



**UNIVERSITÀ
DI PAVIA**

Department of Economics and Management

**Master Programme in International Business and
Entrepreneurship**

**Personalised Digital Nudging and Consumer Choice in
E-Commerce: Evidence from an Online Grocery
Shopping Experiment**

Supervisor:

Prof. Bartosiak Marcin Lukasz

Discussant:

Prof.

**Student: Iraklii Kheladze
ID: 581710**

Academic Year 2025-2026



**UNIVERSITÀ
DI PAVIA**

Dipartimento di Scienze Economiche e Aziendali

**Corso di Laurea magistrale in International Business
and Entrepreneurship**

**Nudging digitale personalizzato e scelta del consumatore
nell'e-commerce: evidenze da un esperimento di spesa
alimentare online**

Relatore:

Chiar.mo Prof. Bartosiak Marcin Lukasz

**Tesi di Laurea di Iraklii Kheladze
Matr. n. 581710**

Anno Accademico 2025-2026

Abstract

This thesis examines the effectiveness of personalised digital nudges in e-commerce, focusing on whether interface-based choice architecture can influence consumer decision-making in a simulated online grocery environment. Building on research on digital nudging, bounded rationality, choice architecture, and behavioural personalisation, the study compares three experimental conditions: a neutral control condition, a generic nudge condition, and a decision-style-based matched nudge condition. The empirical study was implemented as a quantitative online experiment in Qualtrics. Participants completed a reduced GDMS-based decision-style questionnaire, were routed into one of the three experimental conditions, and subsequently made seven binary product choices in a grocery-like shopping task. The primary behavioural outcome was target-option choice, while perceived autonomy was measured as a secondary post-task outcome.

The findings show that generic digital nudging significantly increased target-option choice compared with neutral presentation, indicating that light-touch interface cues can influence purchase-like decisions in e-commerce settings. Contrary to the initial expectation, default preselection significantly outperformed social proof within the generic nudge condition. The matched experimental condition showed a positive but statistically non-significant overall advantage over generic nudging, meaning that the hypothesised superiority of personalised nudging was not confirmed at the aggregate level. However, supplementary analyses suggested that matching may be more effective for specific user groups, particularly respondents with spontaneous decision-making tendencies. Finally, the nudged task versions did not produce a statistically significant reduction in perceived autonomy.

Overall, the thesis concludes that digital nudges can meaningfully affect consumer choice, but their effectiveness is mechanism-dependent and context-sensitive. Personalised digital

nudging should therefore be understood not as a universally superior strategy, but as a conditional design approach requiring empirical validation, careful segmentation, and ethically responsible implementation.

Keywords: digital nudging; personalised nudging; e-commerce; choice architecture; default preselection; social proof; consumer decision-making; perceived autonomy.

Abstract (Italian translation)

Questa tesi esamina l'efficacia dei digital nudge personalizzati nell'e-commerce, concentrandosi sulla capacità dell'architettura delle scelte basata sull'interfaccia di influenzare il processo decisionale dei consumatori in un ambiente simulato di spesa alimentare online. Sulla base della letteratura sul digital nudging, sulla razionalità limitata, sull'architettura delle scelte e sulla personalizzazione comportamentale, lo studio confronta tre ambienti sperimentali: una condizione di controllo neutrale, una condizione con nudge generico e una condizione con nudge personalizzato in base allo stile decisionale. Lo studio empirico è stato realizzato come esperimento quantitativo online tramite Qualtrics. I partecipanti hanno completato un questionario ridotto basato sulla GDMS per rilevare lo stile decisionale, sono stati assegnati a uno dei gruppi sperimentali e hanno successivamente effettuato sette scelte binarie di prodotto in un compito di acquisto simulato. L'esito comportamentale principale era la scelta dell'opzione target, mentre l'autonomia percepita è stata misurata come esito secondario post-compito.

I risultati mostrano che il digital nudging generico ha aumentato significativamente la scelta dell'opzione target rispetto alla presentazione neutrale, indicando che lievi segnali di interfaccia possono influenzare decisioni simili all'acquisto in contesti di e-commerce.

Contrariamente all'aspettativa iniziale, la preselezione di default ha superato significativamente il social proof nella condizione con nudge generico. La condizione personalizzata ha mostrato un vantaggio complessivo positivo ma non statisticamente significativo rispetto al nudging generico; pertanto, la superiorità ipotizzata del nudging personalizzato non è stata confermata a livello aggregato. Tuttavia, le analisi supplementari hanno suggerito che l'abbinamento tra stile decisionale e tipo di nudge potrebbe essere più efficace per specifici gruppi di utenti, in particolare per i rispondenti con tendenze decisionali

spontanee. Infine, le versioni del compito con nudge non hanno prodotto una riduzione statisticamente significativa dell'autonomia percepita.

Nel complesso, la tesi conclude che i digital nudge possono influenzare in modo significativo le scelte dei consumatori, ma che la loro efficacia dipende dal meccanismo utilizzato e dal contesto di applicazione. Il digital nudging personalizzato dovrebbe quindi essere inteso non come una strategia universalmente superiore, bensì come un approccio progettuale condizionale che richiede validazione empirica, segmentazione accurata e implementazione eticamente responsabile.

Parole chiave: digital nudging; nudging personalizzato; e-commerce; architettura delle scelte; preselezione di default; social proof; processo decisionale del consumatore; autonomia percepita.

Acknowledgements

I would like to express my sincere gratitude to Prof. Marcin Łukasz Bartosiak for his guidance, constructive feedback, and academic support throughout the development of this thesis. His insights were highly valuable in refining the direction, structure, and scholarly quality of this research.

I am also very grateful to all participants who took the time to complete the survey and thereby contributed to the empirical foundation of this study.

Last but not least, I would like to thank my family and friends for their encouragement, patience, and support during the writing process.

Table of Contents

ABSTRACT

ABSTRACT (ITALIAN TRANSLATION)

ACKNOWLEDGEMENTS..... 2

LIST OF TABLES 6

LIST OF FIGURES 7

LIST OF ABBREVIATIONS..... 8

CHAPTER 1: INTRODUCTION 9

1.1 CONTEXT OF THE STUDY 9

1.2 RESEARCH PROBLEMS AND OBJECTIVES 11

1.3 SIGNIFICANCE AND IMPORTANCE OF THE STUDY 13

1.4 SCOPE OF THE STUDY AND DELIMITATIONS 14

1.5 STRUCTURE OF THE THESIS 15

1.6 CHAPTER SUMMARY 16

CHAPTER 2: LITERATURE REVIEW 18

2.0 INTRODUCTORY NOTES.....18

2.1 FOUNDATIONS: NUDGING, CHOICE ARCHITECTURE, AND BEHAVIOURAL MECHANISMS.....19

2.2 DIGITAL NUDGING 24

2.3 DIGITAL NUDGING IN E-COMMERCE: CUSTOMER JOURNEY, FOCAL MECHANISMS, AND SHOPPING-RELEVANT EVIDENCE.....26

2.4 PERSONALISATION IN DIGITAL ENVIRONMENTS: BENEFITS, PRIVACY, TRUST, AND CONTROL 34

2.5 PERSONALISED DIGITAL NUDGING: PERSON-NUDGE FIT, TRAIT TARGETING, AND DECISION-STYLE MATCHING.....	37
2.6 PERCEIVED AUTONOMY, REACTANCE, TRANSPARENCY, AND ETHICAL LEGITIMACY	39
2.7 SYNTHESIS: RESEARCH GAPS AND HYPOTHESIS DEVELOPMENT	43
2.8 CONCLUSION	46
CHAPTER 3: METHODOLOGY	49
3.1 INTRODUCTION	49
3.2 RESEARCH DESIGN.....	50
3.3 RESEARCH QUESTIONS AND HYPOTHESES	55
3.4 DATA COLLECTION METHODS	57
3.5 OPERATIONALISATION OF CONSTRUCTS	60
3.6 DATA ANALYSIS.....	63
3.7 ETHICAL CONSIDERATIONS.....	73
3.8 RELIABILITY AND VALIDITY.....	74
3.9 CHAPTER SUMMARY	81
CHAPTER 4: RESULTS.....	82
4.1 INTRODUCTION	82
4.2 ANALYTIC SAMPLE AND PRELIMINARY CHECKS	82
4.3 DESCRIPTIVE BEHAVIOURAL RESULTS	84
4.4 MAIN BEHAVIOURAL HYPOTHESIS TESTS	85
4.5 SUPPLEMENTARY AND ROBUSTNESS ANALYSES	89
4.6 SECONDARY RESULTS: PERCEIVED AUTONOMY	91
4.7 CHAPTER SUMMARY.....	93
CHAPTER 5: DISCUSSION.....	95

5.1 INTRODUCTION	95
5.2 INTERPRETATION OF THE MAIN FINDINGS	96
5.3 THEORETICAL IMPLICATIONS	102
5.4 MANAGERIAL IMPLICATIONS.....	103
5.5 LIMITATIONS	104
5.6 FUTURE RESEARCH.....	106
5.7 CONCLUSION	107
CHAPTER 6: CONCLUSION	109
6.1 INTRODUCTION	109
6.2 SUMMARY OF THE MAIN FINDINGS.....	109
6.3 CONCLUDING ANSWERS TO THE RESEARCH QUESTIONS	110
6.4 SYNTHESIS OF THE THEORETICAL AND MANAGERIAL IMPLICATIONS	112
6.5 FINAL REFLECTIONS AND FUTURE OUTLOOK	113
BIBLIOGRAPHY	115
APPENDICES	124

List of Tables

TABLE 1: CHOICE-ARCHITECTURE TAXONOMY AND MAPPING OF THE FOCAL NUDGES	22
TABLE 2: LEGITIMATE NUDGES AND THEIR DARK-PATTERN LOOKALIKES IN E-COMMERCE INTERFACES.....	42
TABLE 3: EXPERIMENTAL CONDITION ASSIGNMENT LOGIC.....	53
TABLE 4: METHODOLOGICAL PURPOSE OF EACH SURVEY STAGE.....	58
TABLE 5: OPERATIONALISATION OF THE MAIN CONSTRUCTS AND VARIABLES	62
TABLE 6: THE FINAL ANALYTIC SAMPLE	83
TABLE 7: DESCRIPTIVE RESULTS	84
TABLE 8: MAIN HYPOTHESIS TESTING MODELS	86

List of Figures

FIGURE 1: PREDICTED PROBABILITY OF CHOOSING THE TARGET OPTION	87
FIGURE 2: PERCEIVED AUTONOMY BY TASK VERSION	92

List of Abbreviations

AIC - Akaike Information Criterion

ANOVA - Analysis of Variance

BIC - Bayesian Information Criterion

CI - Confidence Interval

GDMS - The General Decision-Making Scale

GLM - Generalised Linear Model

IP - Internet Protocol

OECD - The Organisation for Economic Co-operation and Development

OLS - Ordinary Least Squares

RQ - Research Question

SE - Standard Error

Chapter 1: Introduction

1.1 Context of the study

Digital commerce has become one of the most important environments in which contemporary consumer decisions are made, as an increasing share of search, comparison, and purchasing activity is now mediated through online platforms and interface-based decision environments (Dellaert and Häubl, 2012; Li, Grahl and Hinz, 2022). In such environments, consumers rarely encounter products in a neutral manner, since digital interfaces determine which options are displayed, how they are ordered, what information is made salient, and which alternatives appear easier or more appropriate to select (Häubl and Trifts, 2000; Weinmann, Schneider and vom Brocke, 2016; Jesse, Jannach and Gula, 2021). Accordingly, e-commerce platforms do not merely present products, but function as choice architectures that shape how consumers search, evaluate, and ultimately choose among alternatives (Thaler and Sunstein, 2008; Dellaert and Häubl, 2012; Weinmann, Schneider and vom Brocke, 2016). These developments have stimulated growing scholarly and practical interest in digital nudging, understood as the use of user-interface elements to steer behaviour in predictable ways without removing options or materially altering economic incentives (Thaler and Sunstein, 2008; Weinmann, Schneider and vom Brocke, 2016).

The relevance of digital nudging is grounded in the fact that many online consumer decisions are made under conditions of bounded rationality, limited attention, and imperfect information processing, which increases reliance on heuristics and context cues (Simon, 1955; Tversky and Kahneman, 1974; Kahneman, 2011). In such contexts, seemingly minor interface features may exert disproportionate behavioural effects by influencing attention, reducing cognitive effort, or altering the inferential meaning of available options (Mandel and Johnson, 2002; Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019).

Among the different forms of digital nudging, default preselection and social proof are especially relevant to e-commerce as both can be implemented easily in online interfaces and both have a plausible theoretical basis in consumer decision-making (Johnson and Goldstein, 2003; Chevalier and Mayzlin, 2006; Jachimowicz et al., 2019). Default nudges influence behaviour by making one option the path of least resistance and thereby leveraging effort reduction, inertia, or perceived endorsement, even when alternative options remain available (Johnson and Goldstein, 2003; Jachimowicz et al., 2019; Ortmann et al., 2023). Social-proof cues, by contrast, operate primarily through informational inference, as indicators of what other consumers have chosen or evaluated and can reduce uncertainty and shape perceived product attractiveness or appropriateness (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010; Qiu and Zhang, 2024).

At the same time, the effectiveness of digital nudges is heterogeneous and context-dependent, meaning that not every nudge works equally efficiently across all digital decision environments (Hummel and Maedche, 2019; Mertens et al., 2022; Valenčič et al., 2023). This heterogeneity has directed attention towards personalisation, that is, the tailoring of digital content, communication, or interface logic to characteristics of the individual user rather than applying the same intervention to all users (Awad and Krishnan, 2006; Bleier and Eisenbeiss, 2015; Mills, 2022). Transferred to the digital nudging context, this implies the possibility of personalised digital nudging, in which the type of nudge shown is matched to individual user characteristics such as traits, preferences, or decision-making styles (Peer et al., 2020; Mills, 2022). In principle, such a personalised approach may improve behavioural effectiveness as different users may respond more strongly to different influence mechanisms, rather than to a single generic design applied uniformly across the entire user base (Hirsh, Kang and Bodenhausen, 2012; Peer et al., 2020; Mills, 2022).

However, this promise remains empirically uncertain, particularly in e-commerce settings, where evidence on whether personalised nudges actually outperform generic nudges is still comparatively limited and mixed (e.g., Ingendahl et al., 2020; Peer et al., 2020; de Vries, Bol and van der Laan, 2025). Moreover, the increasing use of digital nudging and personalisation raises broader ethical questions concerning transparency, legitimacy, and perceived autonomy, since behaviourally effective interface design may also be interpreted as intrusive or manipulative if users perceive hidden steering intents (Friestad and Wright, 1994; Kim, Barasz and John, 2019; Michaelsen, 2024b).

Therefore, the present thesis is situated at this intersection of digital nudging, e-commerce choice architecture, and behavioural personalisation, and examines whether digital nudges influence consumer choice, whether one nudge mechanism is more effective than another, and whether a decision-style-based matching rule (Scott and Bruce, 1995) improves behavioural outcomes in a controlled online shopping environment.

1.2 Research problems and objectives

Despite the growing literature on digital nudging, several important gaps remain:

- First, although prior research shows that nudges are effective on average, the evidence is heterogeneous and often difficult to compare across studies because different interventions, contexts, and outcomes are aggregated under broad labels (Hummel and Maedche, 2019; Mertens et al., 2022).
- Second, relatively few studies compare different nudge types head-to-head within the same controlled e-commerce environment, even though such comparisons are necessary to determine whether some mechanisms are more behaviourally potent than others in specific shopping contexts (Hummel and Maedche, 2019; Mertens et al., 2022; Valenčič et al., 2023).

- Third, while personalisation is frequently presented as a promising next step in digital influence design, empirical evidence on whether personalised digital nudging actually outperforms generic nudging remains limited and mixed (e.g., Ingendahl et al., 2020; Peer et al., 2020).
- Fourth, a further problem concerns the normative dimension of digital nudging. Consumer-facing interface interventions should not be judged solely by their behavioural effectiveness, but they should also be evaluated in terms of whether they preserve autonomy, remain transparent, and avoid manipulative or dark-pattern-like design (The Organisation for Economic Co-operation and Development (OECD), 2022; Bruns and Perino, 2023; Michaelsen, 2024b).

Against this background, the overarching research objective of this thesis is to examine the effectiveness of personalised digital nudges in a simulated e-commerce environment, as well as the consequences for legitimacy related to the implementation of such nudges. More specifically, the thesis pursues the following three objectives:

- First, it seeks to assess whether generic digital nudging is behaviourally effective and which type of nudge is more potent: default preselection or social proof.
- Second, it aims to examine whether a nudge personalisation approach can yield stronger behavioural effects than non-personalised generic nudging.
- Third, it seeks to examine the legitimacy consequences of nudge implementation.

Taken together, these research objectives define the analytical focus of the thesis and provide the basis for the theoretical development and empirical investigation presented in the following chapters.

1.3 Significance and importance of the study

The significance of the present study is threefold: theoretical, managerial, and societal. From a theoretical perspective, the thesis contributes to the growing literature on digital nudging by integrating insights from behavioural economics, information systems, marketing, and consumer psychology. More specifically, it extends the discussion from the general effectiveness of digital nudges towards a more differentiated examination of mechanism-specific effects and the potential added value of behavioural matching. In doing so, the study responds to the need for more controlled, context-sensitive, and comparable experimental evidence on digital choice architecture in e-commerce settings (c.f., Mertens et al., 2022; Valenčič et al., 2023).

From a managerial perspective, the topic is highly relevant for digital retailers, platform designers, and managers responsible for online customer experience and conversion optimisation. In this regard, e-commerce firms increasingly rely on interface design to influence how consumers search, compare, and select products, meaning that even relatively small design interventions may have commercially meaningful effects on online behaviour (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Weinmann, Schneider and vom Brocke, 2016). Hence, understanding whether default preselection or social proof is more effective, and whether personalisation improves behavioural outcomes has direct practical relevance for the design of digital retail environments.

From a societal and ethical perspective, the study is important since the growing use of digital nudging raises broader questions of legitimacy, transparency, and user autonomy (Friestad and Wright, 1994; Ryan and Deci, 2000; Bruns and Perino, 2023; Michaelsen, 2024b). Thus, by considering perceived autonomy alongside behavioural effectiveness, the thesis addresses not only whether digital nudges work, but also whether such interventions can remain compatible with a choice-preserving and ethically appropriate digital environment.

Taken together, the study is significant as it addresses a timely problem at the intersection of digital commerce, behavioural influence, and personalisation. It therefore contributes to a more precise understanding of how digital nudges function in e-commerce, when they may be useful in practice, and why their design should be considered not only in behavioural but also in normative terms.

1.4 Scope of the study and delimitations

The present thesis adopts a deliberately bounded scope. Empirically, it focuses on a simulated grocery-like e-commerce environment in which respondents make repeated binary product choices among close substitutes. This setting was chosen since grocery decisions are frequent, comparatively low in involvement, and well suited to controlled manipulation of interface cues in online shopping contexts (c.f., Valenčič et al., 2023). The study hence concentrates on short-run purchase-like decisions in a simplified digital retail context rather than on the full complexity of real-world e-commerce behaviour.

Conceptually, the thesis focuses on two specific digital nudges: default preselection and social proof. Other potentially relevant forms of digital nudging will be discussed in the literature review only insofar as they provide conceptual background. Likewise, the personalisation studied here is intentionally narrow. It does not involve adaptive machine-learning personalisation or dynamic platform optimisation. Instead, it examines a transparent and practically implementable rule-based matching approach based on decision-making style (Scott and Bruce, 1995).

In terms of user heterogeneity, the thesis concentrates on the distinction between dependent and spontaneous decision-making tendencies, operationalised through a reduced GDMS-based instrument (The General Decision-Making Scale; Scott and Bruce, 1995). The study does not attempt to model all possible consumer traits or psychological moderators, even

though such constructs may also be relevant in personalised digital nudging research.

Accordingly, the conclusions will be limited to the specific matching logic implemented in this research design.

The study is also delimited in relation to outcomes. Its primary outcome is behavioural, namely target-option choice, reflecting the thesis's central interest in whether interface-based interventions influence product selection. Perceived autonomy will be included as a secondary perceptual outcome in order to provide an interpretive ethical layer; however, the thesis does not seek to offer a comprehensive welfare analysis or a full account of all normative implications of digital nudging.

This delimitation means that the thesis should be understood as an analytically focused investigation of digital nudging in a controlled e-commerce context. It does not seek to capture the full complexity of real-world online retail environments, nor does it claim to cover all possible forms of digital nudging or personalisation. Rather, the study concentrates on a specific experimental setting, a limited set of focal nudge mechanisms, and a clearly defined form of nudge matching in order to generate theoretically grounded evidence on their behavioural relevance in online consumer choice. Accordingly, the conclusions of the thesis should be interpreted in relation to this defined scope and as a contribution to a more precise understanding of how personalised digital nudges may function in e-commerce.

1.5 Structure of the thesis

Following this introductory Chapter 1, Chapter 2 develops the literature review and conceptual framework. It discusses the foundations of nudging and choice architecture, examines digital nudging in e-commerce, reviews the literature on personalisation, and synthesises the research gaps that motivate the hypotheses.

Chapter 3 presents the methodology. It explains the research design, sampling approach, experimental policy conditions, operationalisation of variables, and data-analysis strategy. Particular attention is given to the implementation of the simulated grocery-shopping task, the nudge matching logic, and the analytical procedures used to test the hypotheses.

Chapter 4 reports the empirical results. It begins with the analytic sample and descriptive statistics, then presents the main behavioural hypothesis tests, supplementary and robustness analyses, and finally the secondary analysis of perceived autonomy.

Chapter 5 discusses the findings in relation to the research questions and the theoretical framework. It interprets the results, outlines the theoretical and managerial implications, reflects on the limitations of the study, and proposes directions for future research.

Chapter 6 concludes the thesis by synthesising the main findings, answering the research questions explicitly, highlighting the broader implications of the study, and offering a final outlook on personalised digital nudging in e-commerce.

The thesis is complemented by appendices, which provide additional material on the experimental instrument, task versions, diagnostic checks, and supplementary statistical outputs.

1.6 Chapter summary

This chapter has introduced the thesis by situating it within the growing importance of digitally mediated consumer choice and the increasing use of behavioural interface design in e-commerce. It has shown that digital nudging constitutes a relevant and timely field of inquiry. The chapter has also clarified the core problem addressed by the thesis, namely whether personalised digital nudging offers added value over generic nudging and neutral presentation in a controlled online shopping context. Furthermore, the chapter has specified the research objectives and explained why the topic is theoretically, managerially, and

ethically important. It has also delineated the study's scope and limitations and outlined the structure of the thesis as a whole. On this basis, the following chapter develops the literature review and conceptual foundation necessary for the subsequent empirical investigation.

Chapter 2: Literature Review

2.0 Introductory notes

2.0.1 Chapter overview and structure

This chapter develops the conceptual and empirical basis for analysing personalised digital nudges in e-commerce. It proceeds from bounded rationality and choice-architecture theory to the evidence on digital nudging, with a focused synthesis of two interventions central to the thesis -defaults and social proof - before turning to personalisation logic and concluding with gap identification and hypothesis development (Johnson and Goldstein, 2003; Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019; Michaelсен, 2024b).

2.0.2 Purpose, scope, and definitional commitments

Digital commerce is fundamentally mediated by interface design. In online retail, the sequence of screens, the ordering of options, and the presence of cues such as preselection or popularity indicators shape what users notice, consider, and ultimately choose (Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019).

In this thesis, ‘nudge’ is treated as a specific class of interventions that steer behaviour in a predictable direction while maintaining freedom of choice and without materially altering economic incentives (Thaler and Sunstein, 2008). This definition follows the libertarian paternalist claim that policy and managerial actors can design environments that help people make better choices as judged by their own standards, yet do so in a manner that preserves option sets (Thaler and Sunstein, 2003).

The review adopts a deliberately bounded domain. The behavioural setting of interest is a short-run purchase-like choice in a grocery-like e-commerce environment, where consumers

make rapid selections among close substitutes under attentional and time constraints (Weinmann, Schneider and vom Brocke, 2016; Michaelsen, 2024a).

A further definitional commitment concerns outcomes. Digital nudges are typically evaluated using behavioural proxies such as click-through, purchase, or uptake (Michaelsen, 2024b; Weijers, Wachner and de Koning, 2024).

2.1 Foundations: nudging, choice architecture, and behavioural mechanisms

2.1.1 From rational choice to bounded rationality in consumer choice

Classical models of consumer choice often assume that decision-makers can process relevant information and select the utility-maximising option given stable preferences. In practice, however, the computational burden of such optimisation is prohibitive in most everyday contexts (Simon, 1955).

Kahneman's dual-process framework provides a complementary micro-foundation for why interface design predictably matters. Fast, automatic processing (System 1) relies on associative cues and heuristics, whereas slow, deliberative processing (System 2) is effortful and often invoked only when stakes or motivation justify it. In many online shopping situations, low deliberation and time pressure mean that System 1 processes are dominant, making users more responsive to salient cues, defaults, and simple social information (Kahneman, 2011).

Heuristics and biases research sharpens the claim that predictable deviations from optimisation can arise from cue-driven judgement. Tversky and Kahneman (1974) show that people rely on heuristics such as representativeness, availability, and anchoring, which can be adaptive under constraints but also generate systematic distortions. In digital environments, interface cues can function as anchors, availability triggers, or signals of typicality, thereby

changing choice shares even when objective option attributes are held constant (Tversky and Kahneman, 1974; Weinmann, Schneider and vom Brocke, 2016).

Moreover, visual and contextual primes embedded in web pages can make particular attributes focal, shifting what consumers treat as relevant and altering their preference ordering. Background imagery and visual context can influence product preferences through mechanisms that differ for novices and experts, illustrating how interface elements impact search and evaluation (Mandel and Johnson, 2002). Interface constraints can also complicate, rather than simplify, decision-making. In this regard, consumers may sometimes overweight small disadvantages of otherwise superior options, increasing conflict and even reversing preferences (Schrift, Netzer and Kivetz, 2011).

2.1.2 Defining “nudge” and libertarian paternalism; core debates

The canonical definition of a nudge emphasises that the intervention changes behaviour by altering the choice context, not by changing option sets or imposing substantial incentives (Thaler and Sunstein, 2008). Moreover, such interventions are pervasive and unavoidable because any presentation of options implicitly structures choice. The normative claim is that deliberate design can be justified when it helps people achieve their own aims while preserving choice and enabling easy opt-out (Thaler and Sunstein, 2003).

Nevertheless, the ethical acceptability of nudging is contested, particularly in consumer settings where commercial actors may have interests that diverge from consumer welfare. Critics argue that nudges can be manipulative, can exploit cognitive vulnerabilities, and can undermine autonomy even when options remain formally available (c.f., Bruns and Perino, 2023; Michaelsen, 2024b). These critiques indicate that welfare cannot be inferred from behaviour alone; it requires measurement of how people experience the influence attempt and

whether they perceive it as legitimate or intrusive (c.f., Bruns and Perino, 2023; Michaelsen, 2024b).

From a consumer psychology perspective, influence attempts activate interpretive processes (Friestad and Wright, 1994). The persuasion knowledge model proposes that consumers learn to recognise persuasion tactics, infer intent, and deploy coping strategies such as resistance, counter-arguing, or avoidance (Friestad and Wright, 1994). When nudges are interpreted as covert persuasion, they may trigger perceptions of manipulation and reduce perceived legitimacy, potentially offsetting behavioural gains in the longer term (Friestad and Wright, 1994; Michaelsen, 2024a).

To position this thesis' scope, it is useful to distinguish ethical nudges from interventions that rely on deception, coercion, or the obstruction of choice. The present review treats nudges as design features that preserve choice sets and avoid false or misleading information (Weinmann, Schneider and vom Brocke, 2016).

2.1.3 Choice-architecture taxonomies: organising mechanisms and interventions

A persistent challenge in the nudging literature is comparability. Without shared categories, empirical findings remain difficult to cumulate, and meta-analytic synthesis becomes vulnerable to mixing interventions that operate through distinct mechanisms. Taxonomies address this by mapping interventions onto mechanism-relevant buckets, thereby supporting theory building and the design of more interpretable experiments (Münscher, Vetter and Scheuerle, 2016).

In this context, a widely used taxonomy groups choice architecture techniques into decision information, decision structure, and decision assistance (Münscher, Vetter and Scheuerle, 2016). Decision information techniques change what information is presented or how it is framed; decision structure techniques change defaults or the set-up of options; decision

assistance techniques support processing through tools or prompts. This taxonomy is valuable for digital contexts because online interfaces can implement all three categories at low marginal cost (Münscher, Vetter and Scheuerle, 2016; Weinmann, Schneider and vom Brocke, 2016).

