https://thereader.mitpress.mit.edu/a-history-of-the-data-tracked-user/ Accessed: 2025-11-04.

[22] Tanya Kant. 2021. *Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your "Ideal User".* MIT Schwarzman College of Computing. https://mit-serc.pubpub.org/pub/identity-advertising-and-algorithmic-targeting/release/2 Accessed: 2025-11-04.

[23] Mark J. Keith, Courtenay Maynes, Paul Benjamin Lowry, and Jeffry Babb. 2014. Privacy Fatigue: The Effect of Privacy Control Complexity on Consumer Electronic Information Disclosure. In *Proceedings of the 35th International Conference on Information Systems (ICIS 2014).* Association for Information Systems (AIS), Auckland, New Zealand, 20. https://www.researchgate.net/publication/268620254_Privacy_fatigue_The_effect_of_privacy_control_complexity_on_consumer_electronic_information_disclosure Accessed: 2025-12-19.

[24] Truman Lee Kelley. 1927. *Interpretation of Educational Measurements.* World Book Company, Yonkers-on-Hudson, NY. https://gwern.net/doc/iq/1927-kelley-interpretationofeducationalmeasurements.pdf Accessed: 2025-12-19.

[25] Alexander Kunze, Stephen J. Summerskill, Russell Marshall, and Ashleigh J. Filtness. 2019. Automation transparency: implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics* 62, 3 (2019), 345–360. doi:10.1080/00140139.2018.1547842 Accessed: 2025-12-19.

[26] Effie L-C. Law, Virpi Roto, Marc Hassenzahl, Arnold P. O. S. Vermeerem, and Joke Kort. 2009. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09).* Association for Computing Machinery, New York, NY, USA, 719–728. doi:10.1145/1518701.1518813 Accessed: 2025-12-19.

[27] Jakob Nielsen. 1994. *10 Usability Heuristics for User Interface Design.* Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/ Accessed: 2025-12-19.

[28] Midas Nouwens, Rolf Bagge, Janus Bager Kristensen, and Clemens Nylandsted Klokmose. 2022. Consent-O-Matic: Automatically Answering Consent Pop-ups Using Adversarial Interoperability. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (New Orleans, LA, USA) *(CHI EA '22).* Association for Computing Machinery, New York, NY, USA, Article 238, 7 pages. doi:10.1145/3491101.3519683 Accessed: 2025-12-19.

[29] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, Article 314, 13 pages. doi:10.1145/3313831.3376321 Accessed: 2025-12-19.

[30] Sebastian A. C. Perrig, Lena Fanya Aeschbach, Nicolas Scharowski, Nick von Felten, Klaus Opwis, and Florian Brühlmann. 2024. Measurement practices in user experience (UX) research: a systematic quantitative literature review. *Frontiers in Computer Science* 6 (2024), 1368860. doi:10.3389/fcomp.2024.1368860 Accessed: 2025-12-19.

[31] Andromachi Psomiadi. 2025. *Essential vs. Non-Essential Cookies: What Sets Them Apart?* Pandectes. https://pandectes.io/blog/essential-vs-non-essential-cookies-what-sets-them-apart/ Accessed: 2025-12-19.

[32] React Contributors. 2025. *Managing State.* Meta Platforms, Inc. https://react.dev/learn/managing-state Accessed: 2025-12-19.

[33] Santander Consumer Finance, S.A. 2025. Santander Consumer Finance Official Website. https://www.santanderconsumer.com/. Accessed: 2024-12-19.

[34] Florian Schaub, Rebecca Balebako, Adam L. Durity, and Lorrie Faith Cranor. 2015. A Design Space for Effective Privacy Notices. In *Eleventh Symposium on Usable Privacy and Security (SOUPS 2015).* USENIX Association, Ottawa, 1–17. https://www.usenix.org/conference/soups2015/proceedings/presentation/schaub Accessed: 2025-12-19.

[35] Irina Shklovski, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. 2014. Leakiness and Creepiness in App Space: User Perceptions of Privacy and Mobile App Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14).* ACM, Association for Computing Machinery, New York, NY, USA, 2347–2356.

doi:10.1145/2556288.2557421 Accessed: 2025-12-19.

[36] Clay Spinuzzi. 2005. The Methodology of Participatory Design. *Technical Communication* 52, 2 (2005), 163–174. https://www.researchgate.net/publication/233564945_The_Methodology_of_Participatory_Design Accessed: 2025-12-19.

[37] Sundhed.dk. 2025. *Sundhed.dk: Den fælles offentlige sundhedsportal.* Danske Regioner, Indenrigs- og Sundhedsministeriet og Kommunernes Landsforening. https://www.sundhed.dk/

[38] Supabase Documentation. 2025. *Edge Functions Architecture.* Supabase Inc. https://supabase.com/docs/guides/functions/architecture Accessed: 2025-12-19.

[39] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un)informed Consent: Studying GDPR Consent Notices in the Field. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (London, United Kingdom) *(CCS '19)*. Association for Computing Machinery, New York, NY, USA, 973–990. doi:10.1145/3319535.3354212 Accessed: 2025-12-19.

[40] Hannah Volk-Jesussek. 2025. *Kruskal-Wallis-Test simply explained.* numiqo e.U., Graz, Austria. https://numiqo.com/tutorial/kruskal-wallis-test Accessed: 2025-12-19.

[41] Zalando SE. 2025. *Zalando: Tøj, sko og accessories online.* Zalando. https://www.zalando.dk/home/

[42] Sebastian Zimmeck, Eliza Kuller, Chunyue Ma, Bella Tassone, and Joe Champeau. 2024. Generalizable Active Privacy Choice: Designing a Graphical User Interface for Global Privacy Control. *Proceedings on Privacy Enhancing Technologies* 2024, 1 (1 2024), 258–279. doi:10.56553/popets-2024-0015 Accessed: 2025-12-19.

Appendix

This appendix provides supplementary data, including detailed statistical analyses, participatory design materials, and full-resolution documentation of the experimental prototypes to support the findings discussed in the main text.

A Further Study Justification

Why did we go with a between subjects study?

- **Reasons for Choosing a Between-Subjects Design:**
  - **No Transfer Across Conditions:** This ensures true data about participants' feelings for each independent variable level, as there is no comparison or learning effect that could alter their comprehension. Participants come in with the same knowledge (or lack thereof) across all levels.
  - **Shorter Session / Reduced Fatigue:** The shorter time commitment for each participant helps prevent survey fatigue, which is particularly beneficial for remote unmoderated testing.
- **Disadvantages (Cons):**
  - Requires More Testers: This design demands a larger sample size compared to within-subjects alternatives.
  - **Potential for Noise/Variability:** It may be more difficult to minimize the noise (variability) in the data set due to individual differences between groups.
- **Importance of Randomization:**
  - Enhances Internal Validity: Randomisation is crucial for internal validity.
  - Prevents Confounding: It avoids the assignment of groups becoming a confounding variable - a hidden variable that could systematically affect the results of the study.
- **Potential alternatives:**
  - Within-Subjects Design (Repeated Measures): Each participant interacts with all CMP variants in a randomized or counterbalanced order. While this approach would potentially reduce noise (variability) and would require fewer participants, it would propose a high risk of transfer or order effects and long sessions lead to fatigue.
  - Longitudinal Study (Sequential/Multi-Phase Design): Such a study would be super valuable, measuring the long-term retention of knowledge, continued use, or change in perceived agency/fatigue. Logistically difficult and expensive on Prolific due to the requirement for follow-up sessions. High participant attrition/dropout is expected, weakening the statistical power of the Phase 2 data. Still introduces the risk of transfer effects if a different condition is introduced in a later phase.

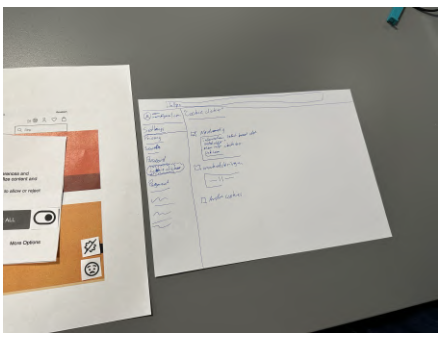
B PD workshop prototypes

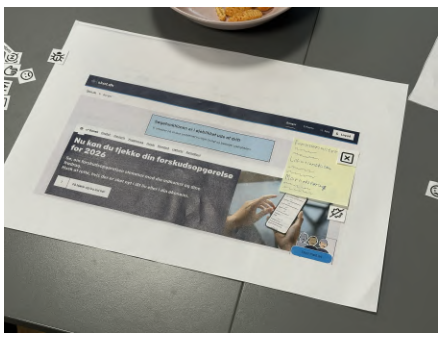
Group 1:

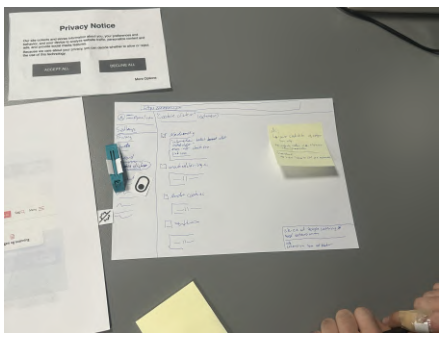
In the first iteration, the participants focused on clarifying the purpose of cookies and helping users understand what they are agreeing to. The participants proposed a browser extension that could explain cookie categories in simple terms and allow users to make their choices directly at the browser level.

In the second iteration, the participants expanded on the need for a clearer structure and a more detailed overview of cookie options. They suggested categorizing cookies more systematically and presenting them in a layout that would be easier for users to navigate.

By the third iteration, the participants were prompted to focus toward automation as a way to reduce user effort. Their prototype suggested a browser extension that could make cookie decisions automatically, removing the need for repetitive choices across websites. Participants envisioned an AI that learns from the user's past cookie decisions, analyzes patterns across visited websites, and to make personalized choices.







Group 2:

In the first iteration, the participants focused on the need for a simple and easy to understand application that explains when and why cookies are used. They suggested allowing users to choose exactly what their cookies may be used for, organized into simple and intuitive categories.

In the second iteration, participants suggested a standardized CMP so all websites share the same interface. They also proposed a gear-icon to show what data is being stored, providing quick transparency. Another key idea was enabling users to revisit and change consent at the browser level through a history-like overview, reflecting a stronger focus on user control and long-term data management.

By the third iteration, the participants focused on real-time transparency. They proposed a system that notifies users when a website collects data, requiring immediate approval or denial. To prevent overload, participants also suggested options to mute or generalize decisions.

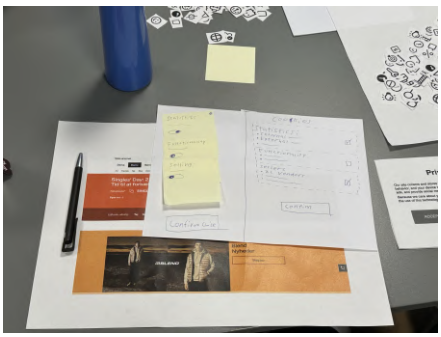
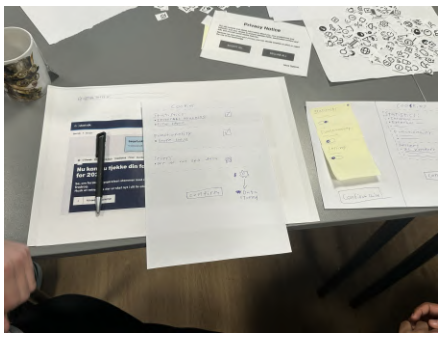


Fig. 11. Group 2, Iteration 1



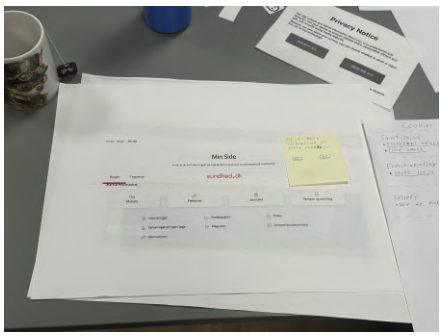


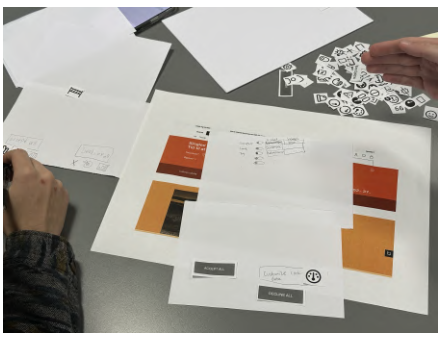
Fig. 12. Group 2, Iteration 2

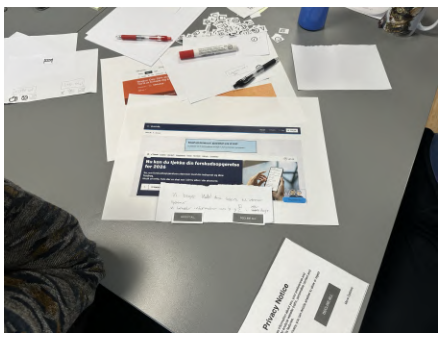
Group 3:

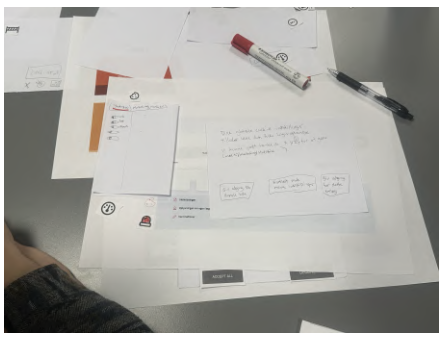
In the first iteration, the participants suggested adding an additional button that leads to a more detailed list of vendors and specific data types the user can allow. This was envisioned as a potential global setting, enabling users to set preferences once rather than repeatedly.

