Selected Essays on Consumer Behavior and Construct Measurement in Market-Based Management

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

SANDRA BARBARA BARINGHORST

2024

Selected Essays on Consumer Behavior and Construct Measurement in Market-Based Management

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

vorgelegt von
SANDRA BARBARA BARINGHORST
München
2024

Referent:	Prof. Dr. Manfred Schwaiger
Korreferent:	Prof. Dr. Jelena Spanjol
Promotionsabschlussberatung:	10. Juli 2024

Table of Contents

List of Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
A_{Ad}	Attitude Toward the Ad
A_{Brand}	Attitude Toward the Brand
AI	Artificial Intelligence
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
B	Beta
BootSE	Standard Errors from the Mean Result of Bootstrapping
C2C	Consumer-to-Consumer
CI	Confidence Interval
C-OAR-SE	Construct definition, Object classification, Attribute classification, Rater identification, Scale formation, and Enumeration and reporting
coeff	Coefficient
COVID-19	Coronavirus Disease 2019
df	Degrees of Freedom
e.g.	for example (exempli gratia)
eco	Ecological
e-purchase	Electronic Purchase
F	Fisher Distribution
Gen Z	Generation Z
H&M	Hennes & Mauritz (clothing company based in Sweden)
HC3	Heteroskedasticity-Consistent Standard Errors
HIV	Human Immunodeficiency Virus
i.e.	that is (id est)

IBM	International Business Machines Corporation
ICC	Intraclass Correlation Coefficient
LLCI	Lower Limit Confidence Interval
M	Mean
Max	Maximum
MI	Multi-Item
Min	Minimum
MSE	Mean Squared Error
N	Total Number of Individuals in the Sample
p	P-value
P2P	Peer-to-Peer
PF	Pearson-Filon Statistic (test for comparing two nonindependent correlations)
PI_{Brand}	Brand Purchase Intention
R	Correlation Coefficient
R-sq	Coefficient of Determination
RQ	Research Question
SD	Standard Deviation
SE	Standard Error
SI	Single-Item
Sig.	Significance
SIPA	Subjective Information Processing Awareness
SPSS	Statistical Package für Social Sciences (statistical software)
t	t-statistic
ULCI	Upper Limit Confidence Interval
WHO	World Health Organization

WOM	Word-of-Mouth
WTP	Willingness to Pay
ZPF	Modified Pearson-Filon Statistic (obtained via Fisher's r-to-z-transformation)

Introduction

In recent years, there have been many significant events with a potential impact on consumer behavior: several wars, the COVID-19 pandemic, inflation, increasing awareness of climate change, and new technologies and digitalization, naming just a few. All of these events naturally have an impact on consumer behavior and force companies to adopt their strategies and rethink in new directions. Through faster trends in innovative technologies and new sources of information, companies are offered novel possibilities to address their target groups and offer opportunities for market-based management.

The first three studies of this dissertation, which contains a total of four essays in the broader field of market-based management, collectively aim to enhance the understanding of consumer behavior. Special regard lies on increasing awareness of sustainability, usage of social media in health-care, and personalized-pricing. The fourth study covers an overarching topic and discusses the usage of single- (SI) versus multi-item (MI) scales in marketing research.

Individuals express significant apprehension about climate change, and they readily link this concern to their own buying decisions. In fact, results from an international survey conducted by McKinsey indicate that 87% of consumers are anxious about the environmental and social consequences of the products they purchase (Bonini & Oppenheim, 2008). Consequently, consumers are striving for more sustainable consumption. However, customers frequently encounter higher prices associated with sustainable products. Only some can afford sustainably produced products or are willing to pay a higher price for them (Deloitte, 2022; Lehmann et al., 2022). To keep pace with this trend, new business models such as the second-hand economy are constantly emerging. Companies can capitalize on these new opportunities and appeal to a larger proportion of consumers (Yrjölä et al., 2021) such as those that were initially not part of their customer base (Abbes et al., 2020) as well as those with a lower willingness to pay (WTP).