For the present thesis, the taxonomy provides a practical mapping. For instance, default nudges are primarily decision structure interventions because they alter the preselected option and thereby affect effort and inertia (c.f., Johnson and Goldstein, 2003; Münscher, Vetter and Scheuerle, 2016). Social proof cues are mainly decision information interventions because they supply informational signals or shift attention without changing the feasible set (c.f., Mandel and Johnson, 2002; Münscher, Vetter and Scheuerle, 2016). Since e-commerce contexts often blend these categories (e.g., a ‘recommended’ preselection), careful operationalisation is required to attribute effects to the intended mechanism. For illustration purposes, **Table 1** summarises the adapted choice-architecture taxonomy (Münscher, Vetter and Scheuerle, 2016) and shows how the two focal nudges - social proof and default preselection - map onto the framework.

Table 1: Choice-architecture taxonomy and mapping of the focal nudges

Choice architecture category	Taxonomy technique	Meaning in digital interfaces	Mapping in this thesis
Decision information	Provide social reference point	The interface presents information about the behaviour or preferences of others, such as popularity signals, ratings, or “most chosen” indicators.	Social proof cue
	Translate information	The interface reframes or simplifies information to affect interpretation and evaluation.	Not examined experimentally
	Make information visible	The interface increases the salience or visibility of specific information.	Not examined experimentally
Decision structure	Change choice defaults	The interface preselects one option so that accepting it requires less effort than actively changing it.	Default preselection
	Change option-	The interface increases or decreases the effort	Not examined

	related effort	required to select particular options.	experimentally
	Change range or composition of options	The interface alters how options are grouped, ordered, or made available.	Not examined experimentally
	Change option consequences	The interface links options to particular costs, benefits, or downstream consequences.	Not examined experimentally
Decision assistance	Provide reminders	The interface prompts users to remember or reconsider a desired option.	Not examined experimentally
	Facilitate commitment	The interface supports self-commitment or public commitment mechanisms.	Not examined experimentally

Note. Adapted from Münscher, Vetter and Scheuerle, 2016.

2.1.4 Defaults: mechanism, robustness, and design contingencies

Defaults are among the most studied choice-architecture interventions because they can exert large effects with minimal implementation cost (Johnson and Goldstein, 2003; Jachimowicz et al., 2019). For instance, in an organ donation context, the magnitude of default effects was highlighted, and plausible mechanisms were described such as effort avoidance, inertia, and the interpretation of preselection as an endorsement (Johnson and Goldstein, 2003). Even when switching costs are objectively low, defaults can shift behaviour as they define the path of least resistance (Johnson and Goldstein, 2003).

Meta-analytic evidence reinforces the robustness of default effects while also emphasising heterogeneity. Jachimowicz et al. (2019) synthesise default effects across decision domains and demonstrate systematic heterogeneity: default effectiveness varies by context and can be accounted for by mechanisms such as ease (switching/effort costs) and endorsement (defaults as implicit recommendations), with susceptibility further shaped by the extent to which preferences are weakly formed or uncertain. In the broader nudging literature, effect sizes are positive on average but widely dispersed, implying that defaults are not uniformly powerful

and that context and implementation details matter for predicting behavioural impact (Jachimowicz et al., 2019; Hummel and Maedche, 2019).

Defaults also sit at the centre of ethical debates because they can be experienced as steering (Steffel, Williams and Pogacar, 2016). In this connection, ethically deployed defaults can be supported by disclosure and by enabling users to articulate preferences, which can increase perceived protection without necessarily eliminating default effects. This suggests that transparency can function as an ethical safeguard while preserving behavioural effectiveness, at least in some contexts (Steffel, Williams and Pogacar, 2016).

Translating default logic to the present thesis requires attention to design contingencies. In consumer shopping, defaults may appear as preselected choices, suggested alternatives, or auto-added items, and users may interpret them as either convenience features or as persuasion attempts (Friestad and Wright, 1994; Steffel, Williams and Pogacar, 2016).

2.2 Digital nudging

2.2.1 The definition of “digital nudging”

Digital nudging extends the choice architecture concept to digital interfaces, defining nudges as design elements in user interfaces that steer behaviour in predictable ways without restricting choice (Thaler and Sunstein, 2008; Weinmann, Schneider and vom Brocke, 2016).

Digital nudging can be positioned as an information systems research agenda focused on how interface artefacts shape decision-making through cognitive and behavioural mechanisms (Weinmann, Schneider and vom Brocke, 2016). This means that, in online contexts, the ‘environment’ is not physical but computationally designed and therefore highly malleable (Weinmann, Schneider and vom Brocke, 2016).

Furthermore, a distinctive feature of digital contexts is scalability: A single interface change can affect large numbers of users, and platforms can iteratively refine interventions through

A/B testing (Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019).

Digital environments also enable fine-grained measurement of behaviour, allowing researchers to connect interface manipulations to micro-level behavioural traces. These properties make digital nudging both practically consequential and methodologically attractive, but they also increase the ethical salience of influence because small effects can accumulate across millions of decisions (Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019).

Digital nudges should be distinguished from broader decision aids and recommender systems. Interactive decision aids structure search and evaluation by reducing effort, supporting comparison, or providing filtering tools (Häubl and Trifts, 2000). Recommendation systems personalise product exposure by predicting relevance (Dellaert and Häubl, 2012). Nudges can overlap with these tools, but the defining feature is the use of lightweight interface cues that steer without substantively changing the information set or the product assortment itself (Häubl and Trifts, 2000; Dellaert and Häubl, 2012).

2.2.2 Evidence on digital nudges: average effects and dispersion

Quantitative synthesis suggests that nudges are effective on average, yet their effects are far from uniform (Hummel and Maedche, 2019). In this regard, nudge effectiveness is conditional on context, implementation, and user characteristics, rather than a stable property of the intervention label alone. Additionally, small interface cues may substantially shift behaviour in digital settings (Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019; Gupta et al., 2024).

From a hypothesis-development standpoint, heterogeneity motivates both cross-nudge comparisons and personalisation logic. If different nudge types operate through different mechanisms, they may differentially influence behaviour in a given shopping task.

Furthermore, if users differ in cue reliance, then a ‘best’ nudge may not exist universally; instead, effectiveness may depend on person-nudge fit. Accordingly, dispersion in prior evidence provides a conceptual bridge from generic digital nudging to personalised interventions (Hirsh, Kang and Bodenhausen, 2012; Hummel and Maedche, 2019).

At the same time, interpreting ‘average effects’ requires caution. Digital nudging studies often differ in operational definitions, outcome metrics, and exposure contexts, creating challenges for aggregation (Hummel and Maedche, 2019). For the present thesis, this implies that strong claims about universal nudge rankings are not warranted; instead, hypotheses should be framed as context-contingent and mechanism-driven (c.f., Hummel and Maedche, 2019).

2.2.3 Transition: from generic to personalised digital nudges

The logic of personalisation follows directly from heterogeneity. If users differ in experience, motivation, and cue sensitivity, then a single generic interface intervention may create variation in nudge effectiveness (Zhu and Zhang, 2010; Hummel and Maedche, 2019). To operationalise matching, the thesis draws on decision-style segmentation. The general decision-making style framework captures stable tendencies such as rational, intuitive, dependent, avoidant, and spontaneous processing (Scott and Bruce, 1995).

2.3 Digital nudging in e-commerce: customer journey, focal mechanisms, and shopping-relevant evidence

2.3.1 E-commerce as a choice architecture across the journey

The e-commerce research may be most productively interpreted as documenting a distributed choice architecture rather than a set of isolated interface effects (Häubl and Trifts, 2000; Dellaert and Häubl, 2012). Online shopping environments shape behaviour across multiple

stages of the customer journey, including search, comparison, product evaluation, basket formation, and checkout (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Li, Grahl and Hinz, 2022). Accordingly, online choice is influenced not merely by the presence of a single cue, but by the cumulative architecture of visibility, ordering, defaults, and informational signals across the purchase process as a whole. For this reason, one should conceptualise interface design as a system of interacting influences rather than as a collection of independent stimuli (Weinmann, Schneider and vom Brocke, 2016; Jesse, Jannach and Gula, 2021).

A further implication of this perspective is that e-commerce platforms do not simply display options, but also actively structure search, consideration, and comparison. Search rankings, recommendation modules, filter settings, and product ordering influence which alternatives enter the consumer's consideration set before any final choice is made (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Li, Grahl and Hinz, 2022). In this sense, the architecture of online shopping affects not only the terminal decision, but also the intermediate cognitive processes through which options are noticed, interpreted, and shortlisted. The e-commerce field therefore suggests that behavioural outcomes are often the end result of earlier architecture effects on exposure, attention, and evaluation rather than the consequence of a single isolated intervention (Häubl and Trifts, 2000; Li, Grahl and Hinz, 2022).

Moreover, effects of digital choice architecture in e-commerce are rarely invariant across settings. Rather, the influence of interface cues depends on their position within the customer journey, their interaction with surrounding design elements, and the decisional demands of the focal task (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Weinmann, Schneider and vom Brocke, 2016). This suggests that one should interpret digital nudges not as universally transferable treatments, but as context-sensitive interventions whose effects are conditional on the architecture in which they are embedded (Weinmann, Schneider and vom Brocke, 2016;

Hummel and Maedche, 2019). Cross-study comparisons in e-commerce remain hence complex: similar labels may refer to interventions that differ materially in timing, placement, and surrounding interface structure (Hummel and Maedche, 2019; Jesse, Jannach and Gula, 2021).

Notably, there is a high concentration of empirical evidence in online grocery and food-choice settings (e.g., Jesse, Jannach and Gula, 2021; Valenčič et al., 2023; de Vries, Bol and van der Laan, 2025). This concentration is theoretically understandable, as grocery choices are frequent, familiar, comparatively low in involvement, and responsive to controlled digital manipulation (Valenčič et al., 2023; Nijssen et al., 2023). At the same time, such settings provide a useful environment for isolating behavioural mechanisms precisely because products are relatively standardised and purchase situations are easy to simulate experimentally (Valenčič et al., 2023). For the present thesis, a controlled grocery-like task is therefore well justified as an analytically appropriate context for identifying short-run behavioural responses to interface cues, rather than as a direct proxy for digital purchasing environments in general.

2.3.2 Social proof cues in online marketplaces

Within the broader e-commerce perspective, social proof may be regarded as an informational intervention that helps consumers cope with uncertainty rather than as a purely interpersonal conformity mechanism (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010). In online marketplaces - where product quality may be often difficult to verify before purchase - ratings, reviews, bestseller labels, and “most chosen” indicators may function as signals that increase the perceived diagnosticity or understanding of characteristics of an option (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010). In this context, social proof may reduce uncertainty through informational social influence, particularly when direct product

knowledge is limited or comparison costs are non-trivial. However, it is important to note that online-review effects (a social proof form) on purchase intention may be meaningful on average, yet they are also contingent on review content, format, product type, and contextual moderators (Qiu and Zhang, 2024).

As said above, social proof effects may vary systematically across product categories, review characteristics, and contextual conditions (Zhu and Zhang, 2010; Qiu and Zhang, 2024).

Namely, the social proof cue effectiveness may depend on whether the signal is perceived as credible, relevant, and informative for the focal decision; hence it is critical to consider how consumers evaluate the quality and trustworthiness of the cue itself (Brand and Reith, 2022; Qiu and Zhang, 2024). In this regard, presentation format or cross-cultural context can shape whether online reviews are treated as believable evidence (Brand and Reith, 2022). A more critical conclusion is that social proof does not function as an automatic behavioural trigger, but as a conditional informational resource whose effects depend on how the cue is interpreted by the consumer.

On the other hand, social proof per se should be distinguished from online reviews as marketplace social proof may appear in multiple forms, including bestseller tags, purchase counters, “trending” labels, and sparse popularity indicators (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010; Valenčič et al., 2023). This is crucial as different social-proof formats are unlikely to operate through identical inferential processes (Ghiassaleh, Kocher and Czellar, 2020; Rosillo-Díaz, Muñoz-Rosas and Blanco-Encomienda, 2024). Richer review formats may support more elaborate evaluation of product attributes, whereas sparse popularity signs such as “best seller” or “top rated” may function more as heuristic shortcuts for screening and narrowing the choice set (Ghiassaleh, Kocher and Czellar, 2020; Wang et al., 2024).

Finally, the persuasive value of social proof is inseparable from issues of authenticity and manipulation (Friestad and Wright, 1994; Mathur et al., 2019). If consumers suspect that popularity signals are fabricated, selectively displayed, or commercially engineered, the informational basis of the cue may be weakened and the intervention may cease to operate as legitimate social proof (Kim, Barasz and John, 2019; Xu, Zhang and Shen, 2025). This motivates a conclusion that social proof should be interpreted as a theoretically plausible but credibility-dependent nudge, the effectiveness of which remains contingent on truthful implementation and the level of uncertainty in the task environment (Zhu and Zhang, 2010; Hummel and Maedche, 2019; Xu, Zhang and Shen, 2025).

2.3.3 Defaults in e-commerce-like micro-decisions

Whereas social proof primarily alters the informational basis of inference, defaults modify the structure of action by making one option the path of least resistance (Johnson and Goldstein, 2003; Jachimowicz et al., 2019). By preselecting one option, defaults may reduce the effort required to accept that option and may simultaneously communicate normality, recommendation, or expected behaviour (Johnson and Goldstein, 2003; Jachimowicz et al., 2019). Notably, defaults remain among the most extensively studied interventions as they can shift behaviour while formally preserving the option set (Jachimowicz et al., 2019; Mertens et al., 2022). However, default effects are not uniform, but vary with switching costs, ambiguity, endorsement inferences, and the strength of prior preferences (Jachimowicz et al., 2019).

In digital consumer settings, “defaults” should not be treated as a single homogeneous intervention class, since the label encompasses materially different architectures, ranging from simple preselected options to more consequential opt-out arrangements (Steffel, Williams and Pogacar, 2016; Jachimowicz et al., 2019). Namely, default-like interventions may appear as preselected delivery options at checkout, as default feature levels in online

product configurators, or as default privacy and information-sharing settings in digital service environments (Herrmann et al., 2011; Craciun, 2018; Nijssen et al., 2023). Importantly, such defaults may be construed differently by users, for example as convenience-enhancing, recommendation-like, or commercially motivated, depending on the surrounding interface and perceived purpose of the design (Steffel, Williams and Pogacar, 2016; Hettler et al., 2025). For instance, consumers may construe a preselection as a convenience feature, a recommendation, a simplification device, or a commercially motivated attempt at steering (Friestad and Wright, 1994; Steffel, Williams and Pogacar, 2016). Hence, defaults may often be effective in digital consumer choice; however, their impact depends on how the preselection is embedded within, and interpreted through, the surrounding interface (Hummel and Maedche, 2019; Nijssen et al., 2023; Hettler et al., 2025).

Furthermore, defaults appear to be especially effective when the target action is simple, immediate, and low in effort, whereas their influence may weaken once active deliberation becomes less costly or the required action becomes more effortful (Behlen, Himmler and Jäckle, 2023; Nijssen et al., 2023; Ortmann et al., 2023). Default effects are at least partly mediated by cognitive-effort considerations: their influence may appear stronger where preselection enables effort minimisation, and weaker where the surrounding task architecture lowers the cost of deliberation (Ortmann et al., 2023). Thus, the important implication is that default effects are contingent on the decisional frictions embedded in the immediate environment.

Lastly, defaults are often examined in conjunction with other interface features, such as labels, rankings, or informational prompts, which complicates attribution (Valenčič et al., 2023; Hettler et al., 2025). This weakens comparability across studies, as apparently similar “default” manipulations may differ materially in placement, reversibility, or visibility. This means that the research corpus is better suited to supporting context-sensitive conclusions

about the effectiveness of defaults than to justifying universal claims about their superiority across digital consumer settings. As a result, default preselection should be rather understood as a theoretically strong yet context-dependent intervention, the effects of which are inseparable from the architecture in which it is embedded (Jachimowicz et al., 2019; Nijssen et al., 2023; Hettler et al., 2025).

2.3.4 Decision aids, recommendation ordering, and the broader interface environment

Defaults and social proof do not operate in isolation, but within a broader interface environment that structures online choice before any focal nudge is encountered (Häubl and Trifts, 2000; Dellaert and Häubl, 2012). In this regard, online shopping environments shape behaviour through decision aids, recommendation systems, rankings, filters, and sorting rules, all of which influence which options become visible, comparable, and ultimately choice-relevant (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Li, Grahl and Hinz, 2022; Wan, Kumar and Li, 2023). Such digital interfaces can influence both the process and the outcome of consumer choice by lowering comparison costs and reshaping search behaviour (Häubl and Trifts, 2000).

Evidently, digital interfaces shape consumer choice not only at the point of final selection, but throughout the preceding stages of search, comparison, and evaluation (Häubl and Trifts, 2000; Li, Grahl and Hinz, 2022). Decision aids, recommender systems, rankings, and sorting mechanisms influence which alternatives become salient and which products enter the consumer's consideration set (Häubl and Trifts, 2000; Li, Grahl and Hinz, 2022; Wan, Kumar and Li, 2023). Consequently, observed purchasing outcomes in e-commerce often reflect the cumulative effect of sequential interface structures, rather than the influence of a single isolated intervention. Thus, e-commerce choice environments may be viewed as layered

architectures that jointly organise exposure, search, and evaluation (Weinmann, Schneider and vom Brocke, 2016; Valta and Maier, 2025).

Moreover, e-commerce interfaces may actively shape which marketplace signals per se become salient, interpretable, and useful for consumers (Cai, Cebollada and Cortiñas, 2023; Wang et al., 2024; Lu et al., 2026). The influence of reviews and related marketplace cues depends not only on their valence or quantity, but also on structural properties such as visibility, readability, multimedia format, and the composition of the displayed review pool. Hence, digital marketplaces may shape demand partly by governing the informational environment itself (Cai, Cebollada and Cortiñas, 2023; Lu et al., 2026).

It is also worth mentioning that the behavioural influence of e-commerce interfaces depends not only on which options or signals are displayed, but also on whether the interface renders the choice environment cognitively manageable for consumers (Groissberger and Riedl, 2017; Wang et al., 2024). In this sense, filtering, sorting, and comparison tools may reduce informational burden transforming large assortments into more compact decision sets.

Recommendation environments may likewise be more readily accepted when their operation appears comprehensible rather than opaque to users (Cramer et al., 2008; Knijnenburg et al., 2012; Li et al., 2024). In this regard, explanations and transparency are consequential because they help users interpret why options are shown and how the recommendation process works, thereby shaping trust and acceptance (Jannach, Jugovac and Nunes, 2019; Li et al., 2024).

Taken together, digital choice architecture in e-commerce shapes consumer behaviour not through isolated cues alone, but through layered interface environments which organise exposure, search, interpretation, and cognitive manageability across the shopping journey (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Weinmann, Schneider and vom Brocke, 2016; Groissberger and Riedl, 2017; Valta and Maier, 2025).

2.4 Personalisation in digital environments: benefits, privacy, trust, and control

2.4.1 Defining personalisation

Personalisation refers to system-driven tailoring of content, presentation, or interactions to an individual based on observed or inferred information (Awad and Krishnan, 2006). It is conceptually distinct from customisation, where the user can actively configure preferences (Ansari and Mela, 2003), and from recommendation, where the system predicts product relevance but may not tailor the interface experience more broadly (Dellaert and Häubl, 2012). In marketing science, personalisation is often analysed in terms of when and where personalised content is delivered and how relevance interacts with consumer attention and context (Awad and Krishnan, 2006; Bleier and Eisenbeiss, 2015).

Importantly, the thesis will not focus on product-set personalisation (Awad and Krishnan, 2006; Bleier and Eisenbeiss, 2015), yet the personalisation concerns which nudge type is presented (default or social proof) rather than which products are offered.

2.4.2 The personalisation-privacy paradox and transparency preferences

In many digital services, consumers trade privacy for relevance (Awad and Krishnan, 2006). In this context, the willingness to be profiled may depend on perceived transparency and on the perceived benefits of personalisation (Awad and Krishnan, 2006). This implies that acceptance of personalised interface interventions may hinge on whether consumers understand why a particular experience is tailored to them and whether they perceive that tailoring as beneficial (Awad and Krishnan, 2006).

Institutional context and regulation shape the feasibility and effectiveness of personalisation (Goldfarb and Tucker, 2011). In this regard, privacy regulation may affect online advertising and the restrictions on targeting can alter both firm behaviour and consumer responses (Goldfarb and Tucker, 2011; Tucker, 2014).

In addition, transparency about targeting offers one route to mitigating intrusiveness concerns (Kim, Barasz and John, 2019; Michaelsen, 2024a). To this extent, explaining why an ad is shown can change perceptions and effectiveness, implying that ‘reason-giving’ can alter how consumers interpret personalised communications (Kim, Barasz and John, 2019). In the context of personalised nudges, similar dynamics may operate: if users infer that they are being targeted based on psychological traits, disclosure could change perceived autonomy or trust, potentially moderating both behavioural and subjective outcomes (Kim, Barasz and John, 2019; Michaelsen, 2024a).

Taken together, the personalisation literature suggests a trade-off. While tailoring may increase relevance and effectiveness by aligning with user needs, it can simultaneously increase perceptions of surveillance and manipulation (Awad and Krishnan, 2006; Komiak and Benbasat, 2006; Bleier and Eisenbeiss, 2015; Kim, Barasz and John, 2019). This motivates treating privacy, trust, and legitimacy as boundary conditions for personalised interventions, even when the empirical focus is behavioural choice (Awad and Krishnan, 2006; Lambrecht and Tucker, 2013; Eckles, Gordon and Johnson, 2018).

2.4.3 Trust formation in e-commerce and recommendation agents

Beyond being a direct antecedent of purchase intention, trust also functions as a processing heuristic that reduces perceived complexity and risk in online decision-making (Gefen, Karahanna and Straub, 2003). When consumers believe that a platform is competent and acts in their interest, they are more willing to rely on cues embedded in the interface - such as recommendations, defaults, or popularity indicators - because these cues are interpreted as informative rather than opportunistic (Friestad and Wright, 1994; Gefen, Karahanna and Straub, 2003; Komiak and Benbasat, 2006; Kim, Barasz and John, 2019).

However, the trust-building role of personalisation is not unconditional: the same tailoring that signals relevance and competence may also elicit suspicion if consumers infer covert profiling or persuasive intent (c.f., Friestad and Wright, 1994; Lambrecht and Tucker, 2013). In this context, transparency about why a suggestion is shown (e.g., “why am I seeing this?” explanations) can serve as a legitimacy cue that stabilises trust perceptions (c.f., Kim, Barasz and John, 2019).

2.4.4 Recapitulation on how personalisation may affect purchase via process channels

Personalised nudges may operate through alignment with the consumer’s decision-making style (Scott and Bruce, 1995). For instance, a default nudge may reduce effort for spontaneous choosers, and a social proof nudge may reduce uncertainty for dependent choosers who rely more on others’ opinions (Scott and Bruce, 1995; Hummel and Maedche, 2019). Similarly to interactive decision aids and recommendation agents (Häubl and Trifts, 2000; Dellaert and Häubl, 2012; Li, Grahl and Hinz, 2022), personalisation mechanisms may also affect purchase by redirecting attention and restructuring search, thereby shaping which options enter the consideration set. Moreover, the expected benefit of personalisation may be most pronounced under higher cognitive load or ambiguity, where decision-making style alignment reduces either search costs (defaults) or informational uncertainty (social proof) relative to a generic nudge (c.f., Simon, 1955; Hummel and Maedche, 2019).

2.5 Personalised digital nudging: person-nudge fit, trait targeting, and decision-style matching

2.5.1 The concept of personalised digital nudging

Personalised digital nudging can be defined as the use of heterogeneity data to tailor digital choice-architecture interventions to individual-level characteristics (e.g., traits, decision styles, or relevant situational differences), rather than applying the same interface design to all users (Weinmann, Schneider and vom Brocke, 2016; Mills, 2022). Importantly, such intervention can be described as nudge since it operates through interface-level steering (e.g., preselection or social information cues) and does not rely on restricting options or materially altering incentives, thereby staying aligned with the libertarian-paternalist constraint (Thaler and Sunstein, 2003; Peer et al., 2020). The central rationale is that “one-size-fits-all” approaches can be suboptimal when response functions vary across users, so personalisation is proposed as a systematic response to relevant heterogeneity (Mills, 2022).

Furthermore, one can distinguish choice personalisation (personalisation within a given nudge, by varying what outcome is nudged towards) and delivery personalisation (personalisation across nudges, by varying which nudge type is assigned to an individual) (Mills, 2022). In digital environments, delivery personalisation may be particularly feasible because platforms can measure or infer user attributes and dynamically deploy different user interface variants at scale, which is consistent with the broader “digital nudging” agenda that treats interface design as a controllable behavioural lever (Weinmann, Schneider and vom Brocke, 2016; Peer et al., 2020). Within the present thesis focus, delivery personalisation (Mills, 2022) is the relevant notion because it corresponds to selecting between defaults and social proof as alternative mechanisms intended to influence the same purchase-like decision (c.f., Ingendahl et al., 2020).

Empirically, personalised assignment can outperform generic assignment when the mapping rule is derived *ex ante* from observed relationships between individual differences and nudge responsiveness, as demonstrated in online security where decision-style-based assignment of password nudges improved outcomes substantially (e.g., Peer et al., 2020). Moreover, nudges may be operationalised through ordering and default settings; however, responsiveness may vary by user characteristics (e.g., expertise), which again suggests that “fit” claims require empirical verification in each domain (Liang and Willemsen, 2022).

2.5.2 Person-message fit and psychological targeting evidence

Evidence on person-message fit supports the general principle that tailoring persuasive content to recipient characteristics can increase effectiveness (Hirsh, Kang and Bodenhausen, 2012). Appeals matching individuals’ personality traits may be more persuasive than mismatched appeals (Hirsh, Kang and Bodenhausen, 2012), suggesting that congruence increases processing fluency or perceived relevance. Although Hirsh, Kang and Bodenhausen (2012) focus on message framing rather than on interface cues, the authors demonstrate a general behavioural rationale for matching interventions to individual differences.

Importantly, the targeting literature also highlights potential threats to internal validity and interpretability (Eckles, Gordon and Johnson, 2018). Field studies of psychologically targeted ads may face substantial identification challenges, including selection effects and correlated design features (Eckles, Gordon and Johnson, 2018). For personalised nudging research, this underscores the value of controlled experiments where assignment to generic versus matched nudges is randomised and where the choice set is held constant; such designs help separate matching effects from confounded personalisation artefacts (Eckles, Gordon and Johnson, 2018).

Finally, translating targeting logic to nudge types requires an explicit mapping from traits or decision styles to mechanisms. To reiterate, one plausible mapping may be that individuals prone to dependent decision making may be more responsive to social proof nudges as they seek social validation; accordingly, individuals prone to spontaneous decision making may benefit from defaults that may reduce effort (c.f., Bearden, Netemeyer and Teel, 1989; Scott and Bruce, 1995).

2.5.3 Measuring decision styles and moderators to operationalise matching

The aforementioned general decision-making style (GDMS) instrument provides a practical method for operationalising individual differences in decision processes (Scott and Bruce, 1995). Scott and Bruce (1995) conceptualise decision style as a learned, habitual pattern of responses that can be measured reliably across individuals. The GDMS dimensions (rational, intuitive, dependent, avoidant, and spontaneous) align with theoretical distinctions between deliberative processing, reliance on others, and rapid choice tendencies, making them suitable as matching variables for nudge selection (Scott and Bruce, 1995).

Additional constructs can complement decision styles by capturing more specific cue sensitivities. Need for cognition measures enjoyment of effortful thought and predicts reliance on central versus peripheral cues; susceptibility to interpersonal influence predicts responsiveness to social information; and need for uniqueness predicts preference for differentiation and potential aversion to popularity cues (Cacioppo and Petty, 1982; Bearden, Netemeyer and Teel, 1989; Tian, Bearden and Hunter, 2001).

2.6 Perceived autonomy, reactance, transparency, and ethical legitimacy

To begin with, personalisation heightens ethical scrutiny because it can enable more precise influence. Even when the intervention remains choice-preserving, tailoring based on

psychological traits may be perceived as exploitative or as taking advantage of vulnerabilities, particularly when the targeting logic is opaque. (Eckles, Gordon and Johnson, 2018; Bruns et al., 2025). Although the empirical focus of this thesis is behavioural effectiveness of the nudges, it is worth discussing whether interface cues undermine users' sense of agency, whether influence attempts are sufficiently transparent, and how to distinguish acceptable nudges from deceptive patterns (Bruns and Perino, 2023; Michaelsen, 2024a; Michaelsen, 2024b; Bruns et al., 2025).

2.6.1 Autonomy and perceived manipulation

Self-determination theory conceptualises autonomy as volition and self-endorsement rather than as mere option availability (Ryan and Deci, 2000). Accordingly, an intervention can be formally choice-preserving yet still be experienced as controlling if users infer external pressure or manipulation (Ryan and Deci, 2000). In consumer contexts, users may interpret marketing actions as influence attempts, infer intent, and may respond with scepticism or resistance when they perceive persuasion (Friestad and Wright, 1994).

Relatedly, autonomy responses to default nudges implementation may not be necessarily uniform (Michaelsen, 2024b). In this regard, people do not always experience default nudges as strongly autonomy-infringing, even when behaviour shifts (Michaelsen, 2024b).