In the second iteration, the participants expanded their focus on transparency by integrating information about how long each website keeps user data. They continued refining the distinction between internal and external vendors, emphasizing the importance of helping users understand where their data goes and for how long it is stored.

In the third iteration, the participants proposed a browser-level global settings panel where users can configure preferences for specific cookie categories. If a website requires a decision that has not yet been set globally, a small prompt would appear, allowing the user to populate the missing preference. Participants also introduced a warning system—such as a yellow or red icon or gauge to indicate when a website cannot comply with the user's established preferences, alerting them before proceeding.







Group 4:

B.1 Categorizing

	Increase Agency	Combat Misconceptions	Reduce Fatigue	Improve Transparency
Fig 1			X	
Fig 2		X		
Fig 3			X	
Fig 4	X	X		
Fig 5		X		
Fig 6	X			
Fig 7				
Fig 8		X		
Fig 9	X			X
Fig 10		X		
Fig 11				
Fig 12	X	X		

C Changes to the prototypes prior to the user test

When designing the variants, it was also crucial to append and take away details, so that our results would prove helpful in comparing design variations and addressing our research questions and hypotheses.

- **Baseline:** The baseline option has been "dumbed down" to contain bare minimum information. Its vocabulary is made even more barebones than before, and any information should be highly technical.
- **Seamful Automation:** The seamful automation option has been simplified to "just work" without elaborate configuration or data being readily available to the user - which you might see in a real product.
- **Seamless Automation:** The seamless automation option was made to be completely invisible after a simplified configuration is completed.
- **Informative:** The informative option was exaggerated with information that might be technically difficult to realise - namely evaluating the consequences of one's choices, but as a proof of concept includes fake data unbeknownst to the user.
- **Revisitability:** The revisitability option was also largely simplified to show bare minimum information to begin with, and only expand on previous choices once the user chooses to revisit them.

D User study survey

Empirical Findings from the Collected Data 1. Overall engagement and task completion are stable across designs

Across all design variants, session completion rates are highly similar. No CMP variant caused noticeable drop-off or premature task abandonment. This indicates that all designs were usable enough to allow participants to complete the browsing task, and that observed behavioral differences are unlikely to be driven by disengagement or attrition.

Implication: Differences in consent behavior and interaction patterns reflect design effects, not failure to complete the task. Graph: Sum of completed_sessions by design_variant

Finding 2 — Decision time differs by CMP design (baseline vs informative vs revisit)

Supported by:

Graph 2.1: Sum of mean_decision_time_seconds by design_variant

Graph 2.2: Sum of mean_decision_time_seconds by design_variant and site_name

What the graphs show:

Informative has the highest mean decision time (22s).

Baseline and Revisit are lower (13–15s), with revisit slightly above baseline in your plot.

Site-level breakdown shows the same pattern across EU_health / Santander / Zalando, so it's not driven by one specific site.

One-line implication (neutral):

"CMP design is associated with different decision times, with informative flows taking longer than baseline and revisit in this dataset."

Graph: Sum of mean_decision_time_seconds by design_variant Graph: Sum of mean_decision_time_seconds by design_variant and site_name

3. Consent outcome distributions shift by design

Even though the baseline CMP makes accepting, rejecting, and customising equally accessible, consent outcomes differ across variants:

Informative designs show higher rates of implicit consent and lower explicit rejection.

Revisit designs show higher explicit interaction (accept, reject, or customise) and lower implicit consent.

Automation variants cluster differently again, suggesting that pre-structured choices influence outcomes without eliminating user action.

Implication: CMP design affects how users express consent, not merely whether they can. This holds even when the baseline already satisfies GDPR choice parity.

Graph: Sum of accepted, rejected, customised, implicit by design_variant

Graph: Sum of accepted_pct, rejected_pct, customised_pct, implicit_pct by design_variant

4. Implicit consent is strongly design-dependent

Implicit consent occurs far more frequently in some designs than others, despite similar completion rates. This suggests that implicit consent is not a by-product of confusion or abandonment, but rather a response to interaction framing.

Designs that foreground explanation or delay explicit choice appear to increase the likelihood of users progressing without making a deliberate selection.

Implication: Implicit consent functions as a behavioral signal of lowered immediacy or decision salience, rather than a failure to provide options.

Graph: Sum of sessions_implicit_consent by design_variant

5. Site order influences interaction, but does not dominate design effects

Site order distributions are balanced across design variants, and while some variation in consent outcomes occurs by site sequence, the relative differences between CMP designs persist across orders.

This indicates that although contextual familiarity and domain expectations matter, CMP design remains a primary driver of interaction patterns.

Implication: Observed effects cannot be explained solely by learning or fatigue across sites.

Graph: Sum of site_order_occurrence by site_order and design_variant

Graphs: accepted_pct by site_order and design_variant rejected_pct by site_order and design_variant customised_pct by site_order and design_variant implicit_pct by site_order and design_variant

5.2 Context matters because the site (e-commerce vs finance vs health) and the position (site 1 vs site 2 vs site 3) can change expectations, attention, and willingness to engage.

Your site_order_occurrence plot shows the different orders are well represented (so you're not accidentally comparing one design mostly on one order).

In the outcome-by-order plots (accepted/rejected/customised/implicit % by site_order × design_variant), you still see the same design "fingerprints" across different orders (e.g., informative tending to look different from baseline/revisit in a similar direction).

So the best conclusion is:

Context contributes noise and some shifts, but the design effect is robust enough that it persists across orders—meaning learning/fatigue alone can't account for the patterns.

If you want a super concrete way to phrase it in the paper:

Site order and domain context influence consent behavior to some degree, but the observed differences between CMP variants are consistent across multiple site sequences, suggesting that CMP design—not only order effects—drives the dominant interaction patterns in this dataset.

If you want to make it even stronger (still without doing RQ answers), run one sanity-check table:

For each design_variant, compute outcome % separately for trial_index 0, 1, 2. If the variant differences persist at each trial index, that's direct evidence context doesn't dominate.

Here's how to tie that paragraph directly to your "sessions_implicit_consent by design_variant" bar chart (the one showing total=108, informative=41, automation=34, baseline=20, revisit=13):

Implicit-consent incidence varies strongly by CMP design. In the "sessions_implicit_consent by design_variant" graph, implicit consent occurs 108 times in total, with the highest counts in informative (41) and automation (34), compared to baseline (20) and revisit (13). Because implicit consent represents progressing without making an explicit accept/reject/customise choice, these higher counts are consistent with greater disengagement under more cognitively demanding or opaque interactions (consent fatigue / learned helplessness). In contrast, lower implicit-consent counts in baseline (GDPR-friendly equal-friction options) and revisit suggest comparatively more explicit engagement with consent choices.

Finding — Session duration increases with CMP "interaction load" (not with drop-off)

Graph used: "Sum of mean_session_duration_seconds by design_variant" (Power BI bar chart)

What the data shows (approx. values from the graph):

Automation with results ≈ 70s

Revisit ≈ 66s

Informative ≈ 64s

Automation ≈ 59s

Baseline ≈ 58s

Conclusion from the pattern: Session duration clearly rises as the CMP variants introduce more to process or more steps to engage with (e.g., extra explanations, extra review/trace UI, or added status/results). In other words, the "heavier" designs are associated with longer end-to-end sessions than the baseline, even though the baseline is GDPR-friendly (accept/reject/customise all equally accessible).

How this ties to the paper (and what it means): In your paper, you explicitly frame time-on-task / time spent as a key in-use behavioural measure captured via logging (alongside task success), and you motivate shorter study sessions to reduce fatigue in remote Prolific studies.

So, this graph supports a clean, non-RQ-specific observation: design choices materially change how long participants spend completing the same three-site task flow. The longer sessions are consistent with added interaction/comprehension work introduced by the CMP interface itself (more content, more decision scaffolding, more "system status/results" elements), rather than participants "getting stuck" on the websites (since the websites are intentionally static and the CMP is the focal interaction).

E The prototyping materials and talking points

The prototyping materials and talking points:

- **Whiteboards and large printouts of selected websites (Zalando, Skat, Sundhed.dk) for shared design space**: They allow for easy iteration, sketching, and grouping of ideas, reinforcing the PD principle of mutual learning. Furthermore, this establishes the study context, ensuring that the design process is grounded in actual practices by providing a visual setting where a CMP would realistically appear.
- **Baseline CMP Printout**: Providing a physical example of the current problematic solution establishes a common starting point for critique, ensuring that all participants are discussing the same "trauma" of online consent. It acts as the mainstream solution that participants are tasked to counterbalance and improve upon, fulfilling a key PD purpose.
- **Pencils and markers**: Basic tools for sketching and writing notes, preferably in multiple colors.
- **Scissors**: Another basic tool, could come in handy when adapting and combining multiple elements and sketches.
- **Blank sheets of paper**: Our final basic tool for sketching, doodling, and/or taking notes during the workshop.
- **Sticky notes (blank)**: They will act as the primary tool for rapidly capturing, moving, and grouping ideas / sketches for UI elements, mirroring the need for quick and dirty prototyping. It is decided that no pre-filled sticky notes are passed out, letting the participants find out how they want to use the notes themselves. A possible alternative approach to post-its would be to map out ideas in text.
- **Cutouts of Common UI Elements (Buttons, Toggles, Links)**: These are used to speed up the design and envisioning stage. Instead of drawing everything, participants can quickly arrange, move, and negotiate the placement of functional components, focusing their cognitive effort on the system logic and information flow, or on whatever they find most interesting.

  Participants receive lots of unique icons to work with, even ones that may seem irrelevant - even to us - because they might surprise us with a unique interpretation.

Basic UI elements: "accept all" / "reject all" buttons, manage preferences links, toggles for different cookie categories (e.g., "analytics," "marketing," "essential"), Generic icons (e.g., clocks for history, magnifying glass for clarity, shields for privacy).

Advanced UI elements to motivate discussions regarding traceability and feedback Seams (Addressing RQ1 and RQ3): My consent history (view/undo) button, trust score/performance gauge, action feedback banners (e.g., "all preferences successfully applied", "this website was declined from sharing your data with *N* vendors", "*X* essential, *Y* non-essential cookies accepted, *Z* declined based on your preferences"), icons for self-expression (e.g. thumbs up / thumbs down).

Advanced UI elements to motivate discussions regarding clarity and abstraction (Addressing RQ2): Plain-language explanation tooltip(s), data type icons.

Advanced UI elements to motivate discussions regarding agency and automation balance (Addressing RQ1): Global preference toggle, one-click override button reinforcing the notion that consent must be easily revisitable and traceable, a categorical status indicator showing the status of major cookie groups.

To ensure that we adhere to the core principles of PD workshops presented above 4.1.1, we prepare some scripts and talking points to be followed on a semi-strict basis - enabling free-flowing and democratised discussion while staying on track.

First, the briefing: "Welcome! We are here to talk about a common source of frustration: online consent pop-ups. Our goal is not to build the final product today, but to hear your personal experiences and explore a wide spectrum of ideas for a solution that truly empowers you, not just automates your choice. We want to design a system that makes you feel in control again. Today's session is an iterative, hands-on collaboration."