Businesses are not just responding to the call for more sustainability, they are also addressing the desire for newer technologies and digitalization. In this evolving landscape, the Internet offers consumers now even more opportunities to receive and disseminate information: from seeking entertainment, utilizing messaging services, making online purchases, to looking for information on health-related topics (Brandt, 2020).

Nowadays, patients are increasingly relying on online sources regarding health-related information, as they frequently refrain from visiting a doctor due to long waiting times for appointments, or feelings of embarrassment. Further, health-care providers transfer more responsibilities to patients through a shift toward greater responsibilization. Patients face the challenges of accessing expertise, acquiring knowledge, and effectively managing complex service systems for resource integration (i.e., combining knowledge and skills to create value). Online health data-sharing platforms provide a new point of contact, allowing consumers to enhance their learning process and make improvements through interactions between consumers and physicians or among consumers (Anderson et al., 2016).

The ever-evolving Internet is also a great help to companies regarding pricing. A market typically consists of consumers with varying levels of WTP, often influenced by their income or personal preferences (Iyer et al., 2002). Due to consumers' different WTPs, companies have used personalized pricing for some time. Personalized pricing entails the practice of assigning unique prices to individual customers for the same product or service (Borgesius & Poort, 2017), allowing to capture a customer's entire WTP (Conitzer et al., 2012). The rise of the Internet has primarily enabled this form of price discrimination by screening customers' personal data (Borgesius & Poort, 2017). Companies are also becoming more creative in this area and are extending the use of (personalized) vouchers and coupons to the Internet.

To gain a deeper understanding of evolving consumer behavior, it is essential to collect data using suitable measurement instruments. Measurement scales have become indispensable in empirical research on consumer behavior and market-based management and are either used as SI or MI scales. Both have their raison d'être but require different amounts of resources. Given the scarcity of resources, quality assessments can help to select appropriate measurement tools. Initial research on reliability and validity is inconclusive, with conflicting outcomes consistently arising in literature. Some studies suggest similar predictive validity for SI and MI measures, while others underscore the superiority of MI measures. An important aspect of measurement accuracy, which has been relatively overlooked in research comparing SI and MI measures, is the concept of test-retest reliability. This measure assesses consistency, ensuring that measurements remain representative and stable over time.

In conclusion, the dissertation aims to explore trends in changing consumer behavior highlighted by a growing concern for sustainability, advancements in technology, and dynamic pricing strategies. While consumers aim for more sustainable choices, challenges arise with higher prices for such products. The emergence of new business models, coupled with advancements in technology and the Internet, provides companies with opportunities to adapt. Health-care dynamics are shifting as patients increasingly turn to online sources, especially social media. Further, digitalization influences pricing strategies, with personalized coupons leveraging consumer data. Finally, the debate on measurement scales in consumer behavior research adds to the ongoing quest for a deeper understanding of this dynamic field. Overall, these elements showcase the multifaceted nature of contemporary consumer behavior. Each study attempts to address and answer research gaps and to provide guidance to practitioners. To contribute to relevant and rigorous research, all four studies use established scientific methods.

Study I, titled "Investigating the Role of Consumer Brand Forgiveness in Second-Hand Consumption", examines consumer reactions toward company-sold second-hand products with a special focus on product failures. This study was presented at the *2023 Global Marketing Conference (GMC)* in Seoul, South Korea.

Existing research regarding this topic has primarily focused on explaining consumer motivations for purchasing second-hand items, but it has largely overlooked the potential repercussions for businesses that sell their own brand's products second-hand. Globally, industry experts anticipate substantial growth in the second-hand market for fashion products (ThredUp, 2022). Prominent fashion brands increasingly recognize this emerging trend and now provide second-hand clothing through their dedicated online platforms (Hubert, 2022). While the demand for second-hand products continues to grow, and an increasing number of companies are venturing into the re-commerce industry, existing research has predominantly concentrated on the reasons behind second-hand consumption. Especially the discussion surrounding more sustainable consumer choices is fueling the increasing popularity of second-hand products (Ek Styvén & Mariani, 2020; Guiot & Roux, 2010). Since pre-owned products have already been used, one could assume that they are also more prone to product failure. For companies, the aftermath of a brand transgression is primarily detrimental and can potentially harm the consumer-brand connection (Aaker et al., 2004).