Additionally, consumers' reactance level to the same intervention may vary under different condition: e.g., a recommendation – which may generally be helpful – may be experienced as more of a threat or intrusion and can cause reactance if it is viewed by a consumer as unsolicited and counter-attitudinal (Fitzsimons and Lehmann, 2004).

2.6.2 Transparency and disclosure

Transparency is frequently proposed as a safeguard: making an intervention visible or explainable can increase perceived ethicality and consumer protection, even when behavioural effects remain (Steffel, Williams and Pogacar, 2016). However, transparency may be multi-dimensional: one may distinguish disclosure-based forms of transparency (what is communicated) from awareness-based transparency (what users notice), which are analytically distinct routes to transparency (Marco and Douglas, 2023). This distinction is crucial as it clarifies that transparency per se can be impeded either because information is not provided (disclosure deficit) or because users do not register it (awareness deficit) (Marco and Douglas, 2023).

Furthermore, disclosure can also be decomposed into what exactly is disclosed – e.g., the default's target behaviour, its purpose, or how defaults work – and these components can have different effects on compliance and evaluation (Paunov, Wänke and Vogel, 2019).

Transparency may also vary in timing (ex-ante vs ex-post) and in whether it is framed as a voluntary/proactive versus mandated disclosure, both of which can shift how the intervention is interpreted and how people respond (Ivanković and Engelen, 2019; Paunov et al., 2022). Similarly, transparency in targeted advertising shows that the content of an explanation (e.g., whether information was stated vs inferred and where it was obtained) shapes acceptability judgements and downstream effectiveness (Kim, Barasz and John, 2019). To summarise, although disclosure is not experimentally manipulated in the present thesis, transparency should remain an important design and governance consideration when implementing defaults and social proof cues in consumer-facing interfaces (Steffel, Williams and Pogacar, 2016; Michaelsen, 2024a; Bruns et al., 2025).

2.6.3 Boundary conditions and dark patterns

Finally, it is important to delimit the intervention class under study. This thesis focuses on choice-preserving interface interventions, specifically a default preselection and a social proof information cue, that steer without restricting the option set (c.f., Thaler and Sunstein, 2003; Weinmann, Schneider and vom Brocke, 2016).

To maintain conceptual and ethical clarity, the thesis treats dark patterns (also dark commercial patterns) as outside the scope, insofar as they rely on deception, coercion, or undue friction that subverts autonomous decision-making (Gray et al., 2018; OECD, 2022). In this regard, illustrative examples include obstruction (e.g., making cancellation or opt-out unnecessarily difficult), sneaking/hidden costs (e.g., late-stage add-ons), and misleading interface claims that distort the informational basis of choice (Gray et al., 2018; Mathur et al., 2019; OECD, 2022). Accordingly, the present study examines legitimate choice architecture effects (defaults and truthful social information) rather than harmful interface practices designed to impair informed choice (Mathur et al., 2019; OECD, 2022).

For clarification purposes, **Table 2** provides a concrete comparison between the legitimate nudges examined in this thesis and comparable interface practices that would qualify as dark patterns and are out of scope of the thesis. The table makes it vivid that the relevant boundary lies not in whether behaviour is influenced, but in whether influence remains transparent, reversible, and consistent with informed choice.

Table 2: Legitimate nudges and their dark-pattern lookalikes in e-commerce interfaces

Interface practice	Legitimate nudge	Dark-pattern counterpart	Difference
Default preselection	A grocery website preselects one delivery slot, but all other slots remain clearly visible and can be changed with one click.	At checkout, an optional insurance product is pre-ticked, and the deselection box is small, visually muted, or hidden in a secondary menu.	Both steer through default effects, but the first preserves easy and informed reversal, whereas the second relies on concealment or friction.

Social proof cue	A product page truthfully states: <i>“Most chosen by customers this week”</i> or <i>“4.6/5 based on 1,248 verified reviews.”</i>	A product page shows fabricated popularity claims such as <i>“500 people are viewing this now”</i> or <i>“Best seller”</i> without a verifiable basis.	Both use social information, but the first provides accurate informational support, whereas the second distorts the informational basis of choice.
Choice support vs obstruction	A website presents a visible <i>“Change selection”</i> or <i>“Opt-out”</i> button next to the preselected option.	A subscription service requires multiple extra screens, confusing labels, or hidden menus to cancel or opt out.	The first keeps alternatives practically available; the second imposes undue effort that undermines effective choice.
Additional charges	The total price, including delivery and service fees, is shown before the final confirmation step.	New charges appear only at the last checkout page, after the user has already invested time and effort in the purchase flow.	The first supports informed choice; the second exploits commitment and surprise through hidden costs.
Recommendation framing	A label states: <i>“Recommended based on customer rating”</i> and the recommendation criteria are consistent with the information displayed.	A label such as <i>“Recommended for you”</i> is shown without explanation, although it is actually paid placement or commercially prioritised ranking.	The first is a transparent cue; the second masks persuasive intent or commercial motive.
Urgency / popularity messaging	A site states: <i>“Only 2 left in stock”</i> when this is factually correct and linked to actual inventory.	A site repeatedly displays <i>“Only 2 left!”</i> or countdown timers that reset, even when stock is not genuinely limited.	The first informs; the second deceives by fabricating urgency to pressure choice.

Note. Adapted from Weinmann, Schneider and vom Brocke (2016); Gray et al. (2018); Mathur et al. (2019); OECD (2022).

2.7 Synthesis: research gaps and hypothesis development

2.7.1 Research gap synthesis

As described above, across behavioural economics, marketing, and information systems, there is robust evidence that interface design may shape consumer choice - defaults, social proof cues, decision aids, and recommendation ordering can shift behaviour by changing effort, attention, and inference (e.g., Häubl and Trifts, 2000; Hummel and Maedche, 2019).

However, digital nudging research still remains methodologically and conceptually heterogeneous, complicating cumulative inference and comparability across intervention forms and contexts (Valenčič et al., 2023). Domain-specific reviews in online grocery environments further emphasise inconsistencies in terminology and reporting, alongside a limited set of well-described user interface implementations, which reinforces the need for more tightly controlled, comparable experimental designs (Valenčič et al., 2023).

The first research gap concerns the comparative value of personalised versus generic digital nudges in purchase-like e-commerce tasks. While person-message fit and psychological targeting research suggests that matching persuasive content to recipient characteristics may increase effectiveness (e.g., Hirsh, Kang and Bodenhausen, 2012), relatively few studies directly test whether matching choice-architecture interventions to individual decision styles outperforms applying the same nudge to everyone (e.g., Peer et al., 2020; Ingendahl et al., 2020). Moreover, the existing, albeit scarce digital nudge personalisation evidence indicates both promise and uncertainty: personalisation may improve outcomes in some large-scale online settings (e.g., online security, Peer et al., 2020), yet in some other domains personalisation effects may not be substantial (e.g., e-commerce, Ingendahl et al., 2020; healthy food choice, de Vries, Bol and van der Laan, 2025). Accordingly, the present thesis addresses a practically consequential question, to what extent and in what ways personalised digital nudges may influence consumer decision-making outcomes in e-commerce compared to non-personalised nudges or no nudges at all.

The second research gap focuses on within-study, head-to-head comparisons of default preselection and social proof cues implemented in the same controlled e-commerce interface (e.g., Jesse, Jannach and Gula, 2021). Although choice-architecture interventions may produce meaningful average effects, meta-analysis cannot straightforwardly adjudicate whether defaults or social proof should be expected to “work better” in a specific shopping

task when different studies vary in user interface implementations, tasks, and outcome definitions (Hummel and Maedche, 2019; Mertens et al., 2022).

On the other hand, recent e-commerce-facing experiments show that cross-nudge comparisons are feasible within the same or closely standardised online-choice environments (e.g., Guath, Stikvoort and Juslin, 2022; Nijssen et al., 2023). However, such designs remain relatively less common in the online shopping literature - particularly in configurations that isolate defaults versus social proof in a purchase-like choice task in grocery or other domains, highlighting the need for more standardised, comparable experimental implementations (Valenčič et al., 2023). In this connection, the thesis will attempt to discover which type of digital nudges (default or social proof) is the most effective at influencing consumer behaviour in the controlled online grocery store environment.

The third research gap is centred around the ethical debates about autonomy, transparency, and legitimacy in consumer-facing nudging research (e.g., Kim, Barasz and John, 2019; Michaelsen, 2024a; Bruns et al., 2025). Although these issues are central in principle, empirical studies frequently prioritise behavioural outcome metrics (e.g., clicks, selections, conversions) over welfare-relevant or perceptual outcomes (Hummel and Maedche, 2019; Valenčič et al., 2023). Autonomy-, trust-, and legitimacy-related perceptions are often treated only implicitly or as secondary measures, although recent e-commerce platform experiments increasingly incorporate explicit user-perception measures alongside behavioural effects (e.g., Hettler et al., 2025). In this regard, the thesis separates behavioural testing from ethical evaluation: it estimates the effects of defaults and social proof on choice outcomes, while formulating implications through established autonomy- and intent-based frameworks and evidence on ethically deployed defaults and legitimacy-oriented nudge design (Ryan and Deci, 2000; Friestad and Wright, 1994; Steffel, Williams and Pogacar, 2016; Janson and Barev, 2026).

2.7.2 Hypotheses development

The aforementioned research gaps motivate the thesis hypotheses.

- **H1**: The implementation of a generic nudge will result in a higher average proportion of target option selections compared to non-nudged individuals.
- **H1a**: Within the generic nudge implementation, the social proof nudge will result in a higher average proportion of target option selections compared to the default nudge.
- **H2**: The implementation of a matched nudge will result in a higher average proportion of target option selections compared to generically nudged individuals.

To reiterate, the **H1** is connected to the evidence that both default and social proof cues may significantly shift choice behaviour towards the nudged option (c.f., Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010; Jachimowicz et al., 2019). Concerning the **H1a**, the theoretical basis is that consumers may rely on other consumers' historical choice or evaluation of the two products within each category (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010). In respect of **H1a**, the social proof cue is expected to yield a larger average effect than the default preselection as the social proof cue may strongly reduce uncertainty via informational social influence (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010). Accordingly, the **H2** is grounded in the person-message fit and psychological targeting evidence (Hirsh, Kang and Bodenhausen, 2012; Weinmann, Schneider and vom Brocke, 2016; Mills, 2022). As mentioned earlier, the person-nudge mapping in this thesis is operationalised through GDMS-based segmentation (Scott and Bruce, 1995).

2.8 Conclusion

This chapter has established the conceptual and empirical foundations for analysing how digital choice architectures shape consumer behaviour in e-commerce environments. Building from bounded rationality and dual-process accounts, the review has shown why interface cues

can exert systematic influence under conditions of limited attention, time pressure, and imperfect preferences, which are characteristic of many online shopping decisions (Simon, 1955; Kahneman, 2011). Within this broader behavioural framework, digital nudging has been positioned as a scalable, interface-mediated form of choice architecture that can be deployed at low marginal cost and evaluated through controlled experimentation (Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019).

Focusing on the two focal interventions of the thesis, the chapter has synthesised that both defaults and social proof cues may meaningfully shift choice by leveraging distinct psychological mechanisms. Namely, defaults exert influence through effort reduction, inertia, and perceived endorsement with robust average effects alongside substantial heterogeneity across contexts and implementations (Johnson and Goldstein, 2003; Jachimowicz et al., 2019). Social proof cues, in contrast, operate primarily through informational inference and uncertainty reduction, with social information being able to influence demand, yet again, with some systematic moderation by product and consumer characteristics (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010). Taken together, both nudges may be expected to alter relative choice shares in a controlled e-commerce task; however, one should not infer universal rankings of “strongest” nudge types.

Furthermore, the chapter developed the logic of personalised digital nudging as a response to heterogeneity in nudge responsiveness. In this regard, tailoring persuasive influence on recipient characteristics may increase its effectiveness (Hirsh, Kang and Bodenhausen, 2012), and decision-style-based assignment rules may improve outcomes in online environments (Peer et al., 2020). In the present thesis, this matching logic is operationalised through segmentation based on the GDMS framework, enabling a mechanism-consistent mapping between decision style and the nudge type that is hypothesised to be most behaviourally effective for that individual (Scott and Bruce, 1995).

Finally, the chapter delineated the ethical boundary conditions that contextualise behavioural effectiveness claims. Even when interventions are choice-preserving, legitimacy judgements can depend on inferred intent, transparency, perceived manipulation, with dark patterns remaining outside the scope of the present work (Ryan and Deci, 2000; Friestad and Wright, 1994; Steffel, Williams and Pogacar, 2016; Gray et al., 2018; OECD, 2022). Overall, the synthesis motivates the thesis hypotheses by (1) explaining behavioural shifts from defaults and social proof (*H1*), (2) motivating a context-contingent expectation of relative effectiveness of the nudges (*H1a*), and (3) showing how decision-style-based matching may outperform random nudge assignment (*H2*), thereby setting the stage for the experimental design and operationalisation in the subsequent chapters.

Chapter 3: Methodology

3.1 Introduction

This chapter sets out the methodological logic of the thesis and specifies how the empirical study is translated from conceptual argument into an executable online experiment in a controlled e-commerce-like environment. The overall objective of the study is to examine 1) whether digital nudges embedded in a simulated grocery e-commerce environment alter product choice, 2) whether social proof outperforms defaults under non-personalised generic assignment, and 3) whether nudge assignment based on a simple form of psychological matching leads to higher behavioural effectiveness than non-personalised assignment. Methodologically, the thesis adopts a quantitative, deductive and causal design. The study uses an online randomised experiment embedded in Qualtrics. Qualtrics is a cloud-based online survey platform designed for the creation, distribution, and analysis of digital questionnaires, including structured survey instruments and experimental choice tasks (Qualtrics, 2026). A controlled online experiment is appropriate because the substantive interest is not merely in describing shopping attitudes, but in identifying whether systematically varied interface cues change the probability that respondents choose a predesignated target option. In line with works on digital nudging, the study treats the online user interface as a digital choice architecture in which seemingly small design features may influence judgement and choice without formally removing alternatives or changing monetary incentives (Weinmann, Schneider and vom Brocke, 2016). Thus, the experiment isolates interface-level changes while holding the product alternatives constant within each category. The methodology adopted in this study also reflects the technical and administrative constraints of the research environment in which the experiment was implemented. As the author's student version imposed a strict upper limit on the number of survey questions, the final instrument had to be restricted to seven repeated grocery-choice categories across the

policy conditions. This may reduce the observational depth of the task and limit the statistical power of the design, particularly with regard to subtle differences between generic and matched nudging policies. The resulting experiment should therefore be understood as a constrained implementation rather than a fully powered confirmatory test. Even so, the design retains the essential causal structure required to compare non-nudged, generically nudged, and psychologically matched nudging conditions in a controlled online setting.

3.2 Research design

3.2.1 Formal consent and attention check

First, all participants entered the survey through a standard consent page and were also required to pass an attention check task ('To show that you are reading the instructions carefully, please select "Strongly disagree" for this item.') in order to proceed to the survey questions (c.f., Oppenheimer, Meyvis and Davidenko, 2009). If a participant did not consent to the study or failed to choose "Strongly disagree" among the multiple-choice options in the attention check task, the participation of this person was automatically finished, and they did not see the following survey tasks accordingly.

3.2.2 GDMS-based questionnaire

Afterwards, participants completed a table indicating how well each statement described their decision-making habits on a 5-point Likert scale from 'strongly disagree' to 'strongly agree' (Likert, 1932). Each statement in the table was related to the dependent and spontaneous decision-making styles based on an adopted reduced version of the GDMS instrument (Scott and Bruce, 1995). The first group of five statements was related to the socially dependent style and the second group of five statements - to the spontaneous one (see Appendix 1). In

order to distinguish between the decision-making styles, the following logic was applied. If a person indicated ‘somewhat agree’ or ‘strongly agree’ among all of the five statements related to the dependent (spontaneous) style while the person’s choices among all of the five statements related to the spontaneous (dependent) style were either ‘strongly disagree’, ‘somewhat disagree’, or ‘neither agree nor disagree’, the person’s decision-making style (i.e. ‘gdms_type’) was categorised as ‘DEPENDENT’ (‘SPONTANEOUS’). In all other cases, the ‘gdms_type’ was categorised as ‘MIXED’, as their decision-making style did not strictly align with either of the two designated styles. Such a deliberately strict categorisation rule was adopted for the GDMS-based grouping, so that only respondents with an unambiguous response pattern were classified as dependent or spontaneous, whereas borderline and overlapping cases were assigned to the mixed category.

3.2.3 Experiment policy conditions and shopping tasks

In the next step, participants were randomly routed into one of three experimental policy conditions (as represented in the dataset by the variable ‘policy_arm’):

- 1) a non-nudge control condition (‘CONTROL’),
- 2) a generic nudge condition (‘GENERIC’),
- 3) a matched nudge condition (‘MATCHED’).

The ‘CONTROL’ constitutes the baseline condition. Participants assigned to ‘CONTROL’ saw seven categories in random order and neutral form and received no interface-based cue. The two product alternatives remained structurally equivalent in price and pack size. Only minor identifying characteristics such as visual representation, origin or producer label differed: e.g., ‘Apples, Gala 1kg – Germany 1,99€’ and ‘Apples, Gala 1kg – Poland 1,99€’ (see Appendix 2). The purpose of the ‘CONTROL’ was to estimate how often the target option is chosen when no default or social endorsement is presented.

The ‘GENERIC’ experimental condition is intended to represent non-personalised choice architecture. Participants were exposed to the same product categories displayed in random order; however, a second randomisation mechanism assigned them to one of two generic nudge types (‘generic_nudge_type’) at participant level. In the ‘SOCIAL’ branch participants saw the social-proof nudged version of all seven product categories: e.g., ‘Apples, Gala 1kg – Germany Our customers’ choice 1,99€’ and ‘Apples, Gala 1kg – Poland 1,99€’ (see Appendix 3). By contrast, in the ‘DEFAULT’ participants saw the default nudged version of the seven product categories: e.g., ‘Apples, Gala 1kg – Germany *Already in your cart* 1,99€’ (this product option appeared preselected in the interface) and ‘Apples, Gala 1kg – Poland 1,99€’ (see Appendix 4). Importantly, for both ‘SOCIAL’ and ‘DEFAULT’ branches, always the same product option among the two within each category was nudged: e.g., the option ‘Apples, Gala 1kg – Germany 1,99€’ was always nudged, and ‘Apples, Gala 1kg – Poland 1,99€’ was never nudged. This architecture is crucial because it produces a clean comparison between social proof and default under random assignment. Since the nudge type in the ‘GENERIC’ condition was not tied to respondents’ psychological classification, any difference between the two generic branches can be interpreted as a treatment contrast rather than an artefact of prior sorting.

The ‘MATCHED’ experimental condition operationalises personalisation as follows. As mentioned above, prior to experimental condition assignment, the GDMS pre-survey classified respondents as ‘DEPENDENT’, ‘SPONTANEOUS’, or ‘MIXED’. Respondents classified as ‘DEPENDENT’ were theorised to be more receptive to social guidance (Bearden, Netemeyer and Teel, 1989; Zhu and Zhang, 2010) and therefore received the social-proof version of the shopping task under the matched experimental condition (i.e., they were always shown the ‘SOCIAL’ branch). Accordingly, respondents classified as ‘SPONTANEOUS’ were theorised to be more receptive to low-friction, action-saving cues

(Johnson and Goldstein, 2003; Jachimowicz et al., 2019) and therefore received the default version of the task (i.e., they were always shown the ‘DEFAULT’ branch). Respondents classified as ‘MIXED’ were not treated as matchable with sufficient confidence. Instead of being completely excluded from the ‘MATCHED’ experimental condition, they were assigned to a generic fallback within this experimental condition. Namely, they were randomly shown either the ‘SOCIAL’ or the ‘DEFAULT’ branch, exactly as in the ‘GENERIC’ experimental condition. This was intended to preserve sample retention and keep the ‘MATCHED’ condition implementable for all respondents while acknowledging that the aforementioned GDMS rule was intentionally very strict. In this regard, the experiment did not test an ideal, fully personalised system with continuous optimisation; it tested whether a transparent and practically deployable decision-style rule can outperform a policy that ignores such information. **Table 3** summarises the implemented experimental architecture and the exact task-routing logic used in the survey

Table 3: Experimental condition assignment logic

Experimental condition	Assignment logic	Task version shown	Interpretive role
1. CONTROL	policy_arm = CONTROL	CONTROL (7 neutral categories)	No-nudge baseline for H1
2. GENERIC	policy_arm = GENERIC and generic_nudge_type = SOCIAL or DEFAULT (randomly)	SOCIAL or DEFAULT (7 categories with respective nudge)	Generic social-proof and default versions; part of H1 and H1a
3a. MATCHED - DEPENDENT	policy_arm = MATCHED and gdms_type = DEPENDENT	SOCIAL (7 categories with social nudge)	Matched social-proof version; part of H2
3b. MATCHED – SPONTANEOUS	policy_arm = MATCHED and gdms_type = SPONTANEOUS	DEFAULT (7 categories with default nudge)	Matched default version; part of H2
3c. MATCHED – Mixed fallback	policy_arm = MATCHED and gdms_type = MIXED; generic_nudge_type = SOCIAL or DEFAULT	SOCIAL or DEFAULT (7 categories with respective nudge)	Fallback branch preserving full sample; <u>not</u> part of H2

Source: Author.

3.2.4 Demographic and autonomy questionnaires

The survey flow further includes a demographic part comprising questions about age, gender, and online shopping frequency (see Appendix 5). This part is followed by a matrix of five statements regarding the perceived autonomy during the shopping task valued on a 5-point Likert scale from ‘strongly disagree’ to ‘strongly agree’ (Likert, 1932; see Appendix 6). The five statements were adapted from autonomy questionnaires from prior research (e.g., Wachner, Adriaanse and De Ridder, 2021).

3.2.5 Important notes on experiment structure

This sequencing of questions is deliberate. Psychological classification and experimental condition assignment occur before any treatment exposure; behavioural choice is then observed during the shopping task; and reflective evaluations are only collected afterwards, thereby reducing the risk that introspective questions contaminate behaviour during the task itself. It is also worth mentioning that participants were required to answer each and every question throughout the survey; and all survey sections (e.g., GDMS matrix, shopping task, demographic questions) were separated from each other without an option to go back to the previous section. This was intended to prevent the participants from e.g., attempting to change their choices in the shopping task after having seen the autonomy post-evaluation matrix.

3.2.6 Concluding notes

In general, the research design may be characterised as a between-subject policy experiment with repeated choice observations within respondents. Experimental condition assignment occurred at participant level. Within the assigned condition, each respondent was exposed to only one displayed task version: ‘CONTROL’, ‘SOCIAL’, or ‘DEFAULT’. Within that task, as mentioned earlier, each respondent completed seven binary product choices. Even though

the final implementation is small, it still enables comparison of behaviour under three substantively distinct decision environments: no nudge, generic nudging, and GDMS-informed nudging.

The shopping environment per se was designed as a scaled-down grocery interface. Each category presented exactly two options that were intended to be as substitutable as possible.

The final implementation retained seven categories that were easy to understand and standardise: tomatoes, apples, bananas, etc. Within each category, price, pack size, and visual quality level were held constant so that the experimental manipulation did not inadvertently become confounded with a superior product profile. Consistent with digital-nudging research, the intervention was attached to the interface and not to intrinsic product quality (Weinmann, Schneider and vom Brocke, 2016).

3.3 Research questions and hypotheses

The research is designed to answer the following broader research questions. The first asks whether nudged choice architectures, taken as a broad class, increase the likelihood that respondents select the focal target option relative to a neutral interface. The second asks whether social proof and default cues differ in effectiveness when they are assigned generically. The third is centred around the idea whether a matching rule based on pre-treatment decision style yields higher behavioural effectiveness than a generic experimental condition that does not use such information. Together, these questions preserve the conceptual progression developed in the literature review: from average nudge effects to differential nudge effectiveness, to the added value of personalisation. Thus, the exact formulation of the research questions is as follows:

- **RQ1:** Does exposure to a generically assigned digital nudge, as opposed to a neutral product presentation, increase the likelihood of selecting the predesignated target option in a simulated grocery e-commerce environment?
- **RQ2:** Within a generic, non-personalised nudge implementation, which nudge type is more effective in increasing target-option choice: social proof or default?
- **RQ3:** Does a GDMS-based matching rule outperform a non-personalised nudge implementation in increasing target-option choice?

The **RQ1** translates into the first hypothesis:

- **H1** proposes that participants assigned to the ‘GENERIC’ experimental condition will show a higher average proportion of the target option selections across product categories than participants assigned to the ‘CONTROL’ experimental condition.

Therefore, **H1** captures the overall behavioural effect of introducing interface cues into the shopping environment.

The **RQ2** concerns the comparative strength of the two nudge types under non-personalised assignment. This becomes the following exact hypothesis:

- **H1a** states that among participants assigned to the ‘GENERIC’ experimental condition, those assigned to the ‘SOCIAL’ task version will show a higher average proportion of the target option selections across product categories than those assigned to the ‘DEFAULT’ task version.

As the ‘GENERIC’ experimental condition randomises the nudge type at participant level, **H1a** is an appropriate test of the relative effectiveness of social proof and defaults in the present study.

The **RQ3** concerns personalisation and translates into the following hypothesis:

- **H2** proposes that among respondents classified as ‘DEPENDENT’ or ‘SPONTANEOUS’, those assigned to the ‘MATCHED’ experimental condition will

show a higher average proportion of target option selections than respondents assigned to the ‘GENERIC’ experimental condition.

Hence, **H2** examines whether decision-style-based nudge assignment yields greater behavioural effectiveness than non-personalised nudge assignment among respondents for whom a clear person-nudge match can be specified (i.e. not ‘MIXED’).

3.4 Data collection methods

3.4.1 Sampling and recruitment

Data collection was conducted as a single-wave online survey experiment implemented in Qualtrics. The study relied on a non-probability convenience sample of adult respondents, recruited primarily through the author’s personal, student and university-related networks. Accordingly, the study should be understood as a resource-constrained online experiment rather than as a population-representative consumer survey. This is methodologically acceptable because the purpose of the research is to identify causal differences between experimentally constructed choice environments rather than to estimate population parameters.

The survey was self-administered and asynchronous. Although the questionnaire was technically accessible across devices, completion on a desktop or laptop was recommended because the product-comparison layout was more clearly displayed on larger screens. Data collection took place between late March and early April 2026.

3.4.2 Survey procedure

To recapitulate, the implemented survey followed a fixed and fully standardised sequence. After entering through the informed-consent page, respondents completed the instructional

attention check, followed by the reduced ten-item GDMS block used for pre-treatment classification into three decision-making types. On this basis, Qualtrics then randomised respondents into one of the three experimental conditions, and where required by the routing logic, assigned the relevant nudge version. Participants subsequently completed the seven-category shopping task within the routed environment, after which they answered the demographic questions and the post-task autonomy evaluation block. This sequencing ensured that classification and assignment occurred before treatment exposure, that behavioural choices were observed before any reflective evaluation, and that all respondents passed through the same procedural structure. The methodological role of each stage is summarised in **Table 4**.

Table 4: Methodological purpose of each survey stage

Stage	Content	Role in design	Key outputs
Consent	Study information and voluntary consent	Ethical entry point and eligibility confirmation	Consent outcome
Attention check	Instructional manipulation check	Data-quality screen before study entry	Pass/fail screening outcome
GDMS block	Reduced 10-item GDMS matrix covering dependent and spontaneous tendencies	Pre-treatment classification for later matched routing	Response profile for pre-treatment classification
Condition assignment	Randomised CONTROL, GENERIC, or MATCHED routing	Participant-level experimental assignment	Assigned experimental condition
Nudge-type assignment	Random SOCIAL / DEFAULT where required	Determines nudge type within GENERIC and within MATCHED fallback cases	Assigned nudge version where applicable
Shopping task	Seven product categories under routed task version	Behavioural outcome collection	7 binary product choices
Demographics	Age, gender, online-shopping frequency	Sample description and balance checks	Demographic descriptors
Post-task autonomy measure	5 Likert-type autonomy items	Secondary perceptual outcome	Post-task autonomy responses

Source: Author.

3.4.3 Data generated by the procedure

The survey generated five types of data. First, it produced screening data through consent and attention-check outcomes. Second, it generated pre-treatment classification data from the GDMS block. Third, it generated the core behavioural data through seven repeated binary product choices within the routed task environment. Fourth, it produced descriptive respondent data through the demographic items. Fifth, it generated post-treatment perceptual data through the five-item autonomy scale.

In procedural terms, the final dataset therefore consists of: (1) survey-entry and exclusion information, (2) GDMS-based respondent classification, (3) experimental assignment information, (4) repeated choice outcomes from the shopping task, (5) demographic descriptors, and (6) post-task autonomy evaluations.

3.4.4 Data-quality safeguards

Two safeguards were built into the survey procedure. First, respondents who did not provide consent were excluded immediately. Second, respondents who failed the instructional manipulation check were also terminated before entering the main study. In addition, the sequential block structure of the survey ensured that post-task questions were only encountered after the behavioural task had been completed. This preserved the intended temporal ordering of classification, assignment, treatment exposure, behavioural response and post-task evaluation.