Secondly, some prompts for contextual inquiry: "Describe a recent, frustrating interaction you had with a cookie banner or privacy pop-up." (Focus on personal anecdotes)

"What information do you feel is currently missing from consent banners that would help you make an informed decision?"
Targets the lack of transparency/clarity.

"When you click "Accept All," what are you really thinking or feeling in that moment?"
Explores "Okay, whatever" effect and learned helplessness.

"How important is it to you to know what a site does with your data after you have given consent, and how could you check that?"
Leads to the topic of traceability and agency.

Third, clarification of goals: "Based on our discussion, let us establish three core priorities for the new system. We aim for a solution that: 1. Increases agency (makes you feel in control); 2. Reduces fatigue (lessens mental effort or saves time); 3. Improves transparency (system operations are visible); 4. Combats misconceptions (is easy to understand)".

Fourth, prototyping: "Your assignment is to redesign the consent experience. You can use the baseline printout, the cutouts, or create entirely new elements with the sticky notes. You will work in small groups".

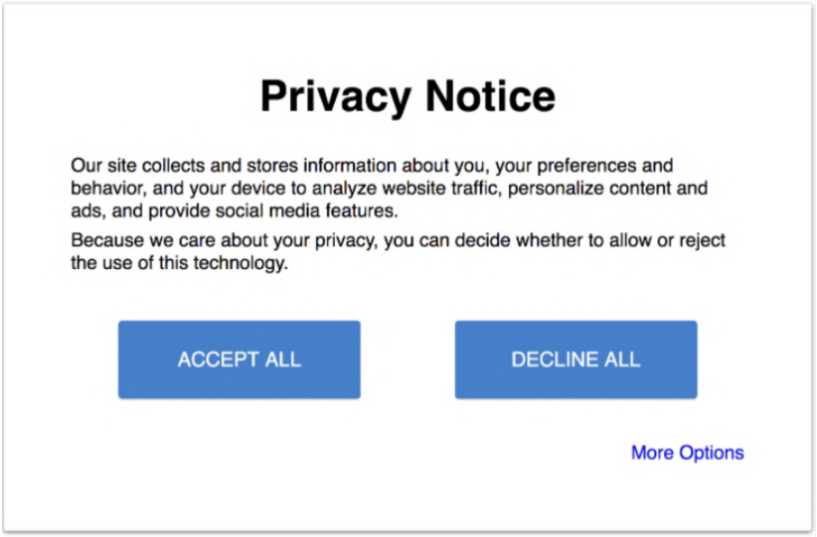
During this phase participants use the materials presented above to create and label their low-fidelity prototypes. Facilitators move between groups, prompting discussion and encouraging playful means of negotiating diverse product visions.

Lastly, pre-next-iteration: "Thank you for these great ideas! Before we conclude, consider these concepts that tackle user practices we have not necessarily explored yet: 1. Global control: What if you only had to set your preferences once in your browser, and it applied to all sites? 2. Ethical rating: What if a small icon showed you the ethical quality of the site's privacy practices (e.g., "good," "manipulative") before you even clicked? Do you think there are some inner-workings of your system you could reveal that could improve the user's sense of autonomy and agency?"

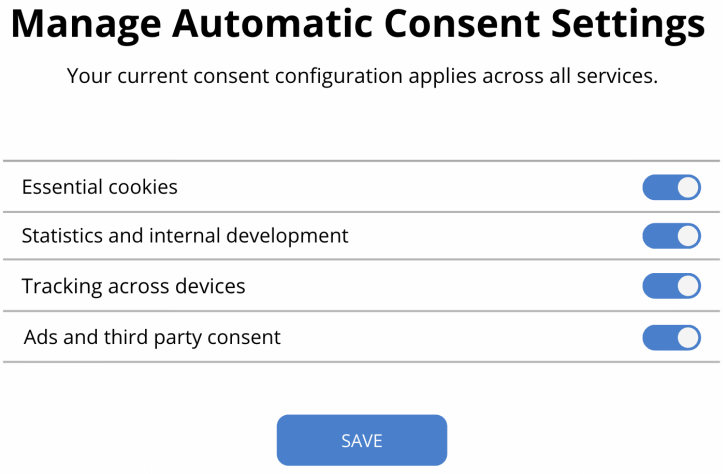
Participants are given another opportunity to revise their prototypes in light of the new concepts. Data is collected, including photographs of sketches, sticky notes, and audio recordings of the discussions for thematic analysis.

F Canva mock-ups

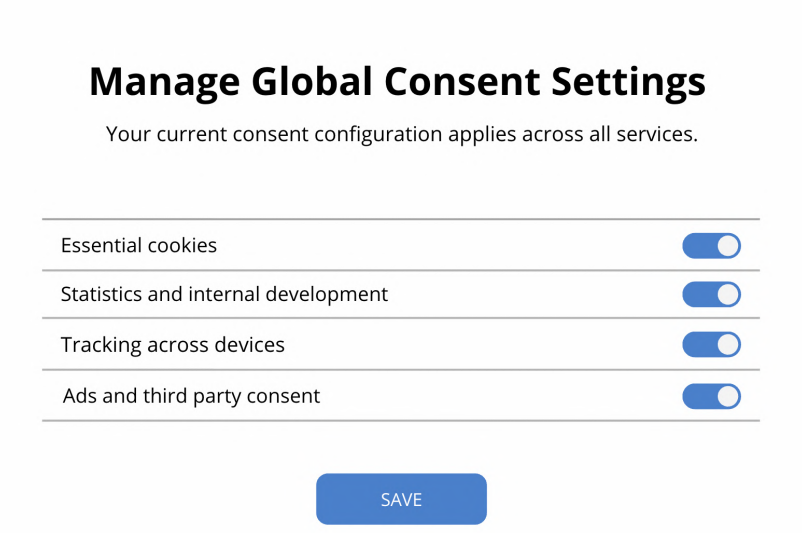
Baseline variant



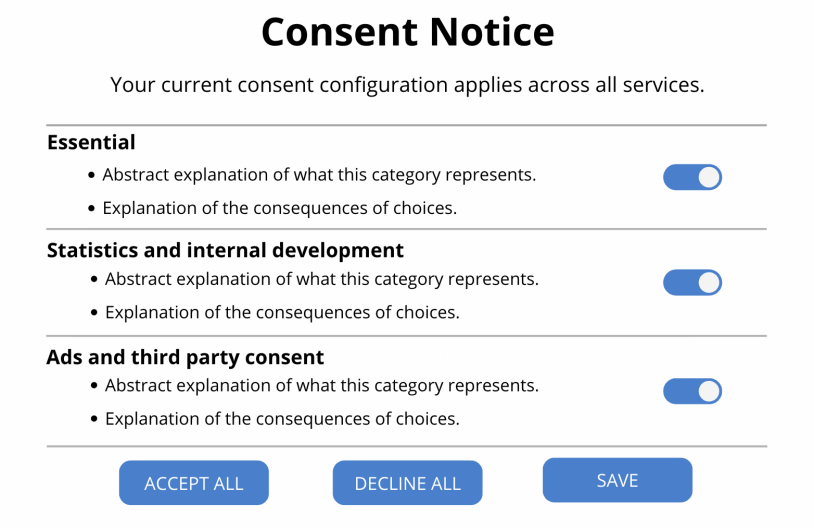
Automatic variants



Global Settings variant (discarded)



Informative variant



Revisitability variant

G Universal UI elements across all prototypes

Welcome page

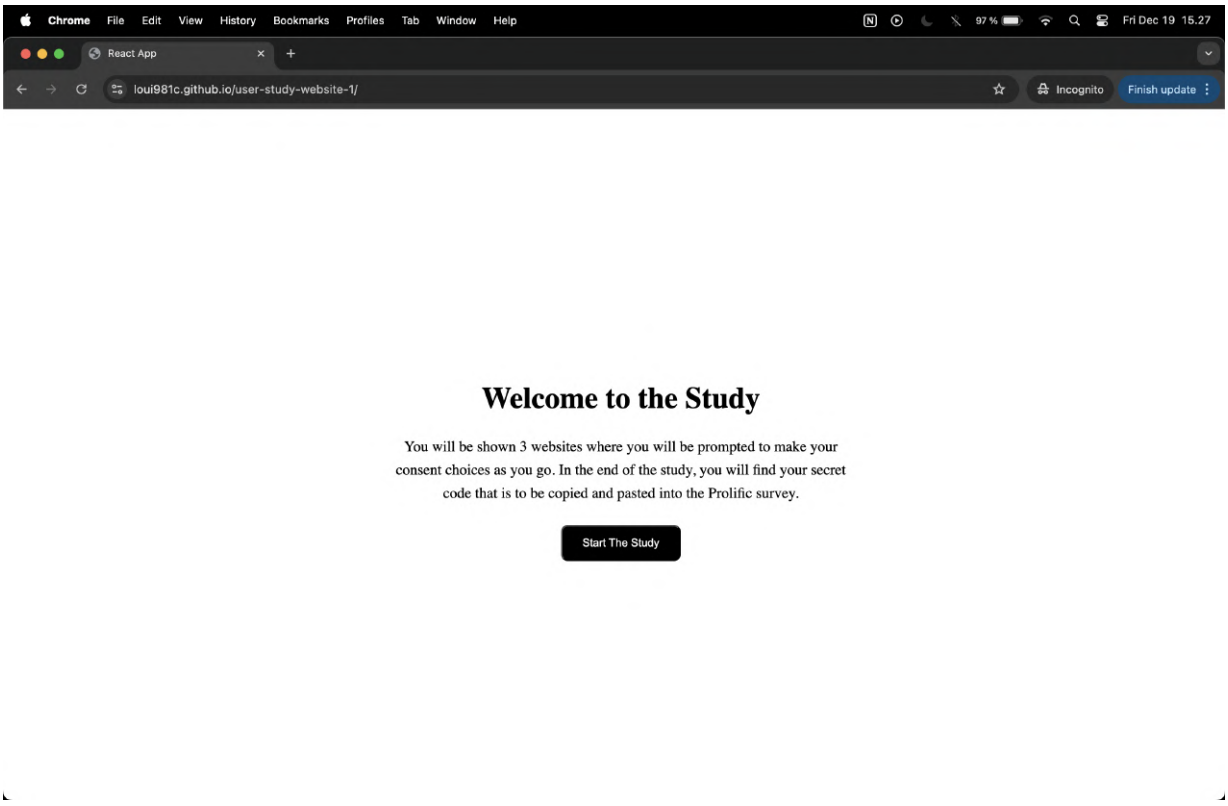


Fig. 13. Study welcome page presented before the experimental tasks.

Artificial loading page

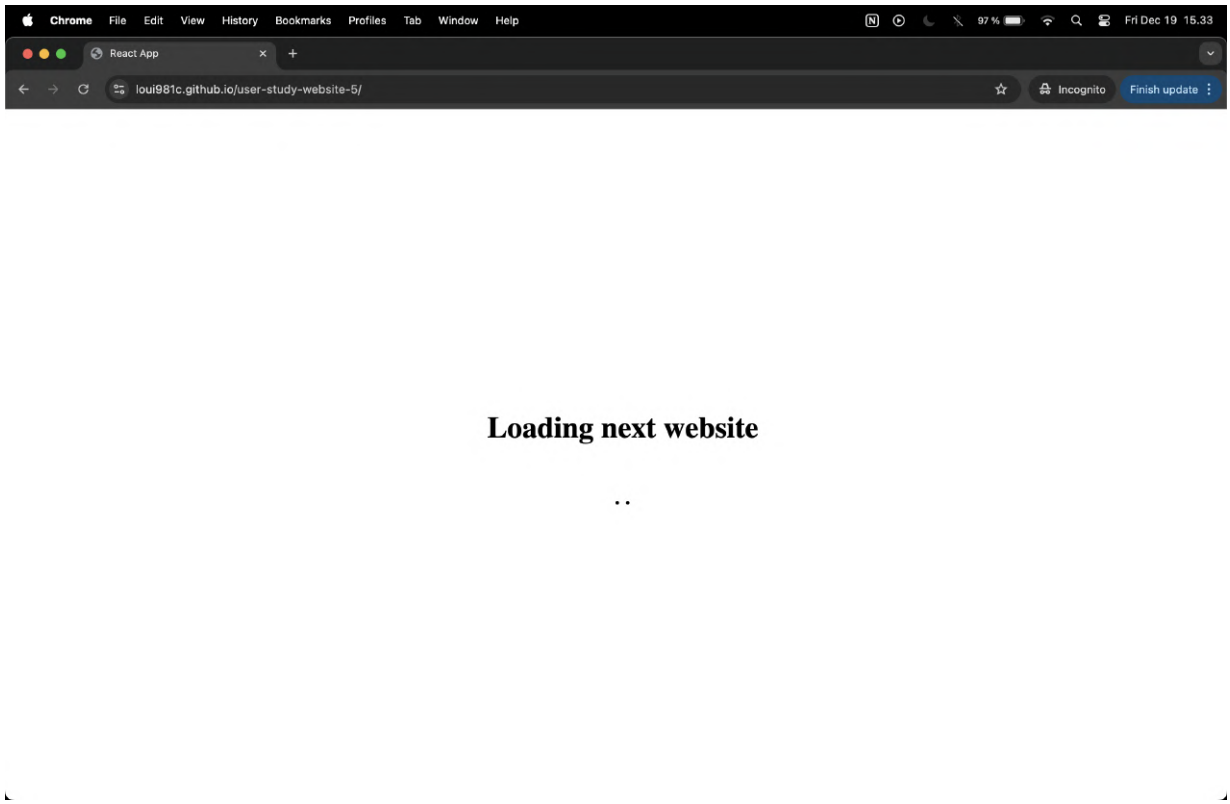


Fig. 14. Artificial loading screen used to standardize transitions between study stages.