The paper consists of two pretests and one main study. The main study uses an experimental design and reveals that consumers are more likely to forgive product failures in second-hand products rather than new ones. These higher forgiveness intentions lead to a reduction in negative word-of-mouth (WOM), an increase in intentions to purchase new products of the transgressing brand, and a more favorable brand attitude. Consumers' underdog perceptions toward pre-owned products contribute to this phenomenon. The underdog effect refers to a phenomenon in which individuals harbor favorable attitudes and support less powerful entities (Kim et al., 2019), often overlooking drawbacks such as their lower chances of success (Goldschmied & Vandello, 2012). However, the study found that this effect is not universally advantageous: Strong passion and determination toward the product can lead to a greater willingness to forgive, but external disadvantages associated with the brand may hinder consumer forgiveness intentions. The results expand the understanding of how product failures affect second-hand items. Given the unique characteristics and consumer motivations associated with second-hand purchases, existing insights from the product failure literature for new products may not be directly applicable. Further, it introduces two factors that mediate the connection between failing second-hand products and subsequent consumer responses.

Study II, “Examination of the Role of Social Media in Health-Care”, co-authored with Manfred Schwaiger and Louisa Weritz, investigates a patient’s intention to visit a doctor after consulting Instagram. The underlying study was presented at the *2023 AM&HCR* conference in Crested Butte, USA.

Over recent years, there has been a substantial increase in online interactions on social media. Initially, informal interactions were prevalent, but now, followers seek specific content and information, with a growing emphasis on health-related information. Research indicates that individuals are already using social media to seek health-related information (Brandt, 2020), and approximately one in three people have engaged in discussions about health-related topics on social platforms (Honigman, 2015). This paper focuses on the social media platform Instagram and examines whether factors such as influencers' personal experiences, medical qualifications, and the number of their followers impact patients' willingness to seek medical advice.

An online experiment was set up to test whether a patient's willingness to see a doctor changes after seeing an Instagram post. Descriptive findings reveal that nearly 28% of patients reconsider visiting a doctor after encountering a post on Instagram.

In an initial descriptive analysis, it was discovered that almost 28% of patients alter their decision whether to consult a doctor following exposure to an Instagram post. Subsequent analyses involving a mixed ANCOVA and a logistic regression revealed that expert advice on Instagram, along with the suggestion to avoid visiting a doctor, heightens the likelihood of patients opting to delay their appointments or refraining from seeing a doctor altogether. However, none of the other factors (i.e., patient influencer or popularity) yielded a significant result. This study underscores the influence of Instagram posts on shaping a patient's perspective and suggests that especially an influencer’s perceived medical expertise has an impact on health-care decisions. Further, the study extends the existing literature on health-care and addresses the need for more research on the effectiveness of health communication through social media platforms.

Study III, “Investigating the Suitability of Customized Coupons for Personalized Pricing”, co-authored with Manfred Schwaiger and Louisa Weritz, examines personalized coupons, a promising yet controversial strategy for revenue increase.

Customized coupons, a specific framing method of personalized pricing, are currently in practical use but lack thorough research. They could offer a solution to negative consumer reactions to personalized pricing. This can be exemplified by the case of the globally operating Japanese lifestyle brand Muji, which demonstrated positive outcomes after implementing personalized coupons (Treasure Data, 2021). Study III first derives a range, in which personalized coupons are perceived to be fair based on consumer’s self-assessment and further investigates this range using an experimental setting.

First, participants were surveyed on their perception of how fair personalized coupons are perceived, providing them with their own personalized voucher and a friend’s customized coupon. Using the van Westendorp method (1976), customers deem personalized coupons fair when a friend receives a voucher for up to 25% compared to their own 10% personalized voucher. Despite the initial assumption that differences within the self-assessed range where coupons differences are tolerated would avoid negative customer reactions, an experimental setting demonstrated otherwise. Even when individual coupon differences fall within the pre-established range, customers exhibit negative reactions, challenging the reliability of self-assessment in indicating perceived fairness. Notably, customers seldom express public complaints even when dissatisfied, and their loyalty remains largely unscathed in most situations. The study implies that if long-term loyalty endures, the expected loss of intangible assets (e.g., corporate reputation) may be less substantial than initially thought.

