UHASSELT
KNOWLEDGE IN ACTION

Faculty of Business Economics

Master of Management

Master's thesis

Sensory Sensitivity in Digital Marketing: A Path to Better Engagement

Auzee Haziqah Rosmadee

Thesis presented in fulfillment of the requirements for the degree of Master of Management, specialization International Marketing Strategy

SUPERVISOR :
dr. Carmen ADAMS

2024
2025

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

Faculty of Business Economics

Master of Management

Master's thesis

Sensory Sensitivity in Digital Marketing: A Path to Better Engagement

Auzee Haziqah Rosmadee

Thesis presented in fulfillment of the requirements for the degree of Master of Management, specialization International Marketing Strategy

SUPERVISOR :

dr. Carmen ADAMS

Acknowledgements

I am deeply grateful to everyone who supported the creation of this thesis. This journey has been more than academic; it has been personal, reflective, and transformative.

First, I sincerely thank my supervisor, Carmen Adams, and the faculty at Hasselt University, whose thoughtful guidance and encouragement helped shape this work into a more inclusive and meaningful study of digital experience.

To the classmates, friends, and personal networks who participated in the survey and generously helped distribute it, thank you. Your trust and time made this research possible. I am also grateful to Katerina Marie Hope for allowing me to use our photo from Kyoto on the study website.

A heartfelt thank you goes to the global neurodivergent community. As a neurodivergent researcher with sensory processing differences, this thesis grew from lived experience and from a desire to make digital spaces more emotionally accessible. I worked with neuro-accessible learning strategies that fit my brain, including storytelling as structure, visual metaphors, and dialogue-based paraphrasing to process complex theory. These methods helped me stay present and engaged while upholding academic standards. I affirm that neurodivergent researchers need not mask or conform to traditional learning styles to produce rigorous academic work. Alternative pathways are valid, ethical, and often essential.

A note of inspiration

This project was also shaped by the creative philosophy of David Lynch, whose reflections on attention and ideas helped me tolerate ambiguity and go deeper during analysis. As he writes, *"Little fish swim on the surface, but the big ones swim down below. If you can expand the container you are fishing in, your consciousness, you can catch bigger fish."* That simple metaphor guided my approach to sensory personalisation: slow down, go deeper, and let the essential signals emerge.

AI and writing-support disclosure

This thesis was developed with the support of transparent, ethical tools. I used OpenAI ChatGPT to brainstorm, clarify ideas, refine passages for academic tone, and surface literature for me to verify, especially in sensory marketing, user experience, and neurodivergent accessibility. Representative prompts included: "Refine this for a specialist audience," "Challenge this claim and suggest counter-arguments with sources to check," and "Explain this section in a playful way to test my understanding." I also used Grammarly to improve clarity, grammar, and tone. These tools acted as digital co-editors, not replacements for critical thinking. All ideas, structure, analyses, and final claims are my own. I reviewed every suggestion for accuracy and appropriateness and verified all references independently.

Neurodivergent learning supports

Throughout this project, I relied on strategies that supported memory, focus, and stamina. I used narrative dialogue and visual storytelling to structure complex sections, and I sometimes prompted AI tools to explain theory in character-driven formats to reinforce comprehension. I also used Bitrunners, a game with programming-like mechanics, to practice conditional reasoning, variable state, and flow logic, translating into better comfort with Qualtrics branching and SPSS syntax. Repeated practice helped me script screening rules, build scales, and specify needed variables and commands.

Technical and research tools

The study websites with adjustable sensory features were prototyped in Framer (free tier). Interfaces were designed toward selected WCAG 2.1 AA contrast and operability checks using SuperColorPalette; this was not a formal conformance audit. Survey deployment and data collection were conducted in Qualtrics. References were managed in Zotero for APA consistency. Milanote.com was also used to organise thesis notes, and the layout was designed to be used as a visual library. SurveySwap.io supported participant recruitment. The Sensory Hypersensitivity Scale (SHS) by Dixon et al. (2016) characterizes sensory traits. Data analysis was conducted using IBM SPSS Statistics.

Website content

All travel descriptions, itineraries, and interface microcopy used in the prototype websites were authored by me and draw on my personal travel experiences and field notes. All text was edited for clarity and realism and does not reproduce third-party marketing materials. Where personal anecdotes could identify individuals, details were generalised to protect privacy.

Image and media credits

Unless otherwise noted, photographs and illustrations in this thesis and on the study websites are my original work. The personal photograph (myself and my friend Katerina) used to visualise Japan on the website appears with permission from Katerina Marie Hope. Prototype pages embedded official promotional videos from Brunei Tourism, the Japan National Tourism Organisation, and VISITFLANDERS, as well as an ambient track by Alex-Productions, all streamed directly from YouTube without download or alteration. Use was exclusively for non-commercial academic illustration within the prototype, where all rights remain with the respective copyright holders. Where required, attribution was provided in the site, and full APA references appear in the References list. Also, images from flaticon.com were used for the thesis poster.

Ethics

This research was conducted in alignment with the ethical procedures of Hasselt University. Participation was anonymous and voluntary, with informed consent obtained before data collection.

To everyone who supported me academically, emotionally, or technologically, thank you. Your presence enabled me to approach this work with depth, integrity, and self-belief. This thesis exists because of your guidance, patience, and shared commitment to inclusive knowledge.

Table of Contents

Chapter 1: Introduction

1.1 Background

Do loud, bright, and dynamic interfaces excite engagement or push users to close the tab (Norman, 2004; Nowakowski, 2021)? Since the late 1990s, "emotional design" has leveraged bold animation, autoplay media, and saturated colour to capture attention (Leiner et al., 2009; Norman, 2004; Sapega, 2020). By the 2010s, however, evidence showed that overstimulation can elevate cognitive load and erode trust, prompting a shift toward "calm" design even as high-intensity patterns persisted in entertainment, e-commerce, and travel (Cyr et al., 2007; Ortiz-Escobar et al., 2023; Stead et al., 2022; Wörfel et al., 2022). Relatedly, research on deceptive patterns highlights how sensory intensity can be used coercively, further undermining trust (Gray et al., 2018; Narayanan et al., 2020).

Crucially, accessibility compliance does not guarantee comfort. The European Accessibility Act and WCAG 2.1 AA emphasise perceivability and operability but do not limit sensory intensity; a site may conform yet still overwhelm via rapid motion, intrusive autoplay, or harsh contrast (European Parliament & Council of the EU, 2019; World Wide Web Consortium, 2018; Motti, 2019; Rodicio, 2024). This compliance–comfort gap matters because sensory sensitivity reflects measurable differences in perception and arousal regulation, not mere preference (Alzahrani et al., 2021; Dixon et al., 2016). For autistic and other neurodivergent users, a single uncontrolled stimulus (e.g., abrupt autoplay) can trigger avoidance and disrupt decision processes; even non-clinical users report lower perceived usability under excessive motion or audio (Lionetti et al., 2019; Nowakowski, 2021; Ortiz-Escobar et al., 2023).

Two frameworks motivate a practical response. The Sensory Hypersensitivity Scale (SHS) provides a validated, non-clinical measure across sensory domains (Dixon et al., 2016), while Self-Determination Theory (SDT) argues that supporting autonomy and environmental regulation fosters well-being and trust (Alberts et al., 2024; Hutmacher & Appel, 2023). Taken together, they position perceptual controls as tools of emotional accessibility. Historically, personalisation has targeted *what* content to show rather than *how* it is perceived; this study tests whether simple, discoverable controls-mute, reduced motion, and a darker theme-measurably improve comfort, trust, engagement, and commercial outcomes.

1.2 Research Problem

The research problem is twofold:

(1) Default high-stimulation interfaces produce avoidable distress and disengagement for a sizeable subset of users, especially those high in sensory sensitivity
(2) Even when sites meet formal accessibility requirements, the absence of perceptual controls depresses comfort, trust, engagement, and purchase intention.

In practice, a travel site may be fully navigable yet still overwhelm via autoplay media, rapid motion, or harsh brightness; without a way to tone these down, perceived autonomy is low (Deci & Ryan, 2000), which undermines trust (Cyr et al., 2007; McKnight et al., 2002) and can be perceived as manipulative or "dark" design (Mathur et al., 2019; Gray et al., 2018). Current frameworks mandate access but not intensity-compliance, which can coexist with overload (European Parliament & Council of the EU, 2019; World Wide Web Consortium, 2018; Rodicio, 2024).

A pragmatic response is user-adjustable sensory controls (e.g., mute, reduced motion, darker theme) to let individuals calibrate presentation without sacrificing stimulation for those who prefer it. What remains uncertain-and thus motivates this study-is whether such controls measurably improve user evaluations, and which users are most likely to use them.

1.3 Research Questions and Hypotheses

Research focus. This study tests whether letting users adjust a website's sensory settings (**brightness, motion, sound**) changes their evaluations (**comfort, trust, engagement, purchase intention, word of mouth**). It also examines who uses these controls (**sensory sensitivity, age, gender, neurotype**) and whether highly sensitive users rate the default high-intensity site lower at baseline. Perceived control was measured once, immediately after the toggle screen, and reported descriptively only. Primary outcomes (**1–5**) are analysed as within-person change scores (**post-baseline**). Each hypothesis is tested independently (no combined path mode will be introduced).

Research questions.

- **RQ1:** Do post-adjustment ratings improve versus baseline within the same participants?
- **RQ2:** Do individual differences (sensory sensitivity, age, gender, neurotype) predict the number of changes selected?
- **RQ3:** Do baseline ratings of the default site differ across sensory-sensitivity tertiles?

Hypotheses.

- **H1 (within-subjects improvement):** Mean change > 0 for
  H1a comfort; **H1b** trust; **H1c** engagement; **H1d** purchase intention; **H1e** word of mouth.
  Rationale: Autonomy over sensory input should raise comfort and downstream appraisals.
- **H2 (predictors of number of changes):**
  H2a higher SHS → more changes; **H2b** older age → more changes; **H2c** gender groups differ; **H2d** neurotype groups differ.

- **H3 (baseline by sensitivity):** Higher-sensitivity users evaluate the default site less favourably at baseline (gradient: high < mid < low).

1.4 Theoretical Framework and Research Design

1.4.1 Theoretical framework

Several theoretical lenses inform this research:

- **Self-Determination Theory (SDT).** Autonomy, competence, and relatedness increase motivation and trust; perceptual controls (mute, reduce motion, darker theme) operationalise autonomy in web contexts (Deci & Ryan, 2000; Hutmacher & Appel, 2023; Alberts et al., 2024).

- **Sensory marketing / S-O-R.** Stimuli shape internal states (arousal, stress, perceived control) and responses; effects often follow an inverted-U (Too little stimulation = **bored**. A calibrated, medium level = **engaged/comfortable** (sweet spot). Too much = **overwhelmed/irritated**.), so calibrated maximal stimulation is optimal (Krishna et al., 2016; Antunes & Veríssimo, 2024; Wang et al., 2024).

- **Digital trust and UX.** Comfort and credible presentation underpin trust; poorly designed or opaque toggles can backfire by lowering perceived control (McKnight et al., 2002; Cyr et al., 2007).

- **Accessibility and ethics.** Standards ensure access but not intensity options; emotional accessibility argues for user-led regulation of sensory load (W3C, 2018; European Parliament & Council of the EU, 2019).

- **Integration.** Sensory personalisation is framed as emotional accessibility: simple, intelligible controls align SDT's autonomy with optimal stimulation to improve experience across diverse users.

1.4.2 Research design

Design. Quasi-experimental within-subjects pre–post. All participants first used Website A (high-intensity default). They then saw a "Customise Your Experience" screen with three binary options: Mute background music, Reduce motion, and Darker theme. Any combination (0–3) could be applied. Second exposure was either A again (no changes) or one of seven variants B1–B7 reflecting the chosen combination. Content and layout were held constant; only sensory intensity changed.

Prototype conditions.
B1 darker; B2 reduced motion; B3 muted audio; B4 darker+motion; B5 darker+audio; B6 motion+audio; B7 all three.

Measures. Baseline and post outcomes (comfort, trust, engagement, purchase intention, word of mouth). Single-item perceived control was collected immediately after applying toggles (descriptive only). SHS (25-item) and demographics at the end.

Flow. Consent → baseline exposure (A) → baseline survey → toggle screen (record CHG_N = 0–3) → perceived-control item → second exposure (A or B1–B7 with on-page confirmation) → post survey → SHS + demographics → optional feedback + debrief.

Rationale for self-selection. Allowing participants to choose mirrors real-world autonomy and addresses who opts in to adjustments (H2). Within-person change addresses H1; baseline group differences across SHS tertiles address H3.

Chapter 2: Literature Review

2.1 Theoretical background

2.1.1 Sensory Sensitivity and Human–Computer Interaction (HCI)

Definition and scope. Sensory processing sensitivity (SPS) is a stable, non-pathological trait in which stimuli are experienced as unusually intense or overwhelming (Aron, 1997). Qualitative accounts describe frequent overstimulation and compensatory strategies such as lowering brightness or muting sound. Although not synonymous with disability, atypical sensitivity is common across neurodivergent conditions, including autism and ADHD (Alzahrani et al., 2021; Bas et al., 2021). In digital contexts, high-stimulation interfaces may deter or exhaust such users long before classic usability issues arise (Nowakowski, 2021).

Prevalence and measurement. SPS is estimated at roughly one-fifth to one-third of the population; gender differences are small, and age effects are mixed (Aron, 1997; Greven et al., 2019). The 25-item Sensory Hypersensitivity Scale (SHS) shows good internal reliability and enables segmentation without clinical labels, supporting analyses of how interface design affects users across sensitivity levels (Dixon et al., 2016).

Implications for digital design. Sensitivity becomes salient when animated, bright, or noisy interfaces provoke discomfort or fatigue (Nowakowski, 2021). Standard accessibility features often focus on assistive-technology compatibility rather than sensory comfort. Accounting for SPS reframes accessibility as not only operability but also emotional safety (Alzahrani et al., 2021; Bas et al., 2021).

2.1.2 Autonomy, self-determination and personalisation

Autonomy support in HCI. Self-Determination Theory (SDT) holds that autonomy, competence, and relatedness are basic psychological needs; designs that support these needs enhance motivation and well-being (Deci & Ryan, 2000). Recent HCI scholarship positions SDT as a basis for responsible, needs-supportive technology (Peters & Calvo, 2024; Hutmacher & Appel, 2023). On the web, autonomy translates into meaningful control over presentation and interaction-e.g., muting sound, reducing motion, or switching colour schemes-so users co-create the experience rather than endure it.

Choice without overload. Adding options can backfire if they are confusing or excessive. Limiting controls to a few intelligible, discoverable toggles reduces decision friction while still conveying respect for user agency (Hutmacher & Appel, 2023). Per SDT, any momentary effort of choosing can be offset by the downstream benefit of a better-fitting experience; this study, therefore, captures perceived control immediately after the choice and evaluates outcomes after exposure.

Relatedness and trust signals. Personalisation also communicates care-"we see you; your comfort matters"-which can strengthen trust alongside social presence cues (Cyr et al., 2007). In an era of AI-generated content, authenticity cues (e.g., human authorship and provenance) may be necessary complements to low-stimulation modes to avoid scepticism (Brauner et al., 2023; Rae, 2024).

2.1.3 The evolution of sensory marketing

Historical roots. Sensory appeals long predate "sensory marketing." Historical analyses trace a shift from disciplining to engaging the senses, as early department stores invited seeing and touching merchandise-an enduring "race to embrace the senses" (Howes, 2017).

Current research and limits. Multisensory campaigns can lift recall, attitudes, and purchase intentions (Krishna et al., 2016; Stead et al., 2022). Online, however, experiences are mostly visual–aural; haptics, smell, and taste remain hard to simulate (Li et al., 2022; Kaushik & Gokhale, 2021). This creates an "intangibility gap" for categories where nonvisual senses matter.

Inverted-U stimulation. The stimulus–organism–response (S-O-R) model predicts that stimuli (e.g., brightness, music) shape internal states and behaviours (Wang et al., 2024). Evidence often follows an inverted-U: moderate arousal engages; excessive arousal triggers withdrawal (Reynolds-McIlnay et al., 2017). For highly sensitive users, the tipping point occurs at a lower intensity. Calibration-not accumulation-of cues is therefore essential (Kaushik & Gokhale, 2021). User-adjustable controls operationalise this calibration, letting individuals dial brightness, motion, or sound to their personal optimum.

2.1.4 Related concepts in HCI

Digital trust and dark patterns. Trust rests not only on function but on how users feel treated (McKnight et al., 2002; Cyr et al., 2007). High-intensity, unsolicited stimuli (e.g., autoplay, flashing urgency) can feel coercive. Research on deceptive design documents short-term manipulation but long-term distrust and avoidance (Gray et al., 2018; Narayanan et al., 2020). Sensory overload can thus be understood as a consent problem, not just an aesthetics problem.

Authenticity and AI. Minimalist interfaces may inadvertently read as templated or AI-made, reducing credibility (Brauner et al., 2023; Rae, 2024). Calm modes should be paired with humanising content and provenance to sustain trust.

Inclusive personalisation. Personalisation is increasingly framed as inclusion. Giving neurodiverse users control over animation and colour intensity improves comfort and engagement (Motti, 2019; Çorlu et al., 2017). Older adults similarly benefit from adjustable layouts and text (Hanson & Crayne, 2005). Features designed for the margins (e.g., dark mode, reduce-motion) often benefit the mainstream (Lazar et al., 2015; Wood et al., 2024).

2.2 Sensory sensitivity in digital experience design

2.2.1 Prevalence and demographics

A sizeable minority-across genders, ages, and cultures-scores high on SPS, including many neurodivergent individuals (Aron, 1997; Greven et al., 2019; Alzahrani et al., 2021; Lopez-Herrejon et al., 2018). Because sensitive users often avoid overstimulating sites rather than complain, usage data and feedback loops can underrepresent their needs. Treating sensitivity as mainstream variability-not niche pathology-aligns design with actual audience diversity.

2.2.2 Barrier and opportunity

For sensitive users, autoplay video, rapid animation, and high-contrast visuals can function like barriers, prompting immediate bounce. The absence of feedback hides the problem. Conversely, simple chill options (dark mode, reduce motion, persistent mute) can unlock participation and are widely adopted beyond target groups (Lazar et al., 2015; Wood et al., 2024).

2.2.3 Commercial digital platforms and sensory load

Digital marketing often escalates stimulation-music, parallax, countdowns-to win attention. This may lift clicks but can degrade long-term trust and retention among sensitive users (Cyr et al., 2007; Nowakowski, 2021). Given the web’s visual–auditory bias (Kaushik & Gokhale, 2021; Li et al., 2022), designers may overcompensate by “maxing out” those channels. User-side controls provide a practical safety valve, balancing reach with respect for comfort.

2.2.4 Emotional safety, trust, and authenticity

Emotional safety-the expectation of no unpleasant sensory surprises-underpins digital trust. Overuse of attention-grabbing effects, especially with coercive calls-to-action, erodes that safety (McKnight et al., 2002; Cyr et al., 2007). Visible, effective controls signal respect and can bolster trust, provided calm modes also retain authenticity cues (Brauner et al., 2023; Rae, 2024).

2.2.5 Deceptive patterns and sensory coercion

Deceptive patterns frequently leverage intensity-pulsing icons, confirmshaming, countdowns-to compress deliberation (Gray et al., 2018; Mathur et al., 2019; Narayanan et al., 2020; Knyzelis, 2024). For sensitive users, the combination of cognitive pressure and sensory assault can trigger flight rather than informed choice. Ethical UX avoids such tactics and favours user-led controls.

2.2.6 AI provenance and authenticity

Because calm UIs can be mistaken for AI-generated or low-effort, transparency about authorship and human curation should accompany sensory personalisation (Brauner et al., 2023; Rae, 2024). In our study context, several participants voiced precisely this concern, underscoring the need to pair "calm" with credibility.

2.3 Ethical and legal considerations

2.3.1 Evolution of digital accessibility

Policy frameworks have progressively codified digital accessibility (U.S. Department of Justice, 2022; European Parliament & Council of the EU, 2016, 2019; World Wide Web Consortium, 2018; European Telecommunications Standards Institute, 2021; W3C Web Accessibility Initiative, n.d.). WCAG and related standards emphasise perceivability and operability (e.g., text alternatives, contrast, keyboard access) but do not constrain sensory intensity. Consequently, a site can comply yet still overwhelm. This compliance comfort gap motivates extending practice toward emotional accessibility.

2.3.2 Emotional accessibility and dark patterns

Emotional accessibility calls for interfaces that avoid distress, manipulation, and overload-protecting users' "emotional sovereignty" (Sapega, 2020; Rafijevas & Razbadauskaite Venske, 2024). Sensory controls operationalise this principle by enabling self-regulation of affective impact. Such measures complement existing safety criteria (e.g., flash thresholds in WCAG) and point to future guidelines that also address intensity and surprise.

2.3.3 Cultural and contextual dimensions

Sensory norms vary by culture and context: exuberant colour and music can read as prestige in one market and as garish in another. People also differ in desire for control; some relish customisation, others prefer strong defaults (Lazar et al., 2015). A "calm-by-default, adjustable-by-choice" strategy balances these differences. Neurodiversity research further suggests targeted options (e.g., motion types, colour combinations) can mitigate discomfort for some groups (Fialkowski & Schofield, 2024; Motti, 2019).

2.4 Integrative framework and research gap

2.4.1 Synthesising autonomy and sensory marketing

Sensory marketing optimises stimulation for persuasion, while personalisation/SDT emphasises user agency. Integrating these streams reframes personalisation as *perceptual* as well as *content*-based: give users control of brightness, motion, and sound to meet heterogeneous optima (Deci & Ryan, 2000; Hutmacher & Appel, 2023; Kaushik & Gokhale, 2021; Krishna et al., 2016). The goal is not a single "best" intensity but a **choice** that yields net benefits across a diverse audience.

2.4.2 Unanswered questions and rationale

Three empirical questions follow. First, do simple sensory controls measurably improve outcomes (comfort, trust, engagement, purchase, word-of-mouth) within users compared to their baseline experience? Second, who uses these controls-do SPS, age, gender, or neurotype predict uptake? Third, do highly sensitive users start at a disadvantage under high-stimulation defaults (Nowakowski, 2021; Dixon et al., 2016; Greven et al., 2019)? A within-subject design with self-selected adjustments directly addresses these gaps.

2.4.3 Contributions of this framework

1. **Choice architecture matters.** The effectiveness of autonomy support depends on *how* and *when* options are offered, not merely on their existence (Hutmacher & Appel, 2023).

2. **Sensory ease ≠ automatic persuasion.** Comfort may not translate into trust or purchase without credibility scaffolding (Cyr et al., 2007; Rae, 2024).

3. **Universal design with targeted benefits.** Features born for sensitivity (dark mode, reduce-motion) generalise broadly (Lazar et al., 2015; Wood et al., 2024).

4. **Integrating S-O-R and SDT.** Calibrated stimuli (S-O-R) and needs-support (SDT) jointly shape outcomes; without competence and relatedness cues, sensory tuning alone may underperform (Deci & Ryan, 2000; Wang et al., 2024).

These propositions set up the methodology (Chapter 3), results (Chapter 4), and discussion (Chapter 5), where this study tests whether user-adjustable sensory controls improve experience on average and for whom.

Chapter 3: Methodology

3.1 Introduction

This chapter describes the design, sampling, and analytic procedures used to investigate whether user-adjustable sensory controls improve comfort, trust, and engagement on a high-stimulation travel website. Consistent with the theoretical framework outlined in Chapters 1 and 2, the study employed a quasi-experimental pre–post design in which all participants first experienced a default high-intensity site (Version A) and then, after baseline measurements, were offered simple perceptual controls to mute background audio, reduce motion effects, and/or switch to a darker theme. Participants could apply any combination of these controls or none at all, according to their preference. They then explored either the *same* high-stimulation site again (if no changes were selected) or one of seven modified versions (B1–B7) reflecting their chosen adjustments. This within-subjects design isolates the effect of presentation intensity while holding content and layout constant, allowing examination of perceptual autonomy without confounding content personalisation.

The decision to allow self-selection of adjustments rather than random assignment reflects the study's interest in *who* chooses to customise their experience. A true randomised controlled trial was considered but deemed not feasible or ethical because forcing sensory changes on unwilling participants would violate the autonomy focus derived from Self-Determination Theory (SDT). Instead, every participant had the opportunity to tailor the site to their liking, which also mirrors real-world scenarios where users can often choose preferences. Nevertheless, by measuring the same outcomes before and after the adjustments within individuals, the design permits strong inferences about within-subject changes (Hypothesis 1). Meanwhile, between-subjects analyses (Hypotheses 2 and 3) examine how individual traits such as sensory sensitivity, age, gender, and neurotype predict adjustment behaviour and baseline evaluations when given the choice. The single-item perceived control measure (taken immediately after the toggle screen) serves as a manipulation check; it is reported descriptively in results but, as planned, was not used in hypothesis testing or included in the causal models.

3.2 Participants and Sampling

3.2.1 Recruitment and inclusion criteria

Participants were recruited through a combination of convenience and snowball sampling methods. Channels included internal university communications at Hasselt University, personal networks (friends, classmates), professional networking posts (e.g., on LinkedIn), targeted community forums (e.g., Reddit threads related to academic surveys, UX, or neurodiversity), WhatsApp groups, and the SurveySwap.io platform. SurveySwap operates on a reciprocity model where researchers earn responses by completing others' surveys, rather than providing monetary incentives. This approach is common in digital research and can yield relatively diverse samples in terms of geographical and background reach (Thornton et al., 2016).

This study followed snowball sampling guidelines to ensure clarity and inclusivity in the recruitment process (Ting et al., 2025), encouraging participants to pass along the survey link voluntarily to anyone who might be interested within their radius. No strict quotas were set for demographics; instead, this study aimed for at least 100 completed responses to provide adequate power for within-subject comparisons (paired statistical tests). Ultimately, 146 individuals provided complete usable data (see below).

Inclusion criteria were implemented to ensure data quality and consistent exposure to the experimental stimuli: (a) use of a desktop or laptop computer (no mobile phones or tablets), (b) using Google Chrome as the web browser, (c) audio enabled on the device (so that background music would be audible), and (d) agreement to maximise the browser window during the site experiences. These criteria were necessary because the presence and effect of motion, colour rendering, and audio playback can differ markedly across devices and browsers; this study needed a uniform technical context so that every participant experienced the site as intended. The Qualtrics survey automatically blocked mobile devices and unsupported browsers to enforce consistency of the sensory presentation. Participants were informed of the requirements at the start of the survey and could opt out or exit if they did not meet them. As a result of these checks, out of 270 survey starts, 146 cases passed all inclusion and quality checks and provided complete baseline, post, and demographic data.

This final sample of 146 is relatively large for an academic user study of this nature and was achieved through the broad online recruitment approach. The sample was international (responses came from multiple countries, given the open online recruitment) and demographically varied, though it leaned toward younger adults (reflecting the channels like university and SurveySwap, which tend to have many student-aged users). There was a roughly balanced gender distribution (with participants identifying as female making up slightly over half), and approximately one-third of participants self-identified as neurodivergent (including autism, ADHD, or other forms of neurodiversity). In terms of age, participants ranged roughly from their late teens to their late 50s, with a median in the mid-20s. While not a probability sample, this group provided a useful cross-section of individuals likely to use a travel website and encompassed both people who might especially benefit from sensory adjustments and those who might not think to use them.

3.2.2 Demographic and neurotype measures

Demographic questions at the end of the survey collected age (as a numerical open response, in years), gender (with options: female, male, non-binary/other, prefer not to say), country of residence, and neurotype identification. Neurotype was self-reported with inclusive, user-centric options: *Autistic*, *ADHD*, *Other neurodivergent condition*, *Neurotypical*, and *Prefer not to say*, with the ability to select multiple options if applicable (acknowledging that neurodivergent conditions can co-occur). For analysis, a binary variable was created to indicate whether a participant was neurodivergent (any selection of autistic, ADHD, or other neurodivergent was coded as neurodivergent = 1, versus neurotypical = 0). This approach to measuring neurotype is inclusive and avoids imposing a single clinical label when someone might identify with multiple categories (e.g., both autistic and ADHD). It recognises that neurodiversity often co-occurs and that identity can be nuanced.

All participants also completed the Sensory Hypersensitivity Scale (SHS) as described in Chapter 2. The SHS is a 25-item instrument covering auditory, visual, tactile, vestibular, and proprioceptive domains of sensitivity. Each item is a statement (e.g., "I find bright lights overwhelming") rated on a 5-point Likert scale from "Strongly Disagree" (1) to "Strongly Agree" (5). Eight items are reverse-worded (indicating insensitivity) and thus were reverse-scored so that higher scores consistently indicate greater sensitivity.

Reverse scoring on a 1 to 5 scale:

$$x_{rev} = 6 - x$$

An overall sensitivity score (SHS_TOTAL) was computed for each participant as the average of all 25 item ratings (after reverse-scoring where appropriate). Higher SHS_TOTAL values indicate greater sensitivity to sensory stimuli.

Total SHS score:

$$SHS\_TOTAL_I = \frac{1}{25}\sum_{j=1}^{25} x^*_{ij}$$

where x^*_{ij} is reverse scored when applicable.

For some analyses (Hypothesis 3), participants were divided into tertiles based on SHS_TOTAL: roughly the lowest one-third of scores labelled "low sensitivity," the middle third "medium," and the top third "high sensitivity." Using tertile groups provides a simple way to compare extremes (high vs. low) while still observing any gradient across three levels. This grouping is somewhat arbitrary but was guided by the distribution of scores and aligns with common practice in HSP research to talk about the highly sensitive versus the rest (Lionetti et al., 2019). This study acknowledges that any cutoff is imperfect, but it serves the purpose of testing H3 (which specifically posited differences among high, medium, and low sensitivity users).

Tertile cut points (define by empirical quantiles)

Let q_1 and q_2 be the 33.3 percent and 66.7 percent quantiles of SHS_TOTAL.