Thank you page with secret UUID code

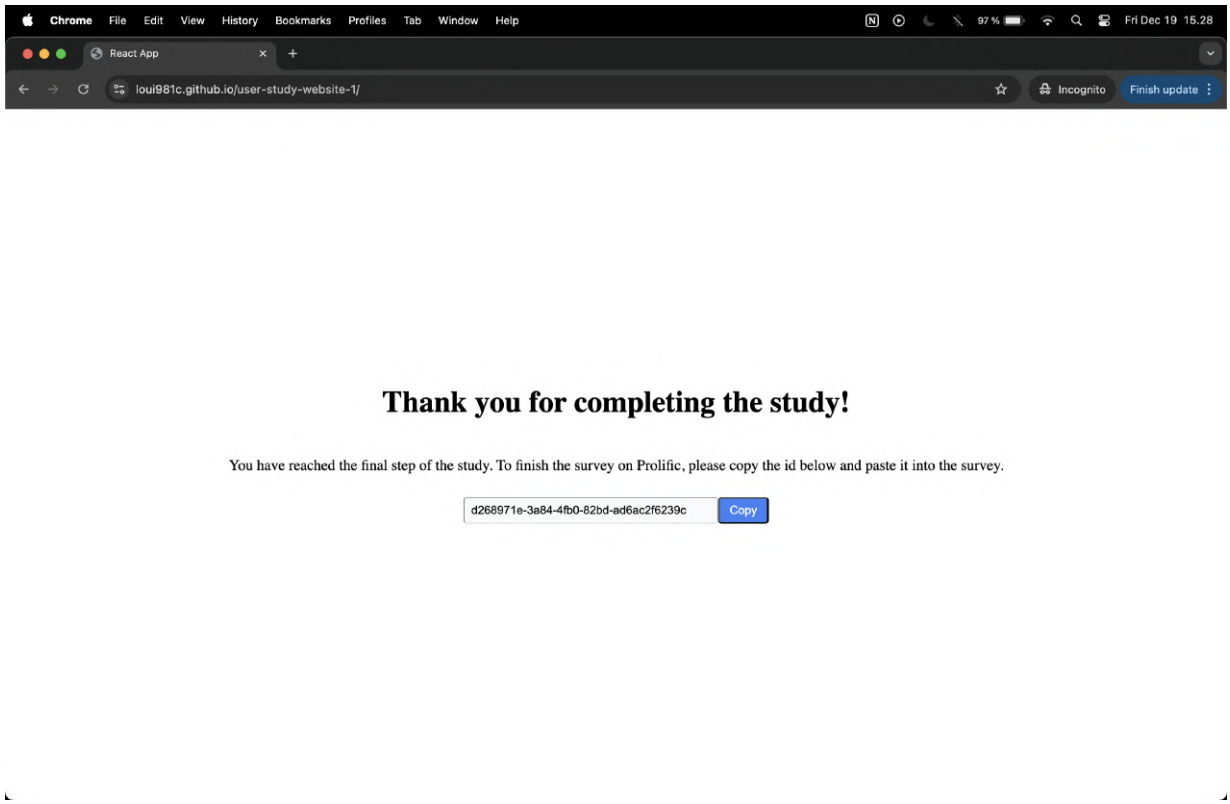


Fig. 15. Study completion page displaying a unique completion code for participants.

H Perrig et al.'s recommendations (1-4)

(1) **Decide if the Research Question Requires Scales**
- **Assessment:** Determine if the research question requires quantitative data (e.g., measuring "how much" or "to what extent") that cannot be answered solely by objective performance metrics (time, errors) or pure qualitative feedback.
- **Goal:** If the goal is to measure subjective states (emotion, satisfaction, aesthetic judgment) or long-term behavioral intentions (agency, trust), a validated scale is necessary.

(2) **Define the measured construct**
- **Clarity:** Precisely define the target UX construct based on established literature (e.g., using the definitions from Law et al.). Is the focus on *pragmatic quality* (e.g., usability), *hedonic quality* (e.g., stimulation), *emotional affect*, or *value/consequence*?
- **Specificity:** Avoid general terms. For example, instead of defining "satisfaction," define the specific facet, such as "post-Use affective satisfaction" or "perceived ease of use."

(3) **Select a scale that corresponds to the chosen definition of the target construct**
- **Validation check:** Select only scales that have been rigorously **validated** and whose psychometric properties (reliability, validity) are clearly reported in the literature (e.g., SUS for Usability, UEQ for hedonic/pragmatic quality, PANAS for affect).
- **Alignment:** Ensure the scale's sub-dimensions directly align with the construct defined in Step 2. For instance, if measuring "stimulation," choose a scale with a dedicated, validated "stimulation" sub-scale.

(4) **Implement the scale**
- **Protocol:** Implement the scale exactly as prescribed by its authors, including the original wording, number of items, response anchors (e.g., 5-point Likert, 7-point Semantic Differential), and scoring protocol (e.g., standardising, inverting items).
- **Reporting:** Ensure proper data reporting and analysis, including calculating the necessary scores (e.g., normalising SUS scores to $0-100$) and reporting the internal consistency (*Cronbach's* α) for the scale within your study context.

I Data Graphs

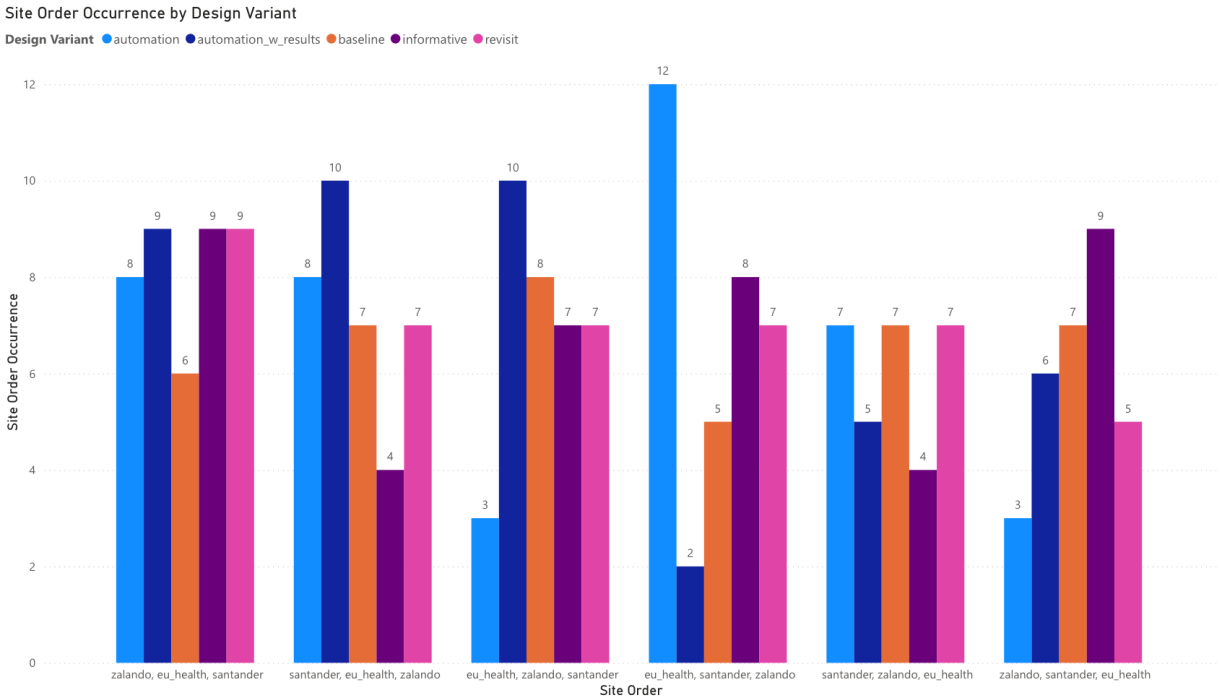


Fig. 16. Distribution of website presentation order across participants.

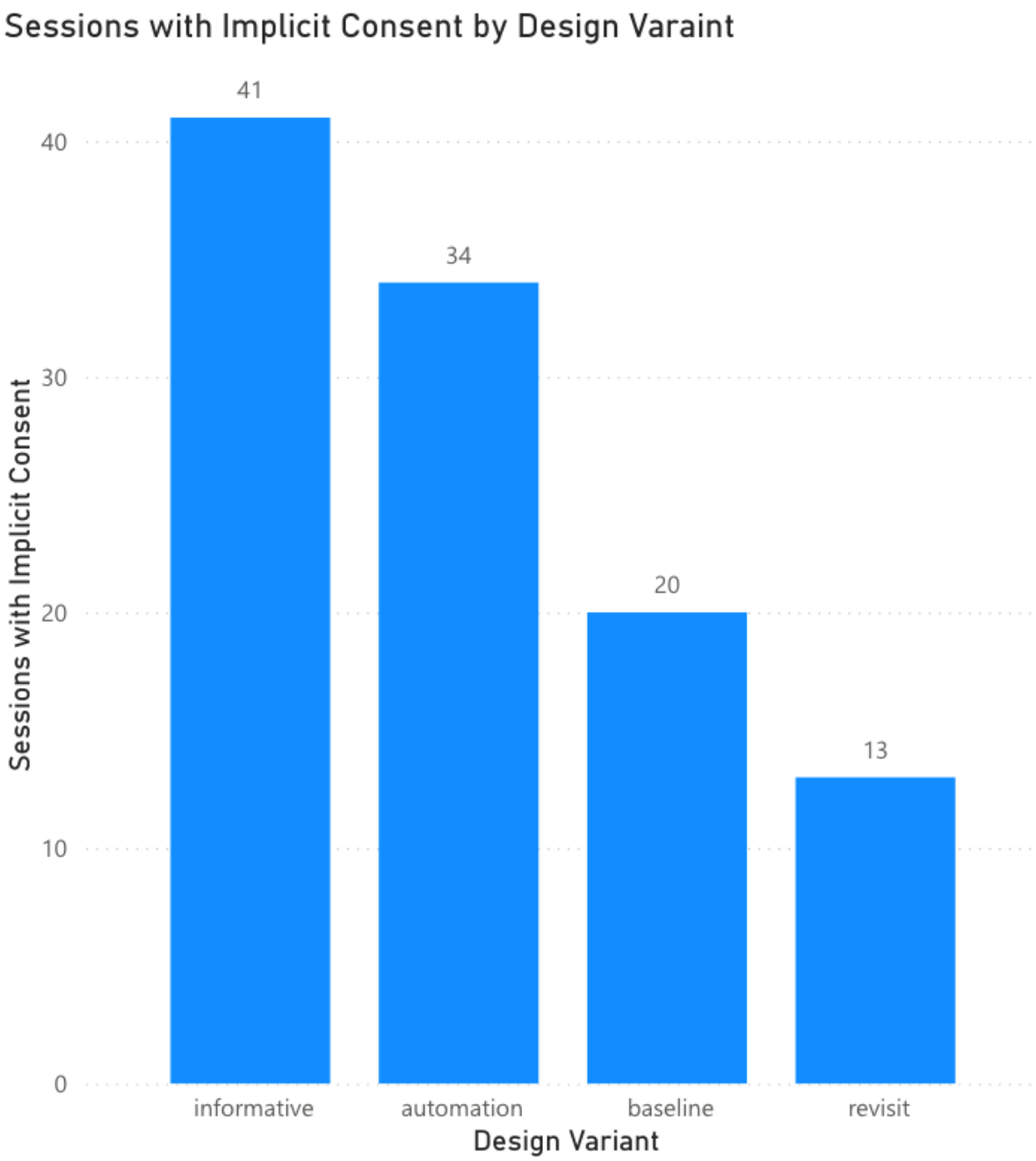


Fig. 17. Share of sessions in which implicit consent was given across CMP variants.