From a managerial perspective, the study prompts considerations regarding balancing financial gains and potential intangible losses. Theoretical implications underscore the adverse effect of personalized pricing on perceived fairness, aligning with equity and distributive justice (Adams, 1965) and social comparison theory (Festinger, 1954).

Study IV, titled “Single- Versus Multi-Item Scales: A Comparison of Test-Retest Reliability”, co-authored with Maximilian Niederberger-Kern and Manfred Schwaiger, deals with reliability issues of SI and MI measurements.

Since the 1970s, MI scales have been used in marketing research, whereas SI scales are more frequent among self-reported facts or demographics. After years of established practices being widely accepted, Bergkvist and Rossiter (2007) rekindled the marketing community's discussion regarding the use of SI and MI measures. Based on Rossiter's (2002) C-OAR-SE procedure, which suggests that constructs with both a singular and concrete object and a concrete attribute do not necessitate MI measures, the authors questioned the prevailing practice of employing MI scales, providing both theoretical and empirical arguments. One aspect of measurement accuracy that has yet to be extensively explored in research concerning the comparison between SI and MI measures is test-retest reliability. Test-retest reliability serves as a measure of consistency, guaranteeing that measurements remain representative and stable over time.

This study uses a within-subjects design to explore participants' attitudes toward different brands and products and their purchase intentions after being exposed to diverse advertisements. It contributes to the ongoing discussion about the reliability of SI scales in academic research, specifically the test-retest context. Test-retest reliability is assessed through the intraclass correlation coefficient (ICC), comparing SI and MI scales over two-week and four-month periods. The results suggest that SI scales perform as well as, or even better than, MI scales over time. It is important to note that this does not advocate for SI scales over MI scales but underscores the importance of careful scale selection, especially in longitudinal research.

References

Aaker, J., Fournier, S., & Brasel, S. A. (2004). When Good Brands Do Bad. *Journal of Consumer Research*, *31*(1), 1–16. https://doi.org/10.1086/383419

Abbes, I., Hallem, Y., & Taga, N. (2020). Second-Hand Shopping and Brand Loyalty: The Role of Online Collaborative Redistribution Platforms. *Journal of Retailing and Consumer Services*, *52*, 1–11. https://doi.org/10.1016/j.jretconser.2019.101885

Adams, J. S. (1965). Inequity In Social Exchange. *Advances in Experimental Social Psychology*, *2*, 267–299. https://doi.org/10.1016/S0065-2601(08)60108-2

Anderson, L., Spanjol, J., Jefferies, J. G., Ostrom, A. L., Nations Baker, C., Bone, S. A., Downey, H., Mende, M., & Rapp, J. M. (2016). Responsibility and Well-Being: Resource Integration under Responsibilization in Expert Services. *Journal of Public Policy & Marketing*, *35*(2), 262–279. https://doi.org/10.1509/jppm.15.140

Bergkvist, L., & Rossiter, J. R. (2007). The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs. *Journal of Marketing Research*, *44*(2), 175–184. https://doi.org/10.1509/jmkr.44.2.175

Bonini, S. M. J., & Oppenheim, J. M. (October 2008). *Helping "Green" Products Grow*. https://acervo-digital.espm.br/clipping/20081110/helping_green_products_grow-1.pdf

Borgesius, F. Z., & Poort, J. (2017). Online Price Discrimination and EU Data Privacy Law. *Journal of Consumer Policy*, *40*(3), 347–366. https://doi.org/10.1007/s10603-017-9354-z

Brandt, M. (2020, March 17). *Sprechstunde bei Doktor Google*. https://de.statista.com/infografik/8497/nutzung-des-internets-fuer-gesundheits-informationen/