Groups:

Low =$\{$SHS_TOTAL < $q_1\}$

Medium = $\{$SHS_TOTAL ≤ $q_2\}$

High =$\{$SHS_TOTAL > $q_2\}$

3.2.3 Data quality and screening

Data quality procedures followed best practices for online experiments and surveys (Galesic & Bosnjak, 2009; DeCamp & Manierre, 2016). The Qualtrics survey flow itself enforced the device/browser criteria noted above (it automatically terminated the survey with a polite message if someone attempted it on a phone, for example). This study also included **instructed-response items** (attention checks) such as "Please select 'Strongly Agree' for this statement" embedded in the questionnaires to detect inattentive or random responding. Automatic logic checks in Qualtrics flagged any duplicate entries from the same browser (cookies) or IP, and implausibly fast completion times suggested the survey was not read carefully.

Responses failing attention checks, showing straight-lined answers on entire multi-item scales (choosing the exact same option for every item, a potential sign of disengagement), or exhibiting extremely short page viewing times were excluded from analysis. These criteria were implemented to ensure that participants engaged meaningfully with both the website and the survey. Data from participants who did not appear to take the task seriously would distort results, especially given the subtlety of some expected effects. The final sample of 146 mentioned above is after excluding those who failed such checks or quit early.

It should be noted that the use of SurveySwap might bias the sample toward younger, more educated participants (Frandsen et al., 2016). Indeed, many respondents were university students or young professionals. This limitation is acknowledged when interpreting the generalisability of results. While our sample included a variety of people, it was not a random cross-section of all internet users. Additionally, given the recruitment methods, participants may have had some interest in UX or marketing (those were the communities tapped in some cases), which could also influence responses. This study did not find evidence of extreme skew in the data due to these factors, but the context should be kept in mind.

In terms of missing data, all 146 included participants provided complete baseline and post-exposure outcome data by design (this study only kept complete cases for paired comparisons). A small number of participants (around 5 out of the initial 151 who had baseline data) missed the post-exposure block due to a Qualtrics routing error; those cases were dropped from the within-subject analyses but their baseline responses were retained for any analyses not requiring the post measures (for example, baseline group comparisons in H3, which ended up with n slightly larger around 151). Importantly, comparing participants with complete data vs. those who missed post data on key baseline variables (comfort, trust, etc.) showed no significant differences, suggesting the missingness was random with respect to our main constructs.

3.3 Materials and Instruments

3.3.1 Experimental websites

Eight interactive prototype websites were built in Framer, a web-based prototyping tool. All versions shared identical layout, text content, navigation structure, and task flow; only **sensory intensity** differed among them. Version A served as the high-stimulation default: it featured bright colours (a light theme with vivid accent colours), several motion animations (e.g., parallax scrolling effects, slide-in transitions), and background media with audio (a gentle music track that autoplayed on the homepage). Versions B1–B7 each applied one or more adjustments corresponding to the three toggle options introduced in Chapter 1, specifically:

1. **B1: Reduced brightness** – a darker theme version of A (e.g., light backgrounds turned to dark grey/black, text to light grey/white, meeting WCAG contrast).

2. **B2: Reduced motion** – animations were slowed or removed relative to A (for instance, parallax effects disabled, content appearing instantly rather than sliding in).

3. **B3: Muted audio** – background music was turned off by default (no sound plays unless a user manually starts a video with sound).

4. **B4: Reduced brightness + Reduced motion** – both of the above adjustments applied.

5. **B5: Reduced brightness + Muted audio** – dark theme plus no background music.

6. **B6: Reduced motion + Muted audio** – animations off/slow plus no music.

7. **B7: All three adjustments** (Reduced brightness + Reduced motion + Muted audio) – essentially a "calmest" mode with dark theme, no motion effects, and silent background.

Each participant who made changes was automatically routed to the appropriate B-version based on their selections (as described in Procedure). If, for example, someone toggled on "Darker mode" and "Mute background music" but left motion unchanged, they would see B5. Someone who toggled all three saw B7, etc. Participants who made no changes simply experienced Version A again. All versions were hosted and accessible via unique URLs, and Qualtrics logic handled the redirection seamlessly.

Design and validation of differences: The colour choices for the default and dark modes were checked using tools like SuperColorPalette and Accessible Web's contrast checker to ensure they met WCAG 2.1 AA contrast ratios for text/background (at least 4.5:1 for normal text, etc.). This was not a formal WCAG compliance audit (that wasn't the focus), but this study wanted to ensure that the darker theme was legitimately easier on the eyes and still fully legible. The prototypes were also manually tested to confirm that navigation, animations, and toggles functioned as intended. All interactive elements (links, buttons) were consistent across versions. The key difference was simply that in the reduced-motion versions, those elements didn't animate, and in muted versions, they had no sound.

It's worth noting that while the prototypes were designed with basic accessibility in mind (e.g., alt text on images, keyboard navigability was checked), the study's primary focus was the *incremental effect* of sensory controls, not achieving full accessibility certification (e.g., compliance with every criterion of EN 301 549 or WCAG for all disabilities). Participants were expected to be sighted and hearing (given the nature of tasks), so visual/auditory impairments were not specifically accounted for in this experiment's design. However, nothing in the design intentionally excluded those users either, aside from needing audio to hear the music (which wouldn't affect answers if unheard, aside from perceived control possibly).

Overall, the materials provided a controlled way to test the impact of toggles: Version A was intentionally a bit *too stimulating* (for some) to create room for improvement, but not so outrageous as to be unusable (this study wanted ecological validity – it looked like a modern travel site, not a crazy spam site). The B versions provided plausible "accessible" alternatives. The study website content itself was carefully written to be neutral and not bias users toward a particular reaction (e.g., it had pleasant travel info, not extremely exciting or extremely boring content, to avoid content sentiment overshadowing sensory factors).

3.3.2 Sensory Hypersensitivity Scale

The **SHS** was administered at the end of the survey (after the second exposure and post-test questions). As noted above, it comprises 25 items rated on a 5-point scale; eight items are reverse-scored so that higher aggregate scores indicate greater sensitivity. The SHS has high internal reliability (Cronbach's $\alpha \approx 0.81$ in the original development study; Dixon et al., 2016) and serves as a self-report tool for assessing an individual's sensitivity across multiple sensory domains. In our sample, this study will report the observed reliability as well. The SHS allowed us to capture individual differences in sensory sensitivity **without** requiring any clinical diagnoses or labels. It provided a continuous variable (and group categorisations as tertiles) to test hypotheses about moderation (H2 and H3). Given that sensitivity might vary by modality, it's noted that the SHS total score is a coarse measure; however, it was adequate for our main planned analyses. (Future work might look at subscale scores for visual vs. auditory sensitivity, for instance, but that was outside this study's scope.) Participants typically took only a few minutes to fill out the SHS, and it was presented simply as "Sensory Preferences Questionnaire" to not prime them about being "highly sensitive" or anything of that sort during the main tasks.

3.3.3 Outcome Measures

Response format and scoring:

Two-item composite

$$C = \frac{x_1 + x_2}{2}$$

Cronbach's alpha, general k items

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{j=1}^{k} s_j^2}{s_T^2}\right)$$

where s_j^2 are item variances, and s_T^2 is the variance of the summed scale.

Cronbach's alpha, two-item identity

$$\alpha = \frac{2r}{1+r}$$

Where r is the inter-item correlation.

Change score definition used later

For participant i, $d_i = B_i - A_I$

All outcomes used 5-point Likert-type items (1 = Strongly disagree to 5 = Strongly agree) and were scored so that higher values indicate a more favorable evaluation. Unless noted, each construct was a two-item composite (mean of its items). Internal consistency for all composites is reported in §4.4 (all α ≥ .80 at baseline and post).

Trust. Two items adapted from McKnight et al. (2002) and Cyr et al. (2007) captured confidence in the site's integrity and competence (e.g., "I feel confident about the information on this website"; "This website seems trustworthy"). Items were averaged to form the Trust score.

Comfort. Two items captured sensory comfort and ease (e.g., "Using this website is comfortable for me"; "The sensory aspects-brightness, motion, sound-feel right for me"). Items were averaged to index perceived sensory comfort during use.

Engagement. Two items reflected attentional absorption and interest (e.g., "I was absorbed in browsing this site"; "This website held my interest"). Wording followed common HCI/marketing practice for concise engagement assessment (e.g., **Kaushik & Gokhale, 2021**), adapted for brevity.

Purchase intention. Two scenario-based items assessed the likelihood of transacting on the site (e.g., "I would be likely to book something on this site"; "I would consider buying a travel product here"). Items were averaged to form Purchase Intention.

Word-of-mouth (WOM) intention. Two items assessed advocacy (e.g., "I would recommend this website to others"; "I think others would benefit from this site"). Items were averaged to form WOM intention.

Overall impression. A global item captured the overall evaluation (baseline **A_IMP**, post **B_IMP**).

Perceived control (CTRL). Immediately after the toggle screen and before the second site visit, participants rated: "Right now, I feel I have control over how this website looks and behaves." CTRL was reported descriptively as an autonomy manipulation check; thus, it was not included in hypothesis tests.

Administration details. Within each block, multi-item scales were shown with a randomized order to reduce order effects. Overall impression appeared first in each block to capture a global evaluation before specific judgments. Composite scores were computed as the item mean at each time point; change scores used in Chapter 4 were calculated as post minus baseline for the same participant. Full item wording appears in **Appendix A**; reliability statistics are reported in **§4.4** (e.g., Trust α ≈ 0.91/0.93; Purchase Intention α ≈ 0.95/0.94).

Scale composition and reliability details appear in Appendix C, Table C0.2. Computation of change scores (D_X = B − A) and all SPSS COMPUTE statements are documented in Table C0.3.

3.3.4 Tools and Software

The following employed several tools and software packages in designing and executing the study:

- **Prototyping:** Framer (free tier) was used to build and host the interactive websites. Framer allowed for quick development of different versions and provided the ability to incorporate toggles and animations without extensive coding. The prototypes were essentially web pages accessible via unique URLs, which this study integrated with Qualtrics.

- **Survey platform:** Qualtrics was used to host the survey and to embed or redirect participants to the site exposures at the appropriate times. Qualtrics' flow logic managed the random assignment to prototypes (based on toggle choices) and recorded responses. It also implemented the inclusion criteria checks.

- **Data analysis:** IBM SPSS Statistics (Version 29) was the primary software for quantitative analyses, as the main analyses were straightforward GLMs and t-tests. This study also made use of SPSS's general linear model procedure and non-parametric tests as needed. Data from Qualtrics was cleaned and processed in SPSS, and graphs were created using either SPSS or Excel for visualization.

- **Reference management:** Zotero was used to manage references and ensure adherence to APA 7th edition format throughout the writing process. All literature cited in this thesis was stored and organized in Zotero, which helped insert citations and generate the reference list.

Other general productivity tools included Excel (for quick data checks) and basic web tools for color contrast, etc. The combination of these tools ensured that the study ran smoothly (Qualtrics/Framer for data collection) and that the analysis was robust (SPSS for statistics). No custom software was needed beyond what is mentioned, and all analysis steps were documented (e.g., SPSS outputs archived, see Appendix C for key outputs and figures supporting the results).

3.4 Procedure

3.4.1 Pre-visit notice and checklist

After clicking the survey link, participants first saw an **introductory page** describing the study in general terms. It was framed as "a survey about a travel website experience" without initially revealing the focus on sensory toggles (to avoid priming expectations). On this page, participants had to confirm the technical requirements: that they were on a desktop/laptop, using Chrome with sound on, and that they would maximize their browser window. They indicated this by checking a box (e.g., "✔ I confirm that I am using Chrome on a computer with sound, and will maximize my window"). This acted as both consent to those conditions and a subtle reminder to follow them. Only after checking this could they proceed to the next page. This checklist approach was intended to maximize compliance with inclusion criteria in practice (since someone could, in theory, use Chrome on a laptop but not have their sound on until reminded).

After confirming the technical setup, participants proceeded to the actual start of the study. There was also a brief set of instructions telling them that they would visit a website and then answer some questions about it, and to imagine they were casually exploring a travel site for information. No mention was made of sensory sensitivity or toggles at this stage.

3.4.2 Initial exposure to Version A

Participants were then automatically directed to **Version A**, the high-stimulation version of the travel website, via an embedded iframe or a new browser tab (depending on what worked best for their setup; Qualtrics can do either-our implementation favored opening the site in a new tab to allow full-screen viewing, then instructing them to return to Qualtrics when done). They were free to explore the multi-page site at their own pace. The site included several pages (e.g., a homepage with an embedded scenic video and music, destination subpages with images and text, etc.), and participants could click around as they wished. There was no forced task (like "find information X") and no enforced time limit; this was to simulate a naturalistic browsing session where a user might just be skimming a travel site out of interest. Most participants spent a few minutes on this. A timer in Qualtrics was running to capture how long they spent on the site before coming back, primarily to identify any extremely short exposures (which might indicate they didn't really look at it).

This initial exposure established a common baseline for everyone: they all saw the "intense" version first, which this study intentionally kept constant. (This study acknowledges that always showing the intense version first could introduce order effects like habituation or contrast, but this was inherent in our design for H1; future work could counterbalance order, see Chapter 5 discussion.)

3.4.3 Baseline measurement

After participants finished browsing Version A, they returned to the survey (either by clicking a prominent "Next" button that was positioned in Qualtrics below the embedded site, or by manually going back to the Qualtrics tab if they opened the site separately – in either case, Qualtrics advanced once the site was closed or 'Next' clicked). They then completed the **baseline questionnaire** about their experience on Website A. First, they gave an *overall impression* rating of the site (A_IMP). Next, they answered a series of Likert-scale questions for comfort, engagement, trust, purchase intention, and word-of-mouth intent, all explicitly referring to "the website you just used." Each construct's items were grouped together on a page, but the order of constructs was randomised to avoid any systematic bias (for each participant, the block order might vary). At this baseline stage, participants were essentially giving their first-impression evaluations *without having had any control* over the site's default starting sensory settings.

3.4.4 Personalisation Toggle Selection

After completing the baseline questions, participants were introduced to customisation adjustment buttons to select from via the survey screen. This page explained that they could now adjust the website's settings to their liking before continuing. It described the three available sensory toggles in plain language:

- *Mute background music:* "Turn off the background music if you don't want to hear audio."
- *Reduce motion effects:* "Slow down or remove animations if you prefer a more static experience."
- *Darker mode:* "Switch to a darker color scheme if you find the current one too bright."

Each option had a toggle button next to it, which by default was set to "No change" (meaning keep it as it was in Version A). If clicked, it would switch to "Adjust." Participants could leave each as-is or change any combination. This design ensured that doing nothing would result in no changes (so nobody was forced to change anything, preserving user autonomy fully).

Participants made their selections and then were led to their routed website variant after submission. The survey recorded exactly which of the three options were adjusted and computed **CHG_N**, the number of changes (0, 1, 2, or 3) for each participant. Qualtrics logic then branched to the appropriate follow-up. A brief loading message appeared as this study redirected them to the appropriate version of the site for their second exposure.

It's worth noting that this stage was presented neutrally, as a standard part of the site experience (this study framed it like "Customize your browsing"). This study did not explicitly tell them "this is the experiment's key manipulation," of course. This study wanted them to treat it as they would if a website provided optional settings (like YouTube offering a theatre mode or dark mode toggle). The interface was designed to be straightforward so they could understand each option without needing further instruction. This study avoided using the word "accessibility" in the participant-facing text to prevent any bias or social desirability effect (some might think it's only for people with disabilities if phrased that way, or might use it out of a sense of doing the right thing rather than personal comfort).

3.4.5 Perceived control

Immediately after applying the toggles (literally on the next Qualtrics screen, which appeared while the site was loading), participants rated their perceived control over the website's appearance and behavior at that moment. This item ("Right now, I feel I have control over how this website looks and behaves") was answered on a 1–5 agreement scale. The purpose, as mentioned, was to see if using the toggles made people *feel* more empowered (this study expected that those who toggled something might report a higher sense of control than those who left all as "no change"). Regardless, this was purely a manipulation check and an insight into user psychology, not tied to hypothesis tests.

Participants have not yet been given any results of their toggles (aside from the actual effect when they saw the site). This study measured perceived control at this juncture, *before* they navigate the adjusted site, to capture the immediate psychological effect of making a choice (distinct from whatever effect the adjusted site might have). This separation helps interpret results in cases, for example, giving choices paradoxically decreases perceived control due to decision effort (a possibility this study considered from SDT theory).

3.4.6 Second exposure

Participants were automatically routed to the version of the site corresponding to their selections (as described in Materials). Those making no changes effectively repeated Version A; those who made one or more changes saw the appropriate Version B1–B7. They were again free to browse as long as they wanted, with the same site content but now presented differently. On each B-version, this study included subtle indicators confirming the active settings. For example, if the background music was muted, a small mute icon on the page header was shown (or text "Audio: Off" in a corner), and if dark mode was on, obviously the colors changed (plus a text toggle showing "Dark

Mode: On"), etc. This was to ensure participants *noticed* that their choices took effect (without them having to hunt or assume). A few participants might not have immediately realized something (especially if they toggled only audio and happened to move to a page without music anyway) – thus the indicators.

After they finished exploring the site with their applied settings, they returned to Qualtrics (via the Next button on that page or by switching back to the survey tab). To confirm the integrity of the manipulation, the survey asked a verification question: "Which of the following adjustments were active during your second visit to the site? (Check all that apply: Adjust Brightness, Reduced motion, Muted audio, None of the above)." This acted as an attention check and data validation – essentially ensuring that participants were aware of what version they saw (and indirectly confirming no technical routing error occurred). All participants answered in line with their toggles (those who toggled none usually checked "None of the above," those who toggled specific ones checked those), so this study was confident that the intended experiences were delivered.

Pass–fail counts for the manipulation verification (user-check vs recorded toggles) appear in Appendix C, Table C13.1.

3.4.7 Post-exposure outcomes, SHS, and demographics

After the second browsing session, participants completed the **post-exposure survey**, which mirrored the baseline survey. They first rated their overall impression of the site they just used (B_IMP). Then they answered the same sets of items for comfort, engagement, trust, purchase intention, and word-of-mouth, this time explicitly referencing "the site you just used *after applying your preferences*" (wording was adjusted to make it clear these were about the second experience). Because this study asked the same questions twice (baseline and post), it could later compute within-person change scores (post minus pre) for each measure.

Following the outcomes, participants proceeded to the **SHS questionnaire** (25 items as described). This was placed here to avoid any bias it might introduce if done earlier (e.g., if someone answered 25 questions about being sensitive to lights and noise *before* using the site, they might realize the study is about that and behave differently). By putting it at the end, this study ensured it didn't prime their experience of the site or their ratings.

After the SHS, standard **demographic questions** were asked in detail in 3.2.2 (age, gender, country, neurotype). Placing demographics last is a common practice to avoid stereotype priming effects, and because these questions can sometimes feel intrusive or boring, it is best to put them after the main tasks.

An **open text box** was used to invite any feedback participants had about the website or the study. This was optional, but many provided comments (as discussed in the results, these comments provided valuable qualitative insights into their experience and perceptions, explaining some of the quantitative patterns).

Finally, there was a separate link to enter an email for the prize draw (a small incentive offered, such as a chance to win a gift card). The link opened a new form unconnected to their survey responses, ensuring response anonymity was maintained. Participants were debriefed on the final page, which explained the study's true purpose (examining the effects of sensory customization on UX outcomes), and they were thanked for their participation. This study also provided contact information for the researchers and a note that if they felt any discomfort during the study (e.g., from animations or something), they could contact us or take a break (though no one reported serious issues).

Throughout, participants could withdraw at any time by closing the browser; partial data from withdrawn participants were not analysed (only complete cases as noted). The debrief also assured them that no deception was used (aside from not telling upfront that sensory settings were the focus, which this study considered mild concealment rather than active deception and is standard to avoid demand characteristics).

3.4.8 Design notes

The study typically took participants around **10–20 minutes** to complete in total, depending on how long they browsed the site. This relatively short duration was intentional to minimize fatigue. No explicit deception was employed; however, as noted, this study *did not mention sensory sensitivity or the true focus* in recruitment or initial instructions to avoid priming effects. The customization screen was framed as a normal preference step on a travel site rather than an "experimental treatment," again to minimize demand characteristics.

This study took measures to make the experience feel realistic and respectful: the site content was professionally presented with original or licensed images to avoid the perception of a low-effort or "scammy" site (though, as shown in the results, some still felt it looked somewhat generic or AI-like). Nonetheless, some participants did comment on its template-like appearance; this feedback is addressed in Chapter 5.

Open-ended feedback fields allowed participants to express any concerns or guesses about the study's purpose. Notably, no participant indicated suspicion that the study was specifically about sensory overload or their sensitivity; most comments were about the site itself (likes/dislikes) or general study design compliments. This suggests that our cover story and procedure succeeded in not overtly revealing our hypotheses, meaning demand bias was likely low.

In summary, the procedure ensured that each participant: (1) experienced a high-stimulation website, (2) provided baseline reactions, (3) had the chance to tailor the site's sensory settings, (4) experienced the adjusted site (or the same site again if they chose nothing), and (5) provided post-adjustment reactions, along with trait and demographic data. The flow was carefully managed so that comparisons could be made within-person (pre vs post) and between-person (differences by those who did vs didn't change, and by trait sensitivity, etc.). The following section outlines how this study planned to analyse these data to test our hypotheses.

3.5 Data analysis plan

All data were analyzed using IBM SPSS (v29). This study used two-tailed tests for all hypotheses, with $\alpha = .05$ as the threshold for statistical significance (unless otherwise noted for specific post-hoc adjustments). Effect sizes are reported to aid interpretation (e.g., Cohen's d for t-tests and partial η^2 for ANOVA/GLM results).

Assumptions and robustness checks are anchored to Appendix C as follows:

C0.2–C0.3 (scale recipe and compute audit), C1–Figure C1 (normality for change scores), C2 (Levene and Welch for H3), C3–Figure C3 (GLM diagnostics and VIFs for H2), C4 (Poisson robustness), C5–Figure C5 (outliers and influence), C6 (reliability), C7 (descriptives and stability), C8–Figure C8 (H1 paired tests and effects), C9 (H2 summary table), C10–Figure C10 (H3 ANOVA and post-hocs), C11 (non-parametric corroboration), C12 (sensitivity analyses), C13 (data quality and timing).

Assumptions and diagnostics cross-walk

Assumption and robustness checks are documented in Appendix C as follows: normality of within-subject change scores for H1 (Table C1) with Q–Q plots (Figure C1); homogeneity tests and Welch adjustments for H3 (Table C2); GLM residual diagnostics and variance inflation factors for H2 (Table C3); Poisson robustness for the CHG_N outcome (Table C4); and outlier screening via studentized residuals for all models (Table C5). These appendix anchors are cited in the relevant Results subsections.

Below is an overview of how each hypothesis was tested

3.5.1 Samples for each hypothesis

All hypothesis tests are two-tailed with alpha = .05. Effect sizes and 95 percent confidence intervals are reported to aid interpretation.

H1 Within-subject improvement

For each outcome, this study defined the individual change score as

$$d_i = B_i - A_i$$

where A_i is the participant's baseline score, and B_i is the post adjustment score.

The mean change and its standard deviation are:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

and

$$s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2}$$

The primary test for H1 is a paired samples t test, implemented as a one sample test on d_i: $t = \frac{\bar{d}}{s_d/\sqrt{n}}$ with degrees of freedom df=n−1.

The 95 percent confidence interval for the mean change is:

$$\bar{d} \pm t_{.975}, n - 1\frac{s_d}{\sqrt{n}}$$

To communicate magnitude, this study reports Cohen's d for paired designs

$$d_z = \frac{\bar{d}}{s_d}$$

and the Hedges small sample correction

$g = J(df)d_z$ where $J(df) = 1 - \frac{3}{4\,df-1}$

If the distribution d_i materially departs from normality, this study corroborates the conclusion with a Wilcoxon signed-rank test. Let W be the signed sum of ranks of non-zero $|d_i|$. Using the large sample normal approximation,

$$Z = \frac{W - \mu_W}{\sigma_W}$$

With

$$\mu_W = \frac{N(N+1)}{4}$$

And

$$\sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

The corresponding effect size is $r = \frac{Z}{\sqrt{n}}$

Interpretation is straightforward. A positive $\bar{d}$ with its confidence interval entirely above zero supports H1 for that outcome. A near-zero or negative $\bar{d}$ value indicates no improvement or a small deterioration, and small $|d_z|$ values signal limited practical change.

Appendix C: C1 (Shapiro–Wilk), Figure C1 (Q–Q), C8 (paired t), Figure C8 (mean Δ with 95% CI), C11.1 (Wilcoxon corroboration), C12.1 (fastest 5 percent trimmed)."

H2 Predictors of number of changes

The behavioral outcome for H2 is the count of adjustments a participant made, CHG_$N_i \in \{0,1,2,3\}$. The primary specification is a general linear model that treats CHG_N as approximately continuous over this short range:

CHG_$N_i = \beta_0 + \beta_1 SHSTOTAL_i + \beta_2 Age_i + \beta_3 Neurodiv_i + \beta_4 Male_i + NonBinary_i + \varepsilon_i$

Gender is coded with Female as the reference category. Neurodiv is 1 for any self reported neurodivergence and 0 for neurotypical. Model fit is summarized by:

$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$

and the corresponding local effect size:

$f^2 = \frac{R^2}{1-R^2}$

Predictor specific tests are reported as omnibus F tests within the GLM. To monitor redundancy among predictors, this study reports variance inflation factors

$VIF = \frac{1}{1-R_l^2}$

where R_l^2 is from regressing the predictor j on the remaining predictors.

Because CHG_N is a count, a Poisson robustness check is also estimated:

$Y_i \sim Poisson(\mu_i)$ with $log(\mu_i) = \beta_0 + x_i\beta$.

In this log link specification, $exp(\beta_k)$ gives the multiplicative change in the expected count for a one unit increase in predictor k.

Support for H2 would be indicated by a significant positive coefficient for SHS_TOTAL (more sensitive equals more changes) and by significant terms for age, gender, or neurotype where hypothesized. Non significant predictors and low R^2 values indicate that adjustment behavior is largely idiosyncratic in this setting.

Appendix C: C3.1 (GLM), C3.2 (VIFs), Figure C3 (residuals), C4 (Poisson), C5 (influence).

H3 Baseline differences by sensitivity

For H3, participants are grouped into Low, Medium, and High sensitivity by tertiles of SHS_TOTAL. Baseline outcomes are compared across these three groups using one way ANOVA. With group means $\overline{Y}_j$, group sizes n_j, total sample N, and grand mean $\overline{Y}$, the sums of squares are

$SS_B = \sum_{j=1}^{g} n_j(\overline{Y}_j - \overline{Y})^2$ and

$$SS_W = \sum_{j=1}^{j} \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_j)^2$$

with mean squares $MS_B = SS_B/(g - 1)$ and $MS_W = SS_W/(N - g)$.

The F statistic is $F = MS_B/MS_W$

Effect sizes for group differences are reported as

$$\eta^2 = \frac{SS_B}{SS_{tot}}$$

$$\eta_p^2 = \frac{SS_B}{SS_B + SS_W}$$, and

$$\omega^2 = \frac{SS_B - (g-1)MS_W}{SS_{tot} + MS_W}.$$

Where the assumption of equal variances is doubtful, Welch ANOVA is used with software provided Satterthwaite degrees of freedom, followed by Games Howell post hoc comparisons. If distributional concerns arise, a Kruskal Wallis check is added with

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{g} n_j \bar{R}_j^2 - 3(N + 1)$$

and the associated effect size

$$\varepsilon^2 = \frac{H - g + 1}{N - g}.$$

H3 is supported when baseline means differ across sensitivity groups in the hypothesized direction, for example High less than Low on comfort at baseline, with meaningful η_p^2 or ω^2 values and confidence intervals that exclude trivially small effects.

Appendix C: C2.1 (Levene), C2.2 (Welch + Games-Howell where needed), C10.1 (ANOVA), C10.2 (post hoc), Figure C10 (means with 95% CI), C11.2 (Kruskal–Wallis corroboration).

3.5.2 Assumptions, outliers, and robustness checks

This study carried out various assumption checks to ensure the validity of our tests:

- **Normality:** For the within-subject change scores and residuals in the GLM/ANOVAs, this study used Shapiro–Wilk tests and Q–Q plots to assess normality. Given our sample size (~146), slight deviations from normality are not too problematic, but this study was on the lookout for major skewness or kurtosis that could affect t-test/ANOVA validity. In the event of non-normal distributions for change scores (especially if outcomes like trust or purchase intentions had ceiling/floor effects), this study planned to report complementary *Wilcoxon signed-rank* test results for H1 as mentioned, and *Kruskal–Wallis* for H3. This study also considered using bootstrap confidence intervals for means if needed. (see C1, Figure C1)

- **Homogeneity of variances:** For ANOVAs in H3, this study checked Levene's test. If p < .05 in Levene's test, this study reported Welch's ANOVA results instead, which do not assume equal variances. For the GLM (H2), this study inspected residual plots to ensure roughly equal variance of residuals across predictor values. If any one predictor with multiple levels (gender) had heteroscedasticity, a robust regression or non-parametric test (like a Kruskal–Wallis on CHG_N by gender) could supplement, but in practice, this study expected CHG_N variances to be similar across groups. (see C2)

- **Linearity:** For the continuous predictors (SHS and age) in the GLM for H2, this study assumed a linear relationship with CHG_N. This study plotted CHG_N vs SHS and vs age to see if perhaps it was non-linear. Given the limited range of CHG_N, a linear fit is probably fine. If there were concerns, categorizing age or using a quadratic term was an option (not ultimately needed)(see Figure C3)

- **Multicollinearity:** This study computed variance inflation factors (VIF) for the regression model in H2. All predictors being fairly distinct constructs (and since this study effect-coded gender, etc.), this study anticipated VIFs < 2, indicating low multicollinearity. If this study had found any issues (eg, age correlating with SHS strongly), it might have separated those analyses or used sequential models. (see C3.2)

- **Outliers:** This study identified potential outliers via studentized residuals for the GLM and ANOVAs. Any case with a studentized residual magnitude > 3 might be considered an outlier. This study checked if results changed meaningfully when excluding any extreme cases. In practice, removing one or two outliers (if found) and re-running analyses was planned as a robustness check. This study reports whether including/excluding them changed conclusions (it did not in our actual analysis; differences were negligible).(see C5, Figure C5)

- **Fast completion times:** As mentioned, this study already removed those who blew through the study implausibly fast (e.g., someone finishing in under 4 minutes clearly didn't read much). As an extra check, this study did a sensitivity analysis excluding the top 5% fastest completers to see if effect patterns remained.(see C12.1)

Additionally, this study planned a more **advanced robustness check**: an *ANCOVA* for H1 if needed, where this study includes baseline scores as a covariate and compares groups by whether they used adjustments or not, to see if improvements held, controlling for baseline differences. (see C12.2)

However, since this study did a within-subject analysis, this was not necessary for H1. However, it could be for exploring the H1 results by subgroups (like did people who changed three things have bigger improvements than those who changed 1?) – This study could do an ANCOVA with number of changes as a factor and baseline as covariate to check interaction, etc.