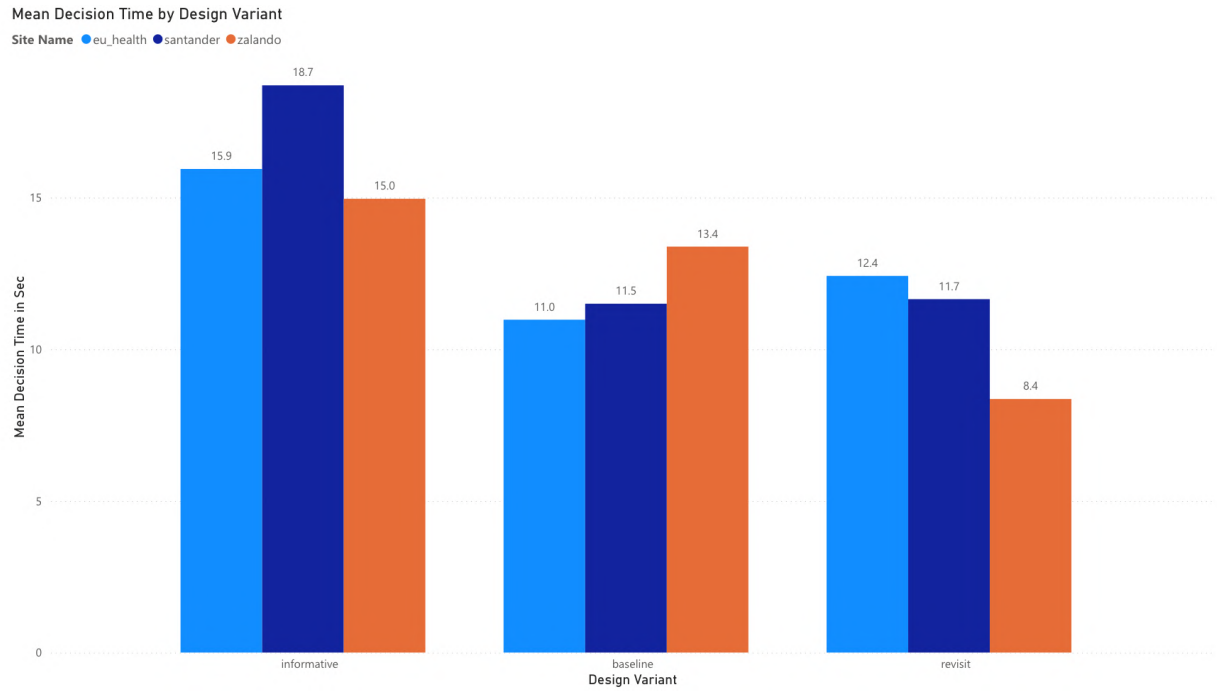


Fig. 18. Mean consent decision time across all websites and CMP variants.

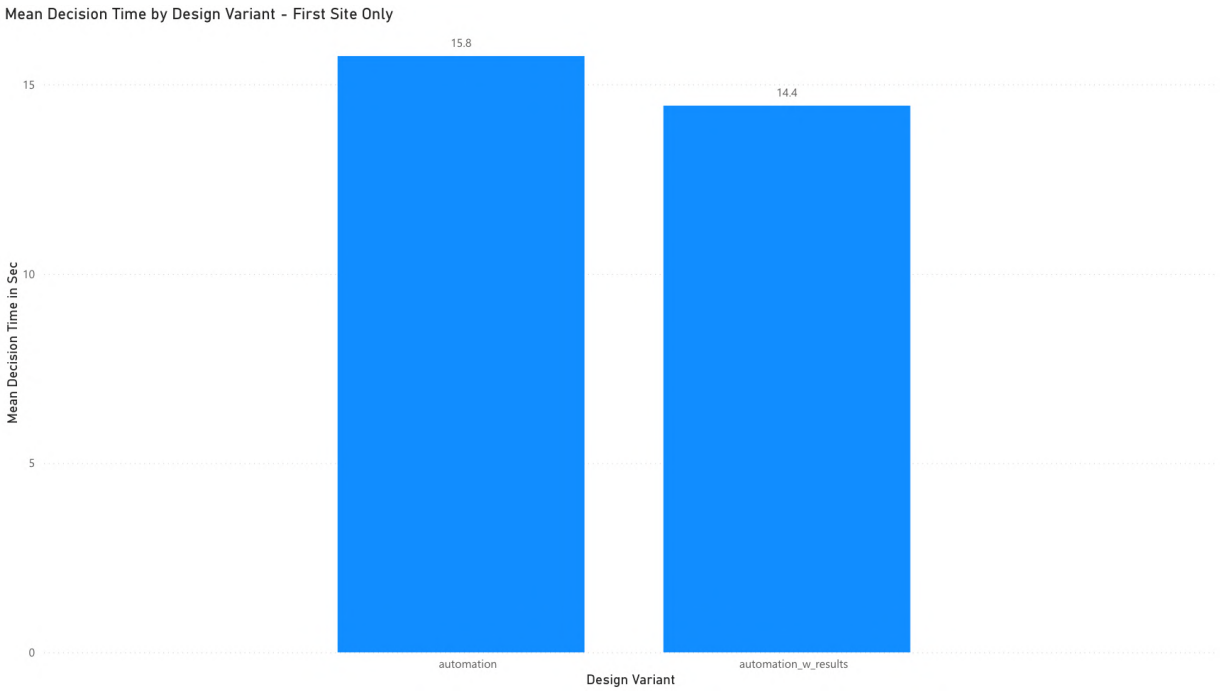


Fig. 19. Mean consent decision time on the first website encountered by participants.

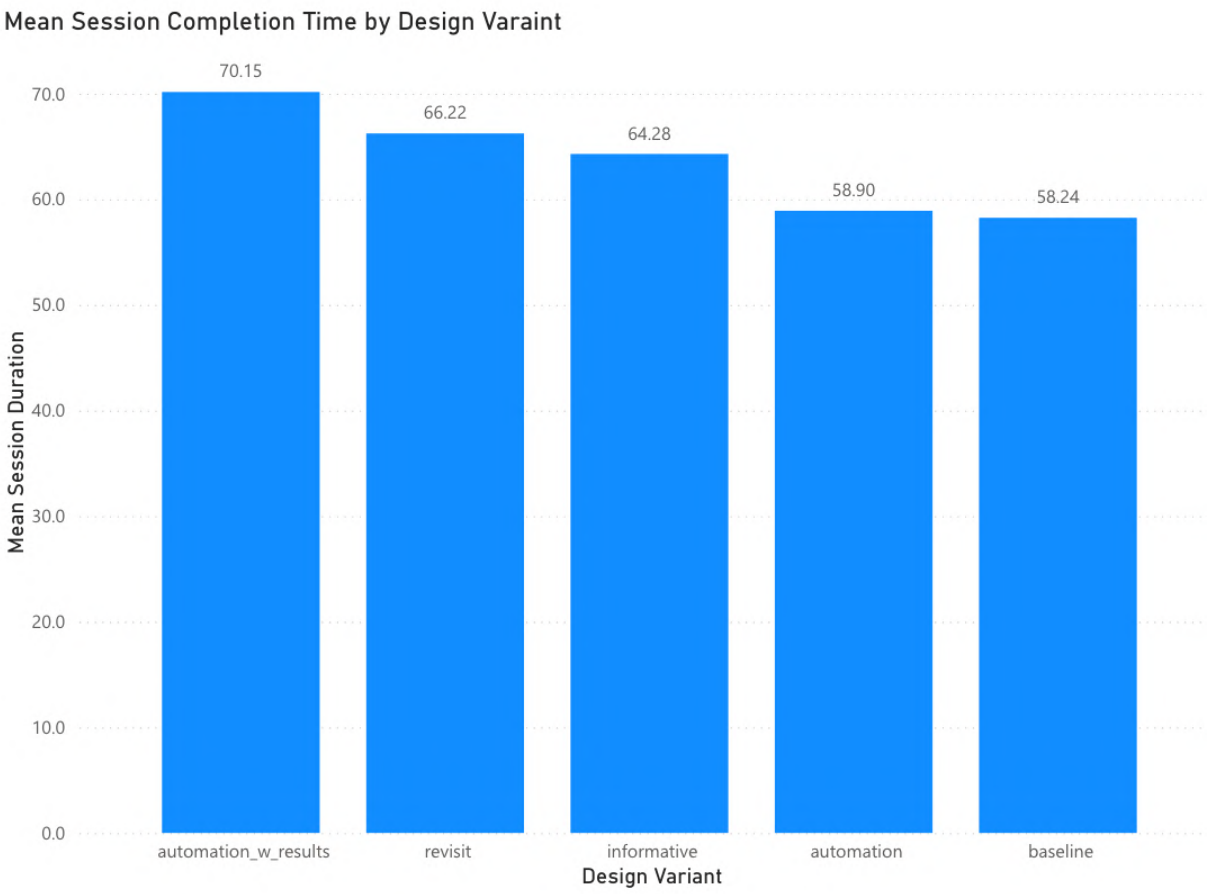


Fig. 20. Mean session duration across CMP variants.

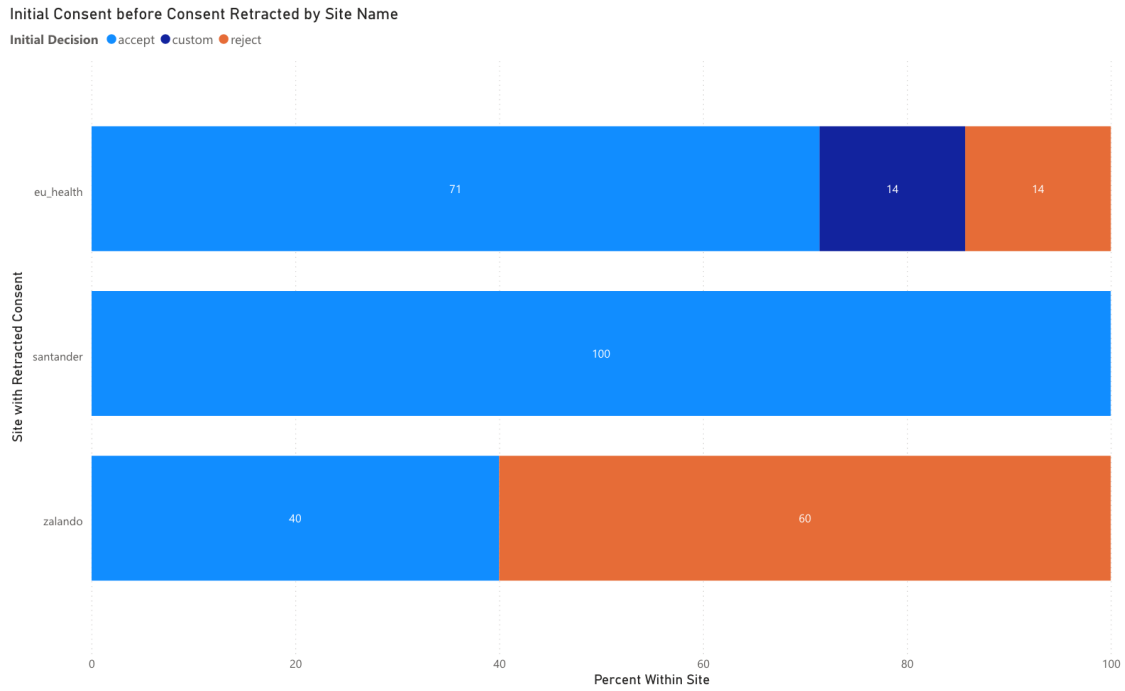


Fig. 21. Relationship between initially granted consent and subsequent consent retraction.