Conitzer, V., Taylor, C. R., & Wagman, L. (2012). Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases. *Marketing Science*, *31*(2), 277–292. http://www.jstor.org/stable/41488023

Deloitte. (October 2022). *Sustainability Trend Under Pressure: With Money Tighter Than Ever, How Can Companies Make a Difference Now?* https://www2.deloitte.com/de/de/pages/consumer-business/articles/studie-nachhaltigkeit-und-verbraucherverhalten.html

Ek Styvén, M., & Mariani, M. M. (2020). Understanding the Intention to Buy Secondhand Clothing on Sharing Economy Platforms: The Influence of Sustainability, Distance from the Consumption System, and Economic Motivations. *Psychology & Marketing*, *37*(5), 724–739. https://doi.org/10.1002/mar.21334

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Goldschmied, N. P., & Vandello, J. A. (2012). The Future is Bright: The Underdog Label, Availability, and Optimism. *Basic and Applied Social Psychology*, *34*(1), 34–43. https://doi.org/10.1080/01973533.2011.637726

Guiot, D., & Roux, D. (2010). A Second-Hand Shoppers' Motivation Scale: Antecedents, Consequences, and Implications for Retailers. *Journal of Retailing*, *86*(4), 383–399. https://doi.org/10.1016/j.jretai.2010.08.002

Honigman, B. (2015). *24 Stats on How Social Media has Impacted the Healthcare Industry*. https://getreferralmd.com/2013/09/healthcare-social-media-statistics/

Hubert, J. (2022, March 24). *Umsatz ausgewählter Anbieter von Secondhandkleidung in Deutschland im Jahr 2019: (in Millionen Euro)*. https://de.statista.com/statistik/daten/studie/1198643/umfrage/umsatz-ausgewaehlter-anbieter-von-secondhandmode/

Iyer, G. R., Miyazaki, A. D., Grewal, D., & Giordano, M. (2002). Linking Web-Based Segmentation to Pricing Tactics. *Journal of Product & Brand Management*, *11*(5), 288–302. https://doi.org/10.1108/10610420210442175

Kim, Y., Park, K., & Stacey Lee, S. (2019). The Underdog Trap: The Moderating Role of Transgression Type in Forgiving Underdog Brands. *Psychology & Marketing*, *36*(1), 28–40. https://doi.org/10.1002/mar.21155

Lehmann, K., Renz, D., & Huber, F. (2022, October 26). *Nun sag', wie hast du's mit der Nachhaltigkeit?* EY Parthenon. https://www.ey.com/de_de/consumer-products-retail/studie-nachhaltigkeit-deutscher-konsument-innen

Rossiter, J. R. (2002). The C-OAR-SE Procedure for Scale Development in Marketing. *International Journal of Research in Marketing*, *19*(4), 305–335. https://doi.org/10.1016/S0167-8116(02)00097-6

ThredUp. (2022). *Resale Report: 2022*. https://www.thredup.com/resale/static/2022-resaleReport-full-92a77020598ceca50f43227326100cc2.pdf

Treasure Data. (2021). *Muji Grows In-Store Revenue with Digital Savvy*. https://get.treasuredata.com/rs/714-XIJ-402/images/muji-case-study.pdf

van Westendorp, P. H. (1976). *NSS-Price Sensitivity Meter (PSM)- A New Approach to Study Consumer Perception of Price.* https://archive.researchworld.com/a-new-approach-to-study-consumer-perception-of-price/

Yrjölä, M., Hokkanen, H., & Saarijärvi, H. (2021). A Typology of Second-Hand Business Models. *Journal of Marketing Management*, *37*(7-8), 761–791. https://doi.org/10.1080/0267257X.2021.1880465

Sandra Baringhorst

I Investigating the Role of Consumer Brand Forgiveness in Second-Hand Consumption

Abstract[1]