In summary, this study approached data analysis with rigor, where parametric tests for planned hypotheses were supplemented by non-parametric tests where suitable, and thorough checks were performed for any violations that might affect interpretation. All significance tests were two-tailed, given that even though our hypotheses were directional, this study wanted to be conservative (and indeed, if effects went opposite, this study wanted to capture that).

3.5.3 Missing data

As noted, an early routing error in Qualtrics caused some participants to miss the post-exposure questions (about 5 out of 151 baseline completers). These cases contributed data to analyses of baseline (and thus to parts of H2 and H3 if they had baseline & traits) but not to H1 or any paired comparisons. For each outcome, this study reports the sample size at baseline and post (they are mostly 151 baseline vs 146 post for each measure). This study examined whether participants with complete vs. incomplete data differed on key baseline variables (comfort, trust, etc., and SHS). There were no significant differences (all $p > .1$), suggesting that the missingness was essentially random with respect to our constructs of interest (likely technical, not because those people had a particular experience). Therefore, this study treated the missing data as ignorable and proceeded with listwise deletion for paired analyses (using $n = 146$ complete cases).

No other missing data occurred on individual items because Qualtrics was set to force responses for all key items (or provide a "prefer not to say" where appropriate), which this study treats as a valid category for gender but didn't include in the analysis. Thus, the data were quite complete.

With the methodology and analysis plan established, this study proceeded to conduct the study and analyze the results. The findings are presented in the next chapter, including descriptive statistics, hypothesis test outcomes, and additional exploratory analyses to illuminate the effects of sensory personalisation.

Exclusion reasons and counts are summarized in Appendix C, Table C13.1.

3.5.4 Design Sensitivity and Bias Mitigations

All participants experienced A then B, and adoption of adjustments was self-selected. Fixed order may induce contrast/habituation; self-selection limits causal attribution. We therefore interpret within-person changes conservatively and report standardized effect sizes with CIs alongside *p*-values. (See **Appendix C** for full diagnostic tables.)

3.5.5 Effect-Size Reporting Plan

For H1 (paired comparisons), we report Cohen's *d* for paired designs (Hedges' correction where available) with 95% CIs. For H2/H3, we report partial η^2 (predictors) and η^2/ω^2 (one-way). Model R^2 and Cohen's f^2 are provided where applicable. Effect sizes are interpreted with conventional benchmarks and practical meaning.

3.6 Expectancy and bias controls

Several measures were implemented to minimise expectancy effects and biases that could threaten the internal validity of the study:

- **Hypothesis masking:** Recruitment materials and instructions deliberately did *not* mention sensory sensitivity, neurodiversity, or the specific focus on toggling sensory settings, in order to avoid priming participants or creating demand characteristics. The customisation screen was presented as if it were a normal preference feature of the website, rather than highlighting it as an "intervention." This way, participants would ideally behave naturally, without trying to conform to what they thought this study was testing.

- **Free exploration vs. task performance:** Participants could browse freely without timed tasks or pressure to achieve something specific. This was intended to reduce any performance anxiety or bias. If this study had given them a task ("find a hotel deal"), some might rush or feel competitive, which could overshadow their sensory experience. By keeping it casual, this study hoped their feedback would be more about genuine comfort/engagement rather than task success.

- **Realistic site content:** The website was populated with reasonably high-quality content and design (e.g., real destination info, attractive images) to avoid users dismissing it as a fake or trivial site. Our intention was to have them react to the sensory aspects in a context that felt plausible. Nonetheless, some did perceive it as somewhat template-like, which this study addresses in the discussion as a limitation. Importantly, this study avoided any extremely polarizing content that could skew results (no very exciting or very boring information, kept it moderate).

- **Post-experiment inquiry:** Through the open-ended feedback at the end, this study effectively conducted a funneled debriefing by giving participants space to write anything unusual they noted or any guesses about the purpose. No participant explicitly guessed our hypotheses (e.g., "I think you wanted to see if sensitive people do X"). This suggests expectancy bias was low. If many had guessed, say, "This is about people's sensitivity," we'd be concerned they might have altered their answers to fit what they think this study wants.

- **Attention checks and participant engagement:** By including the attention check questions and interactive toggling, this study aimed to keep participants engaged and to screen out those who weren't paying attention. Engaged participants are less likely to produce random or biased responses. The fact that this study removed some inattentive cases contributes to the credibility of the data from those remaining.

- **No feedback on performance:** This study didn't give participants any feedback or indications of how they were doing because there was really no right or wrong in this scenario. This avoids any social desirability or "please the experimenter" effect beyond the inherent design.

- **Anonymity and confidentiality:** Participants knew their responses were anonymous (especially since prize draw emails were collected separately). This likely made them more comfortable giving honest feedback about discomfort or trust, etc., without fear of judgment. If someone found the site annoying, they could say so freely. This reduces bias like the Hawthorne effect, as they weren't interacting with an experimenter in person at all.

Taken together, these controls mean that the results can be interpreted with reasonable confidence that they reflect genuine participant experiences rather than artifacts of the experimental procedure or participant guessing. In the results, this study will note if any unexpected pattern might indicate bias (for example, if everyone rated everything high, perhaps due to politeness This study actually sees mixed and nuanced responses, which is a good sign.

3.7 Ethical considerations

The study adhered to Hasselt University's ethical guidelines and GDPR requirements throughout. Participants provided informed consent electronically at the start of the survey (the first page after the info had a consent checkbox). All data were collected anonymously; This study did not ask for names or any identifying information in the survey itself. Those who entered the prize draw did so via a separate form, which was not linked to their survey responses, ensuring that survey data remained non-identifiable. The prize draw data was only used to contact winners and then deleted.

Participants were informed that they could withdraw at any time by simply closing the browser window; if they did so, their partial data would not be used. In practice, a few did leave early (This study saw some drop-offs during the site loading, perhaps), and as stated, those cases were not included in the analysis.

After completing the survey, a **debrief page** explained the study's purpose more transparently: This study disclosed that the research was examining whether giving users some control over sensory elements of a website would affect their comfort, trust, etc. This study explained the rationale (that current web design might overwhelm some users, and we're testing a potential solution). This study provided contact details for the research supervisor and ethics committee in case of questions or concerns. This study also reminded them that their data was anonymous, and they were thanked for their participation.

Throughout the study design, this study incorporated **inclusive language** and respectful phrasing. For instance, this study used the term "neurodivergent" rather than deficit-focused language when asking about neurotype, and made that question optional. This study framed the preference toggles as something *anyone* might use for comfort, rather than implying it's only for the "hypersensitive" (to avoid stigmatizing that trait or making those individuals feel singled out).

This study also took care that the **sensory adjustments themselves would not cause harm or discomfort**. For example, this study did not include any rapid flashing content at all (which could trigger seizures in photosensitive individuals); even Version A was intense but not in violation of known safety thresholds (no strobe effects, etc.). The presence of a mute toggle meant participants had agency to turn off the music if it bothered them; likewise for animations and brightness. In essence, participants could choose not to adjust settings if they found the baseline tolerable, or adjust them if needed, which respects their well-being. This study did not force anyone to endure something extremely uncomfortable (and if the baseline was too much for a participant, presumably they would use the toggles, which is part of the design).

These measures align with ethical principles of **respect**, **beneficence**, and **justice** in research. Respect was shown by acknowledging diverse needs (the toggles themselves are a form of respect for participant comfort, and allowing withdrawal without penalty). Beneficence was addressed by minimizing potential harm (the site wasn't harmful, and the study was relatively short and low-risk, with accommodations as described). Justice in terms of not excluding any group unfairly: aside from needing a computer and Chrome, this study didn't exclude on any other grounds; and in interpretation, this study aims to highlight benefits that could help those often marginalized (like neurodivergent users) while also benefiting all.

In summary, the study was designed and conducted with careful attention to participants' rights and comfort. Ethics approval was obtained prior to recruitment. No adverse events or complaints were reported during the study. With ethical and methodological foundations laid out, this study now proceeds to the results in the next chapter, which detail what this study found when participants engaged with sensory personalisation on the travel website.

Chapter 4: Results

4.1 Overview

This chapter presents the analytical results of the quasi-experimental study described in Chapter 3. We begin with data screening and coding checks, proceed to reliability and descriptive statistics, and then test each hypothesis using the planned inferential procedures. Quantitative results are complemented by thematic insights from open-ended comments, which illuminate why several quantitative effects were small or absent.

All questionnaire items were placed on common **1–5 Likert scales** so that **higher values consistently indicate more positive evaluations** (e.g., for comfort, 5 = "very comfortable"). Two-item constructs (comfort, engagement, trust, purchase intention, word of mouth) were computed as **composite means**; single-item variables (overall impression, perceived control) were analyzed as observed. Internal consistency was assessed via **Cronbach's alpha**. For within-subject effects, this study reports **Cohen's d for paired designs (d_paired)** with 95% CIs; for ANOVA/GLM, we report **partial η^2** (and note model R^2 / f^2 where relevant).

The analytic sample for within-subject tests comprises **n = 146** participants with complete baseline and post measures. Each participant experienced the high-stimulation website (Version A), completed baseline ratings, optionally adjusted sensory settings, then experienced a modified site (Version B) and completed post-adjustment ratings. Below, we detail how these experiences translated into measurable outcomes.

4.2 Data cleaning and screening

Of 270 survey starts, 146 cases met all inclusion criteria and passed quality checks (Chapter 3). Exclusions arose from unsupported devices/browsers (≈15 mobile attempts auto-terminated), failed instructed-response items (~8), straight-lining (~5), extremely short completion times (bottom ~2%), or duplicate entries. A Qualtrics routing error caused a small number of participants to miss the post-exposure block; thus, paired analyses use n = 146, whereas baseline-only analyses use all available baseline cases (up to n ≈ 151). Baseline outcomes and SHS scores did not differ between completers and those missing post data, suggesting missingness was incidental and not outcome-related.

Distributions were acceptable for planned tests. Most outcomes were roughly symmetric around the mid-point (≈3) with no severe skew; where appropriate, non-parametric corroborations (Wilcoxon/Kruskal–Wallis) are reported alongside parametric results.

Appendix C, Table C13.1 lists attention-check fails, straight-lining, device violations, routing errors, and manipulation-check mismatches; Table C13.2 shows the duration distribution and trimming rule.

4.3 Coding and composites

COMFORT_A = MEAN(A_COMFORT, A_PLEASANT);

COMFORT_B = MEAN(B_COMFORT, B_PLEASANT).

TRUST_A = MEAN(A_TRUST_QUALITY, A_TRUST_CONFIDENCE);

TRUST_B = MEAN(B_TRUST_QUALITY, B_TRUST_CONFIDENCE).

ENGAGE_A = MEAN(A_ENGAGED, A_ATTENTION);

ENGAGE_B = MEAN(B_ENGAGED, B_ATTENTION).

PI_A = MEAN(A_PURCHASE_INTENT_1, A_PURCHASE_INTENT_2);

PI_B = MEAN(B_PURCHASE_INTENT_1, B_PURCHASE_INTENT_2).

WOM_A = MEAN(A_WOM_SHARE, A_WOM_BENEFIT);

WOM_B = MEAN(B_WOM_SHARE, B_WOM_BENEFIT).

IMP_A = A_IMP; IMP_B = B_IMP. Change scores: D_X = X_B − X_A.

See Appendix C, Table C0.3 for COMPUTE syntax.

All items were coded **1–5**, with reverse-worded items recoded as needed. Composite construction:

- **Comfort** = mean of two comfort items.
- **Engagement** = the mean of two engagement items.
- **Trust** = mean of two trust items.
- **Purchase intention** = the mean of two purchase items.
- **Word of mouth** = the mean of two WOM items.

Overall impression (A_IMP, B_IMP) and **perceived control (CTRL)** were single items. Within-person **change scores** were computed as **Δ = B − A** (positive = improvement).

4.4 Reliability

Internal consistency for each 2-item composite at baseline and post appears in Appendix C, Tables C6.1 and C6.2.

Despite each scale using only two items, internal consistency was **acceptable–excellent** at both waves (Table 4.1).

Table 4.1. Cronbach's α for two-item outcome scales (baseline A; post B)

Construct	**α (A) n=151**	**α (B) n=146**	**k**
Comfort	.809	.916	2
Engagement	.874	.922	2
Trust	.909	.934	2
Purchase intention	.947	.941	2
Word of mouth	.883	.916	2

Internal consistency. All 2-item composites showed good to excellent internal consistency at baseline and post. Baseline alphas ranged from .809 to .947 and post alphas from .916 to .941 (see Tables C6.1–C6.2). For two-item scales, Cronbach's alpha equals the Spearman-Brown corrected inter-item correlation, $\alpha = 2r/(1+r)$, so reporting both α and the inter-item r demonstrates consistency of the estimates (Field, 2018). Across constructs, inter-item correlations were strong ($r = .680–.899$ at baseline; $r = .845–.889$ post), supporting the use of these concise composites.

4.5 Descriptive statistics

Means on the 5-point scales clustered around ~3.0–3.5 with similar SDs at baseline and post, foreshadowing small average changes (Table 4.2).

Table 4.2. Descriptives at baseline (A) and post-adjustment (B)

Outcome	**N(A)**	**Mean(A)**	**SD(A)**	**N(B)**	**Mean(B)**	**SD(B)**
Comfort	151	3.543	0.924	146	3.510	0.959

Engagement	151	3.523	0.935	146	3.435	0.949
Trust	151	2.974	1.053	146	3.014	1.120
Purchase intention	151	2.921	1.071	146	2.928	1.116
Word of mouth	151	3.159	1.048	146	3.134	1.086

Overall impression showed minimal shift (A: M = 3.33, SD = 0.97; B: M = 3.37, SD = 1.05). **Perceived control** (measured immediately after the toggle screen) was descriptively **higher** among non-adjusters (M = 3.95, SD = 0.65, n = 44) than adjusters (M = 3.59, SD = 0.87, n = 105), suggesting a brief choice-friction effect for those who changed settings. We return to this in §4.12 and Chapter 5.

See Appendix C: C7.1 (descriptives A and B), C7.2 (within-time correlation matrices), and C7.3 (A–to–B stability correlations).

4.6 Hypothesis 1: Within-subject change (B − A)

Paired tests and effect sizes are reported in Appendix C, Table C8; 95% CI plots of Δ appear in Figure C8; Wilcoxon corroboration in Table C11.1.

This study tested within-person change from A to B for the five outcomes, reporting **Δ = B − A** with **d_paired** and 95% CIs. Engagement **decreased** slightly; all other outcomes showed trivially small, non-significant changes (full SPSS tables: Appendix C).

- **Comfort:** ΔM = **−0.072**, 95% CI [−0.197, 0.053]; *t*(145) = **−1.135**, *p* = .258; **d_paired = −0.094**, 95% CI [−0.256, 0.069].

- **Engagement:** ΔM = **−0.127**, 95% CI [−0.224, −0.029]; *t*(145) = **−2.571**, *p* = .011; **d_paired = −0.213**, 95% CI [−0.376, −0.048].

- **Trust:** ΔM = **+0.021**, 95% CI [−0.103, 0.144]; *t*(145) = **+0.328**, *p* = .743; **d_paired = +0.027**, 95% CI [−0.135, 0.189].

- **Purchase intention:** ΔM = **−0.014**, 95% CI [−0.110, 0.082]; *t*(145) = **−0.282**, *p* = .778; **d_paired = −0.023**, 95% CI [−0.186, 0.139].

- **Word of mouth:** ΔM = **−0.038**, 95% CI [−0.152, 0.077]; *t*(145) = **−0.650**, *p* = .517; **d_paired = −0.054**, 95% CI [−0.216, 0.109].

Table 4.3. Within-subject changes (positive Δ = improvement)

Outcome	**ΔM (B−A)**	**95% CIΔ**	**t(145)**	**p**	**d_paired**	**95% CI_d**
Comfort	−0.072	[−0.197, 0.053]	−1.135	.258	−0.094	[−0.256, 0.069]
Engagement	−0.127	[−0.224, −0.029]	−2.571	.011	−0.213	[−0.376, −0.048]
Trust	+0.021	[−0.103, 0.144]	+0.328	.743	+0.027	[−0.135, 0.189]
Purchase intention	−0.014	[−0.110, 0.082]	−0.282	.778	−0.023	[−0.186, 0.139]
Word of mouth	−0.038	[−0.152, 0.077]	−0.650	.517	−0.054	[−0.216, 0.109]

Interpretation. Apart from a small drop in engagement (|*d*| ≈ .21), within-person changes were near zero (|*d*| < .10), indicating these simple sensory toggles did not improve comfort, trust, purchase, or WOM in this prototype. Pre–post correlations were strong (r = .66–.86, all *p* < .001; Appendix C), indicating stable individual rankings across exposures. Wilcoxon signed-rank tests mirrored the t-tests (engagement: *Z* ≈ −2.46, *p* = .014; others *p* > .20).

4.7 Hypothesis 2: Predictors of the number of changes

H2 posited that **SHS**, **age**, **gender**, and **neurotype** would predict how many sensory adjustments participants make (CHG_N ∈ {0,1,2,3}). Usage was common: **0 changes** = 44 (29.1%); **1** = 59 (39.1%); **2** = 32 (21.2%); **3** = 16 (10.6%). Thus, **70.9%** adjusted at least one setting.

A GLM (DV: CHG_N; predictors: SHS_TOTAL, age, gender, neurotype) showed **weak fit** and **no significant predictors** (overall R^2 = .084; adj. R^2 = .024).

Effects:

- **SHS_TOTAL:** $F(1,139) = 0.000$, $p = .990$
- **Age:** $F(1,139) = 0.348$, $p = .556$
- **Gender:** $F(2,139) = 2.400$, $p = .095$
- **Neurotype:** $F(1,139) = 0.156$, $p = .926$

A **Poisson GLM** (count robustness check) corroborated the null pattern. Hence, **H2 is not supported**: broad traits did not predict how many changes users made. The high overall uptake suggests **universal appeal** of the controls (beyond narrowly defined sensitivity groups), while fine-grained differences (0 vs 1 vs 2 vs 3 toggles) appear idiosyncratic or situational.

Full GLM and parameter estimates: Appendix C, Table C3.1; multicollinearity diagnostics: Table C3.2; residual diagnostics: Figure C3; Poisson robustness: Table C4; outlier/influence: Table C5 and Figure C5. The reporting summary table mirrored here is Appendix C, Table C9.

4.8 Hypothesis 3: Baseline differences by sensory sensitivity

H3 predicted that **high-sensitivity** users would rate the intense default site **worse at baseline** than lower-sensitivity users. Participants were split into SHS tertiles (Low/Medium/High; ~n ≈ 49/group). One-way ANOVAs on baseline outcomes yielded **no significant group differences** (Table 4.6).

Homogeneity tests are in Appendix C, Table C2.1; Welch adjustments and Games-Howell where required in Table C2.2; one-way ANOVAs in Table C10.1; post-hoc comparisons in Table C10.2; means with 95% CI by SHS tertile in Figure C10; Kruskal–Wallis corroboration in Table C11.2.

Comfort and purchase intention were marginal but **in the opposite direction** (high-SHS slightly **higher** than low-SHS).

Table 4.6. Baseline differences across SHS tertiles (Low/Medium/High)

Outcome	F(2,143)	p	partial η²
Comfort	2.99	.054	.040
Trust	1.08	.341	.015
Engagement	1.28	.282	.018
Purchase intention	2.93	.057	.039
Word of mouth	1.41	.248	.019

Kruskal–Wallis tests agreed (comfort: $H = 5.48$, $p = .065$; others $p > .20$). **H3 is not supported.** A plausible explanation is that the "intense" default was not extreme enough to depress baseline ratings among high-SHS participants in a brief, low-stakes session; coping strategies and modality mix (visual/auditory focus vs multi-modal SHS) may also have blunted group differences. We return to this in Chapter 5, including a differential susceptibility reading (high-SPS can fare better in benign contexts).

4.9 Thematic insights from open-ended responses

Open-ended comments (≈ half the sample) clarify why quantitative effects were modest:

- **AI/scammy appearance & credibility gaps.** Several participants described the site as "AI-generated," "template-y," or low-effort, reducing **trust** regardless of sensory comfort. Missing contact details, testimonials, and "About us" content were noted. Without credibility scaffolding, trust could not rise even when sensory friction eased.

- **Polarized reactions to dark mode.** Some welcomed darker palettes as relaxing; others found them gloomy. This heterogeneity fits **inverted-U** arousal logic-one person's comfort is another's dullness-underscoring the value of **choice** but warning against assuming a single "calm" optimum.

- **Desire for finer control & better discoverability.** Participants asked for **gradients** (e.g., volume slider; degrees of motion reduction), and several missed the controls at first.

Autonomy support requires **intelligible, timely options**; binary toggles and late discovery limit perceived control and benefits.

- **Content and polish shortfalls.** Comments cited limited depth and a “class project” feel. With animations/audio reduced and novelty gone, **engagement** fell slightly; removing stimulation revealed **content thinness**. Autonomy features cannot compensate for sparse content or a lack of professional finish.

- **Preference persistence.** Users expect preferences to **persist** across visits; the prototype did not store choices beyond the session, weakening ongoing autonomy benefits.

Together, the themes explain **flat trust** (credibility deficits dominated) and the **small engagement drop** (less arousal + limited content). They also point to design upgrades: **graded controls**, **onboarding visibility**, **credibility cues**, **preference persistence**, and holistic accessibility.

4.10 Sensitivity and robustness checks

Assumptions were acceptable for planned tests. Q–Q plots/Shapiro–Wilk showed only minor deviations for change scores; **Wilcoxon** confirmed the engagement decrease ($Z \approx -2.46$, p = .014). For H3, **Levene’s tests** were largely non-significant; where borderline, we reported **Welch’s F** and **Games–Howell** post-hocs-substantive conclusions unchanged. In H2, **VIFs < 2** indicated no multicollinearity; studentized residuals had no |3|+ outliers; model fit remained low (unadj. $R^2 \approx .084$; adj. ≈ .024). Missingness stemmed from routing; completers vs baseline-only did not differ on key baselines/SHS. Two-item alphas (Table 4.1) were high; inter-item rs (.68–.90 baseline; .84–.89 post) supported composite use.

Trimmed H1 results (fastest 5 percent removed) are in Appendix C, Table C12.1. Exploratory ANCOVAs for D_X by CHG_N with baseline covariates are in Table C12.2.

4.11 Summary of findings (linked to hypotheses)

- **H1 (within-subject improvement): Not supported overall.** Comfort, trust, purchase, and WOM showed **near-zero** mean changes; **engagement decreased** slightly ($\Delta \approx -0.13$, $t(145) = -2.57$, p = .011, **d_paired ≈ −0.21**). Pre–post correlations were strong (**r = .66–.86**).

- **H2 (predictors of number of changes): Not supported.** SHS, age, gender, and neurotype **did not** predict CHG_N; model fit was small (Adj. $R^2 \approx .024$).

- **H3 (baseline by SHS tertiles): Not supported** at omnibus level; two outcomes marginal (comfort, purchase) but **opposite** the hypothesized gradient.

- **Context:** Baseline/post means ~3.0–3.5 with similar SDs; two-item alphas were .81–.95 at A and .92–.94 at B.

4.12 Integrative interpretation of quantitative & qualitative evidence

The small or null mean changes with a slight engagement dip align with (a) order/novelty loss on second exposure and (b) an arousal–comfort trade-off when motion/audio is reduced. Qualitative themes explain flat trust: participants questioned authenticity/credibility (template/"AI" look; sparse "About us"; embedded player chrome), so autonomy over sensory load could not by itself raise trust. Further, a non-trivial share did not notice controls initially, and immediate post-choice perceived control was lower among adjusters than non-adjusters, consistent with decision friction blunting short-term autonomy benefits. Overall, the evidence suggests sensory personalization must be paired with credibility cues, content depth, and preference persistence to move outcomes like trust or purchase intention. We elaborate on the theoretical and design implications in Chapter 5.

Chapter 5: Discussion and Conclusion

5.1 Framing the Results: Personalisation as Emotional Accessibility (Revisited)

This study evaluated whether simple sensory controls (mute, reduce motion, darker mode) improve user outcomes over a high-stimulation default, and whether individual differences (SHS, neurotype, age, gender) predict uptake. In line with the pre–post data, comfort, trust, purchase intention, and word-of-mouth did not change significantly; engagement declined slightly on second exposure. SHS, age, gender, and neurotype did not predict how many controls users activated; baseline evaluations of the intense default did not differ reliably across SHS tertiles (Tables 4.2–4.6).

Rather than contradicting the theoretical case for perceptual autonomy, the results clarify its boundary conditions: options ≠ autonomy unless choice architecture is intelligible, low-effort, and paired with credibility scaffolding. In SDT terms, autonomy support is effective only when competence and relatedness are co-supported; otherwise, adding choices can feel like extra work (Deci & Ryan, 2000; Hutmacher & Appel, 2023; Alberts et al., 2024). Similarly, the digital-trust literature implies that distal, consequential outcomes (purchase intention, advocacy) are credibility-gated; they tend not to move until users see signals of integrity, competence, and benevolent intent (McKnight et al., 2002; Cyr et al., 2007). The pattern here-little average change despite accessible controls-fits this integrated view: without visible provenance and professionalism, perceptual comfort alone does not unlock trust-based behavior.

A second framing point is the compliance–comfort gap highlighted earlier. The prototype leaned toward WCAG 2.1 AA for contrast/operability, yet standards do not limit sensory intensity (W3C, 2018). A site can be technically accessible and still feel "too much" or, conversely, feel bland once arousal is reduced. The present results illustrate both sides of that coin: a high-stimulation default that was still navigable, and a calmer variant that, without richer content and credibility cues, felt less absorbing for some users.

5.2 What Changed (and Why) - A Data-First Interpretation

5.2.1 Within-Person Outcomes

The pre–post pattern shows no average gains on comfort, trust, purchase intention, or WOM, and a small decrease in engagement. With $n \approx 146$ and narrow CIs straddling zero, moderate effects are unlikely under this implementation. Two mechanisms are consistent with the data and the literature.

First, novelty decay between first and second exposures: the first visit carries inherent curiosity value, while the second visit-now without animation/audio peaks-can feel less lively, especially when content depth is limited. Short, low-stakes lab sessions are particularly prone to this attenuation (cf. A/B test transients; Kohavi et al., 2009).

Second, the arousal trade-off: the S-O-R tradition and sensory marketing often find inverted-U relationships, where under- and over-stimulation both depress approach responses (Krishna et al., 2016). Reducing motion and audio without compensating for task relevance or narrative richness can shift some users past their optimal arousal into a "too calm" zone-comfortable but a bit less captivated. In SDT terms, autonomy support that is not accompanied by competence cues (clear goals, strong information scent) can fail to translate into energized engagement (Deci & Ryan, 2000; Hutmacher & Appel, 2023).

5.2.2 Perceived Control as a Momentary Friction Point

Descriptively, participants who made no changes reported higher perceived control immediately after the choice step than those who did. This single measure was captured at the peak effort moment (choosing, waiting for the site to reload), which SDT would predict as a temporary dip in autonomy feelings before potential gains appear with use and persistence (Deci & Ryan, 2000). Because the study measured only this "dip" and not a later "recovery," the autonomy→outcome pathway is likely underestimated. Future protocols should measure perceived control twice (post-choice and post-use) and persistent settings between visits to detect delayed benefits (Hutmacher & Appel, 2023; Kohavi et al., 2009).

5.2.3 Individual Differences and the Coarseness Problem

H2's nulls likely reflect **measurement alignment** more than the absence of trait effects. SHS_TOTAL aggregates across modalities (Dixon et al., 2016), whereas the manipulation targeted primarily **visual** (brightness/motion) and **auditory** (music) channels. A total score can dilute modality-specific sensitivity (e.g., an auditory-sensitive participant who is visually robust). Likewise, CHG_N compresses behavior into a 0–3 count, masking intensity (e.g., turning **two** strong controls on could be equivalent in comfort to **three** minor ones for another person). The next iteration should (a) analyze **modality-matched** relations (auditory SHS ↔ mute/volume; visual SHS ↔ motion/brightness), and (b) upgrade controls from binary to **graded** (off/low/med/high; multiple motion profiles), which will increase behavioral variance and statistical power to detect trait × feature alignment (Dixon et al., 2016).

5.2.4 Why No Baseline Gradient by SHS?

H3 expected high-SHS participants to evaluate the intense default worse at baseline. Several contextual features can explain the absence of such a gradient. First, the default was **stimulating yet navigable** and approached AA contrast/operability; WCAG sets floors on perceivability/operability, and the design avoided flashing thresholds (W3C, 2018). High-SPS individuals often cope **well** in "merely busy" environments and struggle chiefly when intensity crosses particular thresholds or when surprise is high (Greven et al., 2019). Second, the exposure was **brief** and **low-stakes**; distal judgments such as trust and purchase intention are known to be **credibility-gated** and to evolve with time and risk (McKnight et al., 2002). Third, **differential susceptibility** implies that high-SPS individuals can fare **better** than others in benign contexts, not only worse in aversive ones (Greven et al., 2019). The site's moderate quality may thus have kept baseline comfort acceptable even for high-SHS users.