J Key activities to achieve these objectives

[36]

- Stage 1 - Initial exploration of work: The first stage is understanding current user practices. What technologies are already in use, what routines are present, and how does the current course-of-action differ between participants? This stage allows facilitators to understand the context in which a solution should fit into. In this stage, we also have some pre-established discussion points:
  - "Describe a recent, frustrating interaction you had with a cookie banner or privacy pop-up."
    Focuses on personal anecdotes.
  - "What information do you feel is currently missing from consent banners that would help you make an informed decision?"
    Targets the lack of transparency/clarity.
  - "When you click "accept all," what are you really thinking or feeling in that moment?"
    Explores "Okay, whatever" effect and learned helplessness.
  - "How important is it to you to know what a site does with your data after you've given consent, and how could you check that?"
    Leads to the topic of traceability and agency.
- Stage 2 - Discovery processes: Designer and users discuss the goals of the system and priorities during an everyday workflow. This stage allows designers and users to clarify and discuss the product goals and values and to achieve a solid common understanding. It is expected that this is the stage where facilitators and users interact most heavily. Depending on when a PD event is facilitated throughout a software development life cycle, staying on track regarding a problem but keeping doors open for solution proposals must be done with outmost care.
- Stage 3 - Prototyping: Users and designers iteratively shape technological artifacts to fit into the envisioned system goals and priorities. This stage allows the formation of tangible, low-fidelity design prototypes. As Hansen et al.'s paper [15] highlights, some tools and materials can be helpful to motivate creative or playful means for negotiating diverse product visions.

K ANOVA Data Analysis Notes

TERMS

Core Statistical Concepts (P-value)

The P-value is the probability of observing a result (like a difference between group means, or an association in a table) as extreme as, or more extreme than, the one you measured, assuming the null hypothesis is true.

Null Hypothesis: There is **no difference** between the groups (e.g., all 5 design variants have the exact same mean for NASA_TLX_Effort, or the design variant has no association with the final consent choice).

If the P-value is low (typically $P < 0.05$ or $P < 0.10$), you **reject the null hypothesis** and conclude the effect is statistically significant (or marginally significant).

Assumptions Test: Levene's Test

Levene's test is used to check a critical assumption of the ANOVA: **homoscedasticity**, or the **equality of variances** across all your groups (design variants).

H0 for Levene's Test: The variance is equal across all groups.

Interpretation: You generally *want* a high P-value ($P > 0.05$). A low P-value (e.g., the P=0.0140 for NASA_TLX_Performance) means the assumption is violated. This suggests the spread of scores is significantly different between design groups, and you should interpret the corresponding ANOVA result with caution.

Difference of Means: ANOVA (Analysis of Variance)

ANOVA tests whether there is a statistically significant difference between the means of your five design variants for any given continuous metric (like EUQ_S_Composite or NASA_TLX_Effort).

C (design variant): this represents the **Between-Group Variation** or the **Effect** of our independent variable. It is the amount of variation in scores that can be attributed to the specific design variant the user experienced. This is the difference you are interested in explaining.

Residual: This represents the Within-Group Variation or the Error/Noise in your model. It is the variation in scores that is not explained by the design variant—i.e., the natural individual differences between participants who all saw the same design.

Df: **Degrees of Freedom.** The number of independent data points used to calculate the statistic. For C(design_variant), df = (Number of Groups - 1) = 5 - 1 = 4. For the Residual, it's based on your total sample size minus the number of groups.

F: **F-statistic.** This is the core test statistic of ANOVA, calculated as a ratio of the two mean squares. A larger F-value indicates that the variation explained by your design is much greater than the unexplained individual variation, suggesting a significant effect.

PR(>F): The **P-value** associated with the F-statistic. This is the overall probability that *at least one* of your design group means is different from the others.

Pairwise Comparisons: Tukey HSD

FWER: **Family-wise Error Rate.** When you run multiple comparisons, the chance of finding a false positive (Type I Error) goes up. Tukey HSD controls the FWER, ensuring that the overall probability of making *any* false discovery across all comparisons stays below your specified alpha (typically 0.05).

Meandiff: The raw difference between the mean scores of the two groups being compared

P-adj: The **adjusted P-value**. This is the P-value for the specific pairwise comparison, but it has been mathematically adjusted to control the FWER. You look at this value to determine significance. If p-adj < 0.05, the difference is significant.

lower/upper: The **confidence interval** for the mean difference. If this range **does not include zero**, the difference is statistically significant. If it includes zero, you cannot rule out that the true mean difference is zero, meaning the comparison is not significant.

Reject: A boolean value (**True/False**) indicating whether the null hypothesis for that specific pair comparison is rejected. True means the two group means are significantly different (because $\mathbf{p-adj < 0.05}$).

Association Test: Chi-Square

CHI-Square statistic: The test statistic, which quantifies the difference between the **Observed Counts** in your Contingency Table and the **Expected Counts** (what you would expect if the design variant and final choice were completely unrelated). A larger statistic indicates a stronger association.

DOF: This defines the shape of the Chi-Square distribution used for the test. For a table with R rows (design variants) and C columns (final choices), DOF = (R-1) x (C-1). In your fixed test (5 designs x 3 choices), DOF = (5-1) x (3-1) = 8.

P-value: The probability of observing the differences you found in your Contingency Table if there were truly no association between the design variant and the final consent choice.

Our result P=0.2462 means we have a 24.62% chance of seeing your data by random chance if the null hypothesis of "no association" is true. Since this is high > 0.10, we **do not reject the null hypothesis** and conclude there is no significant association.

INTRO
We have done a between-subjects experiment comparing the effects of different design variants of consent forms on user behavior and attitudes. Our analysis will therefore focus on comparing the measures, dependent variables, across our different design groups, independent variables.

The analysis tested whether the five different cookie banner design variants resulted in different mean scores for a variety of user experience and behavioral metrics.

RESULTS

Out of 13 dependent variables, one showed a statistically significant difference (P < 0.05), and three others showed a marginal trend (0.05 <= P < 0.10).

Metric (DV)	P-value (ANOVA)	Significance Level	Key Finding / Pairwise Trend
NASA_TLX_Performance	0.0144	**Significant P < 0.05**	**Overall significant difference:** The perceived success of completing the task was affected. **Trend:** The automation_w_results design resulted in *higher* perceived performance than baseline (P=0.077) and informative (P=0.0668). No pair reached the P < 0.05 cutoff, indicating the effect is distributed across multiple pairs.

Perceived_Revisitability	0.0699	**Marginally Significant P < 0.10**	**Strongest Finding:** The baseline design resulted in **significantly lower** perceived revisitability compared to the revisit design P=0.054. This result is extremely close to the P<0.05 cutoff and strongly supports the effectiveness of the explicit revisit design in increasing the perception of choice traceability (H3a).
NASA_TLX_Effort	0.0502	**Marginally Significant P ≈ 0.05**	**Trend:** The Post-hoc test showed a trend suggesting that the informative design required less effort compared to the automation (P=0.1038) and baseline (P=0.1206) designs.
NASA_TLX_Temporal_Demand	0.0585	**Marginally Significant P < 0.10**	**Trend:** The automation design resulted in higher perceived hurry/rush compared to baseline (P=0.0687) and revisit (P=0.0947). This may suggest that users find the black-box nature of

			the automation design stressful.
All Other DVs	P > 0.10	Not Significant	All other core metrics, including **Perceived_Control** (P = 0.8508) and **Perceived_Agency** (P = 0.5949), were **not** significantly affected by the design variant.

Resolution of the Chi-Square Assumption Violation (H3b): The initial Chi-Square test was unreliable due to an assumption violation (minimum expected cell count of 0.19), primarily caused by the low response rate in the "Close the banner" category. This issue was resolved by regrouping the sparse categories into three high-N groups: **'Accepting'** (Accept all/necessary), **'Reject all'**, and **'Other/High Effort'** (Manual choice/Close banner).

Statistic	Value
Chi-Square Statistic	10.2756
Degrees of Freedom (DOF)	8
P-value (FIXED)	0.2462
Minimum Expected Count	2.25

The valid Chi-Square test is **not significant** (P = 0.2462). This indicates that the **design variant does not significantly affect** the distribution of final consent choices users typically make in their daily lives. The data does **not support Hypothesis H3b**, which predicted that the revisit option would increase the propensity for privacy-protective choices like "Reject all."

Metrics for RQ1: Automation, Transparency, and Fatigue Reduction

RQ1 examines how to balance automation (seamlessness) with transparency (seamfulness) to mitigate fatigue while preserving control and agency.

Metric	Theory & Mechanism	Relevance to RQs/Hs
NASA-TLX Subscales (Mental Demand, Temporal Demand, Effort, Frustration, Performance)	The NASA Task Load Index (TLX) quantifies subjective workload. High scores in **Mental Demand, Temporal Demand, and Effort** quantify the user's **cognitive burden** and **consent fatigue**. High **Frustration** measures the emotional cost. **Performance** is the user's perceived success.	**H1a/RQ1:** A successful seamful automation should aim for lower cognitive demand scores while maintaining high Performance. The marginal significance in **Effort** and **Temporal Demand** suggests your designs are affecting the experience of fatigue.
Perceived Agency (Q19)	Assesses the user's belief that the automated system acted on their behalf as they wished, reflecting **autonomy and ownership**. It tests if the user feels like the "pilot" of the automated process.	**H1b/RQ1:** Directly measures the success of "seamful" transparency (like automation_w_results) in ensuring users feel their choices were respected. **Result:** *No significant effect* suggests the transparency mechanism did not sufficiently boost this feeling beyond the baseline.

Perceived Control (Q20)	Measures the user's subjective feeling that they were in command of the decisions and interactions.	**H1a/RQ1:** Directly tests H1a, measuring whether explicit visibility of system status (seamfulness) yields higher feelings of control. **Result:** *No significant effect* suggests the designs failed to boost this feeling.
Time_on_Task_s	The objective measure of the total time spent interacting with the consent banner or task.	**RQ1 (Fatigue):** This is the behavioral proxy for fatigue. **Result:** *No significant effect* suggests the more transparent/complex designs did not meaningfully increase user burden time.

Metrics for RQ2: Comprehension and Data Abstraction

RQ2 focuses on using data abstraction and plain language to improve comprehension and reduce misconceptions.

Metric	**Theory & Mechanism**	**Relevance to RQs/Hs**
Understandability (Q21)	Measures the clarity of the information presented (plain language summaries vs. technical vendor lists).	**H2a/RQ2:** This directly assesses the efficacy of the informative design in using data abstraction. **Result:** *No significant effect* suggests the informative design did not succeed in making the information significantly more

		understandable than other designs.
Confidence_in_Reasons (Q22) & **Understanding_of_Effects** (Q23)	These assess the depth and quality of comprehension (grasping *why* a choice was made and *what consequence* it had).	**H2a/RQ2:** These are direct measures of "high-level understanding" and "reducing misconceptions." **Result:** *No significant effect* suggests the simplified language did not facilitate truly informed intent across the groups.
EUQ_S_Composite (Q11)	The End-User Computing Satisfaction Short-Form measures general user satisfaction and perceived usability.	**RQ2 (Trade-off):** This acts as a check for the trade-off with fatigue. **Result:** *No significant effect* suggests the various designs were all perceived with similar overall satisfaction.
Final_Consent_Choice (Q25)	The behavioral choice (Accept All, Reject All, etc.) that the user usually makes.	**H2b/RQ2 (Privacy Paradox):** Used to assess the **privacy paradox**. The **non-significant Chi-Square** result suggests that the designs did not shift users' privacy choices.

Metrics for RQ3: Agency through Traceability

RQ3 explores whether enabling easy revisitation of choices increases the sense of agency and willingness to exercise privacy-protective choices.

Metric	Theory & Mechanism	Relevance to RQs/Hs
Perceived_Revisitability (Q24)	Measures the extent to which the user believes they can easily go back and adjust their consent choices later. This is the direct operationalization of the **traceability** concept.	**H3a/RQ3:** This is the direct test of the revisit design variant. The **marginally significant result** (baseline vs. revisit, P=0.054) strongly suggests the explicit revisit control successfully combats **learned helplessness** by increasing the perception of choice flexibility.
Final_Consent_Choice (Q25)	The user's typical choice (Accept, Reject, etc.).	**H3b/RQ3:** This is the dependent behavioral measure for H3b. **Result:** The **non-significant Chi-Square** result means that while the revisit design made users *feel* they could change their choice (H3a supported by marginal trend), it did **not** change their actual reported choice behavior (H3b not supported).