Second-hand markets are a rising field for consumers to buy and sell their products. In recent years, more and more companies have entered this re-commerce business. To date, research has mainly examined second-hand consumption regarding consumers' motivations and reactions. The business perspective and consequences of the sales of branded second-hand products have been neglected so far. To fill that research gap, this study proposes an experimental research design that examines consumer reactions after a negative experience with branded second-hand products sold directly by the company that owns the brand. The study finds that consumers are more willing to forgive a transgressing second-hand product than a transgressing new product. Through higher forgiveness intentions, companies benefit from less negative word-of-mouth, higher purchase intentions toward new products of the transgressing brand, and a stronger brand attitude. Further, the underdog effect is proposed as one possible explanation of this phenomenon. It was found that customers have a higher underdog perception toward pre-used products compared to new products. The underdog dimension of passion and determination leads to higher forgiveness intentions, whereas the dimension of external disadvantage rather works as a suppressor for consumer brand forgiveness. Marketers have to keep these findings in mind when designing promotion strategies for their pre-used products. Besides implications for marketers, the study derives avenues for further research.

[1] A previous, mainly conceptual version of this article was submitted as a project study for the Master of Business Research at LMU Munich in March 2022. I am grateful to Julia Wilhelm for pretesting the scenarios as part of her Master's thesis in August 2022. I served as a supervisor for her Master's thesis and provided the research questions, scenarios, and directional guidance for the research framework and measurements based on my project study. For the pretests and main study presented in this paper, I have collected new data and the data analysis was performed exclusively by me.

1 Motivation

In recent years, it has been observed that more and more consumers are following the growing trend of buying second-hand products, especially clothing. Worldwide, the second-hand apparel market is expected to grow three times faster than the overall clothing market. From 2021 to 2026 the global market for pre-used fashion goods is expected to grow by 127% (ThredUp, 2022). This trend has not left Germany unscathed. In a study conducted by Kantar in 2021 with one of the largest resellers in Germany, 67% of all study participants stated that they have bought second-hand products. This represents an increase of eleven percentage points compared with the previous year (Momox, 2020, 2022). Consumers buy second-hand clothes not only for themselves. 18% of the buyers of second-hand fashion stated that they also buy worn clothing for their children (Momox, 2022).

Well-known brands in the fashion industry, such as H&M with their online-store *Sellpy* (Brandt, 2021; Weidemann, 2020), are also becoming aware of this rising trend and offer pre-owned apparel via their own online second-hand platforms (Hubert, 2022). The number of these brand-owned second-hand stores has increased by 275%, from eight stores in 2020 to 30 stores in 2021 (ThredUp, 2022). This increase could be due to the fact that in the past, pre-used goods have mainly been considered to cannibalize turnover of new products (Abbes et al., 2020). However, by selling their pre-used apparel, brands can save advertising costs (Strähle, 2021), attract new customer groups, and benefit from upselling effects (Miller & Brannon, 2022), thereby mitigating the risk of cannibalizing themselves (Ghose et al., 2006).

Although the demand for second-hand products continues to rise and more and more companies are entering the re-commerce business, research to date has focused almost exclusively on the motives for second-hand consumption: The debate about more sustainable consumption is making second-hand products increasingly popular. Besides sustainability aspects and the desire to distance oneself from the consumer society (Ek Styvén & Mariani, 2020; Guiot & Roux, 2010), economic reasons are also important for the decision to purchase pre-owned goods (Roux & Guiot, 2008). Fashionability reasons, the search for unique garments, and nostalgic pleasure also play an important role (Roux & Guiot, 2008). When entering the re-commerce market, companies and brands can take advantage of this knowledge. However, it remains unclear how an entry into the second-hand market might affect the brand, as research so far has mainly focused on the sale of second-hand products via peer-to-peer (P2P) platforms (Roux & Guiot, 2020).