5.3 How the Qualitative Evidence Explains the Nulls

Open-ended responses triangulate the quantitative picture on three fronts.

Credibility headroom was low. Participants repeatedly pointed to an "AI/template" look, visible player chrome, and sparse "About us"/contact details. Trust rarely moves without **provenance** (who made this?), **competence** (do they seem capable?), and **benevolence** (are they on my side?) cues (McKnight et al., 2002; Cyr et al., 2007). Calm visuals cannot compensate for missing credibility scaffolding. Recent work on **AI authorship ambiguity** similarly warns that minimalistic or generic phrasing/layouts are sometimes read as "machine-made," prompting skepticism unless **human oversight and source transparency** are surfaced (Brauner et al., 2023; Rae, 2024). This aligns closely with the comments and explains flat trust.

Preference heterogeneity flattened means. Dark mode and motion reduction elicited **polarized** reactions: relief for some, dullness for others. The inverted-U perspective predicts such splits (Krishna et al., 2016). When effects point in **opposite directions** across subgroups, the **grand mean** moves little even if the feature truly helps a subset-a classic aggregation issue.

Control discoverability and cost. Several non-adjusters didn't notice the controls until prompted; some adjusters reported uncertainty. Autonomy must be **intelligible, timely, and**

low-effort to deliver benefits (Hutmacher & Appel, 2023). If the first experience surprises with audio/motion and the choice arrives later (or feels like extra configuration work), users can experience decision friction. That friction is visible in the immediate **perceived-control** dip among adjusters. In addition, a few comments framed animation and color intensity as bordering on **coercive** persuasion. This resonates with the **dark-patterns** literature: intensity and urgency cues can compress deliberation and undermine trust (Gray et al., 2018; Mathur et al., 2019; Narayanan et al., 2020). Autonomy features must therefore be paired with **explicit non-coercion** design cues.

5.4 Practical Implications: From Binary Toggles to Autonomy-First UX

Taken together, the evidence suggests a systemic rather than point-solution change.

1. **Lead with calm defaults and OS alignment.** Respect **prefers-reduced-motion** and system **dark mode** on first load; never re-enable stimuli without clear consent (W3C, 2018; ETSI, 2021). This meets users where they are and removes decision friction.

2. **Make control intelligible and graded.** Replace coarse toggles with **previews** and **granular sliders** (e.g., volume, motion easing, luminance presets), plus a one-click **Calm Mode** that bundles best-practice settings. Choice should feel reversible, safe, and instantly legible (Hutmacher & Appel, 2023).

3. **Surface credibility up-front.** Clear authorship, team/contact, and policy pages are **prerequisites** for movement in trust/intent (Cyr et al., 2007; McKnight et al., 2002). If calm design is mistaken for automation, add **provenance** and **human-in-the-loop** cues (Brauner et al., 2023; Rae, 2024).

4. **Reduce choice friction.** Place controls **before** any audio/motion plays; offer succinct inline explanations; **persist** preferences across sessions; and provide search-first orientation and visible exits to support **competence** (W3C, 2018; Hutmacher & Appel, 2023).

5. **Measure control twice and at scale.** Collect perceived control at decision time and after use; in field A/Bs, track bounce, return, and conversion trends pre/post autonomy-first changes to detect delayed benefits (Kohavi et al., 2009).

5.5 Theoretical Contributions

- **Choice architecture > choice count.** SDT predicts that autonomy's benefits depend on **how** choice is offered (clarity, timing, reversibility), not simply **that** choices exist (Deci &

Ryan, 2000; Hutmacher & Appel, 2023). This study's nulls under coarse, mid-flow toggles reinforce that insight.

- **Decoupling sensory ease from distal outcomes.** The data show that lowering sensory strain does **not** automatically raise **trust** or **purchase intent** without credibility scaffolds and social presence (Cyr et al., 2007; McKnight et al., 2002). This refines S-O-R predictions: organismic states shaped by stimuli still cascade through **trust gates** before behavior.
- **Temporal dynamics of perceived control.** Autonomy measured only at the **decision** moment can understate benefits. SDT-informed models should incorporate **time** (immediate effort vs. later comfort/fluency), especially when preferences persist (Deci & Ryan, 2000; Hutmacher & Appel, 2023).
- **Universal-design uptake with targeted benefit.** That ~71% of users made ≥1 change supports mainstream placement of autonomy features (Lazar et al., 2015; Wood et al., 2024). Realizing trait-aligned benefits, however, requires **modality matching** between user sensitivities and available controls (Dixon et al., 2016).
- **Integrating SDT with S-O-R.** The present pattern is coherent with S-O-R's arousal calibration (Krishna et al., 2016; Wang et al., 2024) and SDT's needs support: calibrated stimuli plus clear competence cues are jointly necessary for positive downstream appraisals.

5.6 Limitations

Sampling and recruitment. As noted, convenience/reciprocity recruitment (SurveySwap.io) likely skews toward younger, educated, research-savvy participants, with strong English proficiency and above-average digital literacy (Frandsen et al., 2016). Such panels can under-represent users with lower digital confidence and non-Western browsing norms, which are precisely groups for whom sensory autonomy might be most consequential. The desktop-only Chrome requirement further narrows external validity in a mobile-first ecosystem; on handhelds, motion, brightness, and audio are entangled with OS-level affordances (e.g., system dark mode, media volume rockers), so effects may differ when platform controls and site controls cooperate (W3C, 2018; ETSI, 2021). Together, these choices bias the sample and constrain generalizability to real-world audiences and devices.

Diagnostics supporting conservative interpretation: C1 (normality), C2 (homogeneity/Welch), C3–C4 (model fit and count robustness), C5 (influence), and C8/C12.1 (paired effects and trimmed re-runs).

Fixed A then B order, by design. The fixed order was intentional to mirror how most users first meet a site at its default intensity and only later discover preferences; this ecological choice makes the baseline realistic and lets the study ask the SDT question it cares about most: after a default, does giving simple control change the next experience (Deci & Ryan, 2000; Hutmacher & Appel, 2023; Alberts, Lyngs, & Lukoff, 2024). The trade off is familiar in quasi experimental work, where internal validity can be threatened by order and selection effects (Reichardt, 2002). Novelty loss and contrast between first and second exposures cannot be fully separated from any effect of the controls in a fixed sequence, and adoption of changes was self selected, which further cautions against strong causal claims in this design frame (Kohavi, Longbotham, Sommerfield, & Henne, 2009; Reichardt, 2002; Wood et al., 2024). To keep claims conservative, the analysis relied on within person change with confidence intervals, high A to B stability correlations, non parametric corroboration where relevant, and transparent effect size reporting, which are standard recommendations for robust inference with behavioral data (Field, 2018; Cohen, 1977; Ghasemi & Zahediasl, 2012). Taken together, those diagnostics support a limited conclusion that under this implementation there was no detectable average improvement in comfort, trust, purchase, or word of mouth in the immediate next visit, with a small decrease in engagement. The study does not claim that sensory controls never help, only that under a default first, choice second flow like common practice, measurable gains did not emerge in this short window (Tam & Ho, 2006; Wood et al., 2024).

Ecological validity and stakes. The task was a single, brief, low-stakes session with no real purchase risk. Trust and purchase intention are credibility-gated constructs that often change slowly and in response to provenance, guarantees, and accumulated interactions (McKnight et al., 2002; Cyr et al., 2007). Short lab-style exposures can therefore underestimate downstream gains that only emerge with preference persistence or repeated visits. Industrial A/B programs commonly find that effects evolve over days/weeks as novelty and learning equilibrate (Kohavi et al., 2009); by design, we captured only an acute snapshot.

Design threats to causal inference. The fixed A→B order introduces novelty loss/contrast: first exposure may inflate arousal/interest, and the second-especially if calmer-can look less engaging simply because the "newness" has faded. Without counterbalancing or parallel arms, this is indistinguishable from an effect of the controls per se (Kohavi et al., 2009). In addition, adoption of adjustments was self-selected; users who changed settings may differ systematically from those who did not (e.g., in trait curiosity or tolerance for configuration effort). Although within-person comparisons attenuate between-subject bias, complier vs. non-complier differences limit causal claims about "using controls."

Manipulation strength and fidelity. The intervention comprised three binary toggles delivered after baseline. Binary settings risk over- or under-dosing relative to individual comfort ranges, whereas graded controls (e.g., volume sliders, motion easing, luminance presets) would allow closer alignment to personal optima (Dixon et al., 2016). Late presentations also embed a decision cost at the very moment we measure perceived control, likely capturing a temporary autonomy dip rather than a stabilized benefit (Deci & Ryan, 2000; Hutmacher & Appel, 2023). Finally, the prototype necessarily included third-party UI elements (e.g., player chrome) and a deliberately generic aesthetic to isolate sensory factors; these choices inadvertently signaled low provenance, which plausibly capped trust even when sensory friction eased (Cyr et al., 2007; McKnight et al., 2002).

Credibility ambiguity and AI-provenance concerns. Multiple comments described the site as "AI/template-like." Contemporary work shows that ambiguous authorship and generic phrasing/layouts are frequently interpreted as machine-made, prompting heightened skepticism unless human oversight and source transparency are explicit (Brauner et al., 2023; Rae, 2024). In other words, the same minimalism used to isolate sensory variables may have primed distrust, creating a ceiling on trust and intent measures independent of sensory comfort. This context strengthens, rather than weakens, the interpretation that sensory autonomy must be paired with credibility scaffolding to influence distal outcomes (Cyr et al., 2007; McKnight et al., 2002).

Measurement alignment and sensitivity. We employed SHS total as the trait predictor, while the manipulation targeted primarily visual and auditory channels. Aggregating modalities can dilute modality-specific sensitivity (e.g., high auditory sensitivity with average visual sensitivity) and thus attenuate trait–behavior links (Dixon et al., 2016). The behavioral DV CHG_N (0–3) is a coarse count; it compresses nuance (which controls were selected, how strong, in what sequence), limiting variance and model fit. Additionally, perceived control was measured once at the post-choice moment; SDT predicts a time-varying profile in which feelings of autonomy may recover after users experience effective, persistent settings (Deci & Ryan, 2000; Hutmacher & Appel, 2023). Finally, although two-item composites showed high alphas, two-item scales inherently sample a narrow slice of each construct, and self-report in short sessions is vulnerable to common-method and range constraints (Cyr et al., 2007).

Mobile/assistive technology interactions have not been tested. Respecting OS preferences (e.g., prefers-reduced-motion, dark mode) and interactions with assistive technologies (screen readers, high-contrast modes) may amplify or alter the experiential impact of site-level controls. Because we are restricted to desktop Chrome and do not auto-ingest OS settings, the present design cannot speak to zero-click autonomy, which is a promising pathway for reducing configuration friction (W3C, 2018; ETSI, 2021).

Potential dark-pattern spillovers. High-stimulation defaults can border on persuasive saturation; if intensity is perceived as urgency or obstruction, users may infer manipulative intent, which erodes trust (Gray et al., 2018; Mathur et al., 2019; Narayanan et al., 2020). While our design avoided explicit deceptive patterns, the signal of "pushiness" may have persisted for some participants and blunted any positive effect of later autonomy.

Construct boundaries and baseline ceiling. The default was stimulating yet navigable (toward WCAG 2.1 AA), and exposure was brief and risk-free; together, these factors likely bound discomfort at baseline and narrowed between-group differences (W3C, 2018). In line with differential susceptibility, high-SPS individuals do not necessarily fare worse in benign environments; trait differences often emerge only beyond intensity thresholds or over longer, consequential horizons (Greven et al., 2019). Hence, the null H3 is interpretable as a context-bound result rather than a refutation of SPS effects.

Policy and generalizability scope. The European Accessibility Act pushes markets toward baseline accessibility but is agnostic on emotional accessibility and sensory calibration; firms vary widely in implementation maturity (European Parliament & Council of the EU, 2019; W3C, 2018). Our tourism-site context, generic brand, and short horizon mean results should be generalized cautiously to other verticals (e.g., finance, health) where credibility stakes and sensory norms differ.

Statistical considerations. Although powered for small-to-moderate within-subject effects, the combination of coarse DVs, high pre–post correlations, and restricted ranges reduces detectable movement. The GLM for CHG_N produced low R^2/adj. R^2, which is consistent with measurement coarseness and idiosyncratic behavior rather than strong trait drivers. A Poisson GLM corroborated the nulls but could not recover variance that was not present in the behavioral encoding (Kohavi et al., 2009; Dixon et al., 2016).

In sum, these design, measurement, credibility, and context constraints counsel conservative causal claims and help explain the modest outcomes in Chapter 4. Importantly, each limitation maps cleanly to a remedy specified in §5.7 (counterbalancing/parallel arms; graded, modality-matched controls; zero-click OS alignment; two-timepoint autonomy measures; explicit provenance/oversight cues; field A/Bs with behavioral endpoints). Under those conditions, the autonomy–comfort–trust pathway posited by SDT and digital-trust theory is more fairly tested (Deci & Ryan, 2000; McKnight et al., 2002; Cyr et al., 2007; Hutmacher & Appel, 2023).

5.7 Future Research

Zero-click vs. toggle-first.

A direct test of configuration friction is to compare auto-respect of OS preferences (zero-click) with explicit toggling. In the zero-click arm, the site reads prefers-reduced-motion and prefers-color-scheme and honors them on first paint; audio defaults to off unless the OS indicates media permission. In the toggle-first arm, users see an unobtrusive but explicit control panel before any motion/audio (W3C, 2018; ETSI, 2021). Randomly assign new visitors to arms; measure (a) perceived control twice (immediately after the panel/first render, and after content use), (b) time-to-task, bounce, scroll depth, and return rate, and (c) downstream trust and intent where appropriate. Hypothesis: Zero-click reduces early friction and preserves engagement while still improving comfort; toggle-first increases reported autonomy only when controls are intelligible and immediately effective (Deci & Ryan, 2000; Hutmacher & Appel, 2023). In field rollout, use an A/B framework with holdouts and run long enough to stabilize novelty and day-of-week effects (Kohavi et al., 2009).

Modality-matched granularity.

To address the coarseness problem, pair trait subscales with graded controls. Use auditory SHS to predict use/effect of volume sliders and music style/mute, and visual SHS to predict use/effect of motion easing profiles (e.g., full → reduced → off) and luminance presets (Dixon et al., 2016). Present a preview for each level to reduce regret.

Outcomes: comfort, cognitive fluency, trust, engagement; behaviors: level chosen, time spent at level, reversion frequency.

Hypothesis (interaction): trait–feature alignment (e.g., high auditory SHS × lower volume) yields larger improvements than misaligned changes or binary toggles. Model with mixed effects (person as random intercept), report simple slopes and partial η^2. This design also enables estimating optimal personalization curves (inverted-U calibration; Krishna et al., 2016).

Counterbalanced/parallel designs.
To unconfound novelty loss from sensory adjustment, employ (a) counterbalanced within-subjects sequences (A→B vs. B→A; Latin square for page order), and/or (b) parallel between-subjects arms where some users only ever see A or B (Kohavi et al., 2009). Include a third arm "B-at-first-paint (zero-click)" to isolate ordering from configuration effort. Primary contrasts: (i) second-exposure effects with identical content, (ii) first-exposure differences with identical sensory state.

Hypotheses: if the engagement dip is mostly novelty loss, it appears in both A→B and B→A; if it is a sensory-calibration effect, it appears when moving from high to low arousal regardless of order (Krishna et al., 2016). Analyze with repeated-measures ANOVA (sequence × condition) and Welch corrections where needed; preregister contrasts.

Field and longitudinal endpoints.
Short, low-stakes sessions understate trust dynamics (McKnight et al., 2002; Cyr et al., 2007). Deploy autonomy-first changes on a live site and track behavioral endpoints for several weeks: bounce, return within 7/30 days, conversion, assist rate (e.g., newsletter sign-ups), and preference persistence (do users keep their settings?). Use difference-in-differences with a comparable control property, or an A/B platform with cuped baselines to reduce variance (Kohavi et al., 2009). Collect perceived control on day 1 and again on day 7 to capture SDT's temporal dynamics (Deci & Ryan, 2000; Hutmacher & Appel, 2023). Hypothesis: autonomy benefits accumulate as preferences persist and credibility cues are encountered; early dips in engagement normalize as users learn where controls live.

Trust × Sensory factorials.
Cross credibility scaffolds with sensory autonomy in a factorial design. For example, a 2×2 or 3×2 with authorship transparency (none vs. "team + contact" vs. "team + contact + policies") × sensory regime (default high-stim vs. calm defaults with granular controls). Optionally add reviews/ratings as a third trust factor (McKnight et al., 2002; Cyr et al., 2007). Hypotheses: (i) credibility main effects on trust and intent; (ii) sensory main effects on comfort/fluency; (iii) interaction where credibility scaffolding unlocks the influence of comfort on trust/intent (i.e., comfort → trust only when provenance is clear). Analyze with MANCOVA (controlling for baseline affect) and report partial η^2 by factor and indirect effects (comfort → trust → intent) where justified.

AI-provenance transparency.

Qualitative feedback indicated an "AI/template look." Test provenance statements experimentally to mitigate that signal without over-claiming. Conditions could include: no statement; "Drafted by our team, assisted by AI; human-reviewed."; "Written by our editorial team."; plus placement (hero, sidebar, footer) and tone (plain vs. friendly) (Brauner et al., 2023; Rae, 2024). Combine with human presence cues (team photos, named authors, dated updates) drawn from the trust literature (Cyr et al., 2007; McKnight et al., 2002). Outcomes: perceived authenticity, competence, benevolence, and overall trust; behavioral: dwell, scroll to policies, contact clicks. Hypothesis: Explicit, modest human-oversight language placed near decision points increases trust relative to silence, especially when paired with sensory autonomy (interaction with the factorial above).

Design and analysis of hygiene across studies.

Across the programs above, preregister primary endpoints, counterbalance when feasible, and plan a priori power for small effects (e.g., $d \approx .20$) common in UX field work (Kohavi et al., 2009). Use two-time-point autonomy measures, modality-matched SHS predictors (Dixon et al., 2016), and graded controls to increase sensitivity. In reporting, accompany p-values with effect sizes and CIs, and discuss practical thresholds (Cyr et al., 2007; McKnight et al., 2002). Finally, when deploying calm defaults and provenance cues, align with WCAG/EN 301 549 requirements and OS preferences to reduce configuration burden and improve generalizability (W3C, 2018; ETSI, 2021).

Implications for replications.

Replications that aim for stronger causal identification can counterbalance or parallelize exposure order, randomize when controls appear, and measure perceived control twice to separate the decision dip from any benefit after use (Reichardt, 2002; Kohavi et al., 2009; Deci & Ryan, 2000; Alberts et al., 2024). A clean design would compare zero click respect of operating system preferences against toggle first discovery, since respecting system dark mode and reduced motion is recommended in accessibility guidance and standards (World Wide Web Consortium, 2018; European Telecommunications Standards Institute, 2021), while many users adopt personalization and accessibility settings only when friction is low (Wood et al., 2024). Modality matching should pair auditory sensitivity with volume or mute levels and visual sensitivity with motion or brightness presets to test whether trait-matched controls produce larger improvements (Dixon et al., 2016). Because trust and intent often move with credibility scaffolds, crossing sensory controls with visible authorship, reviews, and policy clarity can test interaction effects where comfort and provenance work together (McKnight et al, 2002; Cyr et al, 2007). Finally, field A/B tests can track bounce, return, conversion, and preference persistence over time to connect autonomy features to business metrics in real deployments (Kohavi et al., 2009; Montgomery, 2001). These adjustments do not change the substantive rationale for default first, choice later as a realistic flow, but they remove order as a confound while keeping the question squarely inside real marketing practice (Reichardt, 2002; Kohavi et al., 2009).

5.8 Policy and Ethics: Emotional Sovereignty by Design

The findings strengthen the case for embedding sensory autonomy into practice and policy: disclose intense stimuli; provide granular, persistent controls; align with OS preferences; avoid manipulative urgency/obstruction (dark-pattern tactics); and document preference storage with revocability. This trajectory extends WCAG/EN 301 549 beyond technical operability toward emotional accessibility, with particular relevance for neurodivergent and sensory-sensitive users (W3C, 2018; ETSI, 2021; European Parliament & Council of the EU, 2019). Enforcement contexts that already target deceptive patterns add further impetus to treat sensory coercion as a consent issue, not only an aesthetics issue (Gray et al., 2018; Mathur et al., 2019; Narayanan et al., 2020).

5.9 Conclusion

Circling back to the opening of this thesis, do loud, bright, and dynamic interfaces excite engagement or push users to close the tab? Under a high-stimulation default, simple, binary sensory controls did not raise comfort, trust, purchase intention, or word-of-mouth on average; engagement dipped slightly on second exposure. Qualitative feedback explains why: credibility constraints, choice friction, and heterogeneous preferences. Through an SDT lens, autonomy is not merely the presence of options; it is the experience of intelligible, low-effort, persistent control inside a credible interface. The design mandate is to upgrade personalisation: adopt calm defaults that respect system settings, provide graded controls with previews, persist preferences, and pair these with visible authorship and non-manipulative flows. Properly implemented, and measured on appropriate time horizons, sensory personalisation can act as a compass users set for themselves; but a compass moves people only when the map is trustworthy.

References

Tier 1 — Peer-Reviewed Scholarship (articles, books, peer-reviewed proceedings)

Books & Monographs

- Aron, E. N. (1997). *The highly sensitive person: How to thrive when the world overwhelms you*. Broadway Books.

- Biemer, P. P., de Leeuw, E. D., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L. E., Tucker, N. C., & West, B. T. (2017). *Total survey error in practice*. Wiley.

- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.

- Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2004). *Human–computer interaction* (3rd ed.). Pearson.

- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.

- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis* (2nd ed.). Guilford Press.

- Lazar, J., Goldstein, D. F., & Taylor, A. (2015). *Ensuring digital accessibility through process and policy*. Elsevier.

- Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., & Wolff, S. (2009). A brief history of the internet. *SIGCOMM Computer Communication Review, 39*(5), 22–31. https://doi.org/10.1145/1629607.1629613

- Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Books.

- Nunan, D., Birks, D. F., & Malhotra, N. K. (2020). *Marketing research: Applied insight*. Pearson.

Journal Articles

- Alberts, L., Lyngs, U., & Lukoff, K. (2024). Designing for sustained motivation: A review of self-determination theory in behavior change technologies. *Interacting with Computers*. Advance online publication. https://doi.org/10.1093/iwc/iwae040

- Alzahrani, M., Uitdenbogerd, A. L., & Spichkova, M. (2021). Human–computer interaction: Influences on autistic users. *Procedia Computer Science, 192*, 3860–3869.

- Amaro, A. C., Martinez, L. M., Ramos, F. R., Menezes, K., & Menezes, S. (2023). An overstimulated consumer… *Electronic Commerce Research*. Advance online publication.

- Antunes, I. F. S., & Veríssimo, J. M. C. (2024). A bibliometric review of sensory marketing. *Cogent Business & Management, 11*, 2338879. https://doi.org/10.1080/23311975.2024.2338879

- Balan, U., Mahesh, M., & Mathew, S. K. (2021). Online WOM strategies and consumer decision. *Information Systems Frontiers, 23*(3), 627–647.

- Bas, S., Kaandorp, M., De Kleijn, Z. P. M., Braaksma, W. J. E., Bakx, A. W. E. A., & Greven, C. U. (2021). Adults high in sensory processing sensitivity. *Journal of Clinical Medicine, 10*(21), 4912.

- Beattie, A. (2025). ADHD and digital disconnection. *Media, Culture & Society*. https://doi.org/10.1177/01634437251326482

- Bedi, S. S., Kaur, S., & Lal, A. K. (2017). Web experience & enjoyment. *Global Business Review, 18*(4), 1044–1066.

- Benham, G. (2006). HSP: Stress and physical symptoms. *Personality and Individual Differences, 40*(7), 1433–1440.

- Bevan, N. (2001). International standards for HCI and usability. *International Journal of Human-Computer Studies, 55*(4), 533–552.

- Biswas, D., Szocs, C., & Abell, A. (2019). Vestibular sensations and taste. *Journal of Consumer Research, 46*(4), 708–724.

- Boote, D. N., & Beile, P. (2005). Scholars before researchers. *Educational Researcher, 34*(6), 3–15.

- Brauner, P., Hick, A., Philipsen, R., & Ziefle, M. (2023). Public perception of AI. *Frontiers in Computer Science, 5*, 1113903.

- Celestino, S., Garofano, A., Masiello, B., Izzo, F., & Bonetti, E. (2024). Disability and marketing: A systematic review. *Italian Journal of Marketing*. Advance online publication.

- Chandra, S., Verma, S., Lim, W. M., & Kumar, S. (2022). Personalization in marketing. *Psychology & Marketing, 39*(8), 1604–1622.
- Chen, H.-T., & Lin, Y.-T. (2018). Sensory experience, emotion, buying behavior. *Service Business, 12*, 1–23.
- Chen, N., Watanabe, K., & Wada, M. (2021). Autistic traits & crossmodal correspondences. *Frontiers in Psychology, 12*, 714277.
- Chen, Y. (2023). Content marketing strategies using ML. *Humanities and Social Sciences Communications, 10*, 270.
- Chowdhury, T. G., Khare, A., & Coulter, R. A. (2025). Spillover effects of sensory stimulation. *European Journal of Marketing*.
- Conca, S. D. (2023). Deceptive design in virtual assistants. *Computer Law & Security Review, 49*, 105833.
- Corneille, O., & Gawronski, B. (2024). Self-reports vs implicit measures. *Nature Reviews Psychology, 3*(12), 835–846.
- Couper, M. P., Singer, E., Conrad, F. G., & Groves, R. M. (2008). Risk of disclosure & survey participation. *Journal of Official Statistics, 24*(2), 255–275.
- Cyr, D., Hassanein, K., Head, M., & Ivanov, A. (2007). Social presence & loyalty. *Interacting with Computers, 19*(1), 43–56.
- Dampérat, M., Jongmans, E., Jeannot, F., & Giuffrida, T. (2022). E-personalization and disabilities. *Recherche et Applications en Marketing (English Edition)*.
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits. *Psychological Inquiry, 11*(4), 227–268.
- DeCamp, W., & Manierre, M. J. (2016). Conditional incentives for online surveys. *Survey Practice, 9*(1).
- Dixon, E. A., Benham, G., Sturgeon, J. A., Mackey, S., Johnson, K. A., & Younger, J. (2016). Sensory Hypersensitivity Scale (SHS). *Journal of Behavioral Medicine, 39*(4), 537–550.
- Dominique-Ferreira, S., Viana, M., Prentice, C., & Martins, N. (2022). OTA web design & engagement. In *Advances in tourism, technology and systems* (pp. 341–349).

- Elliot, A. J. (2015). Color and psychological functioning. *Frontiers in Psychology, 6*, 368.

- Engel-Yeger, B., & Dunn, W. (2011). Sensory processing & anxiety. *British Journal of Occupational Therapy, 74*(5), 210–216.

- Fialkowski, B., & Schofield, D. (2024). Considering color in digital interfaces. *Art and Design Review, 12*, 306–329.

- Finnigan, K. A. (2024). Sensory responsive environments. *Land, 13*(5), 636.

- Frandsen, M., Thow, M., & Ferguson, S. G. (2016). Facebook vs traditional recruitment. *JMIR Research Protocols, 5*(3), e161.

- Galesic, M., & Bosnjak, M. (2009). Questionnaire length & response quality. *Public Opinion Quarterly, 73*(2), 349–360.

- Gao, F., & Shen, Z. (2024). Sensory brand experience & loyalty. *Acta Psychologica, 241*, 104127.

- Ghasemi, A., & Zahediasl, S. (2012). Normality tests. *International Journal of Endocrinology and Metabolism, 10*(2), 486–489.

- Gil Urrutia, J. I., Brangier, E., Senderowicz, V., & Cessat, L. (2018). Integrative heuristic inspection. In *Advances in usability and user experience* (pp. 499–512).

- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. In *Proceedings of CHI 2018*.

- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2020). Dark patterns. *Communications of the ACM, 63*(3), 42–49.* **[Flag: see "Verification" below]**

- Greven, C. U., Lionetti, F., Booth, C., Aron, E. N., Fox, E., Schendan, H. E., … & Homberg, J. (2019). Sensory processing sensitivity: Review & agenda. *Neuroscience & Biobehavioral Reviews, 98*, 287–305.

- Guan Lim, C., Lim-Ashworth, N. S. J., & Fung, D. S. S. (2020). Tech-based interventions for ADHD. *Current Opinion in Psychiatry, 33*(6), 610–616.

- Hallikainen, H., Halinen, M., Tervonen, H., Ruusunen, N., & Wang, Y. (2025). Sensory encounters & online reviews. *Psychology & Marketing*.