3.5 Operationalisation of constructs

3.5.1 Behavioural outcome

The primary dependent variable is target-option choice at the category level. For each of the seven product categories, one of the two alternatives was defined ex ante as the focal target option and retained this status across all task environments. In the ‘CONTROL’ task versions, the target option was analytically focal but visually unmarked. In the ‘SOCIAL’ version, it was the option carrying the social-proof cue. In the ‘DEFAULT’ version, it was the option displayed as already in the cart and preselected in the interface. This design ensures comparability across conditions because the same focal option is observed under neutral, social-proof and default presentations.

At category level, the variable is coded 1 = target option chosen and 0 = non-target option chosen. At respondent level, a secondary aggregate measure is computed as target-choice share, defined as the proportion of target options selected across the seven product categories. This respondent-level proportion ranges from 0 to 1 and serves as the main summary outcome for hypothesis testing.

3.5.2 Experimental assignment variables

The main independent variable is ‘policy_arm’, a participant-level categorical variable with three values: ‘CONTROL’, ‘GENERIC’ and ‘MATCHED’. As ‘policy_arm’ is assigned by Qualtrics randomisation, it is exogenous to subsequent shopping choices. The second assignment variable is ‘generic_nudge_type’, coded as ‘SOCIAL’, ‘DEFAULT’ or ‘NA’. It is populated for all respondents in the ‘GENERIC’ condition and for ‘MIXED’ respondents only in the ‘MATCHED’ condition, who are routed through the generic fallback. For matched ‘DEPENDENT’ and ‘SPONTANEOUS’ respondents, the variable is analytically redundant and is retained as missing (‘NA’) or not applicable.

3.5.3 GDMS-based matching variable

The principal pre-treatment classification variable is ‘gdms_type’, coded as ‘DEPENDENT’, ‘SPONTANEOUS’ or ‘MIXED’. This variable is derived from the reduced ten-item GDMS matrix. As described, following the implemented survey logic, a respondent is classified as ‘DEPENDENT’ if all five dependent items receive agreement responses while the spontaneous items do not; a respondent is classified as ‘SPONTANEOUS’ under the converse pattern; all remaining cases are classified as ‘MIXED’. Accordingly, gdms_type functions both as a matching variable in the experimental design and as a subgroup variable in later analysis.

3.5.4 Perceived autonomy

Perceived autonomy is operationalised as a secondary post-task self-report construct. It is measured through a five-item Likert-type matrix administered after the shopping task and after the demographic questions. The items capture closely related aspects of experienced agency, including perceived freedom of choice, perceived decisional control and the extent to which the final choice reflected personal preferences. Each item is measured on a five-point response scale ranging from ‘strongly disagree’ to ‘strongly agree’. Analytically, perceived autonomy is not part of the primary behavioural hypothesis tests. However, it provides an interpretive layer for assessing whether behavioural influence through digital nudging is accompanied by any reduction in the subjective experience of volition and control.

3.5.5 Auxiliary variables and exclusion criterion

Several additional variables are retained for descriptive and robustness purposes. These include age bracket, gender, and online-shopping frequency, all of which are collected after the shopping task. These variables are not central to the causal model but are useful for describing the sample and assessing balance across experimental groups. Finally, attention-

check status is operationalised as a binary screening variable indicating pass

(q_terminate_flag = “empty”) or fail (q_terminate_flag = PoorQuality). Respondents who fail

this check are excluded before entering the main study. **Table 5** summarises the core

construct operationalisation.

Table 5: Operationalisation of the main constructs and variables

Construct / variable	Operational definition	Measurement / coding	Analytical role
Target-option choice (category level)	Whether the predesignated focal option is selected in a given product category	1 = target option chosen; 0 = non-target option chosen. E.g., Apples_Germany = 1 (target); Apples_Poland = 0	Primary dependent variable
Target-choice share (participant level)	Proportion of target options selected by a respondent across the shopping task	Mean proportion across seven categories; range 0-1	Primary respondent-level outcome
policy_arm	Participant-level assignment to the experimental condition	CONTROL / GENERIC / MATCHED	Primary independent variable
generic_nudge_type	Nudge type shown whenever generic assignment is used	SOCIAL / DEFAULT / NA	Secondary independent variable / treatment differentiation variable
gdms_type	Three-way classification derived from the reduced GDMS block	DEPENDENT, SPONTANEOUS, or MIXED	Matching variable / subgroup variable
Perceived autonomy	Subjective sense of freedom, control and preference-consistency during the shopping task	Five Likert-type items from ‘Strongly disagree’ to ‘Strongly agree’	Secondary dependent variable / perceptual outcome variable
Attention-check status	Whether the respondent passed the instructional attention check	Pass q_terminate_flag = ‘empty’; Fail q_terminate_flag = ‘PoorQuality’	Exclusion criterion / data-quality indicator
Age bracket	Respondent age category	Ordinal categorical variable	Descriptive / balance variable
Gender	Self-described gender identity	Nominal categorical variable	Descriptive / balance variable
Online-shopping frequency	Self-reported frequency of online shopping	Ordinal categorical variable	Descriptive / balance variable

Source: Author.

3.6 Data analysis

3.6.1 General analytical strategy

Data analysis was conducted in R on the basis of the final anonymised Qualtrics export. The analytical strategy followed the structure of the research design presented above. Since the experiment combines participant-level assignment with seven repeated binary choice observations within each respondent, the main behavioural analyses were conducted at the choice level rather than relying solely on respondent-level averages. This decision follows directly from the structure of the dataset and from the operationalisation of the primary dependent variable as target-option choice at category level. At the same time, because the methodology also defines target-choice share across the seven categories as a substantively meaningful respondent-level summary measure, additional aggregated analyses were conducted as robustness checks.

Accordingly, the data-analysis strategy consisted of five ordered stages. First, the raw export was cleaned and screened in accordance with the survey logic. Second, the routed task variables were reconstructed into unified behavioural outcome variables suitable for analysis. Third, descriptive statistics and group-balance checks were conducted. Fourth, the main behavioural hypotheses were tested using mixed-effects logistic regression models at the choice level. Fifth, supplementary analyses were estimated for respondent-level robustness, perceived autonomy, and model diagnostics.

The analysis was designed to respect the repeated-measures structure of the data. In substantive terms, the main inferential focus remained on the three behavioural hypotheses *H1*, *H1a*, and *H2*. The autonomy measure and additional sensitivity checks were treated as secondary and interpretive rather than as co-primary outcomes.

3.6.2 Data preparation, cleaning, and analytic sample definition

The raw dataset was first screened according to the procedural exclusion criteria built into the survey. Respondents who did not provide consent or who failed the instructional attention check were excluded by design before entering the main study. In addition, unfinished responses were removed from the final analytic sample. Thus, the retained dataset was restricted to respondents who had consented, passed the attention check, and completed the survey. This procedure reflects the data-quality safeguards described earlier and preserves the intended sequence of consent, screening, pre-treatment classification, assignment, behavioural exposure, and post-task evaluation.

Because the behavioural task was implemented through separate routed blocks, the raw export contained branch-specific task variables for the control, social-proof, and default versions of the seven product categories. Before analysis, these branch-specific variables were harmonised into a unified set of seven binary choice variables, one for each product category. This reconstruction was implemented by coalescing the mutually exclusive routed fields into common variables for the categories (i.e. apples, bananas, butter, eggs, tomatoes, milk, and cucumbers). The resulting unified variables preserved the original coding logic of ‘1 = target option chosen’ and ‘0 = non-target option chosen’. In this way, the final behavioural dataset no longer depended on the original survey branch structure but directly reflected the conceptual outcome defined in Section 3.5.

After reconstruction, a respondent-level summary measure of target-choice share was computed as the number of target options selected divided by seven. This measure ranges from 0 to 1 and was retained for descriptive presentation and robustness analyses. In addition, the routed task version actually shown to each respondent was derived from block occupancy and stored as a separate analytical variable ‘task_version’, distinguishing between

‘CONTROL’, ‘SOCIAL’, and ‘DEFAULT’ exposure. This variable was useful primarily for descriptive checks and for the secondary autonomy analysis.

Several internal consistency checks were conducted before estimation. First, the GDMS classification (Scott and Bruce, 1995) was reconstructed directly from the ten reduced GDMS items using the strict rule specified in Section 3.2.2, and the reconstructed grouping was compared with the stored ‘gdms_type’ variable. Second, the observed task block completion was checked against the theoretically expected routing implied by ‘policy_arm’, ‘gdms_type’, and ‘generic_nudge_type’. Third, the reconstructed unified choice variables were checked to ensure that exactly seven valid behavioural responses were retained for each completed respondent. These steps were intended to verify that the final analysis file was consistent with the implemented survey logic rather than relying uncritically on the raw export structure. On the basis of this preparation, three hypothesis-specific analytic samples were defined. The first comprised respondents assigned to ‘CONTROL’ or ‘GENERIC’ and was used for *H1*. The second comprised respondents in the ‘GENERIC’ experimental condition only and was used for *H1a*. The third comprised only respondents classified as ‘DEPENDENT’ or ‘SPONTANEOUS’ within the ‘GENERIC’ and ‘MATCHED’ experimental conditions and was used for *H2*. ‘MIXED’ respondents were therefore retained in the general descriptive analyses but excluded from the confirmatory *H2* test, in line with the hypothesis formulation.

3.6.3 Data structure, coding, and software implementation

Two related analysis datasets were created. The first was a wide respondent-level dataset containing one row per respondent and including the experimental assignment variables, GDMS classification, demographic variables, autonomy items, the seven unified behavioural choice variables, target count, and target-choice share. The second was a long choice-level dataset in which each respondent contributed seven rows, one for each product category. In

the long dataset, the dependent variable was ‘target_choice’, while category was represented as a categorical factor distinguishing apples, bananas, butter, eggs, tomatoes, milk, and cucumbers. This long format was used for the primary mixed-effects models because it directly represents the repeated binary-choice structure of the experiment.

Categorical variables were recoded as factors prior to modelling. ‘policy_arm’ was coded with the three levels ‘CONTROL’, ‘GENERIC’, and ‘MATCHED’. ‘gdms_type’ was coded as ‘DEPENDENT’, ‘SPONTANEOUS’, and ‘MIXED’. ‘generic_nudge_type’ was retained only where analytically relevant; values that reflected non-applicability rather than substantive treatment assignment were not interpreted as observed nudge exposure. Age bracket and online-shopping frequency were treated as ordered categorical variables for descriptive purposes, whereas gender was retained as a nominal factor. The five autonomy items were converted to numeric values in the same direction prior to reliability assessment and scale construction.

The analysis was implemented in R using packages for data import, transformation, mixed-effects modelling, estimated marginal means, diagnostics, and robust inference. In particular, the workflow relied on ‘haven’, ‘dplyr’, ‘tidyr’, ‘forcats’, ‘ggplot2’, ‘psych’, ‘lme4’, ‘emmeans’, ‘broom’, ‘broom.mixed’, ‘performance’, ‘sandwich’, and ‘lmtest’. The use of a scripted workflow ensured full data analysis reproducibility.

3.6.4 Descriptive statistics and preliminary checks

Descriptive analysis was conducted before hypothesis testing in order to characterise the sample, assess the realised group composition, and inspect the behavioural outcome distributions. First, a sample-flow table was produced to document the number of raw exported cases, exclusions due to attention-check failure or incompleteness, and the size of the

final analytic sample. Second, the retained sample was described in terms of age, gender, online-shopping frequency, GDMS group, and experimental assignment.

Third, descriptive balance checks were carried out across the randomised experimental conditions, and within the ‘GENERIC’ condition across the randomised nudge-type branches. For these checks, cross-tabulations and chi-square tests with simulated p-values were used for categorical variables. These tests were not interpreted as definitive evidence of successful randomisation in a formal sense, but rather as descriptive indicators of whether the realised groups differed materially in observable pre-treatment characteristics.

Fourth, behavioural descriptives were produced both at category level and at respondent level. At category level, raw target-choice rates were summarised by experimental policy condition and by displayed task version in order to inspect whether the seven product categories differed in their baseline ‘difficulty’ or target attractiveness. At respondent level, target-choice share was summarised within the experimental condition groups relevant to **H1**, within the generic ‘SOCIAL’ and ‘DEFAULT’ branches for **H1a**, and within the relevant **H2** subgroup cells defined by experimental policy condition and GDMS type. These descriptive summaries were intended to provide an interpretable overview of the data before turning to model-based inference.

Finally, for the autonomy items, reliability was first assessed and then descriptive statistics were calculated both for the individual items and for the resulting autonomy composite.

Because perceived autonomy was not one of the main behavioural outcomes, these descriptives served primarily to support the later secondary interpretation rather than to motivate a separate multivariate modelling framework.

3.6.5 Main behavioural hypothesis tests

The primary hypothesis tests were conducted using mixed-effects logistic regression at the choice level. This modelling choice follows from three features of the design. First, the primary behavioural outcome is binary, since each category records whether the target option was chosen or not. Second, each respondent contributes seven repeated observations, which implies non-independence within persons. Third, the substantive manipulations of interest operate at participant level, while category-specific baseline differences may remain even after careful task standardisation. Hence, a mixed-effects logistic model allows the analysis to preserve the binary structure of the outcome, account for clustering within respondents, and control for fixed differences across product categories. This is more appropriate than relying exclusively on respondent-level averages or e.g. on Analysis of Variance (ANOVA) type methods that assume a continuous dependent variable.

For all three main behavioural models, a random intercept for respondent was included in order to account for repeated choices within persons, and product category was included as a fixed effect in order to absorb baseline category heterogeneity. The models were estimated with a logit link and fitted by maximum likelihood using the ‘glmer()’ function.

For **H1**, the analysis sample was restricted to respondents in the ‘CONTROL’ and ‘GENERIC’ conditions. The dependent variable was the category-level binary ‘target_choice’, and the main predictor was ‘policy_arm’. The focal coefficient thus tested whether generically assigned nudges increased the odds of selecting the target option relative to neutral control presentation.

For **H1a**, the analysis sample was restricted to respondents in the ‘GENERIC’ experimental condition. Again, the dependent variable was ‘target_choice’, but the focal predictor was ‘generic_nudge_type’. Since this nudge type was randomly assigned at participant level

within the ‘GENERIC’ experimental condition, the model estimates the comparative effectiveness of the two generic nudge types under non-personalised assignment.

For **H2**, the analysis sample was restricted to respondents classified as ‘DEPENDENT’ or ‘SPONTANEOUS’ and belonging to the ‘GENERIC’ or ‘MATCHED’ experimental conditions. The dependent variable again remained ‘target_choice’. The focal predictor was ‘policy_arm’, while ‘gdms_type’ was entered as an additional fixed effect. This specification reflects the fact that **H2** concerns a restricted matchable subsample and compares a matched experimental condition with a non-personalised generic experimental condition rather than testing the psychological subgroup variable itself as a treatment. The inclusion of ‘gdms_type’ hence served to account for subgroup composition within this restricted sample while preserving the confirmatory focus on the experimental condition contrast.

Model results were interpreted primarily through odds ratios, 95% confidence intervals, and p-values for the focal coefficients. In addition, estimated marginal means were computed on the response scale in order to present model-implied predicted probabilities of target-option choice for the key conditions of the hypothesis. These predicted probabilities were used to facilitate substantive interpretation beyond log-odds and odds ratios.

3.6.6 Supplementary and robustness analyses

To assess whether the substantive conclusions depended heavily on one specific level of aggregation, the main mixed-effects models were complemented by respondent-level robustness analyses. Two additional modelling strategies were used. First, aggregated binomial generalised linear models were estimated using the number of target choices out of seven as the dependent response. Second, Ordinary Least Squares models (OLS) were estimated using ‘target_choice_share’ as a continuous respondent-level outcome. These models do not replace the primary mixed-effects specification, but they provide a useful

check of whether the same directional conclusions emerge once the repeated choice observations are collapsed into respondent-level summaries. The latter OLS models were reported with HC3 heteroskedasticity-robust standard errors.

Moreover, the three main behavioural mixed-effects models were re-estimated with demographic control covariates. These adjusted specifications included age group, gender, online-shopping frequency, product category fixed effects, and respondent random intercepts. For the adjusted models, age was grouped into three categories: '18-24', '25-34', and '35+'. This grouping retained the two largest and substantively most interpretable younger age groups while combining the older categories, which contained comparatively few observations, into a single group. Online-shopping frequency was similarly collapsed into three categories: 'Never/Once a month', 'Several times a month', and 'Weekly or more'. This reduced sparse-cell problems and produced a parsimonious control specification appropriate for the modest sample size. Although age and shopping frequency were treated as ordered categorical variables for descriptive purposes, they were entered as nominal control variables in the adjusted regression models to avoid imposing a linear ordinal trend on covariates that were not substantively interpreted. Additionally, the single respondent who selected 'Other / prefer not to say' for gender was excluded from these adjusted specifications only, in order to avoid estimating an unstable one-case control category. To summarise, the demographic-adjusted models were used as robustness checks rather than as replacements for the primary models, since the main hypothesis tests were defined by the randomised experimental contrasts.

In addition, a 'leave-one-category-out' sensitivity analysis was conducted for all three behavioural hypotheses. In this procedure, the main mixed-effects model for each hypothesis was re-estimated seven times, each time omitting one product category. The purpose of this analysis was to examine whether any single category disproportionately drove the estimated

treatment effect. The sensitivity results were summarised through the range of resulting odds ratios and p-values and were treated as an additional robustness exercise.

As **H2** concerns a theoretically structured matching mechanism, two further supplementary analyses were implemented for that hypothesis only. First, an interaction model between ‘policy_arm’ and ‘gdms_type’ was estimated within the restricted **H2** sample to explore whether the relative advantage of matched over generic assignment differed between ‘DEPENDENT’ and ‘SPONTANEOUS’ respondents. Second, a congruence-based model was estimated using a derived binary variable indicating whether the actually displayed nudge version was congruent with the respondent’s pre-treatment GDMS classification. These models should be interpreted as secondary and exploratory.

Finally, a post-hoc sensitivity analysis was conducted for **H2** in order to assess whether the non-significant matched-condition effect may plausibly reflect limited statistical power. This analysis used the observed **H2** coefficient and standard error from the baseline and demographic-adjusted **H2** models to approximate the power available to detect the observed effect and to estimate the odds-ratio threshold required for conventional significance and for approximately 80% power. This analysis was treated as interpretive and diagnostic rather than as a confirmatory hypothesis test.

3.6.7 Secondary autonomy analysis

Perceived autonomy was analysed as a secondary interpretive outcome. First, the internal consistency of the five autonomy items was assessed using Cronbach’s alpha. Provided that the reliability was acceptable, a composite autonomy score was constructed as the arithmetic mean of the five items. This scale-level approach was preferred over separate item-by-item inferential testing, since the autonomy construct was conceptually intended to capture a unified post-task evaluation of agency, freedom, and decisional control.

Descriptively, the autonomy composite was summarised both by displayed task version and by experimental condition. Inferentially, the main secondary autonomy analysis used a linear model in which the autonomy composite was regressed on the displayed task version. This grouping variable was preferred because respondents experienced the task directly as ‘CONTROL’, ‘SOCIAL’, or ‘DEFAULT’, whereas the more abstract distinction between ‘GENERIC’ and ‘MATCHED’ logic was not necessarily salient to them as a phenomenological experience. As with the respondent-level behavioural OLS models, robust standard errors were used for the linear autonomy model. Estimated marginal means and pairwise contrasts were also produced in order to aid interpretation.

3.6.8 Diagnostic checks and reporting conventions

Several standard diagnostic checks were applied to the main mixed-effects models. First, each model was inspected for singular fit, since an overparameterised random-effects structure may produce unstable estimates. Second, Pearson-based dispersion statistics were computed in order to assess whether problematic overdispersion was present. Third, information criteria and log-likelihood values were retained as general indicators of model fit. These diagnostics were not used to compare the three substantive hypothesis models against one another, because the models were estimated on different analytic samples and answer different research questions. Rather, they were used to verify that each main model was statistically appropriate within its own specification.

To summarise, throughout the analysis, statistical inference was based on a two-sided significance level of 0.05. The main behavioural results were reported through odds ratios with 95% confidence intervals and associated p-values. Respondent-level OLS results were reported through coefficient estimates, robust confidence intervals, and p-values. Descriptive

tables were presented in counts, percentages, means, and standard deviations as appropriate to variable type. Figures were used selectively, primarily to display model-estimated predicted probabilities for the main hypotheses and, secondarily, the distribution of perceived autonomy across task versions. The mixed-effects logistic models constituted the principal inferential core of the study, while respondent-level models, *H2*-specific supplementary specifications, autonomy analysis, and sensitivity checks served to contextualise and verify the robustness of the main findings.

3.7 Ethical considerations

The study was designed and implemented in accordance with basic ethical principles applicable to low-risk online behavioural research. Participation was voluntary and began only after respondents had viewed the consent page and actively agreed to proceed. Those who did not consent were automatically routed out of the survey and were not exposed to any subsequent study materials. In the same way, respondents who failed the instructional attention check were excluded before entering the main study, which served both data-quality and procedural-consistency purposes.

From an ethical perspective, the study may be classified as low risk. The experimental task consisted of simulated grocery choices in an e-commerce-like interface and did not expose participants to any kind of physical risk, financial cost, or psychologically invasive content. The questionnaire collected only limited respondent information, namely decision-style responses, basic demographic descriptors, and post-task evaluations of perceived autonomy. No special-category personal data were intentionally collected. Certain technical metadata, such as Internet Protocol (IP) address and approximate geolocation fields, were generated by Qualtrics by default but were not required for the substantive analysis and were removed from the final analytical dataset. The final analytical dataset was processed in anonymised form

(i.e., excluding the technical metadata such as IP addresses and approximate geolocation fields).

The temporal ordering of the survey was also ethically relevant. Behavioural responses were collected before the post-task autonomy evaluation, which reduced the risk that respondents would feel pressured to justify or rationalise their choices while still performing the task.

Moreover, although the experiment manipulated the digital presentation of product options through neutral, social-proof, or default cues, it did not remove alternatives or impose economic consequences on participants. The interventions therefore remained within the scope of legitimate behavioural influence rather than coercion.

The study further sought to balance methodological control with participant fairness. All respondents were presented with a limited number of clearly structured product-comparison tasks, and all required questions were kept directly relevant to the research purpose. Because the survey was self-administered and asynchronous, participants retained the practical freedom to discontinue participation at any time by leaving the survey environment, even though the questionnaire itself was configured without item skipping and without back-navigation in order to preserve experimental integrity.

3.8 Reliability and validity

3.8.1 Reliability

Reliability in the present study concerns the degree to which the implemented procedures and measures are capable of producing stable, internally coherent, and analytically interpretable observations. At this stage, the purpose of the subsection is to clarify the dimensions along which reliability is evaluated in the study. The corresponding empirical results are presented in the subsequent results chapter and in the relevant appendices.

At the procedural level, reliability was strengthened by the use of a fully standardised Qualtrics survey flow. All respondents encountered the same sequence of stages, the same GDMS classification logic, the same assignment architecture within each experimental condition, and the same fixed set of seven product categories within the routed task version. This reduced the scope for interviewer effects, manual administration error, or variation in treatment delivery across participants.

The behavioural task itself was also designed with reliability in mind. Each respondent completed seven structurally parallel binary choices in a controlled grocery-like environment. Across categories, the core comparison logic remained constant: two closely substitutable options were presented, and one predesignated option was treated as the focal target under neutral, social-proof, or default presentation. This repeated-task structure increases the reliability of the behavioural measurement relative to a design based on only a single choice, as the respondent-level target-choice share draws on multiple observations rather than on one isolated decision.

With regard to the questionnaire measures, the reduced GDMS block followed a uniform response format and a deliberately strict classification rule. This strict rule may have reduced the number of respondents classified as clearly ‘DEPENDENT’ or ‘SPONTANEOUS’, yet it increased the reliability of the resulting subgroup assignment by ensuring that only unambiguous response patterns enter those categories. In this sense, reliability for the matching variable is understood primarily as consistency of classification under the implemented rule.

For the autonomy measure, reliability is conceptualised as internal consistency across the five post-task items. As the items were intended to capture related aspects of experienced freedom, control, and preference-consistency, the scale was treated as reliable only insofar as the item set demonstrated acceptable internal coherence. Accordingly, the empirical reliability

assessment of the autonomy measure was conducted during the analysis phase before constructing the composite score, and the corresponding results are reported in the following chapter.

At the same time, some limitations should be acknowledged. First, the study used a reduced ten-item GDMS-based instrument rather than the full original scale (Scott and Bruce, 1995), which may limit psychometric breadth. Second, the experiment was constrained to seven product categories due to practical limitations of the survey environment. While this still permits repeated behavioural observation, it inevitably restricts observational depth.

Reliability in the present study should therefore be understood as adequate for the constrained experimental thesis design rather than as equivalent to a large-scale, completely validated panel instrument.

3.8.2 Validity

Validity concerns the extent to which the implemented design, measures, and inferences correspond to the theoretical constructs and causal claims of the study. In the present thesis, the most relevant dimensions are construct validity, internal validity, and external validity. As with reliability, this subsection sets out the conceptual basis on which these forms of validity are assessed, whereas the empirical results of the corresponding checks are reported in the subsequent results chapter and the appendices.

3.8.2.1 Construct validity

Construct validity was pursued by aligning the implemented variables closely with the conceptual logic developed in the literature review. The primary behavioural construct of interest is the effectiveness of digital nudges in shifting product choice. This was translated into a directly observable measure: whether the predesignated target option was chosen in

each product category. Since the same focal option was retained across control, social-proof, and default environments, differences in target-option choice can be interpreted as differences in response to the interface manipulation rather than as changes in the underlying product pair itself.

The operationalisation of personalisation also reflects construct-validity considerations. The study defines it concretely as a pre-specified matching rule between a respondent's GDMS-based decision-style classification (Scott and Bruce, 1995) and the nudge type shown. In this sense, the matched condition does not claim to represent full-scale adaptive personalisation; instead, it captures a narrower and explicitly rule-based form of psychological matching. This improves conceptual precision by ensuring that the empirical test corresponds closely to the specific theoretical claim under investigation.

The autonomy measure likewise serves as a construct-validity bridge between theory and observation. Rather than inferring autonomy indirectly from behaviour, the study includes a dedicated post-task self-report measure intended to capture respondents' own perception of freedom and decisional control. Although subjective autonomy is not identical to objective behavioural freedom, the inclusion of this scale strengthens the interpretive validity of the study by allowing perceived agency to be examined directly. The empirical assessment of the coherence of this measure is reported later in the results chapter.

Nevertheless, construct validity is not without limits. The reduced GDMS instrument (Scott and Bruce, 1995) provides a simplified classification of dependent and spontaneous tendencies and may not capture the full multidimensional richness of decision-making style. Similarly, the simulated grocery interface represents a stylised rather than exhaustive model of digital consumer choice. The study therefore operationalises the focal constructs in a deliberately parsimonious and testable form rather than claiming exhaustive measurement of all underlying psychological dimensions.

3.8.2.2 Internal validity

Internal validity refers to the credibility of the causal inference that observed differences in target-option choice are attributable to the experimental manipulations rather than to confounding influences. In the present study, internal validity is supported primarily by the experimental architecture. ‘policy_arm’ assignment was randomised at participant level, and within the ‘GENERIC’ condition the assignment of ‘SOCIAL’ vs. ‘DEFAULT’ was also randomised. This substantially reduces systematic selection bias in the *H1* and *H1a* comparisons.

Further internal-validity safeguards arise from the sequencing of the survey. GDMS classification occurred before any treatment exposure, thereby ensuring that the subgroup variable used for matching was not itself influenced by the nudge manipulation. Behavioural choices were recorded before demographics and before the post-task autonomy block, reducing the risk that later reflective questions contaminated task behaviour. The survey was also configured without back-navigation, which limited participants’ ability to revise earlier choices after exposure to subsequent questionnaire sections.

The design of the product pairs additionally supports internal validity. Within each category, price and pack size were held constant, and the alternatives were constructed to be as substitutable as possible. This limits the possibility that treatment effects are spuriously driven by objectively superior product attributes rather than by the interface cue attached to the target option. In this respect, the experiment isolates a choice-architecture effect under controlled comparative conditions.

In empirical terms, internal validity is further examined through data-consistency checks, balance checks across relevant experimental groups, and diagnostic assessment of the main statistical models. These tests are presented in the following results chapter and the corresponding appendices. Their methodological role is to evaluate whether the realised data

structure is consistent with the intended experimental logic and whether the main inferential models are statistically appropriate for the data.

However, internal validity may be stronger for some hypotheses than for others. *H1* and *H1a* rest on direct randomised comparisons and therefore admit the clearest causal interpretation.

H2 is more constrained, as it depends on both random experimental condition assignment and pre-treatment subgroup restriction to respondents classified as ‘DEPENDENT’ or

‘SPONTANEOUS’. The *H2* comparison remains meaningful, yet it is analytically narrower and more sensitive to limited subgroup size. Thus, the internal validity of the ‘MATCHED’ vs. ‘GENERIC’ comparison should be regarded as more conditional than that of the

‘GENERIC’ vs. ‘CONTROL’ and ‘SOCIAL’ vs. ‘DEFAULT’ contrasts.