You need to include a note about the assumption checks for a complete picture.

The ANOVA relies on the assumption of **homogeneity of variance** (that the spread of scores is similar across all design groups), which is checked by Levene's Test. For two of your significant/marginal findings, this assumption was **violated**:

DV	Levene's Test P-value	Interpretation (Add to your notes)
NASA_TLX_Performance	P=0.0140	**VIOLATED.** Since P < 0.05, the assumption of equal variance is violated. This means the overall P=0.0144 for the ANOVA should be interpreted **with caution**. You may need to note that a non-parametric alternative or a corrected ANOVA (like Welch's F-test) might be necessary to confirm this finding.
NASA_TLX_Effort	P=0.0212	**VIOLATED.** Since P < 0.05, the assumption is violated. The marginally significant $\mathbf{P=0.0502}$ should also be interpreted **with caution**.

L Qualitative Data – Inductive Thematic Analysis

Qualitative data - Inductive Thematic Analysis
High-level overview

- 107 chose to not answer
- 93 chose to answer
    - 21 answers essentially meant “No comments”
    - 72 remaining answers will be grouped into 3 categories

Groupings
We first start by grouping the responses in the following categories:

- Behaviour-related responses (17)
- General opinions on CMPs (34)
    - Positive (0)
    - Negative (27)
    - Neutral (7)
- Opinions on our CMPs (21)
    - Baseline (5)
    - Informative (5)
    - Seamful Automation (5)
    - Seamless Automation (3)
    - Revisitability (3)

Important remark: We choose the one most fitting category for each answer, despite some being eligible for multiple.

Behavior when confronted with CMPs (not too important, we have a question for this)
These responses are not super important, since we already have a question for this specifically - however, some interesting comments were made.

Firstly, users act a certain way but do not necessarily trust that their actions are honoured or are unsure about the consequences of their actions:
“I am not sure how my choices will affect the usability/functionality of the sites”.
“[...] to be honest, I don't 100% trust the website”.

Also, learned helplessness shows in several responses, for instance:
“I am used to seeing cookie pop ups and I almost always accept all cookies. It has become a habitual action.”
“I usually don’t bother with those pop-ups and just accept everything.”

Remaining responses often underline common dark-patterns, such as no “reject all” button being visible in most CMPs they are confronted with, all defaults being “accept”, and burying potential refusal under multiple layers of menus.

Responses:

- If there is a button that I can refuse all, I definitely click it to refuse all. But to be honest, I don't 100% trust the website.
- no, i usually accept them without thinking
- If I can reject all I always do, but if I can't, I accept only the necessary ones. I wish all pop ups can make it easy and quick to just reject all or accept the strict minimum necessary. And I hope what we choose is really respected and that they can't technically bypass it.
- They should all be disabled by default and, if you want to, accept the rest.
- I usually accept cookies if I don't find anything unusual about them
- I usually accept them all because I am not sure how my choices will affect the usability/functionality of the sites.
- I am used to seeing cookie pop ups and I almost always accept all cookies. It has become a habitual action. Thank you
- Cookie popups should be short and direct. And the option to refuse all should exist in every popup - something that doesn´t really happens in a lot of sites
- Honestly if the popup is honest i dont mind accepting necessary cookies, but a lot of times the cookie banner is made a way needing 3 extra steps to deny all, if thats the case oh im sure going to find a way to deny all or straight up leave the side and find an alternative
- I usually accept them all, even though I should be reading what all of the options are about.
- I just always reject them.
- I either reject all or accept the bares of bare minimums.
- On most sites I reject them automatically
- I usually reject them all, unless it's very complicated to do so
- I usually dont bother with those pop-ups and just accept everything.
- For some cookie popups we are forced to individually reject the unwanted cookies which is a wrong way to force customers to click on the "accept all" button just because it's an easy option to do
- In the previous question: I usually choose "reject all" if that is an option, but sometimes the only option is "accept necessary"

Feelings towards CMPs

Some users decided to elaborate on their feelings regarding CMPs in general.

There were no positive comments.

The most prominent keywords and phrases in negative comments were the following: “Hate”, “Manipulative”, “Bad/worst”, “Annoying”, “Waste of time”, “Disruptive”, “Only there to comply with EU law”, “A lot of information”.

Some responses underlines issues described in this paper, such as learned helplessness: “My data will be stolen anyway.”

Neutral answers expressed indifference or some understanding regarding “necessary” cookies, and emphasised that the way information is presented is crucial for the user experience:
“I think cookie popups must be clear and concise for the user and we (the users) must always have the choice to reject them if we want to.”

Responses:
Positive

- Nothing to see here.

Negative

- I hate those.
- They're the worst. They are supposed to solve a problem making it annoying
- I consider them annoying even when necessary, but I would like to have a better knowledge of the best way to handle them.
- I feel like they are quite manipulative: especially since all the choices are automatically ticked on and the "accept all" button is placed in such a way that most people will click it without thinking.
- They are very annoying, but I usually accept them, sometimes I go to accept only necessary, but usually accept all
- Worthless
- I saw so many of them that at this point I barely read what's written on them also because it's always the same disclaimer
- Yes, it is ridicullous to use cookie consent forms when every browser would be perfectly capable of having a default answer for cookie settings. It's a humongous waste of time just because burocrats do not understand how technology works/can work.
- I sometimes feel that I’m obliged to provide certain information, and it can be a bit annoying that I have to do it.
- Cookie popups can be a bit tiring because they appear constantly on the screen when you visit a site, and sometimes on mobile they even glitch and block the page, making it hard to scroll or choose whether to accept or reject anything. (Seamless automation)
- I do not like them, generally I would like to just allow cookies that are necessary for smooth functionality if needed nothing else.
- I would like the popups to be more simple and clear.
- Sometimes the site is not accessible unless you pick an option which becomes disruptive when you're looking up for something only for it to stop you in your tracks
- for the user it should be reversed, it should ask which cookies do you want, not which cookies don't you want. No cookies by default
- the popups are irritating and if you dont focus what you are allowing they can take alot of information from you without you even knowing.
- They are annoying, I prefer to not accept any cookie at all if its not necessary
- It's not that I mind essential cookies being used, but I default to reject all always because it's so easy to visit a website just once and then them having your information forever and you completely forget about it. I really dislike the thought of being tracked and everything I do online being monitored. I think privacy is important.

- My data will be stolen anyway.
- i hate them - they should not exist or i should just choose one in browser - reject all and have that choice past me forever.
- I hate them. Very annoying.
- They usually are very info-heavy and ugly, If there was a read more button, it would look more polished.
- Sometimes the site is not accessible unless you pick an option which becomes disruptive when you're looking up for something only for it to stop you in your tracks
- I don't like them. They feel kinda scary. But what can you?
- They are extremely annoying, and some of them you can't reject all easily
- Personally, I believe that cookie banners are on websites only to comply with the law, and the service provider processes the data anyway. Besides you need a cookie or storage access to process "reject all cookies". Many banners don't give option to reject all, because they claim thay have neccessary cookies, while IMO there is no such thing.
- They are annoying but I know that they are required under EU law. I appreciate the ones that allow me to "Reject All" by simply clicking one button rather than going down a long list and having to uncheck everything.
- Only necessary cookies.

Neutral

- They are necessary and acceptable.
- None in particular, they should be unintrusive and do their job at making me understand what to click without any dark pattern attempt at making me accept them.
- I think they are ok
- Needed but not much
- Just casual cookies popup
- I think cookie popups must be clear and concise for the user and we (the users) must always have the choice to reject them if we want to.
- All of them should have option to "accept minimum"

CMP-variant specific comments

Users had a plethora of interesting comments regarding the CMP variant they were presented with. When analysing the results for CMPs extending on the baseline CMP, we focus on the comments regarding the unique extensions themselves (e.g. automation, dynamic explanations, revisitability).

Regarding the baseline, people expressed their distaste regarding the lack of information presented, both regarding what they were consenting to and what the consequences of their choices were:

“[...] I don't actually know what happened: were the cookies accepted or rejected?”

“The one thing that bugs me about most typical cookie consent forms is the lack of consistency [...]”.

A participant has proposed some interesting discussion points when expressing their preference to browser-level CMPs:
“[...] I have manually gone through hundreds of those "legitimate consent" lists and unchecked every single one. What is legitimate consent? Why does an ad bureau have legitimate consent to put cookies on my computer and/or register information about me? It doesn't feel right. [...]”

As for the informative variant, numerous participants expressed their appreciation of our innovative approach of providing them with dynamic information on the pop-up based on their current selections:
“[...] I liked the explanation of the obstructive/innovative etc. terms; it made it easier to understand what exactly am I rating.”
“[...] I liked that the popups in this survey switched between explanations depending on if the cookies are allowed or not, this made it easy to understand what each slider would do.”
“Very easy to understand and accept.”
“They are helpful in letting me choose”.
“I thought the cookie pop ups in the three websites was really easy to understand and I really liked it. I wish all cookie popups were that easy to use and understand.”

For the seamful automation approach, users expressed their gratitude towards this approach compared to the CMPs they meet in their everyday lives:
“I wish there was a system (browser extension?) that would [do what the seamful automation CMP variant does].”

Interestingly, despite the system status visibility has been boiled down to a bare minimum, people did appreciate its simplicity:
“I liked the blue box explaining briefly what I had accepted. It was clear and concise.”

On the contrary, seamless automation resulted in mixed results, some participants even doubting the system’s correct functionality:
“[...] I am confused with that one in this study.”
“Maybe I did my part wrong [...]”
“Only showed up on the first page (Santander), on the other two did not [...].”

Regarding revisitability, not a lot of comments were made, however one user showed distrust regarding the technology they were presented with:
“I do not trust [the CMP variant] in your study”.”

Responses:
Baseline (confusion about consequences)

- As shown in the 3 websites, I dislike the fact that I can click on the X to close the popup when I go in the "More options" menu, as I don't actually know what happened: were the cookies accepted or rejected? As cookies are important in terms of data, while it is more obstructive, I would rather the popup be unable to close until I actively choose an option.

- I don't like this popup but I know that it's necessary to websites to do it.
- The one thing that bugs me about most typical cookie consent forms is the lack of consistency - it'd be nice if all of them offered a very simple 'Reject All' or 'Accept Only Necessary' without having to open any of their extended screen options. A lot of websites do it right, like the way your study has (having the button upfront, no clicking around), but not everyone does! It's especially frustrating on mobile, in my experience.
- I use Consent-O-Matic and it is an absolute game changer! I have manually gone through hundreds of those "legitimate consent" lists and unchecked every single one. What is legitimate consent? Why does an ad bureau have legitimate consent to put cookies on my computer and/or register information about me? It doesn't feel right. Thank you for working on this! I LOVE people who take our privacy and security seriously and who works to spread awareness and/or change things.
- Once I went into more options, the floating buttons on the bottom were outside of the box. This was unexpected.

Informative

- No but i liked the explanation of the obstructive/innovative etc. terms it made it easier to understand what exactly am i rating
- Idk about normal ones, but i liked that the popups in this survey switched between explanations depending on if the cookies are allowed or not, this made it easy to understand what each slider would do
- Very easy to understand and accept
- They are helpful in letting me choose.
- I thought the cookie pop ups in the three websites was really easy to understand and I really liked it. I wish all cookie popups were that easy to use and understand.

Seamful automation

- That sometimes i do not understand the terminology of asked consent therefore i decline, although it could be something not dangerous for my privacy (Seamful automation)
- In my opinion the pop up was a little bit large but was straight forward to make the option. No hidden "crap" like majority of pop ups we find nowadays. Regards.
- Popups are a little bit unnecessary but im okay with these
- Unfortunately, cookie popups ruin the user experience. I wish there was a system (browser extension?) that would allow the user to set their preferred settings once. Then, the system would automatically communicate the preference to all the visited websites without the need to show the popup every time.
- They are most annoying especially if they have long lists to scroll down. I liked the blue box explaining briefly what I had accepted. It was clear and concise.

Seamless automation

- I accept all and I am confused with that one in this study.