- Hamacher, K., & Buchkremer, R. (2022). Online sensory marketing index (OSMI). *Journal of Theoretical and Applied Electronic Commerce Research, 17*(2), 39.
- Hanson, V. L., & Crayne, S. (2005). Personalization of web browsing for older adults. *Universal Access in the Information Society, 4*(1), 1–10.
- Helbling, M., Grandjean, M.-L., & Srinivasan, M. (2024). Multisensory stimulation therapy: Meta-analysis. *Special Care in Dentistry*.
- Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Insufficient effort responding. *Journal of Business and Psychology, 27*(1), 99–114.
- Hutmacher, F., & Appel, M. (2023). Psychology of personalization. *Review of General Psychology, 27*(3), 388–410.
- Hutson, J., & Hutson, P. (2023). Digital inclusion & autism: Legal models. *Journal of Digital Technologies and Law, 1*(4), 851–879.
- Igartua, J.-J., & Hayes, A. F. (2021). Mediation/moderation confusions. *The Spanish Journal of Psychology, 24*, e14.
- Jevremović, M., Gheorghiță, C., Cerić, V., … & Bucea-Manea-Țoniș, R. (2022). Website interactivity impact. *Sustainability, 14*(8), 4507.
- Jiménez-Marín, G., Bellido-Pérez, E., & López-Cortés, Á. (2019). Sensory marketing: Concept & techniques. *Vivat Academia, 147*, 1–20.
- Kaushik, S., & Gokhale, N. (2021). Online sensory marketing: Five-dimensional experiences. *Revista Gestão Inovação e Tecnologias, 11*(3), 2628–2646.
- Kesari, B., & Atulkar, S. (2016). Utilitarian & hedonic shopping values. *Journal of Retailing and Consumer Services, 31*, 22–31.
- Klein, K., Melnyk, V., & Völckner, F. (2021). Background music and visual evaluation. *Psychology & Marketing, 38*(7), 1033–1047.
- Knyzelis, M. (2024). Confirmshaming as deceptive design. *Creativity Studies, 17*(1), 1–13.
- Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2009). Controlled experiments on the web. *Data Mining and Knowledge Discovery, 18*(1), 140–181.

- Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool. *American Psychologist, 70*(6), 543–556.
- Krishna, A., Cian, L., & Sokolova, T. (2016). Sensory marketing in advertising. *Current Opinion in Psychology, 10*, 142–147.
- Krupka, Z. (2023). Sensory marketing & brand perception. *Naše gospodarstvo/Our Economy, 69*(3), 34–45.
- Kumar, L., Rath, S. K., & Santanu, K. (2015). Quality assessment of web services (MARS). In *APSEC 2015* (pp. 388–395). IEEE.
- Kumaradeepan, V., Azam, F., & Tham, J. (2023). Social media & purchase intention: Conceptual paper. *International Journal of Professional Business Review, 8*(4), e01571.
- Le, N., Mai, N. T. B., Nguyen, N. Y. T., & Vu, H. N. T. (2024). Multi-sensory marketing & impulsive buying. *Emerging Science Journal, 8*(4), 1144–1158.
- Le Cunff, A.-L., Martis, B.-L., Glover, C., … & Dommett, E. J. (2024). Cognitive load & neurodiversity online. *Frontiers in Education, 9*, 1437673.
- Lewis, J. R., & Sauro, J. (2021). Usability and UX: Design & evaluation. In *Handbook of human factors and ergonomics* (5th ed., pp. 972–1015). Wiley.
- Li, P., Guo, X., Wu, C., & Spence, C. (2022). Multisensory perception & purchase intent. *Frontiers in Psychology, 13*, 1039875.
- Lionetti, F., Pastore, M., Moscardino, U., … & Pluess, M. (2019). SPS & personality: Meta-analysis. *Journal of Research in Personality, 81*, 196–213.
- Liu, Q., Trevisan, A. H., Yang, M., & Mascarenhas, J. (2022). Digital tech for circular economy. *Business Strategy and the Environment, 31*(4), 1488–1503.
- Lteif, L., van der Sluis, H., Block, L. G., Cian, L., Patrick, V. M., & Scott, M. L. (2025). Marketplace inclusion framework. *Journal of Public Policy & Marketing*.
- Maduku, D. K. (2024). Environmental concerns & anticipated emotions. *Journal of Retailing and Consumer Services, 73*, 103339.
- Maier, M., & Harr, R. (2020). Dark design patterns: End-user perspective. *Human Technology, 16*(2), 170–199.

- Mariani, M. M., & Nambisan, S. (2021). Innovation analytics & experimentation. *Technological Forecasting and Social Change, 172*, 121034.

- Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale. *PACM HCI, 3*(CSCW), 1–32.

- McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Trust measures for e-commerce. *Information Systems Research, 13*(3), 334–359.

- Mejía, V. D., Aurier, P., & Huaman-Ramirez, R. (2021). Assortment size & alignability. *Journal of Retailing and Consumer Services, 61*, 102539.

- Moliner, M. A., & Tortosa-Edo, V. (2023). Multirooming & online CX. *Journal of Research in Interactive Marketing, 17*(1), 1–20.

- Montgomery, A. L. (2001). Quantitative marketing techniques online. *Interfaces, 31*(2), 90–108.

- Motoki, K., & Iseki, S. (2022). Replicability of influential sensory marketing research. *Frontiers in Communication, 7*, 1048896.

- Motti, V. G. (2019). Designing with neurodiverse users. In *SIGDOC 2019*.

- Mouakket, S. (2013). Continuance intentions for online reservation. In *CITS 2013*. IEEE.

- Norman, G. (2010). Likert scales & "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625–632.

- Nowakowski, M. (2021). HSP & information usefulness perception. *Procedia Computer Science, 192*, 3870–3879.

- Ortiz-Escobar, L. M., Chavarria, M. A., Schönenberger, K., … & Velarde, M. R. (2023). UCD standards in AT for visual impairments. *Frontiers in Rehabilitation Sciences, 4*, 1238158.

- Palmer, Z. B., & Palmer, R. H. (2018). Legal/ethical implications of web accessibility. *Business and Professional Communication Quarterly, 81*(3), 305–326.

- Pandey, V., & Tripathi, V. (2025). Four decades of sensory marketing. *International Journal of Consumer Studies*.

- Petit, O., Velasco, C., & Spence, C. (2019). Digital sensory marketing. *Journal of Interactive Marketing, 45*, 42–61.

- Pivin-Bachler, J., van den Broek, E. L., & Gomez, R. (2024). Sensory disabilities & social robot. *Journal of Smart Cities and Society, 3*(1), 3–21.

- Rae, I. (2024). Perceived AI use & content perceptions. In *CHI 2024* (pp. 1–14). ACM.

- Rafijevas, S., & Razbadauskaite Venske, I. (2024). Visual stimuli & online purchase intentions. *Studijos – Verslas – Visuomenė*, 1(IX), 155–161.

- Rao, N. (2020). Cost-effectiveness of incentives for mail surveys. *Survey Practice, 13*(1).

- Reichardt, C. S. (2002). Review of *Experimental and quasi-experimental designs…*. *Social Service Review, 76*(2), 324–327.

- Reynolds-McIlnay, R., Morrin, M., & Nordfält, J. (2017). Brightness contrast & disarray. *Journal of Retailing, 93*(3), 266–281.

- Santos, M. A., Dopico-Parada, A., & Cabanelas, P. (2025). Sensory marketing in slow food. *European Journal of Management and Business Economics*.

- Schmitt, B., Zarantonello, L., & Brakus, J. (2009). Brand experience. *Journal of Marketing, 73*(3), 52–68.

- Schwartzman, R., & Knowles, C. (2022). Sensory-sensitive museum programming. *Curator, 65*(4), 711–733.

- Shams, R., Chatterjee, S., & Chaudhuri, R. (2024). Brand identity & sales strategy. *Journal of Business Research*.

- Sharma, A. P., Starčević, S., & Saha, R. (2024). Omnichannel determinants: SLR. *Journal of Service Theory and Practice*.

- Stead, S., Wetzels, R., Wetzels, M., Odekerken-Schröder, G., & Mahr, D. (2022). Multisensory customer experiences: Bibliometric review. *Journal of Service Research, 25*(4), 551–572.

- Sterne, J. A. C., White, I. R., Carlin, J. B., … & Carpenter, J. R. (2009). Multiple imputation in research. *BMJ, 338*, b2393.

- Storetvedt, M. H. R., Kahlon, S., Berg, K., Sundfjord, I. A., & Nordgreen, T. (2024). Digital mental health for youth with ADHD. *Frontiers in Digital Health, 6*, 1376016.

- Tam, K. Y., & Ho, S. Y. (2006). Web personalization & decision outcomes. *MIS Quarterly, 30*(4), 865–890.

- Thornton, L., Batterham, P. J., Fassnacht, D. B., … & Hunt, S. (2016). Recruiting via Facebook: Systematic review. *Internet Interventions, 4*, 72–81.

- Turjeman-Levi, Y., & Kluger, A. N. (2022). SPS vs sensory-processing theory. *Frontiers in Psychology, 13*, 980464.

- Wang, F., Wang, K., Han, Y., & Cho, J. H. (2024). Design-driven FMCG & purchase intentions. *Humanities and Social Sciences Communications, 11*, 143.

- Wood, R., Griffith, M., Jordan, J. B., Vanderheiden, G., Lazar, J., Kaine-Krolak, M., … & Folmar, D. (2024). Adoption of personalization & accessibility settings. *Universal Access in the Information Society*.

- Wörfel, P., Frentz, F., & Tautu, C. (2022). Sensory experience in marketing: Review & framework. *European Journal of Marketing, 56*(13), 144–181.

Preprints / Unpublished / arXiv / Conference Extended Abstracts

- Bolin, J. H. (2014). Review of Hayes (2013). *Journal of Educational Measurement, 51*(3), 335–337. *(Peer-reviewed book review)*

- Casaca, J., & Miguel, L. (2024). The influence of personalization on consumer satisfaction: Trends and challenges. *Unpublished manuscript*.

- Keil, M. V. (2024). AI-supported UI design for neurodiverse-friendly systems. In *NordDesign 2024*. *(Proceedings; peer-reviewed)*

- Lopez-Herrejon, R. E., Herrera, G., & Sevilla, J. (2018). Customization in interactive technologies for autism. *arXiv preprint*.

- Lopez-Herrejon, R. E., Herrera, G., & Sevilla, J. (2018). Personalization for autism: An open challenge. *arXiv preprint*.

- Mitchell, A., Erfanian, M., Soelistyo, C., Oberman, T., Kang, J., Aldridge, R., Xue, J.-H., & Aletta, F. (n.d.). Soundscape complexity & noise annoyance: Online listening experiment.

Manuscript.

- Peters, D., & Calvo, R. (2024). Self-determination theory and technology design. *Unpublished manuscript*.

- Rae, I. (2024). Perceived AI use… In *CHI 2024* (peer-reviewed proceedings). *(already above under proceedings)*

Tier 2 — Standards, Regulations & Official Guidance (authoritative, non-scholarly)

- European Parliament & Council. (2016). *Directive (EU) 2016/2102* (public sector web/mobile accessibility).

- European Parliament & Council. (2019). *Directive (EU) 2019/882* (European Accessibility Act).

- European Commission. (n.d.). *European Accessibility Act: Union equality strategy*.

- European Telecommunications Standards Institute. (2021). *EN 301 549 V3.2.1: Accessibility requirements for ICT products and services*.

- U.S. Department of Justice, Civil Rights Division. (2022, March 18). *Guidance on web accessibility and the ADA*.

- World Wide Web Consortium. (2018). *Web Content Accessibility Guidelines (WCAG) 2.1*.

- W3C Web Accessibility Initiative. (n.d.). Accessibility conformance testing (ACT) overview.

- W3C Web Accessibility Initiative. (n.d.). Accessibility principles.

- W3C Web Accessibility Initiative. (n.d.). Policies: Belgium.

- W3C Web Accessibility Initiative. (n.d.). Policies: European Union.

- W3C Web Accessibility Initiative. (n.d.). WCAG 2 overview.

Tier 3 — Web Resources & Utilities (contextual / tools / blogs / media; non-scholarly)

- Accessibility Checker. (n.d.). *The state of web accessibility in 2024* [Report].
- Accessible Web. (n.d.). *Color contrast checker – WCAG conformance*.
- Alex-Productions. (2022). *Ambient nature music* [YouTube video].
- Brignull, H., Leiser, M., Santos, C., & Doshi, K. (2023, April 25). *Deceptive patterns* (website). Deceptive.Design.
- Brunei Tourism. (2012, June). *Promotional video* [YouTube video].
- Centre for Sensory Studies. (n.d.). Research agenda pages ("Impact of visual perception…", "The senses and society").
- Datamatics, Lumina. (2025). *How to achieve EAA compliance* (blog).
- Howes, D. (2017). *How capitalism came to its senses* (web essay). Centre for Sensory Studies.
- Interaction Design Foundation (Soegaard, 2023). *UX surveys: The ultimate guide* (tutorial article).
- Japan National Tourism Organization. (2019). *Japan: Where tradition meets the future* [YouTube].
- Jarrett, C. (n.d.). *Definitions chapter: What is a survey?* (Effortmark web resource).
- Metadrop (Rodicio, 2024). *New era of digital accessibility* (blog).
- Nielsen Norman Group. (n.d.). *Articles and videos by Don Norman* (portal).
- SuperColorPalette. (n.d.). *Color palette tool*.
- Userpilot. (2022). *What is a UX survey?* (blog).
- VISITFLANDERS. (2024). *Wonders of Flanders* [YouTube video].
- WhoIsAccessible. (2020). *EN 301 549 and WCAG 2.0/2.1 Level AA* (guide).

- United Nations. (n.d.). *Sustainable Development Goal: Reduce inequality* (SDG page).
- Flaticon. (2025). *Free icons* (asset library).

Appendices

Appendix A: Survey Questions:

UHASSELT

Intro_Consent

Mobile Device Detected – Please Switch to Desktop

This study is best viewed on a desktop or laptop. It looks like you're currently using a mobile device. Some features—like autoplay sound and interactive visuals—may not work correctly.

To continue, **please open this link on a desktop computer using Google Chrome:** https://uhasselt.qualtrics.com/jfe/form/SV_0dFDlrGGlVl2Z0i

Thank you for helping us ensure the quality of this research!

P.S: This survey contains Karma to get free survey responses at SurveySwap.io

Welcome!

Hello and thank you for your interest!
My name is **Auzee Rosmadee**, and I am a **Master of Management student at Hasselt University.**

You are invited to participate in this academic research study as part of my Master's thesis, which investigates how people experience and emotionally respond to using a travel website.

Time Commitment
This survey will take approximately 10 to 15 minutes to complete.

Your Rights & Data Privacy
• Participation is completely voluntary. You may stop the survey at any time without consequences.
• All responses are anonymous. No personal or identifying information will be collected.
• Data will be stored securely and confidentially, and used only for academic purposes related to this research.
• This study complies with the General Data Protection Regulation (GDPR) of the European Union.
• There are no known risks or discomforts associated with participating in this study beyond those typically

experienced during regular internet use.

Contact Information
If you have any questions or concerns, feel free to reach out:
• Student: **Auzee Rosmadee** — auzeehaziqah.rosmadee@student.uhasselt.be

Consent to Participate
By clicking “Next”, you confirm that:
• You are 18 years or older
• You have read and understood the information above
• You voluntarily agree to participate in this study

P.S: This survey contains Karma to get free survey responses at SurveySwap.io

Intro_Disclaimer

⚠ **Important Before You Begin** ⚠

This study includes **interactive visuals, audio, and motion-based elements** that are best experienced on a **desktop or laptop computer** using **Google Chrome**.

🚫 **If you're currently on a mobile device or using a non-Chrome browser, you may:**

- Miss key features (e.g., autoplay audio or subtle animations)
- Encounter layout or navigation issues
- Be unable to proceed past certain sections

To ensure the correct experience, please:

1. Use a **desktop or laptop computer**
2. Use **Google Chrome**

If you're **not on a desktop using Chrome yet**, we kindly recommend switching devices and browsers now.

This helps ensure your experience reflects the true nature of the study and that your input can be counted in the final results.

◯ **Yes**, I'm on a desktop or laptop with sound enabled

◯ **No**, I'm using a mobile device or don't have sound

VISIT_A

Before You Begin - Website Viewing Checklist

To ensure the experience works as intended:

Initial Setup Checklist:

- Use **Google Chrome** on a **desktop or laptop only**
- 🔊 Make sure **sound is manually enabled** in your browser:
    - Click the website link that opens in a new tab

- Click the **View Site Information** icon next to the URL
- Click the **Settings icon (⚙)**
- Under **Permissions → Sound**, set it to **"Allow"**
- Refresh the page after it loads
- Ensure your device **volume is ON** and you're in a **quiet space**

While on the website:

Please take a few minutes to explore it as if you were planning or booking a trip:

- Scroll through the homepage
- Explore the tour options
- Observe the layout, visuals, and content at your own pace

Once you're done, simply return to the survey tab to continue.

P.S.

The website's **text, layout, and content were personally created and provided by the researcher** as part of a Master's thesis.

It is a **non-commercial prototype**, meant solely to simulate a tourism website for academic study purposes. Please kindly note:

This is **not a real travel agency** or a professional commercial platform. All content was developed intentionally to reflect research design conditions.

Your understanding and constructive participation are deeply appreciated

Click here to visit the website

A_EXP

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
At this point, how would you rate your overall impression of the website so far?	○	○	○	○	○

A

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very High
	1	2	3	4	5
I felt comfortable using this website.	○	○	○	○	○
The experience felt emotionally pleasant.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very High
	1	2	3	4	5
I was engaged while browsing this website.	○	○	○	○	○
The website held my attention.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very High
	1	2	3	4	5
I trust the quality of this website.	○	○	○	○	○

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I feel confident using this website to plan something real.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I would consider booking a tour using this website.	○	○	○	○	○
I am interested in purchasing from this website.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I would recommend this website to someone planning a trip.	○	○	○	○	○

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I think others would benefit from this website experience.	○	○	○	○	○

CUSTOMISE

You may personalize your website experience below. These options adjust how the site appears or sounds.

For each feature below, select whether you'd prefer to adjust it, or leave it unchanged based on your experience.

	Adjust	No Change
Reduce Brightness	○	○
Reduce Motion Effects	○	○
Mute Background Music	○	○

YESCUSTOM

Why did you choose the sensory adjustments you selected?

- ○ I prefer calmer websites
- ○ I usually struggle with sound/motion/brightness
- ○ I was curious to try it
- ○ No strong reason
- ○ [ ] Other:

YESCONTROL

Please rate how much you agree with the following statement about the personalization option.

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
I felt like I had control over my browsing experience.	○	○	○	○	○

NOCUSTOM

If you did not select any sensory adjustments, why not?

- I liked the default version
- I didn’t notice the option
- I wasn’t sure how to adjust the settings
- I didn’t feel the need to change anything
- Other: ____

NOCONTROL

Please rate how much you agree with the following statement about the personalization option.

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Even if you did not change anything, did having the option make you feel more in control?	○	○	○	○	○

CUS

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

Based on your selected options, please view the following version of the website:

Click here to visit the website

Opens in a new tab. Please return to continue the survey.

USERCHECK

Just to confirm, which of the following adjustments were applied to the second website version you viewed? You may choose more than one answer.

- ☐ The website had lower brightness
- ☐ The website had reduced motion effects
- ☐ The website had muted background music
- ☐ The second website looked exactly the same as the first one
- ☐ I'm not sure

B_EXP

Please rate the following statements based on your experience with the website:

	Very Low 1	Low 2	Neutral 3	High 4	Very Hig h 5
At this point, how would you rate your overall impression of the website so far?	○	○	○	○	○

B

Now that you've experienced the personalised version of the website, please answer the following questions.

Please rate the following statements based on your experience with the website:

	Very Low 1	Low 2	Neutral 3	High 4	Very Hig h 5
I felt comfortable using this website.	○	○	○	○	○
The experience felt emotionally pleasant.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low 1	Low 2	Neutral 3	High 4	Very Hig h 5
I was engaged while browsing this website.	○	○	○	○	○
The website held my attention.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I trust the quality of this website.	○	○	○	○	○
I feel confident using this website to plan something real.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I would consider booking a tour using this website.	○	○	○	○	○
I am interested in purchasing from this website.	○	○	○	○	○

Please rate the following statements based on your experience with the website:

	Very Low	Low	Neutral	High	Very Hig h
	1	2	3	4	5
I would recommend this website to someone planning a trip.	○	○	○	○	○
I think others would benefit from this website experience.	○	○	○	○	○

SHS

Below you will find a series of statements related to sensory experiences. For each statement, please rate how often you experience each situation using the following scale:

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
I suffer from allergies	○	○	○	○	○
I am allergy-free	○	○	○	○	○
I have a number of allergies	○	○	○	○	○
I often feel too hot in an environment where others don't seem to be bothered	○	○	○	○	○

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
I am easily disturbed by high temperatures	○	○	○	○	○

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
I often feel too cold in an environment where others don't seem to be bothered	○	○	○	○	○
I am easily disturbed by low temperatures	○	○	○	○	○
My eyes are sensitive to sunlight	○	○	○	○	○
I am sensitive to bright light	○	○	○	○	○
I am not really bothered by bright lights	○	○	○	○	○

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Always (5)
I am quite sensitive to pain	○	○	○	○	○
I can tolerate a large amount of pain	○	○	○	○	○
Things that would ordinarily hurt others are not painful to me	○	○	○	○	○
I often react to odors that other do not initially notice	○	○	○	○	○
I seem to notice smells that other people do not	○	○	○	○	○

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
I rarely notice smells	○	○	○	○	○
When I read, it must be totally quiet	○	○	○	○	○
I cannot study or read if there is any conversation or noise around	○	○	○	○	○
I can work even in noisy circumstances	○	○	○	○	○

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
I tend to be a picky eater	○	○	○	○	○

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
There are many foods that taste bad to me	○	○	○	○	○
I can eat almost anything	○	○	○	○	○
I am generally unable to wear clothes made of rough material	○	○	○	○	○
I am sensitive to rough textures	○	○	○	○	○
I can wear almost any kind of fabric without it	○	○	○	○	○

DEMOGRAPHICS

Tell us a bit about yourself!

Age

Gender?

I am

Do you identify as neurodivergent? (e.g., ADHD, autism, dyslexia, etc.)

I am

Any other thoughts or feedback? (Please avoid entering personal or identifying information here.)

VOUCHER

If you'd like to participate in our optional appreciation giveaway, click below:

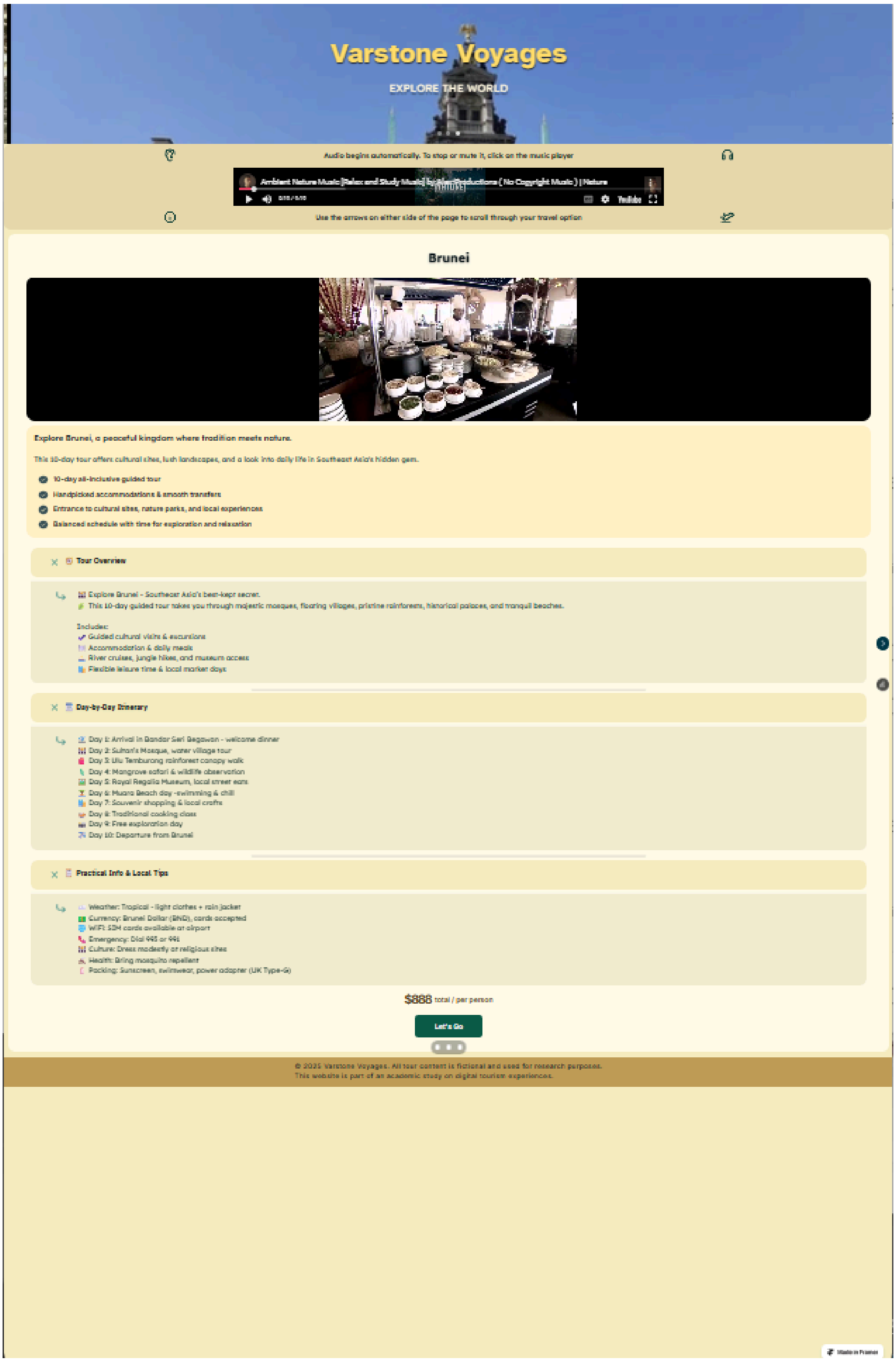
👉 🎟️ **Join to Be Considered for 1 of 5 €10 Vouchers of Your Choice**

Your email (if provided) will be used **only** to contact selected recipients and will be deleted after the selection process.
Your survey responses remain completely anonymous and cannot be linked to your email.
Thank you for supporting inclusive digital design!

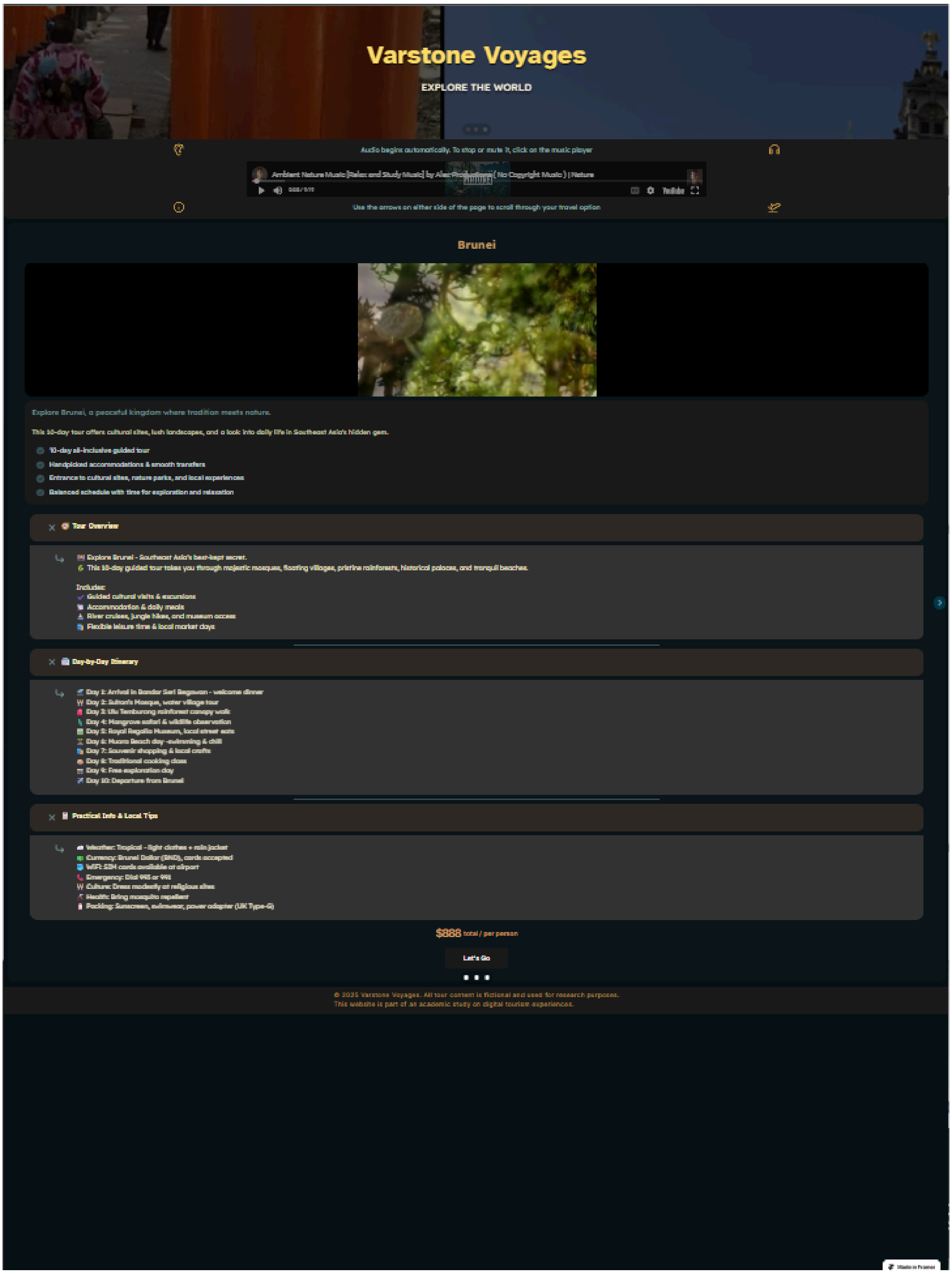
Please click next to complete survey.

Appendix B: Website Screenshot:

CONTROL WEBSITE SCREENSHOT



ADJUST BRIGHTNESS SCREENSHOT



Appendix C: SPSS Outputs

C0.2–C0.3 scale and compute audit; C1–Figure C1 normality; C2 Levene/Welch; C3–Figure C3 GLM diagnostics; C4 Poisson; C5–Figure C5 influence; C6 reliability; C7 descriptives/correlations/stability; C8–Figure C8 paired tests and CIs; C9 H2 summary; C10–Figure C10 H3; C11 non-parametric; C12 sensitivity; C13 data quality and timing.