3.8.2.3 External validity

External validity concerns the degree to which the findings may generalise beyond the implemented study setting. In this regard, the present thesis is intentionally modest. The sample is a non-probability convenience sample recruited primarily through the author’s personal and university-related networks. The resulting respondents are therefore not intended to represent the broader population of online grocery consumers in a statistical sense.

In addition, the experimental environment is a simplified grocery-like simulation rather than a live commercial platform. Real-world e-commerce behaviour may be shaped by additional factors not captured in the present study, such as broader assortment size, brand familiarity, budget constraints, habitual purchasing, platform trust, and repeated exposure over time. The behavioural effects observed in this study should therefore be interpreted as evidence of how digital nudges function under controlled online-choice conditions rather than as direct estimates of marketplace-level effects.

However, external validity is not absent. Grocery products were chosen precisely because they are familiar, low-complexity categories in which online interface nudges can plausibly influence choice under time-efficient decision-making. Hence, the findings may plausibly generalise to comparable digital retail contexts characterised by low-stakes, substitutable product choices and simple interface-level nudges, while broader generalisation should remain cautious.

3.8.3 Overall assessment of methodological quality

Taken together, the study demonstrates an analytically coherent and methodologically transparent design for the purposes of the present thesis. Reliability is approached through standardised administration, repeated behavioural observation, and scale-level assessment of the post-task autonomy measure. Validity is approached through close alignment between theory and operationalisation, through an experimental architecture intended to support causal inference, and through a deliberately cautious understanding of the limits of generalisability. At the methodological level, the study hence offers a defensible basis for the subsequent empirical analysis.

At the empirical level, the adequacy of this design is examined in the following results chapter through the relevant reliability assessments, data-consistency checks, balance checks, model diagnostics, and robustness analyses. Thus, the present subsection should be read as the conceptual and evaluative framework for methodological quality, while the actual diagnostic outcomes are reported alongside the results to which they directly pertain.

Overall, the findings of the thesis should be interpreted neither as psychometric absolutes nor as broad population estimates, but as theoretically informed experimental evidence generated under controlled, transparent, and analytically defensible conditions.

3.9 Chapter summary

This chapter set out the methodological framework of the present thesis and translated the conceptual argument of the study into an executable empirical design. It demonstrated that the research adopts a quantitative, deductive, and causal logic based on an online randomised experiment implemented in a simulated grocery e-commerce environment. In doing so, the chapter specified the experimental architecture and the sequence through which behavioural, demographic, and post-task autonomy data were collected. Furthermore, the chapter defined the research questions, hypotheses, and operationalisation of the core constructs and also explained how the anonymised dataset was prepared and analysed in R. Finally, the chapter addressed the ethical and methodological quality of the study and clarified the strengths and limitations in terms of reliability and validity. Thus, the methodology chapter provides a foundation for the presentation of the empirical results in the following chapter.

Chapter 4: Results

4.1 Introduction

This chapter presents the empirical findings of the experiment in the same sequence as specified in the methodology. It begins with the final analytic sample and preliminary balance checks, then reports descriptive behavioural patterns, the three confirmatory hypothesis tests, supplementary and robustness analyses, and finally the secondary analysis of perceived autonomy. The primary inferential focus remains on the mixed-effects logistic models at the choice level, with respondent-level models and sensitivity analyses used to assess robustness.

4.2 Analytic sample and preliminary checks

The raw export contained 198 cases. After excluding 6 respondents who failed the attention check and 9 unfinished responses, the final analytic sample comprised 183 completed cases, equivalent to 92.4% of the raw export. In line with the pre-specified analysis plan, the *H1* sample consisted of 121 respondents and 847 choice observations, the *H1a* sample of 62 respondents and 434 choice observations, and the *H2* sample of 60 respondents and 420 choice observations. The internal data checks further indicated that the reconstructed GDMS grouping, and the survey routing logic were fully consistent with the stored assignment structure, with no detected GDMS or routing mismatches (see Appendix 7 for the complete internal data checks). **Table 6** summarises the construction of the final analytic sample and the hypothesis-specific respondent- and choice-level subsamples.

Table 6: The final analytic sample

Stage / analytic subset	N
Raw exported cases	198
Excluded: no consent	0
Excluded: attention-check failure	6
Excluded: unfinished responses	9
Final analytic sample	183
<i>H1</i> respondent sample (CONTROL + GENERIC)	121
<i>H1</i> choice rows	847
<i>H1a</i> respondent sample (GENERIC only)	62
<i>H1a</i> choice rows	434
<i>H2</i> respondent sample (GENERIC + MATCHED; DEPENDENT/SPONTANEOUS only)	60
<i>H2</i> choice rows	420

Source: Author.

In descriptive terms, the retained sample was relatively young and moderately frequent in online shopping (i.e., mostly familiar with online shopping). The largest age group was 25–34 years old (43.7%), followed by 18–24 years old (29.0%). Men constituted 54.6% of the sample, women 44.8%, and one respondent selected “other / prefer not to say”. With respect to shopping frequency, the most common categories were “once a month” (38.3%) and “several times a month” (31.1%). Under the strict GDMS rule, 25.7% of respondents were classified as ‘DEPENDENT’, 23.5% as ‘SPONTANEOUS’, and 50.8% as ‘MIXED’ (see Appendices 8 - 10 for the full sample descriptions).

The randomisation checks were broadly satisfactory, though not perfectly balanced on every observable variable. Across the three experimental conditions, no statistically significant differences were observed for gender, shopping frequency, or GDMS type, whereas age distribution differed significantly across experimental conditions, with a simulated chi-square p-value of .018. Within the ‘GENERIC’ experimental condition, however, the ‘SOCIAL’ and ‘DEFAULT’ branches did not differ significantly on age, gender, shopping frequency, or

GDMS type (see Appendices 11 - 12 for the complete balance tests). Given the modest sample size, these checks are best interpreted as descriptive evidence of realised group composition rather than as a formal test of randomisation quality.

4.3 Descriptive behavioural results

At respondent level, the average target-choice share increased monotonically across the three experimental conditions. In the ‘CONTROL’, respondents selected the target option in 50.6% of choices on average. This rose to 63.1% in the ‘GENERIC’ condition and to 71.4% in the ‘MATCHED’ condition (see **Table 7**). The descriptive pattern therefore already suggested that nudged environments shifted behaviour towards the target option, and that the ‘MATCHED’ experimental condition produced the highest raw average target-choice share.

Table 7: Descriptive results

			n	mean	sd	median	iqr	min	max
By experimental condition	CONTROL		59	0.51	0.22	0.43	0.29	0.0	0.86
	GENERIC		62	0.63	0.24	0.71	0.14	0.0	1.0
	MATCHED		62	0.71	0.22	0.71	0.29	0.14	1.0
GENERIC by nudge type	DEFAULT		31	0.72	0.19	0.71	0.29	0.29	1.0
	SOCIAL		31	0.54	0.25	0.57	0.36	0.0	1.0
H2 by experimental condition and GDMS type	GENERIC	DEPENDENT	12	0.76	0.18	0.71	0.32	0.57	1.0
	GENERIC	SPONTANEOUS	17	0.58	0.21	0.71	0.14	0.0	0.71
	MATCHED	DEPENDENT	18	0.71	0.17	0.71	0.14	0.44	1.0
	MATCHED	SPONTANEOUS	13	0.78	0.14	0.71	0.14	0.57	1.0

Source: Author.

Within the ‘GENERIC’ experimental condition, the descriptive contrast between the two nudge types was pronounced. Respondents exposed to the ‘DEFAULT’ branch showed an average target-choice share of 71.9%, whereas those exposed to the ‘SOCIAL’ branch showed an average of 54.4% (see **Table 7**). This raw pattern already ran counter to **H1a**, which had predicted stronger performance for social proof than for default.

The **H2** subsample descriptives were more nuanced. Among ‘DEPENDENT’ respondents, the mean target-choice share was 76.2% in the ‘GENERIC’ condition and 70.6% in the ‘MATCHED’ condition. Among ‘SPONTANEOUS’ respondents, by contrast, the mean was 58.0% in the ‘GENERIC’ experimental condition and 78.0% in the ‘MATCHED’ experimental condition (see **Table 7**). Descriptively, therefore, the overall advantage of the ‘MATCHED’ condition appeared to be driven primarily by the ‘SPONTANEOUS’ subgroup rather than by a uniform effect across both theoretically matchable decision-style groups.

At category level, the same broad pattern was visible. In the ‘CONTROL’ experimental condition, target-choice rates ranged from 30.5% for eggs to 64.4% for apples and butter. In the ‘GENERIC’ condition, they ranged from 56.5% to 69.4%. In the ‘MATCHED’ condition, they ranged from 59.7% to 79.0%. The highest raw rates were observed for butter, tomatoes, and cucumbers under ‘MATCHED’, whereas eggs remained the most “difficult” category across all three experimental conditions alone (see Appendices 13 - 14 for the full description). Even so, the descriptive profile did not suggest that the overall treatment pattern was driven by a single product category.

4.4 Main behavioural hypothesis tests

4.4.1 H1: Generic nudging versus neutral control

H1 predicted that respondents in the ‘GENERIC’ experimental condition would be more likely to select the target option than respondents in the ‘CONTROL’. The mixed-effects logistic regression supported this prediction. The coefficient for ‘policy_armGENERIC’ was positive and statistically significant, $b = 0.579$, $SE = 0.188$, $z = 3.071$, $p = .002$ (see Appendix 15 for model summary). In odds-ratio terms, generic nudging increased the odds of choosing

the target option by a factor of 1.78 relative to the neutral control condition, with a 95% confidence interval from 1.23 to 2.58 (see **Table 8**).

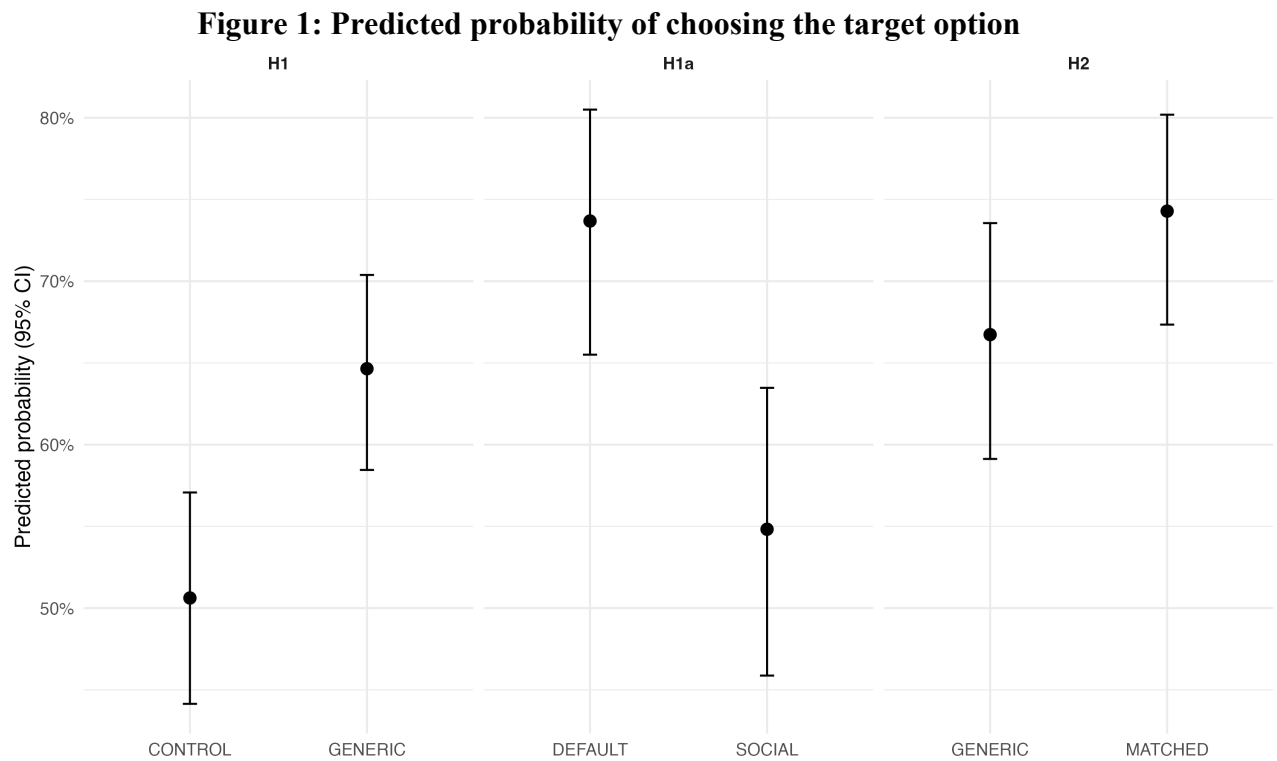
Table 8: Main hypothesis testing models

Hypothesis	Contrast	b	SE	OR	95% CI for OR	z	p- value	Outcome level	N	Support?
H1	GENERIC vs CONTROL	0.579	0.188	1.784	1.233 to 2.581	3.071	0.002	Choice level (mixed-effects logistic model)	121	Yes
H1a	SOCIAL vs DEFAULT	-0.836	0.269	0.433	0.256 to 0.734	-3.110	0.002	Choice level (mixed-effects logistic model)	62	No, opposite direction
H2	MATCHED vs GENERIC	0.365	0.237	1.440	0.905 to 2.292	1.539	0.124	Choice level (mixed-effects logistic model)	60	No

Source: Author.

The model-implied predicted probabilities help to interpret the magnitude of this effect.

Holding product category constant at the model level, the predicted probability of selecting the target option was 50.6% in ‘CONTROL’ and 64.6% in ‘GENERIC’. Thus, the introduction of a generic interface cue was associated with an increase of roughly 14 percentage points in the probability of choosing the target option (27.7% relative probability change). This pattern is also consistent with the respondent-level descriptive means reported above and visually corresponds to the pattern shown in **Figure 1**.



Source: Author.

Taken together, the evidence supports **H1**. In this experimental grocery environment, generic digital nudging significantly increased target-option choice relative to neutral presentation.

4.4.2 **H1a: Social proof vs. default within the ‘GENERIC’ experimental condition**

H1a predicted that, within the ‘GENERIC’ experimental condition, the ‘SOCIAL’ branch would outperform the ‘DEFAULT’ branch. The mixed-effects model showed the opposite.

Using ‘DEFAULT’ as the reference category, the coefficient for

‘generic_nudge_type_analysisSOCIAL’ was negative and statistically significant, $b = -0.836$, $SE = 0.269$, $z = -3.110$, $p = .0019$ (see Appendix 16 for model summary). The corresponding odds ratio was 0.43, with a 95% confidence interval from 0.26 to 0.73 (see **Table 8**). In substantive terms, exposure to the ‘SOCIAL’ cue reduced the odds of selecting the target option relative to exposure to the ‘DEFAULT’ cue.

The estimated probabilities again clarify the direction of the result. The predicted probability of choosing the target option was 73.7% under ‘DEFAULT’ but only 54.8% under ‘SOCIAL’. The gap was therefore close to 19 percentage points in favour of the default nudge (see **Figure 1**). This confirms that, in the present implementation, default preselection was materially more effective than the social-proof cue under non-personalised assignment.

Accordingly, *H1a* was not supported. More precisely, the observed effect ran in the opposite direction to the hypothesis: ‘DEFAULT’ substantially outperformed ‘SOCIAL’ in increasing target-option choice.

4.4.3 H2: Matched vs. generic nudging among ‘DEPENDENT’ and ‘SPONTANEOUS’ respondents

H2 proposed that, among respondents classified as ‘DEPENDENT’ or ‘SPONTANEOUS’, the ‘MATCHED’ experimental condition would outperform the ‘GENERIC’ experimental condition. The confirmatory mixed-effects model did not provide sufficient evidence for this claim at the conventional .05 level. The coefficient for ‘policy_armMATCHED’ was positive, but not statistically significant, $b = 0.365$, $SE = 0.237$, $z = 1.539$, $p = .124$ (see Appendix 17 for model summary). The corresponding odds ratio was 1.44, with a 95% confidence interval from 0.90 to 2.29 (see **Table 8**).

The estimated marginal probabilities nevertheless pointed in the theoretically expected direction. The predicted probability of choosing the target option was 66.7% in the ‘GENERIC’ experimental condition and 74.3% in the ‘MATCHED’ experimental condition within the restricted *H2* sample (see **Figure 1**). Thus, the point estimate suggested a positive behavioural advantage for matching, but the uncertainty around that estimate remained too large for a confirmatory inference.

Therefore, **H2** was not supported in confirmatory terms. The ‘MATCHED’ experimental condition showed a positive but statistically non-significant average effect relative to the ‘GENERIC’ one.

4.5 Supplementary and robustness analyses

The respondent-level robustness models were highly consistent with the choice-level mixed-effects results. For **H1**, the aggregated binomial model yielded a significant positive coefficient for ‘GENERIC’, $p < .001$, and the share-level OLS model likewise showed a significant positive increase in target-choice share of 0.125. For **H1a**, both the aggregated binomial model and the share-level OLS model again indicated a significant disadvantage for ‘SOCIAL’ relative to ‘DEFAULT’, with both p-values below .01. For **H2**, by contrast, neither the respondent-level binomial model nor the share-level OLS model produced a statistically significant advantage for ‘MATCHED’ over ‘GENERIC’ (see Appendix 18 for the complete description).

The demographic-adjusted models produced the same substantive conclusions as the baseline mixed-effects models. For **H1**, the generic nudge effect remained positive and statistically significant. For **H1a**, social proof remained significantly less effective than default preselection. For **H2**, the matched-condition coefficient increased in the adjusted model and approached conventional significance, but it still did not reach the .05 threshold. Thus, the adjusted specifications did not alter the formal hypothesis conclusions: **H1** remained supported, **H1a** remained unsupported, with the effect continuing to run in the opposite direction to the hypothesis, and **H2** remained unsupported in confirmatory terms (see Appendices 19 - 21 for the complete summaries of the adjusted models).

The post-hoc **H2** sensitivity analysis further clarified the interpretation of the non-significant nudge matching effect. In the baseline **H2** model, the observed odds ratio was 1.44, whereas an odds ratio of approximately 1.59 would have been required to reach conventional statistical significance and approximately 1.94 would have been required for 80% power (see Appendix 22). In the demographic-adjusted **H2** model, the observed odds ratio increased to 1.55 and came close to the estimated threshold for conventional significance yet still remained marginal. This indicates that the **H2** test had limited power to detect an effect of the observed magnitude. Consequently, the **H2** result should be interpreted as inconclusive rather than as strong evidence for the absence of a matching effect.

The ‘leave-one-category-out’ sensitivity analysis reinforced the same conclusion. For **H1**, the odds ratio for the ‘GENERIC’ effect remained positive in all seven reduced models and stayed statistically significant throughout, ranging from 1.60 to 2.01. For **H1a**, the odds ratio for ‘SOCIAL’ relative to ‘DEFAULT’ remained below 1 in all seven reduced models and stayed statistically significant throughout, ranging from 0.39 to 0.51. For **H2**, the estimated effect of ‘MATCHED’ remained positive in all reduced models, but none of those re-estimations reached statistical significance (see Appendices 23 - 24 for the complete description). The directional conclusions of the main models therefore did not depend on any single product category.

The exploratory **H2**-specific analyses provided additional nuance to the non-significant average **H2** result. In the interaction model, the interaction term between ‘MATCHED’ and ‘SPONTANEOUS’ was positive and statistically significant, $b = 1.266$, $p = .006$ (see Appendix 25 for the complete model), corresponding to an odds ratio of 3.55 and a 95% confidence interval from 1.43 to 8.78 (see Appendix 26). In other words, the effect of being in the ‘MATCHED’ condition rather than in the ‘GENERIC’ condition was significantly stronger for the ‘SPONTANEOUS’ respondents than for the ‘DEPENDENT’ respondents. In

this regard, the model-implied probabilities showed that among ‘DEPENDENT’ respondents, the predicted probability of choosing the target option was actually slightly lower in ‘MATCHED’ (71.1%) than in ‘GENERIC’ (76.8%). Among ‘SPONTANEOUS’ respondents, however, ‘MATCHED’ increased the predicted probability from 58.2% to 78.6% (see Appendix 27).

The congruence-based model led to a closely aligned interpretation. When the actually displayed nudge was congruent with the respondent’s pre-treatment GDMS classification, the odds of choosing the target option were significantly higher, $b = 0.65$, $SE = 0.265$, $z = 2.457$, $p = .014$ (see Appendix 28 for the complete model), corresponding to an odds ratio of 1.92 and a 95% confidence interval from 1.14 to 3.22 (see Appendix 26).

Taken together, these exploratory results suggest that the non-significant average **H2** result may have concealed subgroup-contingent heterogeneity. Matching did not produce a robust overall gain across both matchable groups combined, but it appears to have worked substantially better for the ‘SPONTANEOUS’ respondents than for the ‘DEPENDENT’ respondents. Since these analyses were pre-specified as supplementary rather than confirmatory, they should be interpreted as qualified evidence of subgroup-contingent fit rather than as a basis for revising the formal hypothesis outcome.

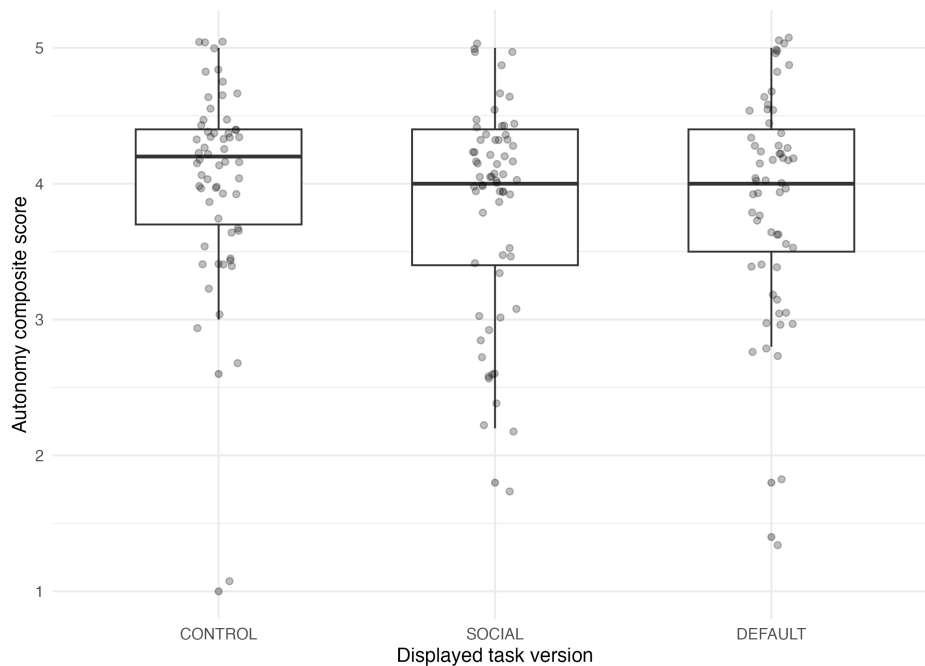
Finally, the main mixed-effects models showed no indication of statistical misspecification. None showed singular fit, and the Pearson-based dispersion ratios ranged from 0.89 to 0.98, providing no indication of problematic overdispersion (see Appendix 29).

4.6 Secondary results: perceived autonomy

The five-item autonomy scale demonstrated good internal consistency, with Cronbach’s alpha of 0.818. This justified the construction of a composite autonomy score for the secondary

analysis. Descriptively, perceived autonomy was relatively high across all displayed task versions. The mean autonomy score was 4.05 in the ‘CONTROL’ task version, 3.85 in the ‘SOCIAL’ task version, and 3.92 in the ‘DEFAULT’ task version, all on a five-point scale. The corresponding means by experimental condition were similarly close: 4.05 in ‘CONTROL’, 3.87 in ‘GENERIC’, and 3.90 in ‘MATCHED’ (see Appendices 30 - 32 for the complete description). As **Figure 2** indicates, although the nudged task versions showed slightly lower raw autonomy means than the control task version, the absolute differences were modest.

Figure 2: Perceived autonomy by task version



Source: Author.

Inferentially, the robust linear model did not indicate statistically significant differences in perceived autonomy across task versions. Relative to ‘CONTROL’, the ‘SOCIAL’ version showed a negative but non-significant coefficient of -0.20, while the ‘DEFAULT’ version showed a non-significant negative coefficient of -0.13 (see Appendix 33). Pairwise estimated

marginal mean contrasts were all non-significant as well, namely, ‘CONTROL’ vs. ‘SOCIAL’; ‘CONTROL’ vs. ‘DEFAULT’; ‘SOCIAL’ vs. ‘DEFAULT’ (see Appendix 34).

Thus, within this experiment, there was no statistically reliable evidence that either the ‘SOCIAL’ or the ‘DEFAULT’ task version reduced perceived autonomy relative to the neutral ‘CONTROL’ presentation. Any autonomy differences remained small and descriptive rather than inferentially meaningful.

4.7 Chapter summary

This chapter produced a clear hierarchy of findings. First, generic digital nudging significantly increased target-option choice relative to a neutral interface, thereby supporting **H1**. Second, the relative nudge comparison contradicted the initial directional expectation: default preselection significantly outperformed social proof, so **H1a** was reversed. Third, the matched condition showed a positive but statistically non-significant average advantage over generic assignment in the confirmatory **H2** model, meaning that **H2** was not supported at the aggregate level.

The supplementary analyses refined that picture. The robustness checks confirmed the **H1** and **H1a** results and did not rescue **H2** as a general confirmatory effect. The demographic-adjusted models further confirmed that the formal conclusions were robust to the inclusion of age, gender, and online-shopping frequency, while the post-hoc **H2** sensitivity analysis indicated that the non-significant matched-condition result should be interpreted as inconclusive rather than as strong evidence for the absence of a matching effect. However, the interaction and congruence analyses suggested that person-nudge fit may still matter in a more conditional sense, especially for ‘SPONTANEOUS’ respondents, for whom personalised assignment appeared substantially more effective than generic assignment.

Finally, the post-task autonomy analysis found no statistically significant reduction in perceived autonomy across displayed task versions. The overall implication is that nudging clearly affected behaviour, default cues were the strongest intervention in this design, and psychologically informed matching showed promise mainly as a subgroup-contingent rather than universally superior strategy.

Chapter 5: Discussion

5.1 Introduction

This chapter interprets the empirical findings of the thesis in light of the research questions, hypotheses, and theoretical framework developed previously. The study set out to examine whether digital nudges in a simulated grocery e-commerce environment increase target-option choice, whether social proof outperforms default preselection under generic assignment, and whether a simple GDMS-based matching rule (Scott and Bruce, 1995) yields greater behavioural effectiveness than non-personalised nudging. The results show a differentiated pattern. First, generic digital nudging significantly increased target-option choice relative to neutral presentation. Second, contrary to the initial directional expectation, default preselection significantly outperformed social proof within the generic nudge condition. Third, the matched assignment did not produce a statistically significant overall advantage over generic assignment in the confirmatory model, although supplementary analyses indicated meaningful subgroup heterogeneity, especially for spontaneously oriented respondents. Finally, no statistically reliable reduction in perceived autonomy was observed across the displayed task versions.

Taken together, these findings support the broader claim that even relatively light-touch interface interventions can shape consumer choice in digital retail environments, yet they also caution against treating all nudge types, or all forms of personalisation, as equally effective. In substantive terms, the study suggests that the behavioural power of digital nudging lies not only in whether a nudge is present, but in the specific mechanism through which it operates, the decisional context in which it is embedded, and the characteristics of the consumer exposed to it.

5.2 Interpretation of the main findings

5.2.1 Generic digital nudging increased target-option choice

The first main finding is that generically assigned nudges increased the probability of selecting the target option relative to the neutral control condition. This result supports *H1* and is consistent with the general proposition that online choice architectures can shift behaviour without altering the objective option set or economic incentives (Thaler and Sunstein, 2008; Weinmann, Schneider and vom Brocke, 2016). In the present experiment, the introduction of either a social-proof cue or a default cue increased target-option choice compared with a neutral presentation, indicating that interface-level steering remained behaviourally consequential even in a simplified two-option grocery task.

This finding fits closely with the core behavioural logic described in the literature review.

Under bounded rationality, consumers often do not undertake fully compensatory evaluation, especially in low-stakes and repetitive purchase-like settings. Instead, they rely on cues that reduce effort or simplify judgement (Simon, 1955; Kahneman, 2011). The present study's grocery-like micro-decisions were intentionally structured to reflect such conditions:

respondents faced repeated low-involvement choices among close substitutes, which is precisely the type of environment in which seemingly small design interventions may exert disproportionate influence. Thus, the result strengthens the thesis argument that digital nudging is a practically effective behavioural mechanism in controlled consumer choice environments (c.f., Weinmann, Schneider and vom Brocke, 2016; Valenčič et al., 2023).