- Maybe I did my part wrong, or it was intended that way, but out of the three examples I was given, only one page offered a pop up with the choices to make. The rest of websites didn't offer any choice to choose or deny the cookies. (Seamless automation)
- Only showed up on the first page (Santander), on the other two did not, so all my user experience is based solely on that one. (Seamless Automation)

Revisitability

- Unlike other sites, I found these banners to be clear and to the point.
- The format a cookies shown on the websites were not very helpful, you pretty much take them or reject them blind. When you enter the more options to see the individual cookies permisions there is nowhere near enough information to make a decision. Even in websites where you get the specifics of who gets the cookies, it doesn't help all that much, you will have no clue who they are or what they will do with them, so you go pretty much blind.
- I do not trust them in your study

M Data Collection

- SESSION_ID – random ID per session
- DESIGN_VARIANT – `baseline`, `automation`, `global_settings`, `informative`, `history`.
- WEBSITE_CONTEXT – `site_1`, `site_2`, `site_3`
- TIMESTAMP – ISO datetime of the interaction.
- EVENT_TYPE – see below
- EVENT_TARGET – see below
- TRIAL_INDEX – what order the website_context is

Event type

`page_loaded` – mock site loaded (optional, but useful as a start marker)
`cmp_shown` – CMP first appears
`cmp_closed` – CMP is closed/hidden (any way)
`choice_made` – user confirms a consent choice (accept/reject/custom)

`button_click` – any button pressed

`toggle_change` – any category/vendor toggle changed (on/off)

`panel_open` – any secondary panel opened (e.g. “More options”, history)

`panel_close` – same panel closed

`notification_shown` – automation/global banner/notice shown

`notification_click` – user clicks that small notification

`feedback_shown` – post-decision feedback bar/summary shown

`link_click` – user clicked a text link (e.g. “view vendors”)

`history_item_update` – user changes something *inside* the history view

`scroll` – user scrolls within CMP (optional, helpful for engagement)

Event target

`cmp_main_banner` – the whole banner container

`cmp_first_layer` – first layer / main view

`cmp_second_layer` – “More options” / detailed view

`btn_accept_all` – “Accept all” button

`btn_reject_all` – “Reject all” button

`btn_more_options` – opens detailed view

`btn_save_custom` – confirm custom/category settings

`btn_close_cmp` – close/X button on CMP

`toggle_necessary`

`toggle_functional`

`toggle_analytics`

`toggle_marketing`

`link_privacy_policy` – “Privacy policy” / “Learn more” link

`notification_auto_handled` – small non-blocking bar/icon that says “privacy assistant handled this”

`btn_open_cmp_from_notification` – button/link on that notification to open the full CMP

`checkbox_use_as_default` – the checkbox/option “Use these settings as my default for similar sites”

`dialog_confirm_global_default` – any confirmation dialog/panel for setting global defaults

`btn_confirm_global_default` – confirm global default in that dialog

`btn_cancel_global_default` – cancel global default

`bar_feedback_summary` – the feedback/summary bar shown after choice

`text_feedback_summary` – main text line (e.g. “We blocked 12 vendors…”)

`link_feedback_vendors` – “View vendors” link in the feedback

`link_feedback_categories` – “View categories” / “Adjust settings” link

`icon_consent_history` – small icon/button that opens the consent history

`panel_consent_history` – the history panel itself

`history_list_sites` – the list container of sites

`history_item_site` – a specific site entry (you may also store a site ID in another field if

needed)

`history_item_category_toggle` – a toggle inside a site's entry (e.g. marketing on/off)

`btn_history_save_changes` – button to save changes made in history

`btn_history_discard_changes` – button to cancel/close without saving

After logging

Decision time – time from `cmp_shown` to first `choice_* or cmp_closed`.

Time on webpage - time from render and click go_to_next

Number of interactions – clicks/scrolls before decision (fatigue proxy).

Revisit behaviour – whether any `history_*` event occurs after initial `choice_*`.

Override behaviour – `choice_automation_overridden` or changes in history.

N Qualtrics

Prolific ID

Q1. **Study on Design Alternatives for Online Consent Forms**

Information about the study and informed consent form

I am Carla Griggio, assistant professor from Aalborg University. Together with my master's students, Bence Szabó, Sara Selman, and Louise Steffens, we are conducting research on design alternatives for online consent forms (or "cookie popups"). We are currently studying what kinds of design aspects in cookie popups help users feel more in control and informed about their privacy choices online. Before you decide, you can talk to anyone you feel comfortable with about the research.

There may be some words that you do not understand. Please feel free to contact us to ask for clarifications.

Purpose of the research and type of activity

This study looks into how Internet users understand, value and feel about different types of consent forms for sharing personal data with websites in the form of cookies. We are designing new technologies for enabling users with more control over their privacy online, and we want to incorporate the insights from this survey into our future designs. The study consists of three tasks, each asking you to visit a website, make a decision about what data you would consent this website to collect, and finally, answering a short set of questions about your experience.

It's important to note that these tasks are aimed to learn about the design of the consent form (cookie popup) you will see, but the websites **will not collect any data regardless of your choices**. However, please complete the task as if these websites would collect data as normal websites do.

Participant selection

You have been invited to this study because you are representative of a group of adult users who have experience with cookie popups compliant with GDPR.

Voluntary Participation

Your participation in this research is entirely voluntary. It is your choice whether to participate or not. You may change your mind later and stop participating even if you agreed earlier.

Procedures and Protocol

The study consists of two parts: First, you will perform three tasks where you will make choices about online consent, each on a different website. Second, you will answer a short set of questions about your experience. We will collect the following data:

1. Demographics (from Prolific: age, gender, employment status, country of residence)
2. Interaction with each cookie popup (e.g., clicks)
2. Questions about your experience

Participants are expected to reflect on each question and provide **thoughtful, personal answers**.

Answers that are generated with AI will be rejected.

Duration

The study will take approximately 5 minutes.

Risks and Benefits

We anticipate no risks from participating in the study. If you participate, you may dedicate time to reflect about what type of privacy decisions you would make when visiting different types of websites.

You will also help the researchers conducting the study advance our understanding of people's privacy choices in online consent forms (cookie popups). This will benefit the technical development of future prototypes.

Reimbursements

You will be awarded £0.75 via Prolific.com after verifying that all questions were answered in a valid way.

Data processing

Your data collected during this study may be shared with researchers at both Aalborg University, for example, in the case that a new researcher joins the research project. These institutions will use the data for the purposes of this research only.

Confidentiality

The data we collect from you will be anonymous, meaning that we will not ask for your name, address, personal ID or anything that identifies you unequivocally. Your data will be associated with a

randomly generated number provided by Prolific.com, which will be deleted after the research project is concluded. All collected data is stored securely in servers hosted by Aalborg University.

Sharing the Results

If you are interested you may contact us at a later stage to learn about the outcomes of the study. We will publish the results in the form of one or more research papers and online articles in order for other interested people to learn from our research. No confidential information will be shared. Publications may include fragments of the text from your answers, but we will remove any information that may identify you. Towards the end of this form we ask for your explicit consent to this: you are welcome to participate in this study even if you do not want your answers to be quoted verbatim in publications.

Right to Refuse or Withdraw

You do not have to take part in this research if you do not wish to do so. You may also stop participating in the research at any time you choose. It is your choice and all of your rights will still be respected. You may withdraw from the study by contacting us to the emails below indicating your Prolific ID. It is important that you understand that if you lose your Prolific ID (e.g., by deleting the account), we cannot withdraw your data from the study.

Who to Contact

If you have any questions about the study at a later stage, please contact:

Principal investigator:

Carla Griggio, Department of Computer Science, Aalborg University,

Denmark, A. C. Meyers Vænge 15, 2450 København, email: cfg@cs.aau.dk.

Q2. Certificate of Consent

I have read the foregoing information. I have had the opportunity to ask questions about it and any questions that I have asked have been answered to my satisfaction.

I consent voluntarily to partcipate as a partcipant in this research and consent the following data to be collected: a) Demographics, and b) interaction with cookie popups and c) questions about your experience

I understand that I can only withdraw from the study by providing my Prolific ID to the researchers

I consent voluntarily to have anonymized fragments of my answers included in publications (quoted verbatim) resulting from this research

Q3. Enter your Prolific ID here:

INTRO

Q4. **Welcome!**

You are invited to participate in a short study about online consent and cookie banners.

You will:

- **Visit 3 websites and interact with the consent pop-up that appears on each one.**
- **Answer a short questionnaire about your experience.**

The study will take approximately 5 minutes.

LINK OPTION 1

Q5. Your task is to **make a choice about online consent** while progressing through 3 different websites.

1. Click **this link** to start.
2. Follow the instructions on the website.

   Important: There is a "NEXT PAGE" button at the bottom of each website!

3. When you are done, you will receive the secret code.

Paste the **secret code here** when you are done:

LINK OPTION 2

Q6. Your task is to **make a choice about online consent** while progressing through 3 different websites.

1. Click **this link** to start.
2. Follow the instructions on the website.

   Important: There is a "NEXT PAGE" button at the bottom of each website!

3. When you are done, you will receive the secret code.

Paste the **secret code here** when you are done:

LINK OPTION 3

Q7. Your task is to **make a choice about online consent** while progressing through 3 different websites.

1. Click **this link** to start.
2. Follow the instructions on the website.

   Important: There is a "NEXT PAGE" button at the bottom of each website!

3. When you are done, you will receive the secret code.

Paste the **secret code here** when you are done:

LINK OPTION 4

Q8. Your task is to **make a choice about online consent** while progressing through 3 different websites.

1. Click **this link** to start.
2. Follow the instructions on the website.

   Important: There is a "NEXT PAGE" button at the bottom of each website!

3. When you are done, you will receive the secret code.

Paste the **secret code here** when you are done:

LINK OPTION 5

Q9. Your task is to **make a choice about online consent** while progressing through 3 different websites.

1. Click **this link** to start.
2. Follow the instructions on the website.

   Important: There is a "NEXT PAGE" button at the bottom of each website!

3. When you are done, you will receive the secret code.

Paste the **secret code here** when you are done:

EUQ-S

Q10. If you are in any doubt, you can about the options below here - otherwise, feel free to skip.

Obstructive - Supportive

Does the consent pop-up slow down or help you quickly continue to the website?

Inefficient - Efficient

The time and effort required to complete your consent choice / assignment.

Conventional - Innovative

Is the design like something you have seen before or does it feel fresh and cleverly designed?

Final - Revisable

Do you perceive your choice as permanent and unchangeable or easily changeable later?

Empowering - Discouraging

Do you feel in control and capable of making an informed choice or do you feel pressured?

Genuine - Manipulative

Does the interface feel honest and transparent or is it trying to trick or steer your choice?

Q11. Please mark the point on the scale that you feel best reflects your impression of the product/system.

Feel free to check the explanations below if you are unsure about what some of the scales represent.

Obstructive ○ ○ ○ ○ ○ ○ ○ Supportive

Complicated	○ ○ ○ ○ ○ ○ ○	Easy
Inefficient	○ ○ ○ ○ ○ ○ ○	Efficient
Clear	○ ○ ○ ○ ○ ○ ○	Confusing
Boring	○ ○ ○ ○ ○ ○ ○	Exciting
Not Interesting	○ ○ ○ ○ ○ ○ ○	Interesting
Conventional	○ ○ ○ ○ ○ ○ ○	Inventive
Final	○ ○ ○ ○ ○ ○ ○	Revisable
Empowering	○ ○ ○ ○ ○ ○ ○	Discouraging
Genuine	○ ○ ○ ○ ○ ○ ○	Manipulative

NASA-TLX

Q12. Please mark the point on the scale that you feel best reflects your impression of the product/system.

Q13. How mentally demanding was the task?

Very Low	○ ○ ○ ○ ○ ○ ○ ○ ○ ○	Very High

Q14. How physically demanding was the task?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q15. How hurried or rushed was the pace of the task?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q16. How succesful were you in accomplishing what you were asked to do?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q17. How hard did you have to work to accomplish your level of performance?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q18. How insecure, discouraged, irritated, stress, and annoyed were you?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q19. To what extent do you believe that the given system acted on your behalf as you wanted it to?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q20. To what extent did you feel that you were in control of the choices you made?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q21. How understandable was the information presented in the pop-ups?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q22. How confident are you of the reasons behind the choices you made?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q23. How well did you understand the effects of the choices you made?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

Q24. To what extent do you feel like you could go back and change your consent choices later?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Very High

CLOSING QUESTION

Q25. What do you usually do for cookie banners in your daily life?

- Accept all
- Reject all
- Accept necessary
- Manually choose which ones to accept
- Close the banner

Comments

Q26. Would you like to share any other comments about cookie popups?

Powered by Qualtrics