How to read this appendix.

Each item lists: label, caption, variables, and what to paste from SPSS. Labels C1–C5 satisfy the assumption and diagnostics cross-walk referenced in Chapter 3.5; subsequent items house the core results and supporting descriptives. The variable map below fixes the names used across outputs so captions match Chapter 4.

C0. Data audit and variable map

Table C0.1. Variable map and derived fields (links §3.3, §3.5).
Variables.

- Outcomes at baseline: COMFORT_A, TRUST_A, ENGAGE_A, PI_A, WOM_A, A_IMP
- Outcomes post: COMFORT_B, TRUST_B, ENGAGE_B, PI_B, WOM_B, B_IMP
- Change scores: D_COMFORT, D_TRUST, D_ENGAGE, D_PI, D_WOM, D_IMP = Post − Pre
- Trait and covariates: SHS_TOTAL, AGE, GENDER(F,M,NB), NEUROTYPE(0,1)
- Behavior: CHG_N in {0,1,2,3}

Codebook

A_COM

		Value
Standard Attributes	Position	128
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	151
	Missing	119
Central Tendency and Dispersion	Mean	3.543
	Standard Deviation	.9237
	Percentile 25	3.000
	Percentile 50	3.500
	Percentile 75	4.000

A_ENG

		Value
Standard Attributes	Position	129
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	151
	Missing	119
Central Tendency and Dispersion	Mean	3.523
	Standard Deviation	.9351
	Percentile 25	3.000
	Percentile 50	4.000
	Percentile 75	4.000

A_TRU

		Value
Standard Attributes	Position	130
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	151
	Missing	119
Central Tendency and Dispersion	Mean	2.974
	Standard Deviation	1.0532
	Percentile 25	2.000
	Percentile 50	3.000
	Percentile 75	4.000

A_PI

		Value
Standard Attributes	Position	131
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	151
	Missing	119
Central Tendency and Dispersion	Mean	2.921
	Standard Deviation	1.0710
	Percentile 25	2.000
	Percentile 50	3.000
	Percentile 75	4.000

A_WOM

		Value
Standard Attributes	Position	132
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	151
	Missing	119
Central Tendency and Dispersion	Mean	3.159
	Standard Deviation	1.0478
	Percentile 25	2.500
	Percentile 50	3.000
	Percentile 75	4.000

A_IMP

		Value	Count	Percent
Standard Attributes	Position	16		
	Label	A_IMP		
	Type	Numeric		
	Format	F40		
	Measurement	Scale		
	Role	Input		
N	Valid	156		
	Missing	114		
Central Tendency and Dispersion	Mean	3.39		
	Standard Deviation	.941		
	Percentile 25	3.00		
	Percentile 50	4.00		
	Percentile 75	4.00		
Labeled Values	1	Very low	3	1.1%
	2	Low	28	10.4%
	3	Neutral	43	15.9%
	4	High	69	25.6%
	5	Very high	13	4.8%

B_COM

		Value
Standard Attributes	Position	133
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	3.510
	Standard Deviation	.9586
	Percentile 25	3.000
	Percentile 50	3.500
	Percentile 75	4.000

B_ENG

		Value
Standard Attributes	Position	134
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	3.435
	Standard Deviation	.9492
	Percentile 25	3.000
	Percentile 50	3.500
	Percentile 75	4.000

B_TRU

		Value
Standard Attributes	Position	135
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	3.014
	Standard Deviation	1.1203
	Percentile 25	2.000
	Percentile 50	3.000
	Percentile 75	4.000

B_PI

		Value
Standard Attributes	Position	136
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	2.928
	Standard Deviation	1.1157
	Percentile 25	2.000
	Percentile 50	3.000
	Percentile 75	4.000

B_WOM

		Value
Standard Attributes	Position	137
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	3.134
	Standard Deviation	1.0864
	Percentile 25	2.500
	Percentile 50	3.000
	Percentile 75	4.000

B_IMP

		Value	Count	Percent
Standard Attributes	Position	44		
	Label	B_IMP		
	Type	Numeric		
	Format	F40		
	Measurement	Scale		
	Role	Input		
N	Valid	99		
	Missing	171		
Central Tendency and Dispersion	Mean	3.48		
	Standard Deviation	.941		
	Percentile 25	3.00		
	Percentile 50	4.00		
	Percentile 75	4.00		
Labeled Values	1	Very low	1	0.4%
	2	Low	15	5.6%
	3	Neutral	31	11.5%
	4	High	39	14.4%
	5	Very high	13	4.8%

D_COM

		Value
Standard Attributes	Position	138
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	-.072
	Standard Deviation	.7656
	Percentile 25	-.500
	Percentile 50	.000
	Percentile 75	.500

D_TRU

		Value
Standard Attributes	Position	139
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	.021
	Standard Deviation	.7563
	Percentile 25	-.500
	Percentile 50	.000
	Percentile 75	.500

D_ENG

		Value
Standard Attributes	Position	140
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	-.127
	Standard Deviation	.5954
	Percentile 25	-.500
	Percentile 50	.000
	Percentile 75	.000

D_PI

		Value
Standard Attributes	Position	141
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	-.014
	Standard Deviation	.5871
	Percentile 25	.000
	Percentile 50	.000
	Percentile 75	.000

D_WOM

		Value
Standard Attributes	Position	142
	Label	<none>
	Type	Numeric
	Format	F8.1
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	-.038
	Standard Deviation	.7000
	Percentile 25	-.500
	Percentile 50	.000
	Percentile 75	.000

SHSTOTAL

		Value
Standard Attributes	Position	114
	Label	<none>
	Type	Numeric
	Format	F8.2
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	73.8219
	Standard Deviation	8.30346
	Percentile 25	68.0000
	Percentile 50	75.0000
	Percentile 75	80.0000

REALAGE

		Value
Standard Attributes	Position	115
	Label	<none>
	Type	Numeric
	Format	F8.2
	Measurement	Scale
	Role	Input
N	Valid	146
	Missing	124
Central Tendency and Dispersion	Mean	26.2603
	Standard Deviation	6.91929
	Percentile 25	22.0000
	Percentile 50	24.0000
	Percentile 75	28.0000

GENDER_1

		Value	Count	Percent
Standard Attributes	Position	119		
	Label	Gender		
	Type	Numeric		
	Format	F40		
	Measurement	Scale		
	Role	Input		
N	Valid	146		
	Missing	124		
Central Tendency and Dispersion	Mean	1.64		
	Standard Deviation	.597		
	Percentile 25	1.00		
	Percentile 50	2.00		
	Percentile 75	2.00		
Labeled Values	1	Male	62	23.0%
	2	Female	75	27.8%
	3	Other	9	3.3%
	4	Prefer not to say	0	0.0%

NEURO_1

		Value	Count	Percent
Standard Attributes	Position	118		
	Label	Neurodivergence		
	Type	Numeric		
	Format	F40		
	Measurement	Scale		
	Role	Input		
N	Valid	146		
	Missing	124		
Central Tendency and Dispersion	Mean	2.02		
	Standard Deviation	.679		
	Percentile 25	2.00		
	Percentile 50	2.00		
	Percentile 75	2.00		
Labeled Values	1	Yes	31	11.5%
	2	No	82	30.4%
	3	Unsure	32	11.9%
	4	Prefer not to say	1	0.4%

BRIGHT_FLAG

		Value	Count	Percent
Standard Attributes	Position	124		
	Label	Brightness change flag		
	Type	Numeric		
	Format	F8.2		
	Measurement	Nominal		
	Role	Input		
Valid Values	.00	No change	109	40.4%
	1.00	Change	42	15.6%
Missing Values	System		119	44.1%

MOTION_FLAG

		Value	Count	Percent
Standard Attributes	Position	125		
	Label	Motion change flag		
	Type	Numeric		
	Format	F8.2		
	Measurement	Nominal		
	Role	Input		
Valid Values	.00	No change	98	36.3%
	1.00	Change	53	19.6%
Missing Values	System		119	44.1%

AUDIO_FLAG

		Value	Count	Percent
Standard Attributes	Position	126		
	Label	Audio change flag		
	Type	Numeric		
	Format	F8.2		
	Measurement	Nominal		
	Role	Input		
Valid Values	.00	No change	75	27.8%
	1.00	Change	76	28.1%
Missing Values	System		119	44.1%

CHG_N

		Value	Count	Percent
Standard Attributes	Position	127		
	Label	Number of changes selected (0-3)		
	Type	Numeric		
	Format	F8.2		
	Measurement	Scale		
	Role	Input		
N	Valid	151		
	Missing	119		
Central Tendency and Dispersion	Mean	1.1325		
	Standard Deviation	.95691		
	Percentile 25	.0000		
	Percentile 50	1.0000		
	Percentile 75	2.0000		
Labeled Values	.00	0 changes	44	16.3%
	1.00	1 change	59	21.9%
	2.00	2 changes	32	11.9%
	3.00	3 changes	16	5.9%

Table C0.2. Scale composition and reliability plan (links §3.3.3, §4.4).

Item lists for each 2-item composite, alpha formula reference, item-to-total notes.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.809	.809	2

Inter-Item Correlation Matrix

	A_COMFORT	A_PLEASANT
A_COMFORT	1.000	.680
A_PLEASANT	.680	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.874	.876	2

Inter-Item Correlation Matrix

	A_ENGAGED	A_ATTENTION
A_ENGAGED	1.000	.779
A_ATTENTION	.779	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.909	.910	2

Inter-Item Correlation Matrix

	A_TRUST_QUALITY	A_TRUST_CONFIDENCE
A_TRUST_QUALITY	1.000	.834
A_TRUST_CONFIDENCE	.834	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.947	.947	2

Inter-Item Correlation Matrix

	A_PURCHASE_INTENT_1	A_PURCHASE_INTENT_2
A_PURCHASE_INTENT_1	1.000	.899
A_PURCHASE_INTENT_2	.899	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.883	.884	2

Inter-Item Correlation Matrix

	A_WOM_SHARE	A_WOM_BENEFIT
A_WOM_SHARE	1.000	.792
A_WOM_BENEFIT	.792	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.916	.916	2

Inter-Item Correlation Matrix

	B_COMFORT	B_PLEASANT
B_COMFORT	1.000	.845
B_PLEASANT	.845	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.922	.923	2

Inter-Item Correlation Matrix

	B_ENGAGED	B_ATTENTION
B_ENGAGED	1.000	.857
B_ATTENTION	.857	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.934	.934	2

Inter-Item Correlation Matrix

	B_TRUST_QUALITY	B_TRUST_CONFIDENCE
B_TRUST_QUALITY	1.000	.876
B_TRUST_CONFIDENCE	.876	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.941	.941	2

Inter-Item Correlation Matrix

	B_PURCHASE_INTENT_1	B_PURCHASE_INTENT_2
B_PURCHASE_INTENT_1	1.000	.889
B_PURCHASE_INTENT_2	.889	1.000

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.916	.916	2

Inter-Item Correlation Matrix

	B_WOM_SHARE	B_WOM_BENEFIT
B_WOM_SHARE	1.000	.845
B_WOM_BENEFIT	.845	1.000

Table C6.1. Internal consistency of baseline composites (n=151)

Construct	Items (baseline)	Inter -item r	Cronbach’s α	Note
Comfort	A_COMFORT, A_PLEASANT	.680	.809	α = 2r/(1+r) check: 2*.680/(1+.680)=.809
Engagement	A_ENGAGED, A_ATTENTION	.779	.874	2*.779/(1+.779)=.876 (matches within rounding)
Trust	A_TRUST_QUALITY, A_TRUST_CONFIDENCE	.834	.909	2*.834/(1+.834)=.910
Purchase intention	A_PURCHASE_INTENT_1, A_PURCHASE_INTENT_2	.899	.947	2*.899/(1+.899)=.947
Word-of-mouth intention	A_WOM_SHARE, A_WOM_BENEFIT	.792	.883	2*.792/(1+.792)=.884

Case processing summary: Valid n=151 of 270 starts, listwise deletion across scale items.

Table C6.2. Internal consistency of post composites (n=146)

Construct	Items (post)	Inter-item r	Cronbach’s α	Note
Comfort	B_COMFORT, B_PLEASANT	.845	.916	2*.845/(1+.845)=.915

Engagement	B_ENGAGED, B_ATTENTION	.857	.922	2*.857/(1+.857)=.923
Trust	B_TRUST_QUALITY, B_TRUST_CONFIDENCE	.876	.934	2*.876/(1+.876)=.934
Purchase intention	B_PURCHASE_INTENT_1, B_PURCHASE_INTENT_2	.889	.941	2*.889/(1+.889)=.942
Word-of-mouth intention	B_WOM_SHARE, B_WOM_BENEFIT	.845	.916	2*.845/(1+.845)=.915

Case processing summary: Valid n=146 of 270 starts, listwise deletion across scale items.

Note: *For two-item composites, α equals the Spearman-Brown corrected inter-item correlation, α = 2r/(1+r). Values above .80 are generally considered good for research use (Field, 2018).*

Table C0.3. Computation audit for change scores (links §3.5.1).
Definition D_X = X_B − X_A for all outcomes.

SPSS SYNTAX:

```
COMPUTE D_COM = B_COM - A_COM.

COMPUTE D_TRU = B_TRU - A_TRU.

COMPUTE D_ENG = B_ENG - A_ENG.

COMPUTE D_PI  = B_PI  - A_PI.

COMPUTE D_WOM = B_WOM - A_WOM.

COMPUTE D_IMP = B_IMP - A_IMP.

EXECUTE.
```

C1. Normality of within-subject change scores for H1

Table C1. Shapiro-Wilk tests for change scores D_X (links §4.6, §3.5 cross-walk).

Tests of Normality

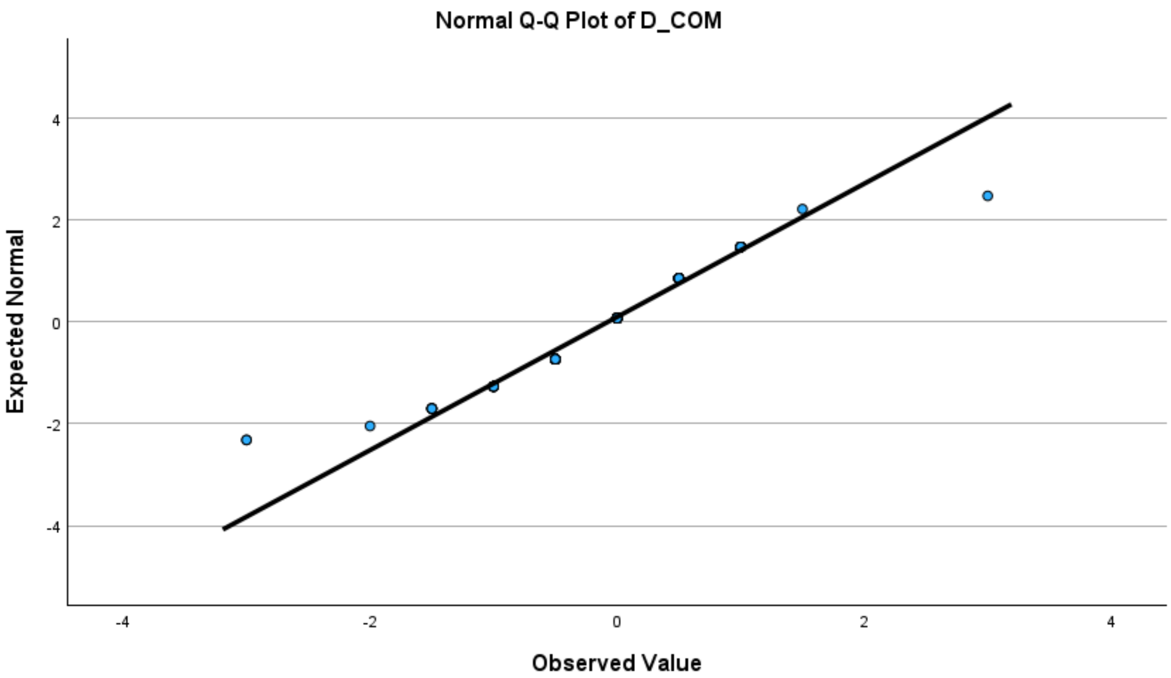
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
D_COM	.216	146	<.001	.899	146	<.001
D_TRU	.229	146	<.001	.900	146	<.001
D_ENG	.290	146	<.001	.860	146	<.001
D_PI	.290	146	<.001	.842	146	<.001
D_WOM	.268	146	<.001	.883	146	<.001

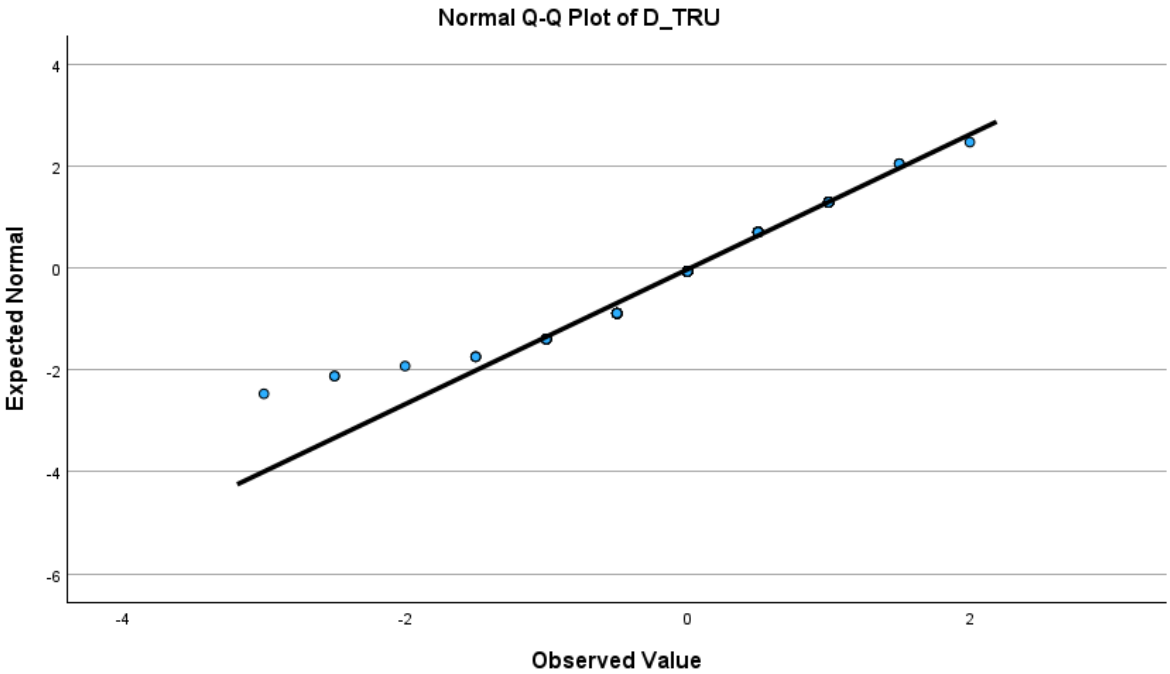
a. Lilliefors Significance Correction

Variables. D_COMFORT, D_TRUST, D_ENGAGE, D_PI, D_WOM, D_IMP

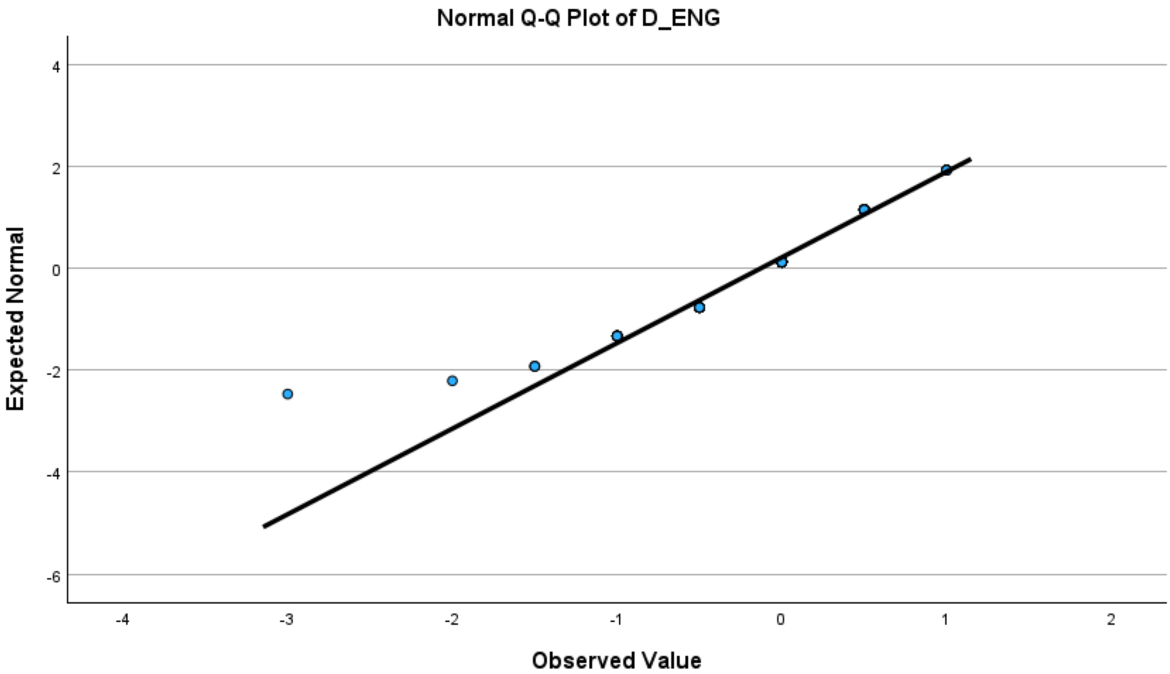
Analyze → Descriptive Statistics → Explore... Dependent List: all D_*; Plots: Normality plots with tests.

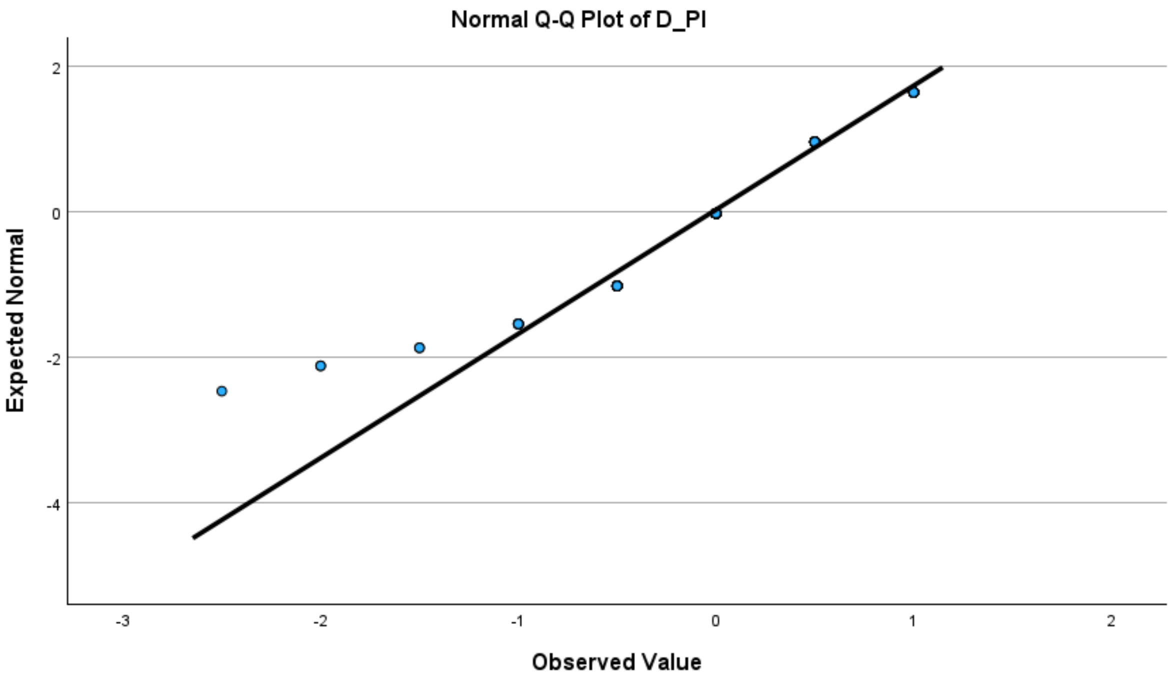
Figure C1. Q-Q plots of change scores D_X (links §4.6).

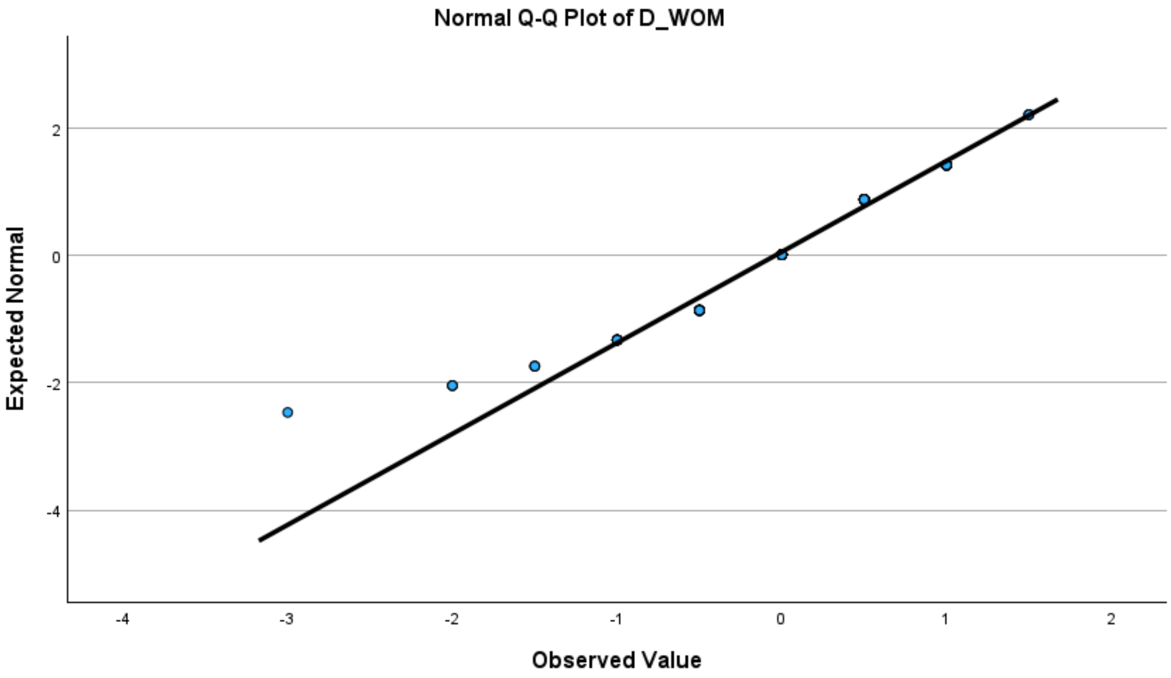




Normal Q-Q Plot of D_TRU
Expected Normal
Observed Value
4
2
0
-2
-4
-6
-4
-2
0
2


Normal Q-Q Plot of D_ENG
Expected Normal
Observed Value
4
2
0
-2
-4
-6
-4
-3
-2
-1
0
1
2


Normal Q-Q Plot of D_PI
Expected Normal
Observed Value
2
0
-2
-4
-3
-2
-1
0
1
2



C2. Homogeneity and Welch adjustments for H3

Table C2.1. Levene tests for baseline outcomes across SHS tertiles (links §4.8).

Variables. COMFORT_A, TRUST_A, ENGAGE_A, PI_A, WOM_A by SHS_TERTILE(Low, Mid, High)

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
A_COM	Based on Mean	.801	2	143	.451
	Based on Median	.946	2	143	.391
	Based on Median and with adjusted df	.946	2	136.282	.391
	Based on trimmed mean	.799	2	143	.452
A_ENG	Based on Mean	1.130	2	143	.326
	Based on Median	.891	2	143	.413
	Based on Median and with adjusted df	.891	2	140.711	.413
	Based on trimmed mean	1.027	2	143	.361
A_TRU	Based on Mean	.084	2	143	.919
	Based on Median	.183	2	143	.833
	Based on Median and with adjusted df	.183	2	142.901	.833
	Based on trimmed mean	.145	2	143	.865
A_PI	Based on Mean	.165	2	143	.848
	Based on Median	.019	2	143	.981
	Based on Median and with adjusted df	.019	2	137.759	.981
	Based on trimmed mean	.148	2	143	.862
A_WOM	Based on Mean	.053	2	143	.948
	Based on Median	.128	2	143	.880
	Based on Median and with adjusted df	.128	2	140.955	.880
	Based on trimmed mean	.063	2	143	.939

Table C2.2. Welch ANOVA where variances are unequal (links §4.8).