However, the effect should not be overstated. The study demonstrates a positive average effect of generic nudging within this specific experimental configuration. However, this does not imply that all nudges are equally effective across all digital retail settings. Average nudge effects may be positive yet heterogeneous, and their magnitude depends on context, implementation, and user characteristics (c.f., Hummel and Maedche, 2019; Mertens et al.,

2022). The present study should therefore be read as providing controlled evidence that digital nudges can work in e-commerce-like settings, without establishing a universal behavioural law.

5.2.2 Default preselection outperformed social proof

The second and theoretically most notable finding is that default preselection significantly outperformed social proof under generic assignment, thereby contradicting *H1a*. Instead of the expected superiority of social proof, the default nudge generated markedly higher target-option choice.

This result is counter-expected, though plausible when interpreted through the mechanisms outlined in the literature review. Social proof operates primarily through informational inference: it provides consumers with a behavioural signal about others and may reduce uncertainty when product quality or appropriateness is difficult to assess (Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010). Default preselection, by contrast, changes the structure of action directly by making one option the path of least resistance (Johnson and Goldstein, 2003; Jachimowicz et al., 2019). In the study's shopping task, characterised by simple binary choices, low switching costs, and repeated selection among highly substitutable products, it is conceivable that effort minimisation and passive acceptance were behaviourally more potent than social information. This interpretation aligns with prior theory suggesting that defaults may be especially effective when the target action is simple, immediate, and low in effort (Nijssen et al., 2023; Ortmann et al., 2023).

The observed inferiority of social proof does not necessarily imply that social proof is weak in general. Rather, it suggests that in this specific implementation, the social-proof cue may have functioned as a weaker intervention than default preselection. This may also be because the cue used in the experiment was a relatively sparse popularity signal rather than a richer review

environment. As discussed in the literature review, different forms of social proof do not operate identically: ratings, reviews, bestseller labels, and “most chosen” indicators vary in diagnosticity and credibility, and their persuasiveness depends on whether consumers treat them as relevant and trustworthy signals (Ghiassaleh, Kocher and Czellar, 2020; Qiu and Zhang, 2024; Wang et al., 2024). A sparse cue used in the study ‘our customers’ choice’ may have provided less inferential value than detailed reviews or broader marketplace evidence could have done, particularly when the two options are otherwise closely matched. In such a setting, a default may simply impose a stronger immediate behavioural architecture than a socially framed label.

Accordingly, the reversal of *H1a* should be interpreted not as a failure of theory, but as an informative correction to a context-contingent expectation. The literature had already cautioned against universal rankings of nudge strength (Hummel and Maedche, 2019; Valenčič et al., 2023), and the present results empirically reinforce that warning. For grocery-like micro-decisions implemented through simple interface cues, default preselection appears to be the stronger intervention.

5.2.3 Personalisation did not produce a confirmatory overall advantage

The third main finding concerns *H2*. Although the matched condition showed a positive directional advantage over generic assignment, that advantage was not statistically significant in the confirmatory mixed-effects model. Strictly speaking, the study therefore does not provide confirmatory support for the statement that GDMS-based matching outperforms non-personalised nudging on average.

This non-significant result is important since it disciplines the personalisation narrative. The literature review positioned personalised digital nudging as promising albeit empirically uncertain. Evidence from adjacent domains suggests that matching can improve outcomes;

however, the existing literature also indicates that personalisation effects are not guaranteed and may vary substantially by domain and matching rule (c.f., Ingendahl et al., 2020; Peer et al., 2020; de Vries, Bol and van der Laan, 2025). Hence, the present study is broadly consistent with that mixed evidence base: personalisation showed promise, but not enough to justify a strong average-effect claim in this sample.

However, the confirmatory null should not be interpreted as evidence that person-nudge fit is irrelevant. The supplementary analyses suggest a more nuanced pattern. Matching appeared substantially more effective for respondents with spontaneous tendencies than for those with dependent tendencies; and congruent exposure was associated with significantly higher target-option choice. In predicted-probability terms, the personalised condition raised target-option choice markedly for spontaneous respondents, while no comparable gain appeared for dependent respondents.

This subgroup-contingent pattern is theoretically meaningful. The matched experimental condition assumed that dependent respondents would respond more strongly to social proof and spontaneous respondents more strongly to defaults. The latter mapping appears more consistent with the data than the former. One possible interpretation is that the spontaneous-default match captured a mechanism with relatively strong behavioural leverage: defaults reduce friction and support fast, low-effort action, which fits the idea of a spontaneous decision tendency. By contrast, the dependent-social mapping may have been weaker because the social-proof cue itself was comparatively less effective and/or insufficiently credible to have been able to activate the expected reliance on ‘others’ choice’. In other words, the issue may not have been the logic of matching as such, but the asymmetry in strength between the two nudges used in the matching architecture. If one matched option is intrinsically more potent than the other in the given context, then a simple matched-versus-generic comparison may dilute true fit effects.

A second explanation concerns statistical power and design dilution. The methodology explicitly notes that the final instrument was constrained by survey-length limits and should be understood as a constrained implementation rather than a fully powered confirmatory test. Moreover, the strict GDMS rule classified more than half of respondents as mixed, and those respondents were assigned to fallback generic routing within the ‘MATCHED’ experimental condition. While they were excluded from the formal **H2** test, the broader architecture still reflects the difficulty of operationalising trait-based personalisation with a highly restrictive segmentation rule and limited sample size. The confirmatory **H2** model was therefore estimated on only 60 respondents and 420 choice observations, which may have reduced precision for detecting effects that may have been real yet limited.

In this connection, the demographic-adjusted robustness model and the post-hoc sensitivity analysis reinforce this interpretation. After controlling for age, gender, and online-shopping frequency, the matched-condition estimate increased and approached conventional significance, although it remained marginal rather than statistically significant at the .05 level. Furthermore, the sensitivity analysis indicated that the restricted **H2** sample had limited approximate power to detect an effect of the magnitude observed. Thus, the non-significant **H2** result should not be interpreted as clear evidence that GDMS-based matching lacks behavioural value. Rather, the result is more appropriately understood as inconclusive at the aggregate level, with suggestive evidence that matching may matter under specific subgroup conditions.

Overall, the **H2** finding suggests that simple psychological matching should not be presented as uniformly superior to generic nudging. By contrast, the more precise conclusion is that person-nudge fit may matter conditionally, and that its behavioural value depends on the strength of the underlying cue, the quality of the segmentation rule, and the decision context in which the matching is applied.

5.2.4 No significant reduction in perceived autonomy

The final major finding is that neither social proof nor default preselection produced a statistically significant reduction in perceived autonomy relative to the control condition.

Although the nudged conditions showed slightly lower mean autonomy scores descriptively, these differences were small and non-significant.

This result is theoretically consequential as it tempers a common ethical concern in the nudging literature. Critics frequently argue that nudges, especially defaults, may be autonomy-undermining even when they preserve formal choice (Bruns and Perino, 2023; Michaelsen, 2024b). However, the present findings suggest that in a low-stakes e-commerce-like setting, light-touch reversible interface nudges can alter behaviour without generating a detectable decrease in respondents' experienced agency. This is consistent with recent works showing that not all defaults are experienced as strongly autonomy-threatening, particularly when opt-out remains easy and the intervention is not perceived as coercive (e.g., Michaelsen, 2024b; Weijers, Wachner and de Koning, 2024).

The result also corresponds to the manner the experiment was implemented. The interventions were transparent at the interface level, did not conceal alternatives, did not impose additional friction on the non-target option, and operated in a relatively low-risk consumer context.

These features place the manipulations within the domain of legitimate choice architecture rather than “dark patterns” or coercive design (c.f., Gray et al., 2018; OECD, 2022). In that sense, the autonomy finding supports the thesis's conceptual distinction between ethical nudges and manipulative commercial interfaces. Even though the finding does not prove that digital nudges are ethically unproblematic in general, it does suggest that behaviour change and perceived autonomy loss do not inevitably co-occur.

5.3 Theoretical implications

The study makes several theoretical contributions to the digital nudging literature. First, it provides controlled empirical support for the proposition that digital nudges can shape consumer choice in e-commerce-like settings. This reinforces information-systems and behavioural-economics claims that interface design is not merely a neutral presentation layer, but a causal component of decision-making environments (c.f., Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019).

Second, the findings refine theory on cross-nudge comparison. The present results indicate that defaults and social proof nudges are not interchangeable and that their relative effectiveness depends on context. In low-complexity binary grocery choices, action-structuring interventions may dominate information-based cues. This strengthens calls for mechanism-sensitive rather than label-based theorising about digital nudges (c.f., Münscher, Vetter and Scheuerle, 2016; Hummel and Maedche, 2019).

Third, the study contributes to the literature on personalised digital nudging by showing that average personalisation effects may conceal subgroup heterogeneity. The non-significant confirmatory **H2** result, combined with suggestive supplementary evidence for spontaneous respondents and congruent exposure, indicates that person-nudge fit should not be conceptualised as a monolithic treatment effect. Instead, the relevant theoretical unit may be the interaction between a specific trait configuration and a specific nudge mechanism. This perspective moves the discussion away from asking whether “personalisation works” in the abstract and towards asking under what conditions, for which users, and through which behavioural channels it functions.

Fourth, the study adds to the ethical literature by showing that behavioural effectiveness may not necessarily entail a measurable reduction in perceived autonomy. This supports a more differentiated understanding of legitimacy in digital nudging: choice-preserving interface

interventions may be behaviourally consequential while remaining phenomenologically acceptable, at least in low-stakes retail contexts and when implemented without deception or obstruction.

5.4 Managerial implications

The findings also carry practical implications for digital retailers, marketplace designers, and managers responsible for conversion optimisation. First, the results suggest that relatively simple interface nudges can produce meaningful behavioural shifts. Managers do not necessarily need complex algorithmic personalisation systems to influence choice in routine, low-involvement product categories. Even a generic nudge may outperform neutral presentation when the surrounding choice environment is appropriately controlled.

Second, where the goal is to steer consumers towards a focal option in simple binary or near-binary decisions, default preselection may be more effective than sparse social-proof cues.

From a managerial perspective, this means that defaults may be particularly useful in contexts such as delivery-slot selection, feature preselection, add-on recommendation, or low-stakes product comparison. However, this implication must be qualified by ethical boundaries. The present study examined a reversible and relatively transparent default, not a friction-laden or hidden opt-out. Managers should therefore avoid translating “defaults work” into the use of obstructive pre-ticks or disguised commercial steering. The practical lesson is to use defaults as convenience-oriented architecture, not as “dark-pattern” design.

Third, the personalisation results imply that simple rule-based matching should be approached cautiously. Firms should not assume that any psychological segmentation will automatically outperform generic nudging. Personalisation appears more promising when the segmentation rule is behaviourally meaningful, when the nudge mechanisms differ clearly, and when one

can validate subgroup-specific response patterns empirically. In practice, this favours iterative experimentation over premature deployment of psychologically targeted nudges at scale.

Fourth, the autonomy results offer a managerial governance insight. In low-stakes settings, truthful and reversible nudges may influence behaviour without materially damaging users' perceived sense of choice autonomy. This supports a design philosophy in which firms can pursue behavioural optimisation while preserving user agency, provided that interventions remain transparent, non-deceptive, and easy to override.

5.5 Limitations

Several limitations should be acknowledged. First, the study relied on a non-probability convenience sample recruited primarily through personal and university-related networks.

This limits population-level generalisability and means that the findings should not be interpreted as representative of online grocery consumers at large. The relatively young sample profile further narrows the external scope of the results.

Second, the experiment was conducted in a simulated grocery-like environment rather than on a live e-commerce platform, which strengthens internal control but limits ecological validity.

In particular, the shopping task relied on repeated binary choices between closely substitutable alternatives, with price, pack size, and visual quality held constant. This design was useful for isolating the behavioural effect of the interface cues, but it necessarily simplified the complexity of real-world online shopping. Actual e-commerce environments may typically involve broader assortments, stronger brand effects, price variation, product rankings, filters, reviews, budget constraints, platform trust, repeated exposure, habit formation, and basket-level trade-offs. Consequently, the observed effects should be interpreted as short-run causal responses under controlled binary-choice conditions rather than

as direct estimates of how the same nudges would perform in a full-scale online grocery platform.

Third, the study used only seven product categories and was explicitly constrained by survey-platform limitations. This reduced observational depth and likely limited statistical power, especially for *H2*. In this regard, the post-hoc sensitivity analysis confirmed that the restricted *H2* sample had limited power to detect the observed nudge matching effect. Consequently, the non-significant *H2* result should be interpreted cautiously, as a Type II error cannot be ruled out.

Fourth, the GDMS-based segmentation rule was deliberately strict. While this increased subgroup clarity, it also produced a large mixed category and reduced the size of the clearly matchable subgroups. It is therefore possible that the personalisation test underestimated the potential of matching by relying on a restrictive classification procedure rather than a more continuous or probabilistic assignment rule.

Fifth, the operationalisations of the focal nudges were necessarily narrow. The default cue was a preselection framed as ‘already in the cart’, whereas the social-proof cue was a relatively sparse statement of customer choice. These are defensible operationalisations, but they do not exhaust the broader categories of “default” or “social proof”. Hence, the findings should be interpreted as evidence about these specific implementations, not about all possible variants of the two nudge families.

Sixth, the autonomy analysis was secondary and post-task. Although the five-item scale showed good internal consistency, autonomy was measured after the behavioural task and in a relatively low-stakes context. This may have reduced sensitivity to subtle ethical discomfort or delayed perceptions of manipulation. The non-significant result therefore should not be read as closing the autonomy question definitively.

Seventh, there was an age imbalance across experimental conditions in the realised sample, even though most other balance checks were satisfactory. Demographic-adjusted models were therefore estimated as robustness checks and produced the same substantive conclusions as the baseline models. Nevertheless, the imbalance reinforces the need for cautious interpretation and replication with larger, better-balanced samples.

5.6 Future research

The present findings point to several priorities for future research. First, future studies should replicate the design with larger and more diverse samples. This is especially important for testing personalisation effects, which appear likely to be smaller and more heterogeneous than generic nudge effects. Particularly, future studies should also conduct an ex-ante power analysis and oversample clearly classifiable decision-style groups, since the present study suggests that personalisation effects may require substantially larger matchable subsamples than generic nudge effects.

Second, future research should compare richer forms of social proof with default preselection. For instance, review aggregates, verified-review counts, bestseller badges, or dynamically explained popularity signals may differ substantially from a sparse ‘customers’ choice’ label. Such work would help determine whether the present default advantage was driven by mechanism strength in general or by the specific operationalisation of social proof used in the present study.

Third, future studies should examine person-nudge fit using more refined segmentation strategies. Rather than forcing respondents into strict GDMS categories, researchers could use continuous GDMS scores to capture variation in decision-making tendencies more precisely. For example, dependent and spontaneous tendencies could be modelled as separate continuous predictors or used to create probabilistic matching rules instead of binary

subgroup assignments. Future studies could also combine GDMS scores with additional constructs such as need for cognition or need for uniqueness. This would allow more precise testing of whether matching improves outcomes, for whom, and under which psychological conditions.

Fourth, research should move beyond single-exposure experiments and examine repeated exposure over time. This is particularly relevant for questions of learning, habituation, persuasion knowledge, and long-term trust. A nudge that works immediately may not retain its effect once users recognise the steering attempt.

Fifth, future research should integrate transparency more directly into the experimental design. Since the present study did not manipulate disclosure, it cannot assess whether explicit explanation of the nudge or of the personalisation logic would alter either behavioural effectiveness or perceived autonomy. Given the ethical importance of transparency in the literature, this remains one of the key next steps.

Sixth, subsequent studies should test similar architectures in more realistic live-platform or field settings. In particular, one could also use more ecologically realistic shopping tasks with larger assortments, multi-attribute product comparisons, search and filtering functions, and basket-level outcomes rather than isolated binary choices. This would help bridge the gap between internal validity and external validity and would show whether the present effects may still exist under the noise and complexity of actual online retail behaviour.

5.7 Conclusion

This thesis set out to examine the effectiveness of personalised digital nudges in e-commerce and to compare generic nudging, cross-nudge performance, and simple psychological matching in a controlled online grocery environment. The discussion of the results leads to the following principal conclusions.

First, digital nudges clearly function: generically assigned interface cues increased target-option choice relative to neutral presentation. Second, default preselection was more effective than social proof in the present design, indicating that action-structuring cues may exert stronger influence than sparse informational cues in low-complexity purchase-like decisions. Third, the hypothesised overall superiority of matched nudging was not confirmed, but the supplementary evidence suggests that matching may still matter in subgroup-contingent ways, especially for spontaneous decision-makers. Fourth, the behavioural influence of the nudges was not accompanied by a statistically significant reduction in perceived autonomy, which suggests that behaviourally effective digital steering need not necessarily be experienced as autonomy-undermining when it remains choice-preserving and non-deceptive.

To summarise, the study contributes to the literature by showing that digital nudging in e-commerce should be understood neither as universally effective nor as uniformly problematic. Digital nudging effects should be viewed as substantial, mechanism-dependent, and context-sensitive. The broader implication is that the question is not simply whether digital nudges work, but which nudge works, for whom, under what conditions, and with what ethical consequences.

Chapter 6: Conclusion

6.1 Introduction

This chapter concludes the present thesis by drawing together the conceptual foundations, methodological approach, and empirical findings of the study. The thesis set out to examine the effectiveness of personalised digital nudges in e-commerce, with particular attention to three issues: whether digital nudges increase target-option choice relative to neutral presentation, whether social proof outperforms default preselection under generic assignment, and whether a simple GDMS-based matching rule produces higher behavioural effectiveness than non-personalised nudging. In this way, the study addressed a research gap at the intersection of digital nudging, e-commerce choice architecture, and behavioural personalisation.

6.2 Summary of the main findings

The study was based on the premise that e-commerce environments function as digitally constructed choice architectures that can shape consumer behaviour through interface-level cues (Münscher, Vetter and Scheuerle, 2016; Weinmann, Schneider and vom Brocke, 2016; Mertens et al., 2022). Building on research on bounded rationality and heuristic decision-making (Simon, 1955; Kahneman, 2011), digital nudging (Weinmann, Schneider and vom Brocke, 2016), and person-nudge fit in personalised influence contexts (Hirsh, Kang and Bodenhausen, 2012; Mills, 2022), the thesis developed the expectation that digital choice architectures may systematically influence purchase-like decisions and that some consumers may respond more favourably to specific nudge types than others. These propositions were tested through an online randomised experiment in a simulated grocery e-commerce environment consisting of three experimental conditions.

The empirical findings yield a differentiated picture. First, generic digital nudging significantly increased the probability of selecting the target option compared to neutral presentation (supporting **H1**). Second, the comparison between the two generic nudge types showed that default preselection significantly outperformed social proof, thereby contradicting the original directional expectation of the **H1a**. Third, the matched experimental condition did not significantly outperform the generic one in the confirmatory **H2** model. However, demographic-adjusted robustness checks and the post-hoc **H2** sensitivity analysis indicated that this non-significant result should be interpreted cautiously, while supplementary analyses suggested that matching may have been more effective for some respondents than for others, particularly for the spontaneous subgroup. Finally, although the nudged conditions showed slightly lower autonomy scores descriptively, these differences were small and statistically non-significant.

Overall, the findings suggest that digital nudges are capable of influencing online consumer choice, but that their effectiveness depends on the mechanism used and the context in which they are embedded. The results also indicate that the behavioural value of personalisation should be understood more cautiously and conditionally than strong personalisation claims may imply.

6.3 Concluding answers to the research questions

6.3.1 RQ1

The first research question asked whether exposure to a generically assigned digital nudge, as opposed to a neutral product presentation, may increase the likelihood of selecting the predesignated target option in a simulated grocery e-commerce environment. The findings provide a clear affirmative answer. Respondents in the ‘GENERIC’ condition were significantly more likely to select the target option than respondents in the ‘CONTROL’

condition. This demonstrates that even small interface interventions can exert meaningful behavioural influence in a controlled digital shopping context. Accordingly, the study supports the broader proposition that digital nudging is an effective choice-architecture instrument in e-commerce-like environments (Weinmann, Schneider and vom Brocke, 2016; Hummel and Maedche, 2019; Valenčič et al., 2023).

6.3.2 RQ2

The second research question asked which nudge type is more effective within a generic, non-personalised implementation: social proof or default. The result runs counter to the initial expectation formulated in *H1a*, yet it is theoretically interpretable in light of the shopping-task environment. In the implemented grocery-like micro-decisions, the default nudge appears to have exerted stronger behavioural leverage than the sparse social-proof cue. Thus, the answer to **RQ2** is that, within this specific experimental context, default preselection was the more effective digital nudge.

6.3.3 RQ3

The third research question asked whether a GDMS-based nudge matching rule (Scott and Bruce, 1995) may outperform a non-personalised nudge implementation in increasing target-option choice. The answer to this question is more complex. The implemented GDMS-based matching rule did not yield a robust overall advantage over generic nudging in the confirmatory model. However, demographic-adjusted robustness checks and the post-hoc *H2* sensitivity analysis suggest that this null result should be interpreted cautiously, since the matched-condition estimate remained positive and the restricted *H2* sample had limited power. Thus, **RQ3** is most properly answered by stating that matching was not confirmed as generally superior, but may have conditional value for specific user groups.

6.4 Synthesis of the theoretical and managerial implications

The findings of this thesis generate implications at both the theoretical and practical levels.

Theoretically, the study reinforces the digital-nudging argument that online interfaces are active choice architectures rather than neutral presentation channels. The positive effect of generic nudging supports the view that relatively light-touch interface interventions can influence purchase-like decisions in e-commerce settings.

Moreover, the study refines theory by showing that nudge effects are mechanism-dependent. The finding that default preselection outperformed social proof suggests that different nudges should not be treated as interchangeable and that their relative effectiveness depends on the decision context and the way they are implemented. In addition, the non-significant overall result for the matched experimental condition indicates that simple psychological matching should not be assumed to be universally superior to generic nudging. Instead, the supplementary findings suggest that person-nudge fit may matter in a more conditional and subgroup-specific way.

From a managerial perspective, the results suggest that digital retailers can influence routine online choices through comparatively simple interface interventions and do not necessarily require highly sophisticated personalisation systems to generate behavioural effects. In the present setting, default preselection appeared more effective than sparse social-proof cues, which implies that carefully designed defaults may be especially useful in low-involvement decision environments. However, such interventions should remain transparent, easily reversible, and clearly distinct from manipulative design practices.

Finally, the results suggest that managers should approach behavioural personalisation cautiously. The evidence does not justify assuming that any simple matching rule will automatically outperform generic nudging. Rather, personalisation appears to require empirical validation and careful testing across user groups and contexts. At the same time, the

absence of a statistically significant reduction in perceived autonomy suggests that behaviourally effective nudges need not necessarily undermine users' subjective sense of agency when they remain choice-preserving and legitimate.

6.5 Final reflections and future outlook

A broader reflection arising from this thesis is that personalised digital nudging in e-commerce should be understood in a more differentiated way than strong general claims may suggest. The findings show that digital nudges can clearly influence consumer choice, yet their effectiveness depends on the mechanism used, the context in which they are embedded, and the users exposed to them. In this respect, the study suggests that the crucial question is not simply whether digital nudges work, but which nudge works, for whom, and under what conditions.

The thesis also points to an important ethical consideration. In the present low-stakes and choice-preserving shopping environment, behaviourally effective nudges were not accompanied by a statistically significant reduction in perceived autonomy. Without resolving the broader ethical debate, it suggests that behavioural influence and autonomy loss do not necessarily coincide.

Overall, the thesis concludes that personalised digital nudging should be viewed as a context-sensitive behavioural instrument rather than as a universally superior strategy. Digital nudges can work effectively in e-commerce, but not all nudges are equally effective, and personalisation may add value only under specific conditions. The broader implication is that effective digital choice architecture requires not only behavioural impact, but also careful theoretical reasoning, empirical validation, and responsible design.

Thus, future research should extend the present study under broader and more realistic conditions. In particular, larger and more diverse samples would be needed to assess whether

the subgroup-specific patterns observed here remain stable beyond the current constrained design. Future studies should also compare richer implementations of social proof and default nudges, apply more fine-grained matching approaches, and examine repeated rather than single-exposure settings. This would help clarify whether the present findings are specific to the implemented design or reflect more generalisable patterns of digital choice architecture in e-commerce.

Finally, research on personalised digital nudging should increasingly combine behavioural effectiveness with questions of transparency, legitimacy, and longer-term user response. As digital platforms become more capable of tailoring interface interventions at scale, future scholarship will need to examine not simply which nudges work, but also under which conditions they remain acceptable, repeatedly effective, and beneficial from the user's perspective. In this sense, the present thesis establishes a foundation for a more nuanced research agenda, evaluating digital choice architecture not only for its immediate behavioural impact, but also for its contextual limitations and ethical implications.

Bibliography

- Ansari, A. and Mela, C.F. (2003) 'E-Customization', *Journal of Marketing Research*, 40(2), pp. 131-145.
- Awad, N.F. and Krishnan, M.S. (2006) 'The personalisation privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalisation', *MIS Quarterly*, 30(1), pp. 13-28.
- Bearden, W.O., Netemeyer, R.G. and Teel, J.E. (1989) 'Measurement of consumer susceptibility to interpersonal influence', *Journal of Consumer Research*, 15(4), pp. 473-481.
- Behlen, L., Himmler, O. and Jäckle, R. (2023) 'Defaults and effortful tasks', *Experimental Economics*, 26(5), pp. 1022-1059.
- Bleier, A. and Eisenbeiss, M. (2015) 'Personalised online advertising effectiveness: The interplay of what, when, and where', *Marketing Science*, 34(5), pp. 669-688.
- Brand, B.M. and Reith, R. (2022) 'Cultural differences in the perception of credible online reviews – The influence of presentation format', *Decision Support Systems*, 154, p. 113710.
- Bruns, H. and Perino, G. (2023) 'The role of autonomy and reactance for nudging: Experimentally comparing defaults to recommendations and mandates', *Journal of Behavioral and Experimental Economics*, 106, p. 102047.
- Bruns, H., Fillon, A., Maniadis, Z. and Paunov, Y. (2025) 'Comparing transparent and covert nudges: A meta-analysis calling for more diversity in nudge transparency research', *Journal of Behavioral and Experimental Economics*, 116, p. 102350.
- Cacioppo, J.T. and Petty, R.E. (1982) 'The need for cognition', *Journal of Personality and Social Psychology*, 42(1), pp. 116-131.
- Cai, X., Cebollada, J. and Cortiñas, M. (2023) 'Impact of seller- and buyer-created content on product sales in the electronic commerce platform: The role of informativeness,

- readability, multimedia richness, and extreme valence', *Journal of Retailing and Consumer Services*, 70, pp. 103-141.
- Chevalier, J.A. and Mayzlin, D. (2006) 'The effect of word of mouth on sales: Online book reviews', *Journal of Marketing Research*, 43(3), pp. 345-354.
- Craciun, G. (2018) 'Choice defaults and social consensus effects on online information sharing: The moderating role of regulatory focus', *Computers in Human Behavior*, 88, pp. 89-102.
- Cramer, H., Evers, V., Ramlal, S., Van Someren, M., Rutledge, L., Stash, N., Aroyo, L. and Wielinga, B. (2008) 'The effects of transparency on trust in and acceptance of a content-based art recommender', *User Modeling and User-Adapted Interaction*, 18(5), pp. 455-496.
- Dellaert, B.G.C. and Häubl, G. (2012) 'Searching in choice mode: Consumer decision processes in product search with recommendations', *Journal of Marketing Research*, 49(2), pp. 277-288.
- De Vries, R., Bol, N. and van der Laan, N. (2025) 'Being transparent about personalization: The case of personalized digital 'just-in-time' nudges for healthier food choice', *Computers in Human Behavior Reports*, 21, p.100894.
- Eckles, D., Gordon, B.R. and Johnson, G.A. (2018) 'Field studies of psychologically targeted ads face threats to internal validity', *Proceedings of the National Academy of Sciences of the United States of America*, 115(23), pp. E5254-E5255.
- Fitzsimons, G.J. and Lehmann, D.R. (2004) 'Reactance to recommendations: When unsolicited advice yields contrary responses', *Marketing Science*, 23(1), pp. 82-94.
- Friestad, M. and Wright, P. (1994) 'The persuasion knowledge model: How people cope with persuasion attempts', *Journal of Consumer Research*, 21(1), pp. 1-31.

- Gefen, D., Karahanna, E. and Straub, D.W. (2003) 'Trust and TAM in online shopping: An integrated model', *MIS Quarterly*, 27(1), pp. 51-90.
- Ghiassaleh, A., Kocher, B. and Czellar, S. (2020) 'Best seller!? Unintended negative consequences of popularity signs on consumer choice behavior', *International Journal of Research in Marketing*, 37(4), pp. 805-820.
- Goldfarb, A. and Tucker, C. (2011) 'Privacy regulation and online advertising', *Management Science*, 57(1), pp. 57-71.
- Gray, C.M., Kou, Y., Battles, B., Hoggatt, J. and Toombs, A.L. (2018) 'The dark (patterns) side of UX design', in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. New York: ACM, Paper 534, pp. 1-14.
- Groissberger, T. and Riedl, R. (2017) 'Do online shops support customers' decision strategies by interactive information management tools? Results of an empirical analysis', *Electronic Commerce Research and Applications*, 26, pp. 131-151.
- Guath, M., Stikvoort, B. and Juslin, P. (2022) 'Nudging for eco-friendly online shopping – Attraction effect curbs price sensitivity', *Journal of Environmental Psychology*, p.101821.
- Gupta, P., Kim, Y.J., Glikson, E. and Woolley, A.W. (2024) 'Using digital nudges to enhance collective intelligence in online collaboration: Insights from unexpected outcomes', *MIS Quarterly*, 48(1), pp. 393-408.
- Herrmann, A., Goldstein, D.G., Stadler, R., Landwehr, J.R., Heitmann, M., Hofstetter, R. and Huber, F. (2011) 'The effect of default options on choice—Evidence from online product configurators', *Journal of Retailing and Consumer Services*, 18(6), pp. 483-491.
- Hettler, F.M., Schumacher, J.-P., Anton, E., Eybey, B. and Teuteberg, F. (2025) 'Understanding the user perception of digital nudging in platform interface design', *Electronic Commerce Research*, 25, pp. 2097-2134.

- Hirsh, J.B., Kang, S.K. and Bodenhausen, G.V. (2012) 'Personalized persuasion: Tailoring persuasive appeals to recipients' personality traits', *Psychological Science*, 23(6), pp. 578-581.
- Hummel, D. and Maedche, A. (2019) 'How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies', *Journal of Behavioral and Experimental Economics*, 80, pp. 47-58.
- Häubl, G. and Trifts, V. (2000) 'Consumer decision making in online shopping environments: The effects of interactive decision aids', *Marketing Science*, 19(1), pp. 4-21.
- Ingendahl, M., Hummel, D., Maedche, A. and Vogel, T. (2020) 'Who can be nudged? Examining nudging effectiveness in the context of need for cognition and need for uniqueness', *Journal of Consumer Behaviour*, 20(2), pp. 324-336.
- Ivanković, V. and Engelen, B. (2019) 'Nudging, transparency, and watchfulness', *Social Theory and Practice*, 45(1), pp. 43-73.
- Jachimowicz, J.M., Duncan, S., Weber, E.U. and Johnson, E.J. (2019) 'When and why defaults influence decisions: A meta-analysis of default effects', *Behavioural Public Policy*, 3(2), pp. 159-186.
- Jannach, D., Jugovac, M. and Nunes, I. (2019) 'Explanations and user control in recommender systems', in Augstein M., Herder E. and Wörndl W. (Eds) *Personalized Human-Computer Interaction*. Berlin: Walter de Gruyter GmbH, pp. 133-156.
- Janson, A. and Barev, T.J. (2026) 'Making the default more legitimate – the role of autonomy and transparency for digital privacy nudges', *Information Technology & People*, 39(1), pp. 428-455.
- Jesse, M., Jannach, D. and Gula, B. (2021) 'Digital nudging for online food choices', *Frontiers in Psychology*, 12, pp. 1-12.

- Johnson, E.J. and Goldstein, D. (2003) 'Do defaults save lives?', *Science*, 302(5649), pp. 1338-1339.
- Kahneman, D. (2011) *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Kim, T., Barasz, K. and John, L.K. (2019) 'Why am I seeing this ad? The effect of ad transparency on ad effectiveness', *Journal of Consumer Research*, 45(5), pp. 906-932.
- Knijnenburg, B.P., Willemsen, M.C., Gantner, Z., Soncu, H. and Newell, C. (2012) 'Explaining the user experience of recommender systems', *User Modeling and User-Adapted Interaction*, 22, pp. 441-504.
- Komiak, S.Y.X. and Benbasat, I. (2006) 'The effects of personalization and familiarity on trust and adoption of recommendation agents', *MIS Quarterly*, 30(4), pp. 941-960.
- Lambrecht, A. and Tucker, C. (2013) 'When does retargeting work? Information specificity in online advertising', *Journal of Marketing Research*, 50(5), pp. 561-576.
- Li, X., Grahl, J. and Hinz, O. (2022) 'How do recommender systems lead to consumer purchases? A causal mediation analysis of a field experiment', *Information Systems Research*, 33(2), pp. 620-637.
- Li, Y., Deng, X., Hu, X. and Liu, J. (2024) 'The effects of e-commerce recommendation system transparency on consumer trust: Exploring parallel multiple mediators and a moderator', *Journal of Theoretical and Applied Electronic Commerce Research*, 19(4), pp. 2630-2649.
- Liang, Y. and Willemsen, M.C. (2022) 'Promoting music exploration through personalized nudging in a genre exploration recommender', *International Journal of Human-Computer Interaction*, pp.1-24.
- Likert, R. (1932) 'A technique for the measurement of attitudes', *Archives of Psychology*, 140, pp. 1-55.

- Lu, Y., Zou, P., Huo, D., Chen, Y. and Li, W. (2026) 'Review system design and sales: How interface visibility moderates the effect of platform-generated default reviews', *Journal of Theoretical and Applied Electronic Commerce Research*, 21(3), 89, pp. 1-23.
- Mandel, N. and Johnson, E.J. (2002) 'When web pages influence choice: Effects of visual primes on experts and novices', *Journal of Consumer Research*, 29(2), pp. 235-245.
- Marco, G.D. and Douglas, T. (2023) 'Nudge transparency is not required for nudge resistibility', *Ergo an Open Access Journal of Philosophy*, 10(5), pp. 135-157.
- Mathur, A., Acar, G., Friedman, M.J., Lucherini, E., Mayer, J., Chetty, M. and Narayanan, A. (2019) 'Dark Patterns at Scale: Findings from a crawl of 11k shopping websites', *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Article 81, pp. 1-32.
- Mertens, S., Herberz, M., Hahnel, U. J. J., and Brosch, T. (2022) 'The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains', *Proceedings of the National Academy of Sciences of the United States of America*, 119(1), p. e2107346118.
- Michaelsen, P. (2024a) 'Transparency and nudging: An overview and methodological critique of empirical investigations', *Behavioural Public Policy*, 8(4), pp. 807-817.
- Michaelsen, P. (2024b) 'Experiencing default nudges: Autonomy, manipulation, and choice satisfaction as judged by people themselves', *Behavioural Public Policy*, 8(1), pp. 85-106.
- Mills, S. (2022) 'Personalized nudging', *Behavioural Public Policy*, 6(1), pp. 150-159.
- Münscher, R., Vetter, M. and Scheuerle, T. (2016) 'A review and taxonomy of choice architecture techniques', *Journal of Behavioral Decision Making*, 29(5), pp. 511-524.
- Nijssen, S.R.R., Pijs, M., van Ewijk, A. and Müller, B.C.N. (2023) 'Towards more sustainable online consumption: The impact of default and informational nudging on consumers' choice of delivery mode', *Cleaner and Responsible Consumption*, 11, pp. 100-135.

- The Organisation for Economic Co-operation and Development (OECD) (2022) 'Dark commercial patterns', *OECD Digital Economy Papers*, No. 336, OECD Publishing, Paris.
- Oppenheimer, D.M., Meyvis, T. and Davidenko, N. (2009) 'Instructional manipulation checks: Detecting satisficing to increase statistical power', *Journal of Experimental Social Psychology*, 45(4), pp. 867-872.
- Ortmann, A., Ryvkin, D., Wilkening, T. and Zhang, J. (2023) 'Defaults and cognitive effort', *Journal of Economic Behavior & Organization*, 212, pp. 1-19.
- Paunov, Y., Wänke, M. and Vogel, T. (2019) 'Ethical defaults: which transparency components can increase the effectiveness of default nudges?', *Social Influence*, 14(3-4), pp. 104-116.
- Paunov, Y., Vogel, T., Ingendahl, M. and Wänke, M. (2022) 'Transparent by choice: Proactive disclosures increase compliance with digital defaults', *Frontiers in Psychology*, 13, pp. 1-8.
- Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A. and Frik, A. (2020) 'Nudge me right: Personalizing online security nudges to people's decision-making styles', *Computers in Human Behavior*, 109, p. 106347.
- Qiu, K. and Zhang, L. (2024) 'How online reviews affect purchase intention: A meta-analysis across contextual and cultural factors', *Data and Information Management*, 8(2), p. 100058.
- Qualtrics (2026) The Leading Research & Experience Software. Available at: <https://www.qualtrics.com/> (Accessed: 20 March 2026).
- Rosillo-Díaz, E., Muñoz-Rosas, J. F. and Blanco-Encomienda, F. J. (2024) 'Impact of heuristic-systematic cues on the purchase intention of the electronic commerce consumer through the perception of product quality', *Journal of Retailing and Consumer Services*, 81, p. 103980.

- Ryan, R.M. and Deci, E.L. (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, 55(1), pp. 68-78.
- Schrift, R.Y., Netzer, O. and Kivetz, R. (2011) 'Complicating choice', *Journal of Marketing Research*, 48(2), pp. 308-326.
- Scott, S.G. and Bruce, R.A. (1995) 'Decision-making style: The development and assessment of a new measure', *Educational and Psychological Measurement*, 55(5), pp. 818-831.
- Simon, H.A. (1955) 'A behavioral model of rational choice', *Quarterly Journal of Economics*, 69(1), pp. 99-118.
- Steffel, M., Williams, E.F. and Pogacar, R. (2016) 'Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation', *Journal of Marketing Research*, 53(5), pp. 865-880.
- Thaler, R.H. and Sunstein, C.R. (2003) 'Libertarian paternalism', *American Economic Review*, 93(2), pp. 175-179.
- Thaler, R. H. and Sunstein, C. R. (2008) *Nudge: Improving decisions about health, wealth, and happiness*. United States: Yale University Press.
- Tian, K.T., Bearden, W.O. and Hunter, G.L. (2001) 'Consumers' need for uniqueness: Scale development and validation', *Journal of Consumer Research*, 28(1), pp. 50-66.
- Tucker, C.E. (2014) 'Social networks, personalized advertising, and privacy controls', *Journal of Marketing Research*, 51(5), pp. 546-562.
- Tversky, A. and Kahneman, D. (1974) 'Judgment under uncertainty: Heuristics and biases', *Science*, 185(4157), pp. 1124-1131.
- Valenčič, E., Beckett, E., Collins, C.E., Seljak, B.K. and Bucher, T. (2023) 'Digital nudging in online grocery stores: A scoping review on current practices and gaps', *Trends in Food Science & Technology*, 131, pp. 151-163.

- Valta, M. and Maier, C. (2025) 'Digital nudging: A systematic literature review, taxonomy, and future research directions', *A quarterly newsletter of SIGBIT*, New York, NY: ACM, 56(1), pp. 101-125.
- Wachner, J., Adriaanse, M.A. and De Ridder, D.T.D. (2021) 'The effect of nudges on autonomy in hypothetical and real life settings', *PLOS ONE*, 16(8), e0256124.
- Wan, X., Kumar, A. and Li, X. (2023) 'How do product recommendations help consumers search? Evidence from a field experiment', *Management Science*.
- Wang, L., Che, G., Hu, J. and Chen, L. (2024) 'Online review helpfulness and information overload: The roles of text, image, and video elements', *Journal of Theoretical and Applied Electronic Commerce Research*, 19(2), pp. 1243-1266.
- Weijers, R.J., Wachner, J. and de Koning, B.B. (2024) 'The effect of a default nudge on experienced and expected autonomy: A field study on food donation', *Journal of Behavioral Decision Making*, 37(4), e2404.
- Weinmann, M., Schneider, C. and vom Brocke, J. (2016) 'Digital nudging', *Business & Information Systems Engineering*, 58(6), pp. 433-436.
- Xu, C., Zhang, Z. and Shen, Y. (2025) 'Research on the relationship between the credibility of consumer online reviews and purchase decisions under international law: An empirical analysis based on e-commerce platform data', *Journal of Organizational and End User Computing*, 37(1), pp. 1-44.
- Zhu, F. and Zhang, X. (2010) 'Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics', *Journal of Marketing*, 74(2), pp. 133-148.

Appendices

Appendix 1: The GDMS-based questionnaire (Scott and Bruce, 1995)

* Please indicate how well each statement describes your typical behaviour.					
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Before making a decision, I often seek advice from other people.	☐	☐	☐	☐	☐
I feel more confident about a decision when someone I trust supports it.	☐	☐	☐	☐	☐
When a decision is important, I prefer to discuss it with others first.	☐	☐	☐	☐	☐
I frequently ask others for their opinions before deciding.	☐	☐	☐	☐	☐
I tend to rely on guidance from others when I am uncertain.	☐	☐	☐	☐	☐
I usually make decisions quickly.	☐	☐	☐	☐	☐
I prefer to decide immediately rather than delay a decision.	☐	☐	☐	☐	☐
I feel uncomfortable when I have to postpone a decision.	☐	☐	☐	☐	☐
I tend to decide "in the moment."	☐	☐	☐	☐	☐
I often make decisions without spending much time weighing options.	☐	☐	☐	☐	☐

Appendix 2: The seven product categories in the ‘CONTROL’ task version

*Which product would you add to your cart?

Bananas, 1kg - Colombia
0,99€



Bananas, 1kg - Ecuador
0,99€



*Which product would you add to your cart?

Apples, Gala 1kg - Germany
1,99€



Apples, Gala 1kg - Poland
1,99€



*Which product would you add to your cart?

Tomatoes, Cherry 500g - Italy
1,99€



Tomatoes, Cherry 500g - Spain
1,99€



*Which product would you add to your cart?

Eggs, 10-pack - Farm A, Italy
2,19€

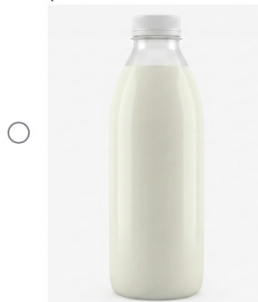


Eggs, 10-pack - Farm B, Italy
2,19€

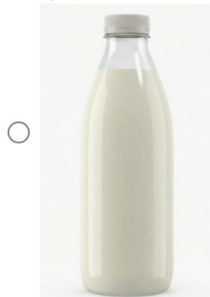


*Which product would you add to your cart?

Milk, bottled, 1l - Farm A, Italy
1,79€



Milk, bottled, 1l - Farm B, Italy
1,79€



*Which product would you add to your cart?

Cucumbers, 3-pack - Spain
1,19€



Cucumbers, 3-pack - Netherlands
1,19€



*Which product would you add to your cart?

Butter, 250g - Creamery A, Italy
0,99€



Butter, 250g - Creamery B, Italy
0,99€



Appendix 3: The seven product categories in the ‘SOCIAL’ task version

* Which product would you add to your cart?

Bananas, 1kg - Colombia
0,99€



Bananas, 1kg - Ecuador
Our customers' choice
0,99€



* Which product would you add to your cart?

Cucumbers, 3-pack - Spain
Our customers' choice
1,19€



Cucumbers, 3-pack - Netherlands
1,19€



* Which product would you add to your cart?

Butter, 250g - Creamery A, Italy
Our customers' choice
0,99€



Butter, 250g - Creamery B, Italy
0,99€



* Which product would you add to your cart?

Milk, bottled, 1l - Farm A, Italy
1,79€



Milk, bottled, 1l - Farm B, Italy
Our customers' choice
1,79€



* Which product would you add to your cart?

Apples, Gala 1kg - Germany
Our customers' choice
1,99€



Apples, Gala 1kg - Poland
1,99€



* Which product would you add to your cart?

Tomatoes, Cherry 500g - Italy
Our customers' choice
1,99€



Tomatoes, Cherry 500g - Spain
1,99€



* Which product would you add to your cart?

Eggs, 10-pack - Farm A, Italy
2,19€



Eggs, 10-pack - Farm B, Italy
Our customers' choice
2,19€



Appendix 4: The seven product categories in the 'DEFAULT' task version

Which product would you add to your cart?

Eggs, 10-pack - Farm A, Italy
2,19€



Eggs, 10-pack - Farm B, Italy
2,19€
Already in your cart



Which product would you add to your cart?

Apples, Gala 1kg - Germany
1,99€
Already in your cart



Apples, Gala 1kg - Poland
1,99€



Which product would you add to your cart?

Cucumbers, 3-pack - Spain
Already in your cart
1,19€



Cucumbers, 3-pack - Netherlands
1,19€



Which product would you add to your cart?

Butter, 250g - Creamery A, Italy
0,99€
Already in your cart



Butter, 250g - Creamery B, Italy
0,99€



Which product would you add to your cart?

Milk, bottled, 1l - Farm A, Italy
1,79€



Milk, bottled, 1l - Farm B, Italy
Already in your cart
1,79€



Which product would you add to your cart?

Bananas, 1kg - Colombia
0,99€



Bananas, 1kg - Ecuador
0,99€
Already in your cart



Which product would you add to your cart?

Tomatoes, Cherry 500g - Italy
Already in your cart
1,99€



Tomatoes, Cherry 500g - Spain
1,99€



Appendix 5: The demographic questionnaire in the experiment

* How old are you?

- ☐ 18-24 years old
- ☐ 25-34 years old
- ☐ 35-44 years old
- ☐ 45-54 years old
- ☐ 55-64 years old
- ☐ 65+ years old

* How do you describe yourself?

- ☐ Male
- ☐ Female
- ☐ Non-binary / third gender
- ☐ Prefer to self-describe
- ☐ Prefer not to say

* How often do you shop online?

- ☐ Never
- ☐ Once a month
- ☐ Several times a month
- ☐ Every week
- ☐ Several times a week

Appendix 6: The perceived autonomy evaluation question matrix

*Please indicate how well each statement describes how you experienced the present study.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I felt free to choose the product I wanted.	☐	☐	☐	☐	☐
I felt that the final choice was entirely up to me.	☐	☐	☐	☐	☐
I was able to choose whichever option I preferred.	☐	☐	☐	☐	☐
I felt in control of my decision during the shopping task.	☐	☐	☐	☐	☐
My choice reflected my own preferences.	☐	☐	☐	☐	☐

Appendix 7: Internal consistency and routing check

Diagnostic	n_control_block	n_social_block	n_default_block	n	gdms_type	gdms_type_reconstructed	gdms_match_check	route_match_check	policy_arm	generic_nudge_type_raw	task_version	unified_nonmissing_count
Branch occupancy	0.0	0.0	7.0	59								
	0.0	7.0	0.0	65								
	7.0	0.0	0.0	59								
GDMS reconstruction crosstab				47	DEPENDENT	DEPENDENT						
				93	MIXED	MIXED						
				43	SPONTANEOUS	SPONTANEOUS						
GDMS reconstruction match summary				183			TRUE					
Route match summary				183				TRUE				
Routing crosstab				17	DEPENDENT				CONTROL		CONTROL	
				29	MIXED				CONTROL		CONTROL	
				13	SPONTANEOUS				CONTROL		CONTROL	
				5	DEPENDENT				GENERIC	DEFAULT	DEFAULT	
				7	DEPENDENT				GENERIC	SOCIAL	SOCIAL	
				17	MIXED				GENERIC	DEFAULT	DEFAULT	
				16	MIXED				GENERIC	SOCIAL	SOCIAL	
				9	SPONTANEOUS				GENERIC	DEFAULT	DEFAULT	
				8	SPONTANEOUS				GENERIC	SOCIAL	SOCIAL	
				18	DEPENDENT				MATCHED		SOCIAL	
				15	MIXED				MATCHED	DEFAULT	DEFAULT	
				16	MIXED				MATCHED	SOCIAL	SOCIAL	
				13	SPONTANEOUS				MATCHED		DEFAULT	
Unified choice non-missing check				183								7.0

Appendix 8: Overall sample description

variable	level	n	pct	stat
age_label	18-24 years old	53	0.2896174863387978	53 (29%)
age_label	25-34 years old	80	0.4371584699453552	80 (43.7%)
age_label	35-44 years old	20	0.1092896174863388	20 (10.9%)
age_label	45-54 years old	17	0.09289617486338798	17 (9.3%)
age_label	55-64 years old	12	0.06557377049180328	12 (6.6%)
age_label	65+ years old	1	0.00546448087431694	1 (0.5%)
gender_label	Male	100	0.546448087431694	100 (54.6%)
gender_label	Female	82	0.44808743169398907	82 (44.8%)
gender_label	Other / prefer not to say	1	0.00546448087431694	1 (0.5%)
shopping_frequency_label	Never	4	0.02185792349726776	4 (2.2%)
shopping_frequency_label	Once a month	70	0.3825136612021858	70 (38.3%)
shopping_frequency_label	Several times a month	57	0.3114754098360656	57 (31.1%)
shopping_frequency_label	Every week	29	0.15846994535519127	29 (15.8%)
shopping_frequency_label	Several times a week	23	0.12568306010928962	23 (12.6%)
gdms_type	DEPENDENT	47	0.2568306010928962	47 (25.7%)
gdms_type	SPONTANEOUS	43	0.23497267759562843	43 (23.5%)
gdms_type	MIXED	93	0.5081967213114754	93 (50.8%)
policy_arm	CONTROL	59	0.3224043715846995	59 (32.2%)
policy_arm	GENERIC	62	0.33879781420765026	62 (33.9%)
policy_arm	MATCHED	62	0.33879781420765026	62 (33.9%)
task_version	CONTROL	59	0.3224043715846995	59 (32.2%)
task_version	SOCIAL	65	0.3551912568306011	65 (35.5%)
task_version	DEFAULT	59	0.3224043715846995	59 (32.2%)

Appendix 9: Sample description by experimental condition

variable	level	CONTROL	GENERIC	MATCHED
age_label	18-24 years old	17 (28.8%)	18 (29%)	18 (29%)
age_label	25-34 years old	27 (45.8%)	24 (38.7%)	29 (46.8%)
age_label	35-44 years old	6 (10.2%)	3 (4.8%)	11 (17.7%)
age_label	45-54 years old	8 (13.6%)	9 (14.5%)	0 (0%)
age_label	55-64 years old	1 (1.7%)	7 (11.3%)	4 (6.5%)
age_label	65+ years old	0 (0%)	1 (1.6%)	0 (0%)
gender_label	Male	36 (61%)	34 (54.8%)	30 (48.4%)
gender_label	Female	23 (39%)	27 (43.5%)	32 (51.6%)
gender_label	Other / prefer not to say	0 (0%)	1 (1.6%)	0 (0%)
shopping_frequency_label	Never	1 (1.7%)	1 (1.6%)	2 (3.2%)
shopping_frequency_label	Once a month	22 (37.3%)	25 (40.3%)	23 (37.1%)
shopping_frequency_label	Several times a month	22 (37.3%)	19 (30.6%)	16 (25.8%)
shopping_frequency_label	Every week	6 (10.2%)	8 (12.9%)	15 (24.2%)
shopping_frequency_label	Several times a week	8 (13.6%)	9 (14.5%)	6 (9.7%)
gdms_type	DEPENDENT	17 (28.8%)	12 (19.4%)	18 (29%)
gdms_type	SPONTANEOUS	13 (22%)	17 (27.4%)	13 (21%)
gdms_type	MIXED	29 (49.2%)	33 (53.2%)	31 (50%)

Appendix 10: Sample description within the GENERIC experimental condition by task version

variable	level	DEFAULT	SOCIAL
age_label	18-24 years old	9 (29%)	9 (29%)
age_label	25-34 years old	11 (35.5%)	13 (41.9%)
age_label	35-44 years old	0 (0%)	3 (9.7%)
age_label	45-54 years old	5 (16.1%)	4 (12.9%)
age_label	55-64 years old	5 (16.1%)	2 (6.5%)
age_label	65+ years old	1 (3.2%)	0 (0%)
gender_label	Male	15 (48.4%)	19 (61.3%)
gender_label	Female	16 (51.6%)	11 (35.5%)
gender_label	Other / prefer not to say	0 (0%)	1 (3.2%)
shopping_frequency_label	Never	1 (3.2%)	0 (0%)
shopping_frequency_label	Once a month	15 (48.4%)	10 (32.3%)
shopping_frequency_label	Several times a month	9 (29%)	10 (32.3%)
shopping_frequency_label	Every week	3 (9.7%)	5 (16.1%)
shopping_frequency_label	Several times a week	3 (9.7%)	6 (19.4%)
gdms_type	DEPENDENT	5 (16.1%)	7 (22.6%)
gdms_type	SPONTANEOUS	9 (29%)	8 (25.8%)
gdms_type	MIXED	17 (54.8%)	16 (51.6%)

Appendix 11: Balance checks across experimental conditions

variable	test	statistic	p_value
age_label	Chi-square with simulated p-value	20.25236056513964	0.017798220177982202
gender_label	Chi-square with simulated p-value	3.9441796796863544	0.38536146385361464
shopping_frequency_label	Chi-square with simulated p-value	6.76776466688611	0.5771422857714229
gdms_type	Chi-square with simulated p-value	2.2022398390682207	0.7038296170382962

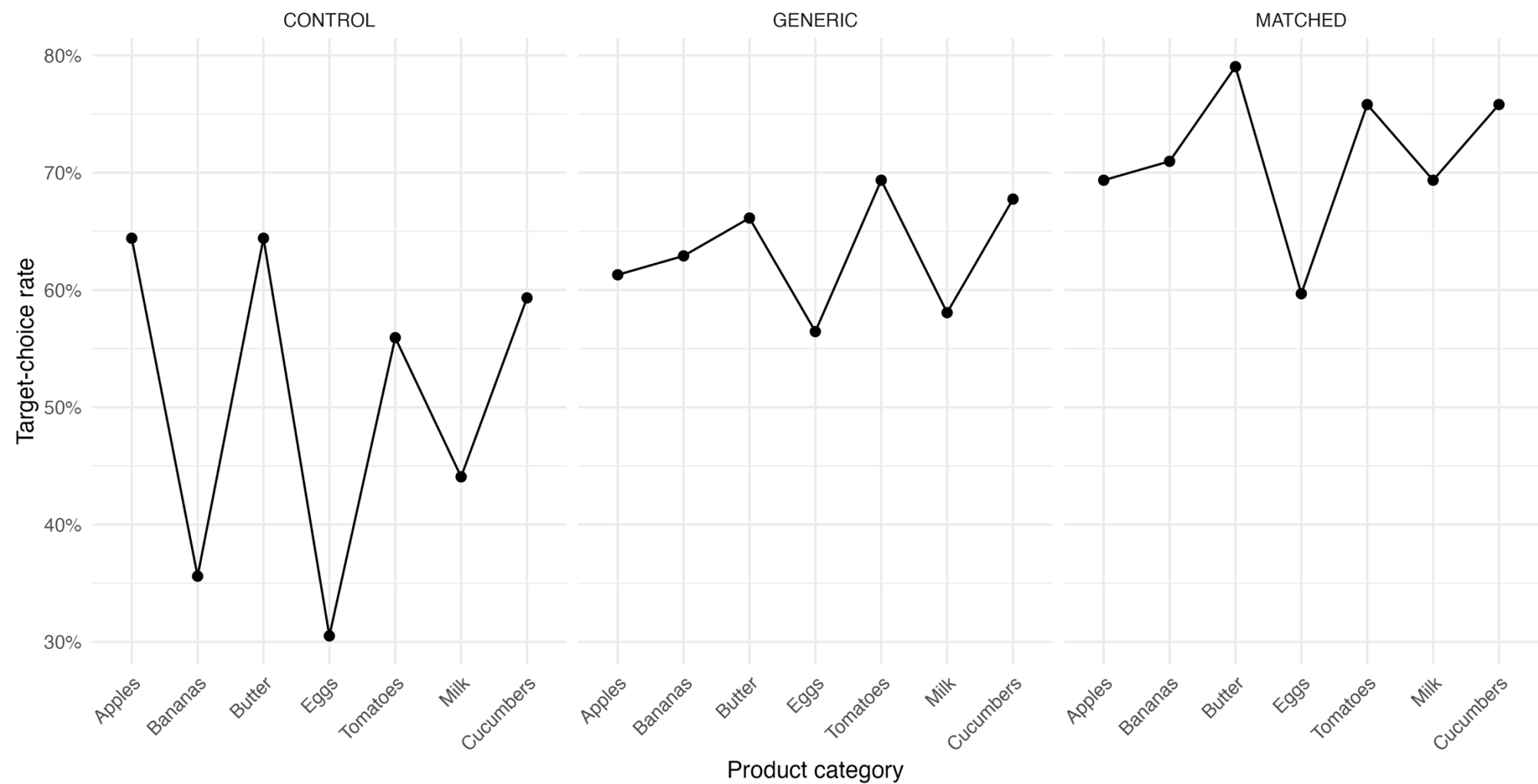
Appendix 12: Balance checks within the GENERIC experimental condition

variable	test	statistic	p_value
age_label	Chi-square with simulated p-value	5.563492063492062	0.3663633636636365
gender_label	Chi-square with simulated p-value	2.396514161220044	0.30386961303869614
shopping_frequency_label	Chi-square with simulated p-value	3.552631578947368	0.47455254474552544
gdms_type	Chi-square with simulated p-value	0.4224598930481283	0.8863113688631137