Robust Tests of Equality of Means

		Statistic[a]	df1	df2	Sig.
A_COM	Welch	2.745	2	95.070	.069
A_ENG	Welch	1.144	2	95.201	.323
A_TRU	Welch	1.045	2	94.443	.356
A_PI	Welch	2.934	2	94.603	.058
A_WOM	Welch	1.345	2	95.004	.265

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable		(I) SHS tertiles: Low Mid High	(J) SHS tertiles: Low Mid High	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
A_COM	Tukey HSD	Low SHS	Mid SHS	-.2693	.1817	.302	-.700	.161
			High SHS	-.4271*	.1768	.044	-.846	-.008
		Mid SHS	Low SHS	.2693	.1817	.302	-.161	.700
			High SHS	-.1578	.1825	.663	-.590	.274
		High SHS	Low SHS	.4271*	.1768	.044	.008	.846
			Mid SHS	.1578	.1825	.663	-.274	.590
	Games-Howell	Low SHS	Mid SHS	-.2693	.1834	.311	-.706	.167
			High SHS	-.4271	.1822	.055	-.861	.007
		Mid SHS	Low SHS	.2693	.1834	.311	-.167	.706
			High SHS	-.1578	.1716	.630	-.567	.251
		High SHS	Low SHS	.4271	.1822	.055	-.007	.861
			Mid SHS	.1578	.1716	.630	-.251	.567
A_ENG	Tukey HSD	Low SHS	Mid SHS	-.2092	.1857	.499	-.649	.231
			High SHS	-.2780	.1807	.276	-.706	.150
		Mid SHS	Low SHS	.2092	.1857	.499	-.231	.649
			High SHS	-.0689	.1865	.928	-.511	.373
		High SHS	Low SHS	.2780	.1807	.276	-.150	.706
			Mid SHS	.0689	.1865	.928	-.373	.511
	Games-Howell	Low SHS	Mid SHS	-.2092	.1855	.500	-.651	.233
			High SHS	-.2780	.1883	.307	-.726	.170
		Mid SHS	Low SHS	.2092	.1855	.500	-.233	.651
			High SHS	-.0689	.1740	.917	-.483	.346
		High SHS	Low SHS	.2780	.1883	.307	-.170	.726
			Mid SHS	.0689	.1740	.917	-.346	.483
A_TRU	Tukey HSD	Low SHS	Mid SHS	-.0190	.2154	.996	-.529	.491
			High SHS	-.2778	.2096	.383	-.774	.219
		Mid SHS	Low SHS	.0190	.2154	.996	-.491	.529
			High SHS	-.2589	.2164	.457	-.771	.254
		High SHS	Low SHS	.2778	.2096	.383	-.219	.774
			Mid SHS	.2589	.2164	.457	-.254	.771
	Games-Howell	Low SHS	Mid SHS	-.0190	.2129	.996	-.526	.488
			High SHS	-.2778	.2089	.382	-.775	.219
		Mid SHS	Low SHS	.0190	.2129	.996	-.488	.526
			High SHS	-.2589	.2201	.470	-.783	.266
		High SHS	Low SHS	.2778	.2089	.382	-.219	.775
			Mid SHS	.2589	.2201	.470	-.266	.783
A_PI	Tukey HSD	Low SHS	Mid SHS	-.3444	.2148	.248	-.853	.164
			High SHS	-.4933	.2091	.051	-.988	.002
		Mid SHS	Low SHS	.3444	.2148	.248	-.164	.853
			High SHS	-.1489	.2159	.770	-.660	.362
		High SHS	Low SHS	.4933	.2091	.051	-.002	.988
			Mid SHS	.1489	.2159	.770	-.362	.660
	Games-Howell	Low SHS	Mid SHS	-.3444	.2097	.233	-.844	.155
			High SHS	-.4933	.2105	.055	-.994	.008
		Mid SHS	Low SHS	.3444	.2097	.233	-.155	.844
			High SHS	-.1489	.2192	.776	-.671	.373
		High SHS	Low SHS	.4933	.2105	.055	-.008	.994
			Mid SHS	.1489	.2192	.776	-.373	.671
A_WOM	Tukey HSD	Low SHS	Mid SHS	-.2418	.2141	.497	-.749	.265
			High SHS	-.3396	.2083	.236	-.833	.154
		Mid SHS	Low SHS	.2418	.2141	.497	-.265	.749
			High SHS	-.0978	.2151	.892	-.607	.412
		High SHS	Low SHS	.3396	.2083	.236	-.154	.833
			Mid SHS	.0978	.2151	.892	-.412	.607
	Games-Howell	Low SHS	Mid SHS	-.2418	.2139	.498	-.751	.268
			High SHS	-.3396	.2115	.248	-.843	.164
		Mid SHS	Low SHS	.2418	.2139	.498	-.268	.751
			High SHS	-.0978	.2102	.888	-.599	.403
		High SHS	Low SHS	.3396	.2115	.248	-.164	.843
			Mid SHS	.0978	.2102	.888	-.403	.599

*. The mean difference is significant at the 0.05 level.

C3. GLM diagnostics and VIFs for H2

Table C3.1. GLM predicting CHG_N from traits and demographics (links §4.7).
Model. CHG_N = b0 + b1 SHS_TOTAL + b2 AGE + Gender dummies + NEUROTYPE + e

Between-Subjects Factors

		Value Label	N
Gender	1	Male	62
	2	Female	75
	3	Other	9
Neurodivergence	1	Yes	31
	2	No	82
	3	Unsure	32
	4	Prefer not to say	1

Parameter Estimates

Dependent Variable: Number of changes selected (0-3)

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Intercept	1.568	1.494	1.050	.296	-1.386	4.522	.008	1.050	.181
SHSTOTAL	.000	.011	.012	.990	-.021	.021	.000	.012	.050
REALAGE	.007	.012	.590	.556	-.017	.031	.003	.590	.090
[GENDER_1=1]	-.774	.486	-1.592	.114	-1.735	.187	.019	1.592	.352
[GENDER_1=2]	-.646	.479	-1.349	.180	-1.593	.301	.013	1.349	.268
[GENDER_1=3]	0[a]	.	.	.	.	.	.	.	.
[NEURO_1=1]	.207	1.295	.160	.873	-2.355	2.769	.000	.160	.053
[NEURO_1=2]	.284	1.441	.197	.844	-2.566	3.135	.000	.197	.054
[NEURO_1=3]	.242	1.003	.242	.810	-1.742	2.227	.000	.242	.057
[NEURO_1=4]	0[a]	.	.	.	.	.	.	.	.
[GENDER_1=1] * [NEURO_1=1]	.029	.892	.032	.974	-1.736	1.794	.000	.032	.050
[GENDER_1=1] * [NEURO_1=2]	-.228	1.083	-.211	.833	-2.371	1.914	.000	.211	.055
[GENDER_1=1] * [NEURO_1=3]	0[a]	.	.	.	.	.	.	.	.
[GENDER_1=1] * [NEURO_1=4]	0[a]	.	.	.	.	.	.	.	.
[GENDER_1=2] * [NEURO_1=1]	-.411	.879	-.468	.641	-2.151	1.328	.002	.468	.075
[GENDER_1=2] * [NEURO_1=2]	-.498	1.079	-.462	.645	-2.633	1.636	.002	.462	.074
[GENDER_1=2] * [NEURO_1=3]	0[a]	.	.	.	.	.	.	.	.
[GENDER_1=3] * [NEURO_1=1]	0[a]	.	.	.	.	.	.	.	.
[GENDER_1=3] * [NEURO_1=2]	0[a]	.	.	.	.	.	.	.	.
[GENDER_1=3] * [NEURO_1=3]	0[a]	.	.	.	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

b. Computed using alpha = .05

Descriptive Statistics

Dependent Variable: Number of changes selected (0-3)

Gender	Neurodivergence	Mean	Std. Deviation	N
Male	Yes	1.2500	1.21543	12
	No	1.0270	.89711	37
	Unsure	1.2500	.96531	12
	Prefer not to say	1.0000	.	1
	Total	1.1129	.95993	62
Female	Yes	.9412	1.08804	17
	No	.9091	.85775	44
	Unsure	1.3571	.84190	14
	Total	1.0000	.91533	75
Other	Yes	2.0000	.00000	2
	No	2.0000	.	1
	Unsure	2.0000	1.26491	6
	Total	2.0000	1.00000	9
Total	Yes	1.1290	1.11779	31
	No	.9756	.87455	82
	Unsure	1.4375	.98169	32
	Prefer not to say	1.0000	.	1
	Total	1.1096	.96220	146

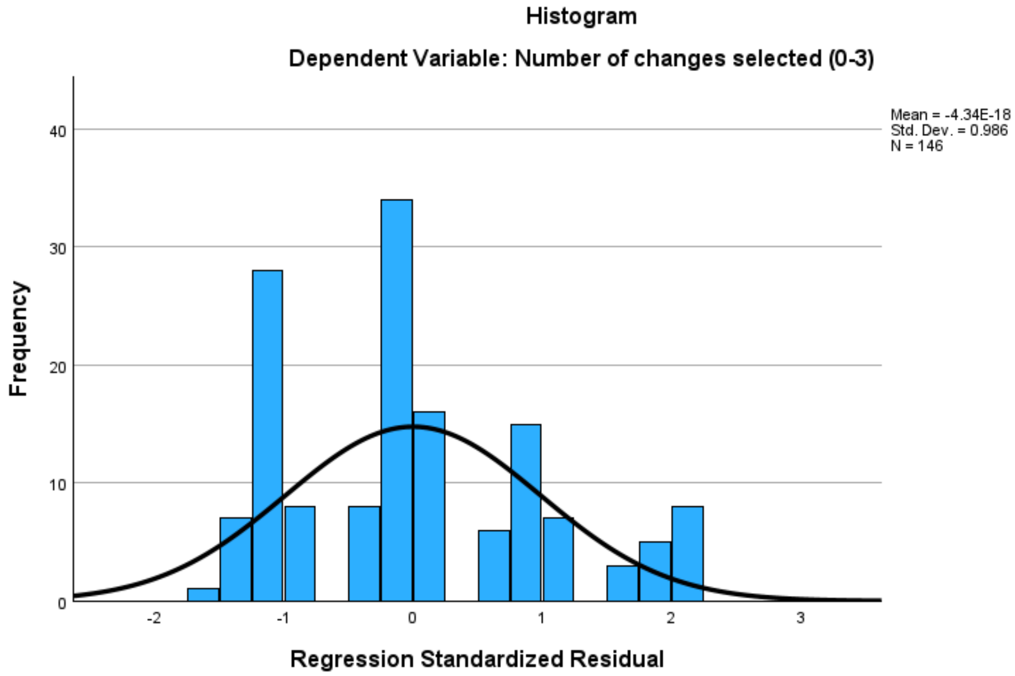
Table C3.2. Multicollinearity diagnostics (VIF) for H2 (links §4.7).

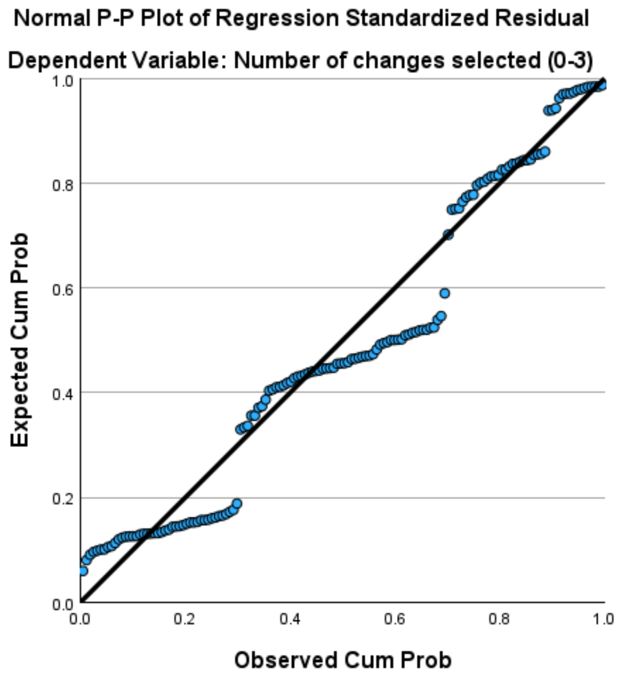
Coefficients[a]

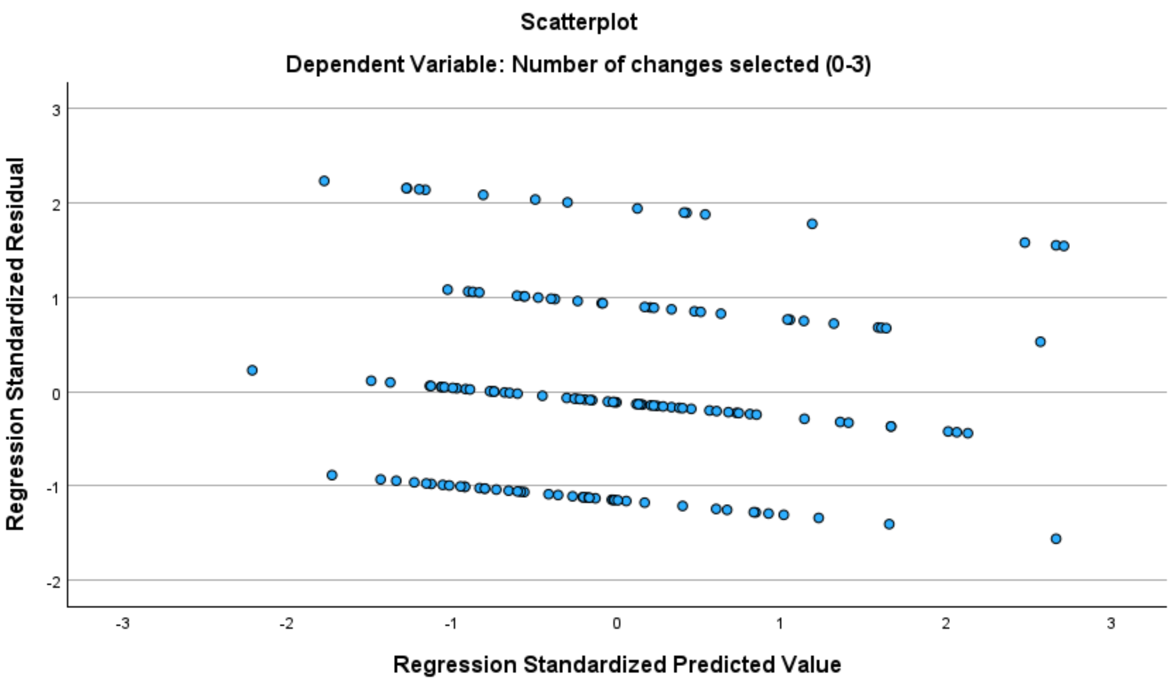
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	-.080	.860		-.093	.926		
	SHSTOTAL	.007	.010	.056	.676	.500	.997	1.003
	REALAGE	.007	.012	.050	.598	.550	.973	1.027
	Gender	.134	.135	.083	.994	.322	.990	1.010
	Neurodivergence	.151	.119	.106	1.263	.209	.976	1.025

a. Dependent Variable: Number of changes selected (0-3)

Figure C3. Residual diagnostics for GLM on CHG_N (links §4.7).







C4. Poisson robustness for CHG_N

Table C4. Poisson regression of CHG_N with log link (links §4.10).

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.643	1.3686	-1.040	4.325	1.441	1	.230
[Gender=1]	-.740	.3445	-1.415	-.064	4.609	1	.032
[Gender=2]	-.865	.3436	-1.539	-.192	6.342	1	.012
[Gender=3]	0[a]	.	.	.	.	.	.
[Neurodivergence=1]	.128	.9438	-1.721	1.978	.019	1	.892
[Neurodivergence=2]	.035	.9403	-1.808	1.878	.001	1	.970
[Neurodivergence=3]	.349	.9469	-1.507	2.205	.136	1	.712
[Neurodivergence=4]	0[a]	.	.	.	.	.	.
SHSTOTAL	-.001	.0098	-.020	.018	.006	1	.937
REALAGE	.006	.0116	-.017	.029	.275	1	.600
(Scale)	.848[b]	.0992	.674	1.066			

Dependent Variable: Number of changes selected (0-3)
Model: (Intercept), Gender, Neurodivergence, SHSTOTAL, REALAGE

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Goodness of Fit[a]

	Value	df	Value/df
Deviance	123.743	138	.897
Scaled Deviance	146.000	138	
Pearson Chi-Square	123.743	138	.897
Scaled Pearson Chi-Square	146.000	138	
Log Likelihood[b]	-195.091		
Akaike's Information Criterion (AIC)	408.182		
Finite Sample Corrected AIC (AICC)	409.505		
Bayesian Information Criterion (BIC)	435.034		
Consistent AIC (CAIC)	444.034		

Dependent Variable: Number of changes selected (0-3)
Model: (Intercept), Gender, Neurodivergence, SHSTOTAL, REALAGE

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Note: Check overdispersion: Pearson Chi-Square df ratio considerably > 1 suggests overdispersion.

C5. Outlier and influence screening across models

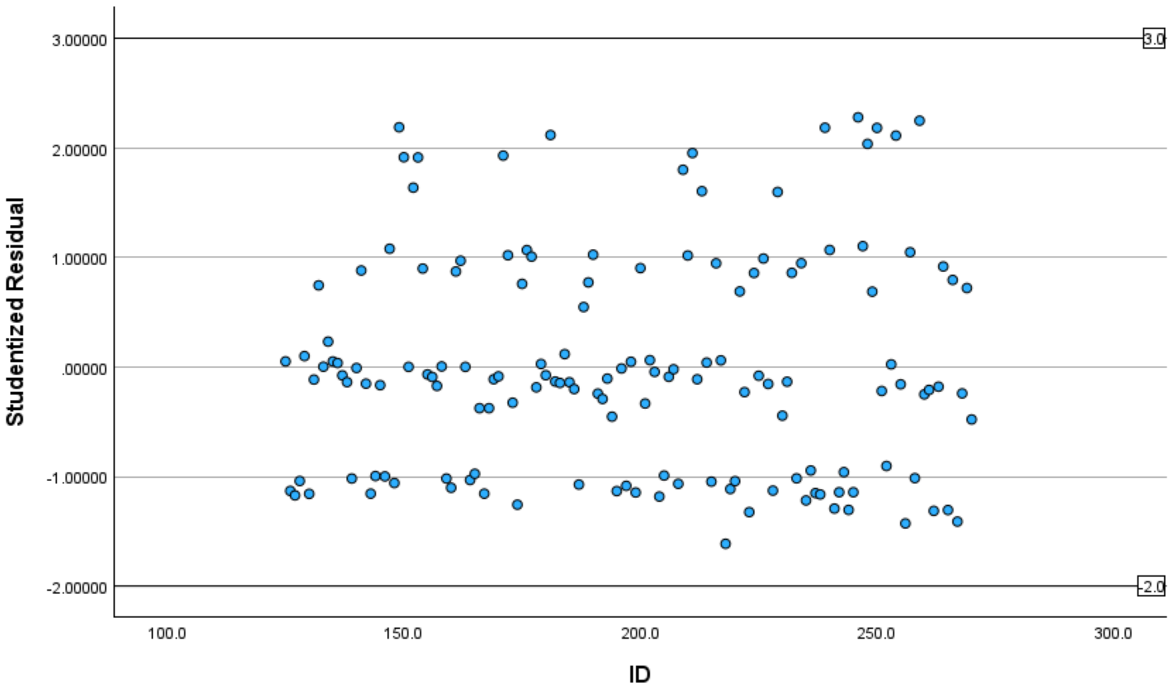
Table C5. Studentized residuals, leverage, and Cook's D for H2 GLM (links §4.10).

Criteria. |Studentized residual| > 3, Cook's D > 4/n, leverage > 2p/n

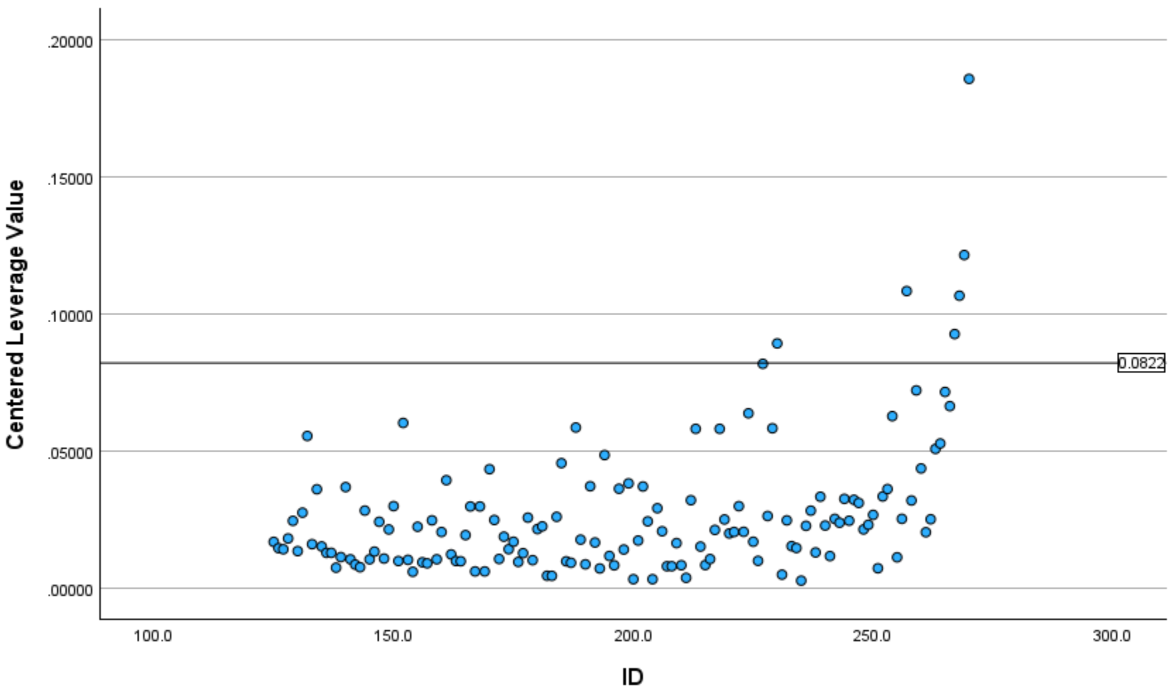
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Studentized Residual	146	-1.61341	2.27886	.0012892	1.00464507
Cook's Distance	146	.00000	.08677	.0077289	.01287481
Centered Leverage Value	146	.00285	.18577	.0273973	.02559038
Valid N (listwise)	146				

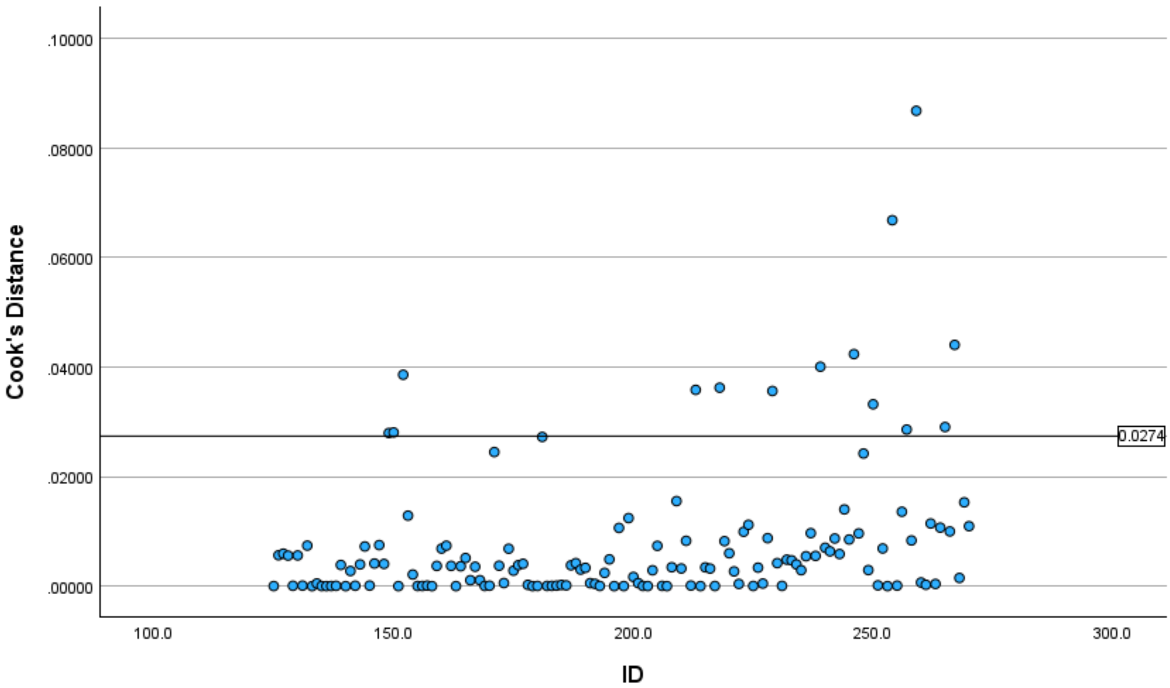
Figure C5. Influence plot by Studentized residuals, leverage, and Cook's D (links §4.10).



3.00000
2.00000
1.00000
.00000
-1.00000
-2.00000
3.0
-2.0
Studentized Residual
100.0
150.0
200.0
250.0
300.0
ID



.20000
.15000
.10000
.05000
.00000
0.0822
Centered Leverage Value
100.0
150.0
200.0
250.0
300.0
ID



C6. Reliability

Table C6.1. Internal consistency at baseline (links §4.4).

Variables. Each 2-item composite at time A

Item Statistics

	Mean	Std. Deviation	N
A_COMFORT	13.52	.992	151
A_PLEASANT	13.56	1.024	151

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.809	.809	2

Item Statistics

	Mean	Std. Deviation	N
A_ENGAGED	13.54	.951	151
A_ATTENTION	13.51	1.032	151

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.874	.876	2

Item Statistics

	Mean	Std. Deviation	N
A_TRUST_QUALITY	13.03	1.083	151
A_TRUST_CONFIDENCE	12.92	1.117	151

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.909	.910	2

Item Statistics

	Mean	Std. Deviation	N
A_PURCHASE_INTENT_1	12.99	1.092	151
A_PURCHASE_INTENT_2	12.85	1.106	151

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.947	.947	2

Item Statistics

	Mean	Std. Deviation	N
A_WOM_SHARE	13.07	1.129	151
A_WOM_BENEFIT	13.25	1.085	151

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.883	.884	2

Table C6.2. Internal consistency post (links §4.4).

Item Statistics

	Mean	Std. Deviation	N
B_COMFORT	13.53	.991	146
B_PLEASANT	13.49	1.005	146

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.916	.916	2

Item Statistics

	Mean	Std. Deviation	N
B_ENGAGED	13.45	.947	146
B_ATTENTION	13.42	1.023	146

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.922	.923	2

Item Statistics

	Mean	Std. Deviation	N
B_TRUST_QUALITY	13.02	1.148	146
B_TRUST_CONFIDENCE	13.01	1.166	146

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.934	.934	2

Item Statistics

	Mean	Std. Deviation	N
B_PURCHASE_INTENT_1	12.99	1.148	146
B_PURCHASE_INTENT_2	12.86	1.148	146

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.941	.941	2

Item Statistics

	Mean	Std. Deviation	N
B_WOM_SHARE	13.08	1.130	146
B_WOM_BENEFIT	13.18	1.133	146

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.916	.916	2

C7. Descriptives and stability

Table C7.1. Descriptive statistics for all outcomes at A and B (links §4.5).

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
A_COM	151	1.0	5.0	3.543	.9237
A_ENG	151	1.0	5.0	3.523	.9351
A_TRU	151	1.0	5.0	2.974	1.0532
A_PI	151	1.0	5.0	2.921	1.0710
A_WOM	151	1.0	5.0	3.159	1.0478
B_COM	146	1.0	5.0	3.510	.9586
B_ENG	146	1.0	5.0	3.435	.9492
B_TRU	146	1.0	5.0	3.014	1.1203
B_PI	146	1.0	5.0	2.928	1.1157
B_WOM	146	1.0	5.0	3.134	1.0864
Valid N (listwise)	146				

Table C7.2. Correlations among outcomes at baseline and post (links §4.5).

Correlations

		A_COM	A_ENG	A_TRU	A_PI	A_WOM
A_COM	Pearson Correlation	1	.726**	.561**	.645**	.630**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001
	N	151	151	151	151	151
A_ENG	Pearson Correlation	.726**	1	.578**	.651**	.605**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001
	N	151	151	151	151	151
A_TRU	Pearson Correlation	.561**	.578**	1	.775**	.773**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001
	N	151	151	151	151	151
A_PI	Pearson Correlation	.645**	.651**	.775**	1	.812**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001
	N	151	151	151	151	151
A_WOM	Pearson Correlation	.630**	.605**	.773**	.812**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	
	N	151	151	151	151	151

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		B_COM	B_ENG	B_TRU	B_PI	B_WOM
B_COM	Pearson Correlation	1	.776**	.621**	.652**	.654**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001
	N	146	146	146	146	146
B_ENG	Pearson Correlation	.776**	1	.609**	.653**	.701**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001
	N	146	146	146	146	146
B_TRU	Pearson Correlation	.621**	.609**	1	.864**	.871**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001
	N	146	146	146	146	146
B_PI	Pearson Correlation	.652**	.653**	.864**	1	.843**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001
	N	146	146	146	146	146
B_WOM	Pearson Correlation	.654**	.701**	.871**	.843**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	
	N	146	146	146	146	146

**. Correlation is significant at the 0.01 level (2-tailed).

Table C7.3. A to B stability correlations for each outcome (links §4.5).

Correlations

		A_COM	A_ENG	A_TRU	A_PI	A_WOM	B_COM	B_ENG	B_TRU	B_PI	B_WOM
A_COM	Pearson Correlation	1	.726**	.561**	.645**	.630**	.662**	.624**	.503**	.564**	.556**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	N	151	151	151	151	151	146	146	146	146	146
A_ENG	Pearson Correlation	.726**	1	.578**	.651**	.605**	.626**	.796**	.522**	.568**	.615**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	N	151	151	151	151	151	146	146	146	146	146
A_TRU	Pearson Correlation	.561**	.578**	1	.775**	.773**	.423**	.456**	.760**	.739**	.743**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001	<.001	<.001	<.001	<.001	<.001
	N	151	151	151	151	151	146	146	146	146	146
A_PI	Pearson Correlation	.645**	.651**	.775**	1	.812**	.528**	.606**	.724**	.856**	.721**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001	<.001
	N	151	151	151	151	151	146	146	146	146	146
A_WOM	Pearson Correlation	.630**	.605**	.773**	.812**	1	.547**	.567**	.741**	.759**	.786**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001	<.001
	N	151	151	151	151	151	146	146	146	146	146
B_COM	Pearson Correlation	.662**	.626**	.423**	.528**	.547**	1	.776**	.621**	.652**	.654**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001
	N	146	146	146	146	146	146	146	146	146	146
B_ENG	Pearson Correlation	.624**	.796**	.456**	.606**	.567**	.776**	1	.609**	.653**	.701**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001	<.001
	N	146	146	146	146	146	146	146	146	146	146
B_TRU	Pearson Correlation	.503**	.522**	.760**	.724**	.741**	.621**	.609**	1	.864**	.871**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001	<.001
	N	146	146	146	146	146	146	146	146	146	146
B_PI	Pearson Correlation	.564**	.568**	.739**	.856**	.759**	.652**	.653**	.864**	1	.843**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		<.001
	N	146	146	146	146	146	146	146	146	146	146
B_WOM	Pearson Correlation	.556**	.615**	.743**	.721**	.786**	.654**	.701**	.871**	.843**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	
	N	146	146	146	146	146	146	146	146	146	146

**. Correlation is significant at the 0.01 level (2-tailed).

C8. H1 paired tests and effects

Table C8. Paired t tests on change scores with effect sizes (links §4.6).