Appendix 13: Target choice share at category level and by experimental condition

policy_arm	category	n_choices	target_rate	target_count	non_target_count
CONTROL	Apples	59	0.6440677966101694	38	21
CONTROL	Bananas	59	0.35593220338983056	21	38
CONTROL	Butter	59	0.6440677966101694	38	21
CONTROL	Eggs	59	0.30508474576271183	18	41
CONTROL	Tomatoes	59	0.5593220338983051	33	26
CONTROL	Milk	59	0.440677966101695	26	33
CONTROL	Cucumbers	59	0.5932203389830508	35	24
GENERIC	Apples	62	0.6129032258064516	38	24
GENERIC	Bananas	62	0.6290322580645161	39	23
GENERIC	Butter	62	0.6612903225806451	41	21
GENERIC	Eggs	62	0.5645161290322581	35	27
GENERIC	Tomatoes	62	0.6935483870967742	43	19
GENERIC	Milk	62	0.5806451612903226	36	26
GENERIC	Cucumbers	62	0.6774193548387096	42	20
MATCHED	Apples	62	0.6935483870967742	43	19
MATCHED	Bananas	62	0.7096774193548386	44	18
MATCHED	Butter	62	0.7903225806451613	49	13
MATCHED	Eggs	62	0.5967741935483871	37	25
MATCHED	Tomatoes	62	0.7580645161290324	47	15
MATCHED	Milk	62	0.6935483870967742	43	19
MATCHED	Cucumbers	62	0.7580645161290324	47	15

Appendix 14: Target choice share at category level and by experimental condition (visual representation)



Appendix 15: *H1* model

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial (logit)

Formula: target_choice ~ policy_arm + category + (1 | response_id)

Data: dat_h1_long

Control: glmer_ctrl

AIC	BIC	logLik	-2*log(L)	df.resid
1125.8	1168.5	-553.9	1107.8	838

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.8991	-0.9430	0.5744	0.8048	1.8097

Random effects:

Groups	Name	Variance	Std.Dev.
--------	------	----------	----------

response_id	(Intercept)	0.3992	0.6318
-------------	-------------	--------	--------

Number of obs: 847, groups: response_id, 121

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.2864920751	0.2257526916	1.269	0.20442
policy_armGENERIC	0.5787196622	0.1884174210	3.071	0.00213 **
categoryBananas	-0.6015720897	0.2760427398	-2.179	0.02931 *
categoryButter	0.1191545543	0.2814846000	0.423	0.67207
categoryEggs	-0.8609561799	0.2776690345	-3.101	0.00193 **
categoryTomatoes	0.0000004049	0.2795338426	0.000	1.00000
categoryMilk	-0.5278239488	0.2759232070	-1.913	0.05576 .
categoryCucumbers	0.0393595346	0.2801309203	0.141	0.88826

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	p_GENE	ctgryBn	ctgryBt	ctgryE	ctgryT	ctgryM
plc_GENERIC	-0.404						
categoryBnns	-0.624	-0.019					
categoryBttr	-0.615	0.004	0.502				
categoryEggs	-0.619	-0.028	0.516	0.499			
categoryTmts	-0.619	0.000	0.506	0.497	0.503		
categoryMlk	-0.624	-0.017	0.517	0.502	0.515	0.507	
ctgryCcmbrrs	-0.618	0.001	0.505	0.496	0.502	0.499	0.505

Appendix 16: *H1a* model

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial (logit)
 Formula: target_choice ~ generic_nudge_type_analysis + category + (1 | response_id)
 Data: dat_h1a_long
 Control: glmer_ctrl

AIC	BIC	logLik	-2*log(L)	df.resid
564.4	601.1	-273.2	546.4	425

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.0961	-0.9916	0.5380	0.7142	1.3229

Random effects:

Groups	Name	Variance	Std.Dev.
response_id	(Intercept)	0.399	0.6317

Number of obs: 434, groups: response_id, 62

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.93610	0.32216	2.906	0.00366 **
generic_nudge_type_analysisSOCIAL	-0.83596	0.26878	-3.110	0.00187 **
categoryBananas	0.07729	0.39227	0.197	0.84380
categoryButter	0.23586	0.39605	0.596	0.55149
categoryEggs	-0.22617	0.38792	-0.583	0.55986
categoryTomatoes	0.40128	0.40116	1.000	0.31716
categoryMilk	-0.15156	0.38865	-0.390	0.69656
categoryCucumbers	0.31759	0.39842	0.797	0.42539

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	g__S0	ctgryBn	ctgryBt	ctgryE	ctgryT	ctgryM
gn__SOCIAL	-0.445						
categoryBnns	-0.602	-0.003					
categoryBttr	-0.593	-0.010	0.492				
categoryEggs	-0.617	0.010	0.502	0.496			
categoryTmts	-0.581	-0.017	0.486	0.482	0.490		
categoryMlk	-0.614	0.007	0.501	0.496	0.507	0.489	
ctgryCcmbrrs	-0.587	-0.014	0.489	0.485	0.493	0.479	0.493

Appendix 17: *H2* model

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial (logit)

Formula: target_choice ~ policy_arm + gdms_type + category + (1 | response_id)

Data: dat_h2_long

Control: glmer_ctrl

AIC	BIC	logLik	-2*log(L)	df.resid
520.1	560.5	-250.0	500.1	410

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.2132	-1.1435	0.5405	0.6462	1.0736

Random effects:

Groups	Name	Variance	Std.Dev.
--------	------	----------	----------

response_id	(Intercept)	0.09999	0.3162
-------------	-------------	---------	--------

Number of obs: 420, groups: response_id, 60

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.5750	0.3338	1.722	0.085 .
policy_armMATCHED	0.3649	0.2371	1.539	0.124
gdms_typeSPONTANEOUS	-0.2438	0.2371	-1.028	0.304
categoryBananas	0.5888	0.4145	1.420	0.156
categoryButter	0.3193	0.4007	0.797	0.425
categoryEggs	-0.4339	0.3820	-1.136	0.256
categoryTomatoes	0.4952	0.4092	1.210	0.226
categoryMilk	0.2360	0.3972	0.594	0.552
categoryCucumbers	0.4952	0.4092	1.210	0.226

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	p_MATC	g_SPON	ctgryBn	ctgryBt	ctgryE	ctgryT	ctgryM
plc_MATCHED	-0.410							
g_SPONTANEO	-0.428	0.170						
categoryBnns	-0.546	0.012	-0.008					
categoryBttr	-0.566	0.007	-0.004	0.458				
categoryEggs	-0.596	-0.011	0.007	0.477	0.494			
categoryTmts	-0.553	0.010	-0.006	0.449	0.463	0.483		
categoryMlk	-0.571	0.005	-0.003	0.461	0.477	0.499	0.467	
ctgryCcmbrrs	-0.553	0.010	-0.006	0.449	0.463	0.483	0.454	0.467

Appendix 18: Aggregated binomial model and the share-level OLS model

hypothesis	model	contrast	effect_metric	estimate	conf_low	conf_high	std_error	statistic	p_value
H1	Aggregated binomial GLM	Intercept	Odds ratio	1.02	0.84	1.24	0.1	0.25	0.81
H1		GENERIC vs CONTROL		1.67	1.27	2.2	0.14	3.67	<0.01
H1a		Intercept		2.56	1.91	3.46	0.15	6.22	<0.01
H1a		SOCIAL vs DEFAULT		0.47	0.31	0.69	0.2	-3.75	<0.01
H2		Intercept		2.19	1.49	3.25	0.2	3.96	<0.01
H2		MATCHED vs GENERIC		1.42	0.93	2.18	0.22	1.62	0.1
H2 (add.)		SPONTANEOUS vs DEPENDENT		0.79	0.51	1.21	0.22	-1.09	0.27
H1	Share-level OLS	Intercept	Coefficient	0.51	0.45	0.56	0.03	17.84	<0.01
H1		GENERIC vs CONTROL		0.13	0.04	0.21	0.04	3.01	<0.01
H1a		Intercept		0.72	0.65	0.79	0.03	20.74	<0.01
H1a		SOCIAL vs DEFAULT		-0.18	-0.29	-0.06	0.06	-3.05	<0.01
H2		Intercept		0.68	0.59	0.78	0.05	14.67	<0.01
H2		MATCHED vs GENERIC		0.07	-0.03	0.17	0.05	1.49	0.14
H2 (add.)		SPONTANEOUS vs DEPENDENT		-0.05	-0.15	0.05	0.05	-1.01	0.32

Appendix 19: *H1* model (adjusted for demographic covariates)

Generalized linear mixed model fit by maximum likelihood (Laplace

Approximation) [glmerMod]

Family: binomial (logit)

Formula:

target_choice ~ policy_arm + age_control + gender_control + shopping_frequency_control +
category + (1 | response_id)

Data: dat_h1_long_adj

Control: glmer_ctrl

AIC	BIC	logLik	-2*log(L)	df.resid
1128.6	1194.8	-550.3	1100.6	826

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.7798	-0.9449	0.5741	0.7978	1.7127

Random effects:

Groups	Name	Variance	Std.Dev.
--------	------	----------	----------

response_id	(Intercept)	0.3567	0.5973
-------------	-------------	--------	--------

Number of obs: 840, groups: response_id, 120

Fixed effects:

	Estimate	Std. Error
(Intercept)	0.163194598	0.298230039
policy_armGENERIC	0.541954457	0.185923725
age_control25-34	-0.014319559	0.229672834
age_control35+	-0.089405066	0.247685251
gender_controlFemale	0.240341186	0.192012179
shopping_frequency_controlSeveral times a month	0.080823497	0.216632906
shopping_frequency_controlWeekly or more	0.121954307	0.236007523
categoryBananas	-0.601439089	0.276094115
categoryButter	0.118954686	0.281334822
categoryEggs	-0.861475722	0.277843638
categoryTomatoes	0.000001001	0.279409015
categoryMilk	-0.527600798	0.275944136
categoryCucumbers	0.039298277	0.279998175
	z value	Pr(> z)
(Intercept)	0.547	0.58423
policy_armGENERIC	2.915	0.00356 **
age_control25-34	-0.062	0.95029
age_control35+	-0.361	0.71813
gender_controlFemale	1.252	0.21068
shopping_frequency_controlSeveral times a month	0.373	0.70908
shopping_frequency_controlWeekly or more	0.517	0.60534
categoryBananas	-2.178	0.02938 *
categoryButter	0.423	0.67242
categoryEggs	-3.101	0.00193 **
categoryTomatoes	0.000	1.00000
categoryMilk	-1.912	0.05588 .
categoryCucumbers	0.140	0.88838

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix 20: *H1a* model (adjusted for demographic covariates)

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
 Family: binomial (logit)
 Formula:
 target_choice ~ generic_nudge_type_analysis + age_control + gender_control + shopping_frequency_control + category + (1 | response_id)
 Data: dat_h1a_long_adj
 Control: glmer_ctrl

AIC	BIC	logLik	-2*log(L)	df.resid
561.7	618.5	-266.8	533.7	413

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.1410	-0.9491	0.5275	0.7411	1.3839

Random effects:

Groups	Name	Variance	Std.Dev.
	response_id (Intercept)	0.2395	0.4894

Number of obs: 427, groups: response_id, 61

Fixed effects:

	Estimate	Std. Error
(Intercept)	0.90723	0.41208
generic_nudge_type_analysisSOCIAL	-0.78960	0.25741
age_control25-34	-0.33201	0.31472
age_control35+	-0.26303	0.33112
gender_controlFemale	0.51739	0.26454
shopping_frequency_controlSeveral times a month	-0.05355	0.30298
shopping_frequency_controlWeekly or more	-0.10788	0.32158
categoryBananas	0.07740	0.39307
categoryButter	0.23608	0.39678
categoryEggs	-0.22679	0.38890
categoryTomatoes	0.40145	0.40183
categoryMilk	-0.15192	0.38958
categoryCucumbers	0.31780	0.39912
	z value	Pr(> z)
(Intercept)	2.202	0.02770 *
generic_nudge_type_analysisSOCIAL	-3.067	0.00216 **
age_control25-34	-1.055	0.29145
age_control35+	-0.794	0.42698
gender_controlFemale	1.956	0.05049 .
shopping_frequency_controlSeveral times a month	-0.177	0.85971
shopping_frequency_controlWeekly or more	-0.335	0.73728
categoryBananas	0.197	0.84389
categoryButter	0.595	0.55186
categoryEggs	-0.583	0.55978
categoryTomatoes	0.999	0.31776
categoryMilk	-0.390	0.69656
categoryCucumbers	0.796	0.42589

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix 21: *H2* model (adjusted for demographic covariates)

```
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) [glmerMod]
Family: binomial ( logit )
Formula:
target_choice ~ policy_arm + gdms_type + age_control + gender_control +
shopping_frequency_control + category + (1 | response_id)
Data: dat_h2_long_adj
Control: glmer_ctrl
```

AIC	BIC	logLik	-2*log(L)	df.resid
519.0	579.3	-244.5	489.0	398

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.6084	-1.1257	0.5369	0.6713	1.2345

Random effects:

Groups	Name	Variance	Std.Dev.
response_id	(Intercept)	0.001783	0.04223

Number of obs: 413, groups: response_id, 59

Fixed effects:

	Estimate	Std. Error
(Intercept)	0.86244	0.41810
policy_armMATCHED	0.43880	0.22826
gdms_typeSPONTANEOUS	-0.08602	0.23221
age_control25-34	-0.40742	0.28429
age_control35+	-0.27415	0.31060
gender_controlFemale	0.20721	0.23507
shopping_frequency_controlSeveral times a month	-0.35359	0.29168
shopping_frequency_controlWeekly or more	-0.46509	0.27238
categoryBananas	0.59029	0.41526
categoryButter	0.32041	0.40148
categoryEggs	-0.43678	0.38343
categoryTomatoes	0.49658	0.40999
categoryMilk	0.23690	0.39806
categoryCucumbers	0.49658	0.40999

	z value	Pr(> z)
(Intercept)	2.063	0.0391 *
policy_armMATCHED	1.922	0.0546 .
gdms_typeSPONTANEOUS	-0.370	0.7110
age_control25-34	-1.433	0.1518
age_control35+	-0.883	0.3774
gender_controlFemale	0.882	0.3780
shopping_frequency_controlSeveral times a month	-1.212	0.2254
shopping_frequency_controlWeekly or more	-1.708	0.0877 .
categoryBananas	1.422	0.1552
categoryButter	0.798	0.4248
categoryEggs	-1.139	0.2546
categoryTomatoes	1.211	0.2258
categoryMilk	0.595	0.5517
categoryCucumbers	1.211	0.2258

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

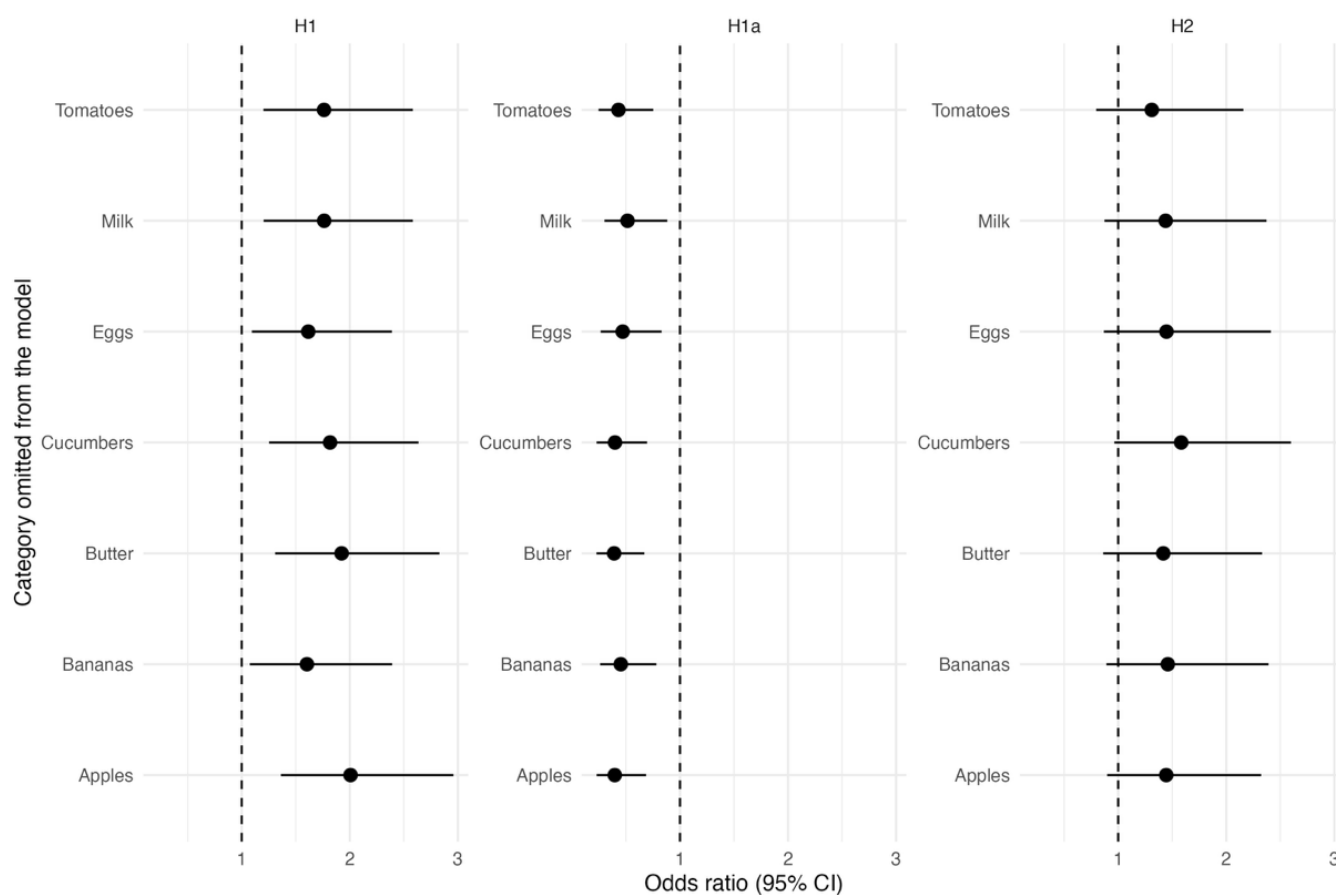
Appendix 22: *H2* model post-hoc sensitivity analysis

specification	estimate_log_odds	std_error	n_respondents	n_choice_rows	odds_ratio	z_observed	p_value	noncentrality_observed	approx_power_observed_effect	min_log_odds_for_p05	min_or_for_p05	log_odds_for_80_power	or_for_80_power
H2 baseline	0.3648712691617599	0.23711135244044773	60	420	1.440328581508771	1.5388182193992588	0.1238486806642658	1.5388182193992588	0.33705801647935185	0.46472971110886085	1.5915839432852317	0.6642876600439326	1.9431058762718414
H2 demographic-control adjusted	0.43880485338198166	0.22826007729405182	59	413	1.550852614213143	1.9223898396244712	0.05455672308955761	1.9223898396244712	0.48506533646627026	0.44738153060467034	1.5642109798141992	0.6394900584323391	1.8955140314112366

Appendix 23: ‘Leave-one-category-out’ sensitivity analysis

hypothesis	n_reduced_models	min_or	max_or	min_p	max_p	all_same_direction	n_significant_categ
H1	7	1.6034168547087746	2.0073032140684735	4.32593243266773e-4	0.020714802101806127	TRUE	7
H1a	7	0.39049698143108974	0.5147164013058807	0.00065239968720229725	0.015881283160352373	TRUE	7
H2	7	1.3105685641212803	1.5828160005067602	0.06962815161291183	0.28816200146241605	TRUE	0

Appendix 24: ‘Leave-one-category-out’ sensitivity analysis (visual representation)



Appendix 25: Additional *H2* interaction model

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial (logit)
 Formula: target_choice ~ policy_arm * gdms_type + category + (1 | response_id)
 Data: dat_h2_long
 Control: glmer_ctrl

AIC	BIC	logLik	-2*log(L)	df.resid
514.7	559.1	-246.3	492.7	409

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.2715	-1.1232	0.5254	0.6442	1.1913

Random effects:

Groups	Name	Variance	Std.Dev.
response_id	(Intercept)	0.01664	0.129

Number of obs: 420, groups: response_id, 60

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.9516	0.3647	2.610	0.00907 **
policy_armMATCHED	-0.2933	0.3303	-0.888	0.37451
gdms_typeSPONTANEOUS	-0.8660	0.3258	-2.658	0.00785 **
categoryBananas	0.5904	0.4153	1.422	0.15513
categoryButter	0.3202	0.4014	0.798	0.42498
categoryEggs	-0.4348	0.3825	-1.137	0.25568
categoryTomatoes	0.4966	0.4100	1.211	0.22583
categoryMilk	0.2367	0.3979	0.595	0.55193
categoryCucumbers	0.4966	0.4100	1.211	0.22583
policy_armMATCHED:gdms_typeSPONTANEOUS	1.2657	0.4628	2.735	0.00624 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	pl_MATCHED	g_SPON	ctgryBn	ctgryBt	ctgryE	ctgryT	ctgryM	ctgryC
plc_MATCHED	-0.571								
g_SPONTANEO	-0.578	0.641							
categoryBnns	-0.492	-0.006	-0.022						
categoryBttr	-0.514	-0.004	-0.013	0.458					
categoryEggs	-0.554	0.007	0.019	0.477	0.494				
categoryTmts	-0.500	-0.006	-0.019	0.449	0.463	0.483			
categoryMlk	-0.520	-0.003	-0.010	0.461	0.477	0.499	0.467		
ctgryCcmbrrs	-0.500	-0.006	-0.019	0.449	0.463	0.483	0.454	0.467	
p_MATCHED:_	0.407	-0.716	-0.706	0.022	0.013	-0.020	0.019	0.010	0.019

Appendix 26: Supplementary hypothesis testing models

hypothesis	term	odds_ratio	conf_low	conf_high	std.error	statistic	p.value
H2 interaction	(Intercept)	2.5899319588831036	1.2672930471896753	5.2929727394299535	0.9444815498953054	2.609538647184732	0.009066440431364129
H2 interaction	policy_armMATCHED	0.7457665221098648	0.390330451183444	1.424863737414278	0.24634296118758633	-0.888051214100306	0.37451320334291355
H2 interaction	gdms_typeSPONTANEOUS	0.4206405750670064	0.22213779109188853	0.7965258523684081	0.13702851556093218	-2.6583143870788746	0.007853259609466364
H2 interaction	categoryBananas	1.804719119983489	0.7996671335309754	4.072958566713205	0.7494921579438878	1.4216494756459421	0.15512803424283764
H2 interaction	categoryButter	1.3774210052606375	0.6272299024925353	3.024869538575299	0.5528444383383984	0.7978157384605069	0.4249774267318775
H2 interaction	categoryEggs	0.6473723259554455	0.30586304899935285	1.370191429739688	0.24765225765404497	-1.1366716242641546	0.25567558579908245
H2 interaction	categoryTomatoes	1.6430496890134767	0.7356637930505889	3.6696277648418127	0.6736092245215981	1.2111814972420845	0.2258258524192036
H2 interaction	categoryMilk	1.2670437062527333	0.5809161737687757	2.7635652544142557	0.5041344129141563	0.5948651817624632	0.551933592933068
H2 interaction	categoryCucumbers	1.6430494202895587	0.7356642863531324	3.669624103810296	0.6736084151154925	1.2111823555713919	0.22582552353173493
H2 interaction	policy_armMATCHED:gdms_typeSPONTANEOUS	3.545733045332046	1.4313724143815298	8.783334583258611	1.641037390630203	2.734851515709968	0.00624084208844278
H2 congruence	congruent_exposure	1.91546225575608	1.1405850170119314	3.216766482553041	0.5066468059195356	2.457277693354043	0.013999442560284488

Appendix 27: Predicted probabilities by gdms_type

policy_arm	gdms_type	prob	SE	df	asympt.LCL	asympt.UCL
GENERIC	DEPENDENT	0.7676850723936008	0.04729184783730201	Inf	0.6627470219347138	0.8474851504952435
MATCHED	DEPENDENT	0.7113486236458914	0.041696069810310564	Inf	0.6233846365455215	0.7858260067367233
GENERIC	SPONTANEOUS	0.5815912700924148	0.04661483090067482	Inf	0.4884665111700895	0.669242371595107
MATCHED	SPONTANEOUS	0.786122861089938	0.04408828576510067	Inf	0.6873515573471027	0.8600440888621042

Appendix 28: Additional *H2* congruence model

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: target_choice ~ congruent_exposure + gdms_type + category + (1 | response_id)
Data: dat_h2_long
Control: glmer_ctrl
```

AIC	BIC	logLik	-2*log(L)	df.resid
516.7	557.1	-248.4	496.7	410

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.1430	-1.1220	0.5346	0.6427	1.2333

Random effects:

Groups	Name	Variance	Std.Dev.
response_id	(Intercept)	0.05462	0.2337

Number of obs: 420, groups: response_id, 60

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.2481	0.3698	0.671	0.502
congruent_exposure	0.6500	0.2645	2.457	0.014 *
gdms_typeSPONTANEOUS	-0.2373	0.2287	-1.037	0.300
categoryBananas	0.5891	0.4148	1.420	0.156
categoryButter	0.3194	0.4008	0.797	0.425
categoryEggs	-0.4334	0.3819	-1.135	0.256
categoryTomatoes	0.4954	0.4095	1.210	0.226
categoryMilk	0.2361	0.3973	0.594	0.552
categoryCucumbers	0.4954	0.4095	1.210	0.226

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	cngrn_g_SPON	ctgryBn	ctgryBt	ctgryE	ctgryT	ctgryM
cngrnt_xpsr	-0.578						
g_SPONTANEO	-0.384	0.117					
categoryBnns	-0.501	0.021	-0.007				
categoryBttr	-0.515	0.012	-0.004	0.457			
categoryEggs	-0.532	-0.017	0.007	0.477	0.494		
categoryTmts	-0.506	0.018	-0.006	0.448	0.463	0.483	
categoryMlk	-0.519	0.009	-0.003	0.461	0.477	0.499	0.467
ctgryCcmbrs	-0.506	0.018	-0.006	0.448	0.463	0.483	0.454 0.467

Appendix 29: Main models diagnostics

model	singular	n obs	aic	bic	logLik	Pearson ChiSq	ratio	df	p value
H1	FALSE	847	1125.79730 85036915	1168.47261475556	-553.8986542518458	746.0394153928776	0.8902618322110711	838	0.9897674642656 654
H1a	FALSE	434	564.407020 0506401	601.0644208575437	- 273.20351002532004	380.93550916463886	0.8963188450932679	425	0.9387018826651 827
H2	FALSE	420	520.077539 2033688	560.480086316143	- 250.03876960168444	401.51416486738214	0.9793028411399565	410	0.6083739881876 005

Appendix 30: Perceived autonomy by each item

item	mean	sd	nonmissing
autonomy_1	3.994535519125683	1.056103494343441	NA
autonomy_2	3.9836065573770494	0.9107140649449933	NA
autonomy_3	4.032786885245901	0.9252361702917106	NA
autonomy_4	3.8469945355191255	1.0101567906657958	NA
autonomy_5	3.8251366120218577	0.915088332926822	NA

Appendix 31: Perceived autonomy by task version

task version	n	mean	sd	median	iqr	min	max
CONTROL	59	4.0508474576271185	0.6711731575116175	4.2	0.7000000000000002	1	5
SOCIAL	65	3.8492307692307692	0.7566627131942352	4	1.0000000000000004	1.8	5
DEFAULT	59	3.91864406779661	0.7675474996581277	4	0.9000000000000004	1.4	5

Appendix 32: Perceived autonomy by experimental condition

policy arm	n	mean	sd	median	iqr	min	max
CONTROL	59	4.0508474576271185	0.6711731575116175	4.2	0.7000000000000002	1	5
GENERIC	62	3.867741935483871	0.7867288977408006	4	0.9500000000000006	1.4	5
MATCHED	62	3.8967741935483873	0.7375119319813832	4.1	1.0000000000000004	1.8	5

Appendix 33: Perceived autonomy: robust linear model

analysis	term	estimate	conf_low	conf_high	std.error	statistic	p.value
Autonomy ~ task_version	(Intercept)	4.05084745762712	3.876947901585929	4.224747013668312	0.08812935314404734	45.96479280865723	2.155095001520175e-101
	task_versionSOCIAL	-0.2016166883963492	-0.4567112218897471	0.0534778450970487	0.1292775941419343	-1.5595640507897568	0.12061939609033717
	task_versionDEFAULT	-0.13220338983050858	-0.39638194860960907	0.13197516894859193	0.13388122448233328	-0.9874677374791556	0.3247388929817867

Appendix 34: Perceived autonomy: pairwise estimated marginal mean contrasts of the task versions

contrast	estimate	SE	df	t.ratio	p.value
CONTROL - SOCIAL	0.2016166883963492	0.1319599351710361	180	1.5278628936504817	0.2803176288901912
CONTROL - DEFAULT	0.13220338983050858	0.13511480135917167	180	0.9784523124085882	0.5914198522222319
SOCIAL - DEFAULT	-0.06941329856584061	0.13195993517103596	180	-0.5260179801988583	0.858703842000459