Variables. D_COMFORT, D_TRUST, D_ENGAGE, D_PI, D_WOM, D_IMP

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	A_COM	3.582	146	.9002	.0745
	B_COM	3.510	146	.9586	.0793
Pair 2	A_ENG	3.562	146	.9095	.0753
	B_ENG	3.435	146	.9492	.0786
Pair 3	A_TRU	2.993	146	1.0537	.0872
	B_TRU	3.014	146	1.1203	.0927
Pair 4	A_PI	2.942	146	1.0643	.0881
	B_PI	2.928	146	1.1157	.0923
Pair 5	A_WOM	3.171	146	1.0495	.0869
	B_WOM	3.134	146	1.0864	.0899

Paired Samples Correlations

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	A_COM & B_COM	146	.662	<.001	<.001
Pair 2	A_ENG & B_ENG	146	.796	<.001	<.001
Pair 3	A_TRU & B_TRU	146	.760	<.001	<.001
Pair 4	A_PI & B_PI	146	.856	<.001	<.001
Pair 5	A_WOM & B_WOM	146	.786	<.001	<.001

Paired Samples Test

		Paired Differences					t	df	Significance	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				One-Sided p	Two-Sided p
					Lower	Upper				
Pair 1	A_COM - B_COM	.0719	.7656	.0634	-.0533	.1972	1.135	145	.129	.258
Pair 2	A_ENG - B_ENG	.1267	.5954	.0493	.0293	.2241	2.571	145	.006	.011
Pair 3	A_TRU - B_TRU	-.0205	.7563	.0626	-.1443	.1032	-.328	145	.372	.743
Pair 4	A_PI - B_PI	.0137	.5871	.0486	-.0823	.1097	.282	145	.389	.778
Pair 5	A_WOM - B_WOM	.0377	.7000	.0579	-.0768	.1522	.650	145	.258	.517

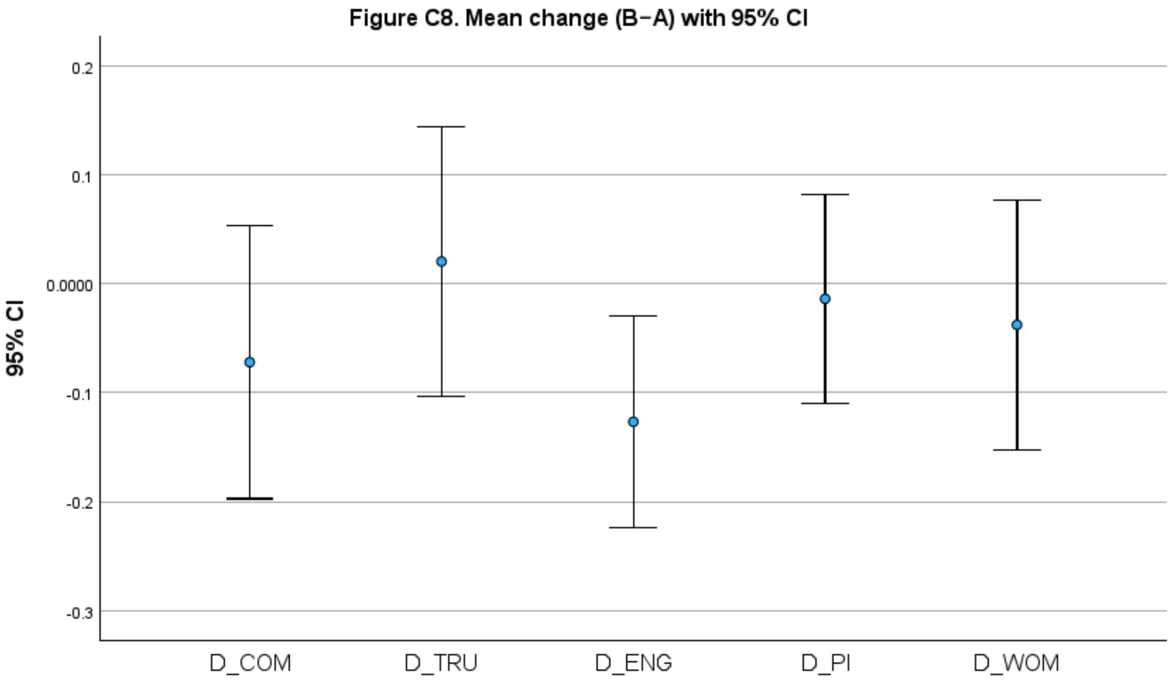
Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	A_COM - B_COM	Cohen's d	.7656	.094	-.069	.256
		Hedges' correction	.7696	.093	-.068	.255
Pair 2	A_ENG - B_ENG	Cohen's d	.5954	.213	.048	.376
		Hedges' correction	.5985	.212	.048	.375
Pair 3	A_TRU - B_TRU	Cohen's d	.7563	-.027	-.189	.135
		Hedges' correction	.7602	-.027	-.188	.134
Pair 4	A_PI - B_PI	Cohen's d	.5871	.023	-.139	.186
		Hedges' correction	.5901	.023	-.138	.185
Pair 5	A_WOM - B_WOM	Cohen's d	.7000	.054	-.109	.216
		Hedges' correction	.7036	.054	-.108	.215

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Note. Report Cohen d for paired designs: $d_z = Mean(D_X) / SD(D_X)$.

Figure C8. Mean change with 95% CI for each outcome (links §4.6).



C9. H2 main results table

Table C9. GLM summary for predictors of CHG_N (links §4.7).

Panel A. Omnibus tests (UNIANOVA "Tests of Between-Subjects Effects")

Predictor	df_effect	df_error	F	p	partial η²
SHS_TOTAL	1	139	0.000	.990	0.000
Age	1	139	0.348	.556	0.003
Gender (female, male, non-binary/other)	2	139	2.400	.095	0.033

Neurotype (neurodivergent vs neurotypical)	1	139	0.156	.926	0.001

Panel B. Model fit

Metric	Value
R^2	.084
Adjusted R^2	.024
Sample used in model (listwise)	n = 146

Panel C. Outcome distribution (CHG_N)

Changes made	n	%
0	44	29.1
1	59	39.1
2	32	21.2
3	16	10.6
Total	**151**	**100.0**

Notes.

1. DV = CHG_N ∈ {0,1,2,3}. Predictors: SHS_TOTAL (continuous), Age (continuous), Gender (3 levels), Neurotype (binary). Factors were coded with indicator dummies.

2. partial η² was computed from the omnibus F as $\eta_p^2 = \frac{F.df_{effect}}{F.df_{effecr}+df_{error}}$

3. Panel C reports the raw usage distribution described in §4.7; the GLM itself used n = 146 complete cases (df_error = 139).

4. No effect reached p < .05; Gender approached conventional thresholds (p = .095) with a small effect size.

C10. H3 baseline differences by sensitivity

Table C10.1. One-way ANOVA across SHS tertiles at baseline (links §4.8).

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
A_COM	Between Groups	4.709	2	2.354	2.985	.054
	Within Groups	112.805	143	.789		
	Total	117.514	145			
A_ENG	Between Groups	2.111	2	1.055	1.281	.281
	Within Groups	117.834	143	.824		
	Total	119.945	145			
A_TRU	Between Groups	2.387	2	1.193	1.076	.344
	Within Groups	158.606	143	1.109		
	Total	160.993	145			
A_PI	Between Groups	6.457	2	3.229	2.926	.057
	Within Groups	157.798	143	1.103		
	Total	164.255	145			
A_WOM	Between Groups	3.081	2	1.541	1.406	.248
	Within Groups	156.638	143	1.095		
	Total	159.719	145			

Tests of Between-Subjects Effects

Dependent Variable: A_COM

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	4.709[a]	2	2.354	2.985	.054	.040
Intercept	1870.830	1	1870.830	2371.607	<.001	.943
SHS_TERT	4.709	2	2.354	2.985	.054	.040
Error	112.805	143	.789			
Total	1991.000	146				
Corrected Total	117.514	145				

a. R Squared = .040 (Adjusted R Squared = .027)

Tests of Between-Subjects Effects

Dependent Variable: A_ENG

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2.111[a]	2	1.055	1.281	.281	.018
Intercept	1849.291	1	1849.291	2244.242	<.001	.940
SHS_TERT	2.111	2	1.055	1.281	.281	.018
Error	117.834	143	.824			
Total	1972.000	146				
Corrected Total	119.945	145				

a. R Squared = .018 (Adjusted R Squared = .004)

Tests of Between-Subjects Effects

Dependent Variable: A_TRU

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2.387[a]	2	1.193	1.076	.344	.015
Intercept	1302.272	1	1302.272	1174.133	<.001	.891
SHS_TERT	2.387	2	1.193	1.076	.344	.015
Error	158.606	143	1.109			
Total	1469.000	146				
Corrected Total	160.993	145				

a. R Squared = .015 (Adjusted R Squared = .001)

Tests of Between-Subjects Effects

Dependent Variable: A_PI

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	6.457[a]	2	3.229	2.926	.057	.039
Intercept	1263.242	1	1263.242	1144.779	<.001	.889
SHS_TERT	6.457	2	3.229	2.926	.057	.039
Error	157.798	143	1.103			
Total	1427.750	146				
Corrected Total	164.255	145				

a. R Squared = .039 (Adjusted R Squared = .026)

Tests of Between-Subjects Effects

Dependent Variable: A_WOM

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3.081[a]	2	1.541	1.406	.248	.019
Intercept	1466.603	1	1466.603	1338.909	<.001	.904
SHS_TERT	3.081	2	1.541	1.406	.248	.019
Error	156.638	143	1.095			
Total	1628.000	146				
Corrected Total	159.719	145				

a. R Squared = .019 (Adjusted R Squared = .006)

Table C10.2. Post hoc comparisons by outcome (links §4.8).

A_COM

	SHS tertiles: Low Mid High	N	Subset for alpha = 0.05 1
Tukey HSD[a,b]	Low SHS	51	3.353
	Mid SHS	45	3.622
	High SHS	50	3.780
	Sig.		.050

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 48.520.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

A_ENG

	SHS tertiles: Low Mid High	N	Subset for alpha = 0.05 1
Tukey HSD[a,b]	Low SHS	51	3.402
	Mid SHS	45	3.611
	High SHS	50	3.680
	Sig.		.290

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 48.520.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

A_TRU

	SHS tertiles: Low Mid High	N	Subset for alpha = 0.05 1
Tukey HSD[a,b]	Low SHS	51	2.892
	Mid SHS	45	2.911
	High SHS	50	3.170
	Sig.		.398

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 48.520.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

A_PI

	SHS tertiles: Low Mid High	N	Subset for alpha = 0.05 1
Tukey HSD[a,b]	Low SHS	51	2.667
	Mid SHS	45	3.011
	High SHS	50	3.160
	Sig.		.057

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 48.520.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

A_WOM

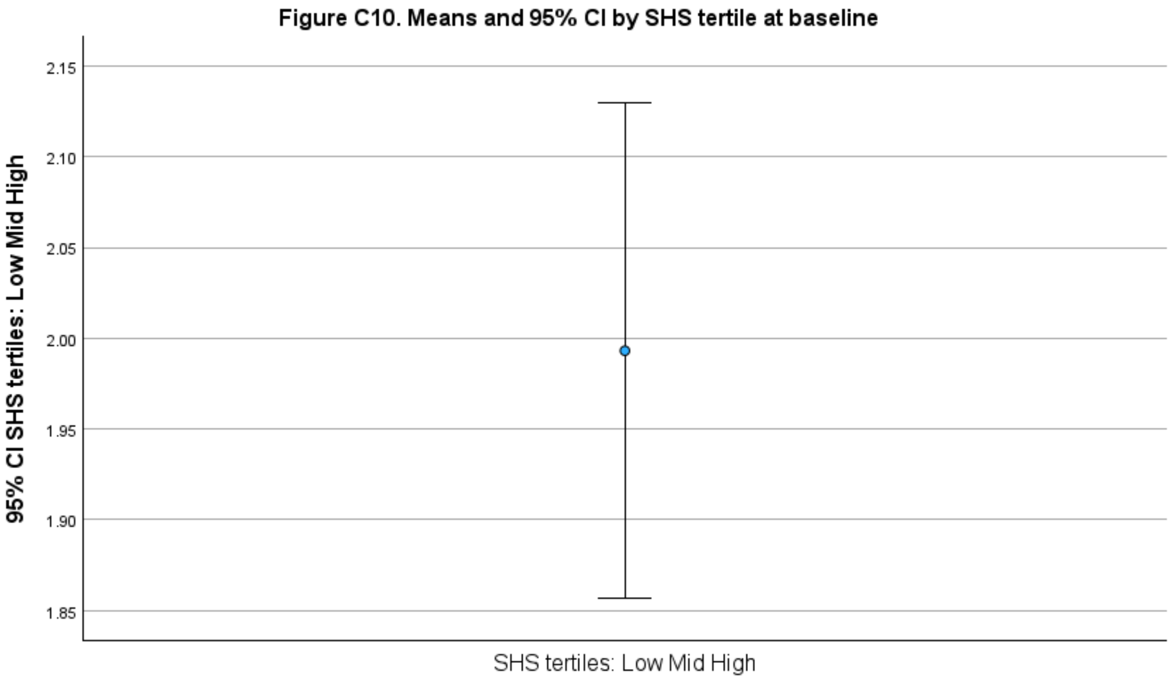
	SHS tertiles: Low Mid High	N	Subset for alpha = 0.05 1
Tukey HSD[a,b]	Low SHS	51	2.980
	Mid SHS	45	3.222
	High SHS	50	3.320
	Sig.		.250

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 48.520.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Figure C10. Means and 95% CI by SHS tertile at baseline (links §4.8).



C11. Non-parametric corroboration

Table C11.1. Wilcoxon signed-rank tests for D_X (links §4.10).

Ranks

		N	Mean Rank	Sum of Ranks
B_COM - A_COM	Negative Ranks	47[a]	44.05	2070.50
	Positive Ranks	39[b]	42.83	1670.50
	Ties	60[c]		
	Total	146		
B_ENG - A_ENG	Negative Ranks	43[d]	40.49	1741.00
	Positive Ranks	29[e]	30.59	887.00
	Ties	74[f]		
	Total	146		
B_TRU - A_TRU	Negative Ranks	38[g]	41.71	1585.00
	Positive Ranks	46[h]	43.15	1985.00
	Ties	62[i]		
	Total	146		
B_PI - A_PI	Negative Ranks	32[j]	34.66	1109.00
	Positive Ranks	35[k]	33.40	1169.00
	Ties	79[l]		
	Total	146		
B_WOM - A_WOM	Negative Ranks	37[m]	37.93	1403.50
	Positive Ranks	36[n]	36.04	1297.50
	Ties	73[o]		
	Total	146		

a. B_COM < A_COM

b. B_COM > A_COM

c. B_COM = A_COM

d. B_ENG < A_ENG

e. B_ENG > A_ENG

f. B_ENG = A_ENG

g. B_TRU < A_TRU

h. B_TRU > A_TRU

i. B_TRU = A_TRU

j. B_PI < A_PI

k. B_PI > A_PI

l. B_PI = A_PI

m. B_WOM < A_WOM

n. B_WOM > A_WOM

o. B_WOM = A_WOM

Test Statistics[a]

	B_COM - A_COM	B_ENG - A_ENG	B_TRU - A_TRU	B_PI - A_PI	B_WOM - A_WOM
Z	-.884[b]	-2.477[b]	-.913[c]	-.194[c]	-.298[b]
Asymp. Sig. (2-tailed)	.377	.013	.361	.846	.766

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.

Table C11.2. Kruskal-Wallis tests across SHS tertiles at baseline (links §4.10).

Ranks

	SHS tertiles: Low Mid High	N	Mean Rank
A_COM	Low SHS	51	63.78
	Mid SHS	45	74.20
	High SHS	50	82.78
	Total	146	
A_ENG	Low SHS	51	67.96
	Mid SHS	45	73.28
	High SHS	50	79.35
	Total	146	
A_TRU	Low SHS	51	70.08
	Mid SHS	45	70.28
	High SHS	50	79.89
	Total	146	
A_PI	Low SHS	51	62.28
	Mid SHS	45	75.28
	High SHS	50	83.34
	Total	146	
A_WOM	Low SHS	51	66.39
	Mid SHS	45	74.63
	High SHS	50	79.73
	Total	146	

Test Statistics[a,b]

	A_COM	A_ENG	A_TRU	A_PI	A_WOM
Kruskal-Wallis H	5.314	1.947	1.802	6.580	2.666
df	2	2	2	2	2
Asymp. Sig.	.070	.378	.406	.037	.264

a. Kruskal Wallis Test

b. Grouping Variable: SHS tertiles: Low Mid High

C12. Sensitivity analyses

Table C12.1. Fastest 5 percent removed: H1 paired tests re-run (links §4.10).

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	A_COM	3.582	146	.9002	.0745
	B_COM	3.510	146	.9586	.0793
Pair 2	A_ENG	3.562	146	.9095	.0753
	B_ENG	3.435	146	.9492	.0786
Pair 3	A_TRU	2.993	146	1.0537	.0872
	B_TRU	3.014	146	1.1203	.0927
Pair 4	A_PI	2.942	146	1.0643	.0881
	B_PI	2.928	146	1.1157	.0923
Pair 5	A_WOM	3.171	146	1.0495	.0869
	B_WOM	3.134	146	1.0864	.0899

Paired Samples Test

		Paired Differences							Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	One-Sided p	Two-Sided p
Pair 1	A_COM - B_COM	.0719	.7656	.0634	-.0533	.1972	1.135	145	.129	.258
Pair 2	A_ENG - B_ENG	.1267	.5954	.0493	.0293	.2241	2.571	145	.006	.011
Pair 3	A_TRU - B_TRU	-.0205	.7563	.0626	-.1443	.1032	-.328	145	.372	.743
Pair 4	A_PI - B_PI	.0137	.5871	.0486	-.0823	.1097	.282	145	.389	.778
Pair 5	A_WOM - B_WOM	.0377	.7000	.0579	-.0768	.1522	.650	145	.258	.517

Paired Samples Correlations

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	A_COM & B_COM	146	.662	<.001	<.001
Pair 2	A_ENG & B_ENG	146	.796	<.001	<.001
Pair 3	A_TRU & B_TRU	146	.760	<.001	<.001
Pair 4	A_PI & B_PI	146	.856	<.001	<.001
Pair 5	A_WOM & B_WOM	146	.786	<.001	<.001

All three paired-test blocks (Stats, Test, Correlations) for the filtered sample.

n (retained) = [value], fastest 5 percent excluded based on percentile of total duration.

Table C12.2. Exploratory ANCOVA for change by number of changes (links §4.10).

Model. D_X as DV, CHG_N as factor, baseline X_A as covariate.

Tests of Between-Subjects Effects

Dependent Variable: D_COM

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	12.300[a]	4	3.075	5.964	<.001	.145
Intercept	6.814	1	6.814	13.216	<.001	.086
A_COM	7.609	1	7.609	14.759	<.001	.095
CHG_N	2.091	3	.697	1.352	.260	.028
Error	72.695	141	.516			
Total	85.750	146				
Corrected Total	84.995	145				

a. R Squared = .145 (Adjusted R Squared = .120)

Tests of Between-Subjects Effects

Dependent Variable: D_ENG

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	5.718[a]	4	1.429	4.412	.002	.111
Intercept	1.643	1	1.643	5.071	.026	.035
A_ENG	3.050	1	3.050	9.412	.003	.063
CHG_N	2.262	3	.754	2.327	.077	.047
Error	45.688	141	.324			
Total	53.750	146				
Corrected Total	51.406	145				

a. R Squared = .111 (Adjusted R Squared = .086)

Tests of Between-Subjects Effects

Dependent Variable: D_TRU

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	6.126[a]	4	1.531	2.811	.028	.074
Intercept	5.456	1	5.456	10.016	.002	.066
A_TRU	6.055	1	6.055	11.114	.001	.073
CHG_N	.164	3	.055	.101	.960	.002
Error	76.813	141	.545			
Total	83.000	146				
Corrected Total	82.938	145				

a. R Squared = .074 (Adjusted R Squared = .048)

Tests of Between-Subjects Effects

Dependent Variable: D_PI

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2.769[a]	4	.692	2.068	.088	.055
Intercept	1.219	1	1.219	3.640	.058	.025
A_PI	1.580	1	1.580	4.720	.031	.032
CHG_N	1.038	3	.346	1.033	.380	.022
Error	47.203	141	.335			
Total	50.000	146				
Corrected Total	49.973	145				

a. R Squared = .055 (Adjusted R Squared = .029)

Tests of Between-Subjects Effects

Dependent Variable: D_WOM

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	6.028[a]	4	1.507	3.268	.013	.085
Intercept	4.194	1	4.194	9.096	.003	.061
A_WOM	5.240	1	5.240	11.363	<.001	.075
CHG_N	.465	3	.155	.336	.799	.007
Error	65.015	141	.461			
Total	71.250	146				
Corrected Total	71.043	145				

a. R Squared = .085 (Adjusted R Squared = .059)

C13. Data quality and participation checks

Table C13.1. Attention check failures and exclusions (links §4.2).

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
ADJ_TOTAL * REPORTED_TOTAL	270	100.0%	0	0.0%	270	100.0%

ADJ_TOTAL * REPORTED_TOTAL Crosstabulation

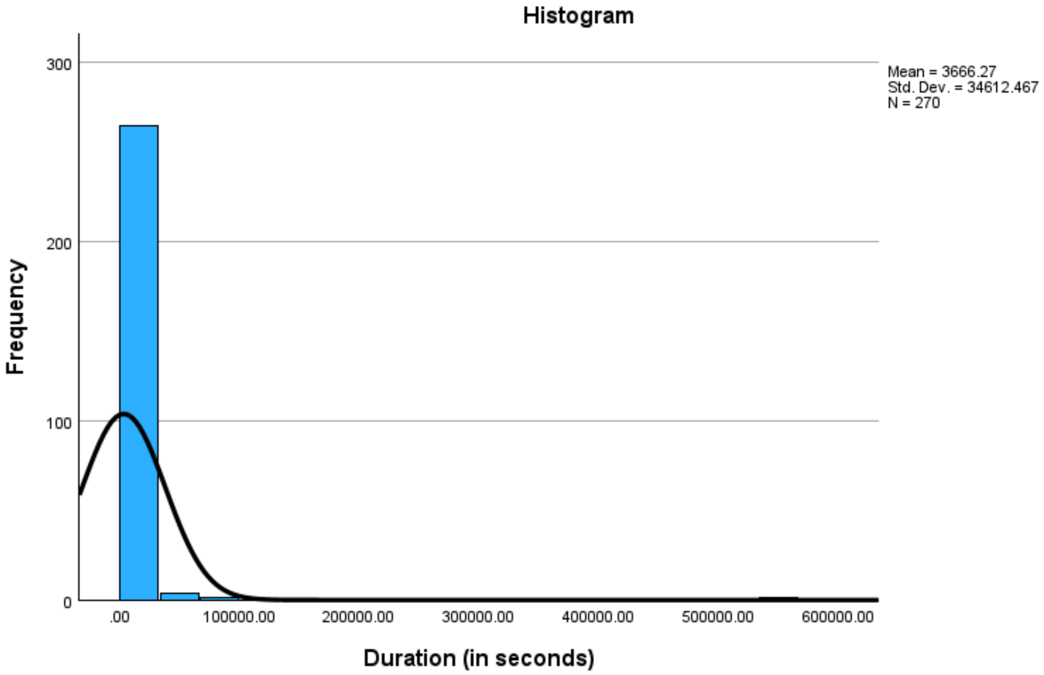
			REPORTED_TOTAL				
			.0	1.0	2.0	3.0	Total
ADJ_TOTAL	.0	Count	119	0	0	0	119
		% within ADJ_TOTAL	100.0%	0.0%	0.0%	0.0%	100.0%
	3.0	Count	0	4	8	4	16
		% within ADJ_TOTAL	0.0%	25.0%	50.0%	25.0%	100.0%
	4.0	Count	8	10	13	1	32
		% within ADJ_TOTAL	25.0%	31.3%	40.6%	3.1%	100.0%
	5.0	Count	9	36	12	2	59
		% within ADJ_TOTAL	15.3%	61.0%	20.3%	3.4%	100.0%
	6.0	Count	33	8	3	0	44
		% within ADJ_TOTAL	75.0%	18.2%	6.8%	0.0%	100.0%
Total		Count	169	58	36	7	270
		% within ADJ_TOTAL	62.6%	21.5%	13.3%	2.6%	100.0%

Table C13.2. Completion time distribution and trimming rule (links §3.2.3, §4.2).

Statistics

Duration (in seconds)

N	Valid	270
	Missing	0
Mean		3666.2704
Std. Deviation		34612.46735
Minimum		1.00
Maximum		557500.00
Percentiles	1	1.7100
	2	2.0000
	5	3.0000
	10	6.0000
	25	22.7500
	50	261.5000
	75	575.2500
	90	1130.7000
	95	3490.9000
	98	42959.2200
	99	55080.9300



Time trim: bottom/top 2% flagged (1=exclude)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retain	258	95.6	95.6	95.6
	Exclude (2% tails)	12	4.4	4.4	100.0
	Total	270	100.0	100.0	

Fastest 5% flagged (1=exclude)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retain	256	94.8	94.8	94.8
	Exclude (fastest 5%)	14	5.2	5.2	100.0
	Total	270	100.0	100.0	

Qualitative data:

Why did you choose the sensory adjustments you selected? - Other: - Text

	N	%
	131	94.2%
I just don't like random music coming from one of my tabs. Also music coming from a website really makes me think of poor skyblogs I made in 2010	1	0.7%
I like the audio at a low level and then the motion and brightness to a minimum. Sometimes the brightness is necessary to show the colour of the scenery	1	0.7%
I think sound should be an option	1	0.7%
It could be distracting at times	1	0.7%
Some animations weren't smooth	1	0.7%
sometimes I'm in an environment where I don't necessarily want the sound to play in the background	1	0.7%
This website looks to old school style.	1	0.7%
Too much going on at once, colours are not calming or make it easy to browse. Don't prefer sound on websites	1	0.7%

If you did not select any sensory adjustments, why not? - Selected Choice

	N	%
I liked the default version	11	7.9%
I didn't notice the option	11	7.9%
I wasn't sure how to adjust the settings	3	2.2%
I didn't feel the need to change anything	15	10.8%
Other:	3	2.2%
Missing System	96	69.1%

If you did not select any sensory adjustments, why not? - Other: - Text

	N	%
	136	97.8%
A dark mode would be nice, or colours as darker	1	0.7%
Didn't think those were the issues with the website	1	0.7%
i want changes like changing the bg color and making it more attractive and may be adding images of places in bg	1	0.7%

Any other thoughts or feedback? (Please avoid entering personal or identifying information here.)

	N	%
	120	86.3%
1) I was never diagnosed, but there is a possibility I might be slightly autistic. 2) the website seemed overall ok, although the thing that bothered me and definitely made it look less professional to me was the heavy usage of emojis when listing trip attractions	1	0.7%
Can user given more options to choose for the background music while exploring the website ?	1	0.7%
Darker mode or colours on the website would be nice. Font easy to read for dyslexics. Perhaps a little more variation in layout? Different indents, etc. Symbols were helpful.	1	0.7%
I feel like the logo and the fact that the videos are played using an embedded YouTube player reduce the quality and my trust in the website. However the novelty factor is what guides me towards wanting to complete a purchase.	1	0.7%
I felt uncomfortable answering this survey	1	0.7%
I liked the website. 2 things: 1. I am not realy a person who books all-inclusive holiday, I like to pick all things seperately, so I probably would not buy on the website. 2. To appear more trusted, the website could link in a final for to external information like Health offices for vaccinations or tourist shops for local activities.	1	0.7%
Maybe potentially some more visuals of the places in the form of pictures. Other than that nothing else I can really think of	1	0.7%
N/A	1	0.7%
NA	1	0.7%
no	1	0.7%
No	1	0.7%

The discrepancy between the first and second pages was too big for me. I am sensitive to light, but I strongly dislike dark view. I felt that the benefits of the page suffer greatly from the dark almost murky view. I felt that I couldn't really see what was presented to me. So I lost interest. I did like the ambient music and visuals! It has a meditative, ASMR type of feel, which makes vacation planning more relaxed.	1	0.7%
The website could have been improved with some fake reviews as testimonials. Without testimonials it made it feel less real. I liked the colour scheme, however the top banner with the visuals was not always very captivating - the alignment of them seemed off at times e.g. only seeing the top of a building. As a sensory experience I really enjoyed this website, it wasn't too involved or busy and was very clear about what to expect. I liked that no automatic chat box popped up (these make me stressed) - however an option for "contact us" would have also made me feel more comfortable/safe/secure.	1	0.7%
The website feels unreliable because it looks like a scam site.	1	0.7%
The website interface doesnt look as attractive when I look at it the second time	1	0.7%
the website looked like it was made with ai which lowered the trustworthiness of the site	1	0.7%
Volume adjustment is needed tho Slight in-depth explanation of the days of the tours are needed	1	0.7%
What got me the most unsettled about the website is how ChatGpt looking the information is, it makes it look like a scam	1	0.7%
while I am not really oversensitized, a lot of my family is	1	0.7%