Second-hand products differ from new products in various characteristics: Pre-used products are cheaper (Bardhi & Arnould, 2005; Guiot & Roux, 2010), might evoke nostalgic feelings, and are perceived as being original and authentic (Guiot & Roux, 2010). They could therefore encourage consumers to engage in indulgent consumption behavior (Parguel et al., 2017). Previously successful marketing strategies for new products cannot be applied to second-hand products without further ado. The peculiarities of second-hand products influence not only marketing strategies, consumer purchase intentions, and word-of-mouth (WOM) behavior (Lo et al., 2019) but can also impact the brand if a product commits a failure. Product failures are only a matter of time (Hassey, 2019) and are most likely experienced by all brands in the course of their existence (Fetscherin & Sampedro, 2019). Consequences for the brand after a transgression are mainly destructive and can negatively affect the consumer-brand relationship (Aaker et al., 2004), purchase intentions, brand image, and WOM behavior (Grégoire & Fisher, 2008; Smith & Bolton, 1998). These findings primarily relate to new products. However, it initially seems obvious that the same pattern should apply to transgressing second-hand products, especially as these products are more susceptible to defects and therefore militate against a company's entry into the re-commerce business. Nevertheless, this tendency has not yet been observed in the market. In fact, the trend toward second-hand consumption continues to grow (ThredUp, 2022). Hence, it is worthwhile to examine whether consumer outcomes after negative experiences with second-hand products differ from those of negative experiences with new products:

> ***RQ 1:*** *Do consumer reactions differ if a product failure occurs with second-hand products compared to a failure occurring with new products?*

The assessment of the severity of a misconduct performed by a pre-used product could differ from that of new products due to a violation of basic needs. Second-hand products are perceived as being riskier than new goods (Bezançon et al., 2019) and a fear of contagion might deter consumers from purchasing pre-owned goods (Bardhi & Arnould, 2005). Consumers therefore purchase second-hand products with different initial expectations and are already aware of this fact when deciding to purchase pre-used products. Therefore, consumers might judge a transgression by second-hand products differently and hence, be more forgiving.

Although forgiveness plays an important role in interpersonal (McCullough et al., 1998) as well as corporate and brand relationships, scarce research on forgiveness intentions, especially in the role as a mediator, has been conducted (Tsarenko & Tojib, 2015). What is known, however, is that forgiveness is an important mechanism in restoring brand relationships (Christodoulides et al., 2021). Consumers are less likely to switch to competing brands (Fetscherin & Sampedro,

2019) or to engage in malign WOM behavior (Casidy et al., 2021; Fetscherin & Sampedro, 2019) after forgiving a transgressing brand. Factors influencing forgiveness intentions are ambiguous (Yao et al., 2017) and have, so far, mainly been examined in the usage of new products. It remains unclear whether forgiveness intentions depend on the type of product (Fetscherin & Sampedro, 2019) or whether it matters if the product is a second-hand product. The second research question therefore investigates whether a pre-used product increases the likelihood of consumers engaging in forgiving behavior after a transgression:

> ***RQ 2:*** *Are consumers more forgiving of negative product experiences with second-hand products than with new products?*

The motives of purchasing second-hand products resemble those of supporting an underdog. The underdog effect describes a phenomenon where consumers exhibit positive attitudes toward a weaker entity, while ignoring certain disadvantages of the same (Goldschmied & Vandello, 2012; Kim et al., 2019). Unexpected wins can trigger higher satisfaction or mitigate negative experiences with the underdog (Vandello et al., 2007). Such positive emotions can increase forgiveness intentions (e.g., Hegner et al., 2017; Rahman et al., 2021; Schnebelen & Bruhn, 2018) and therefore serve as mediators for the relationship between the transgressing product and the following forgiveness intentions. This should be especially true for fast fashion products, as these garments are subject to strong fashion cycles, have a limited lifetime, and tend to be in the low-price segment. Higher priced products with longer life cycles, such as luxury products, are excluded from the study as they demonstrate a distinct status among customers; hence, different mechanisms might play a role, and the underdog effect might not be applied equivalently.

The underdog effect is enjoying growing interest in the marketing and consumer behavior literature. So far, the underdog effect has mainly been examined in the form of storytelling or underdog brands (Delgado-Ballester, 2021). Research has also looked at different product types' (e.g., hedonic products versus functional products) (Li & Zhao, 2018) outcomes regarding consumer forgiveness intentions toward a transgressing underdog brand (Kim & Park, 2020; Kim et al., 2019). However, research on the underdog effect in consumer behavior literature remains scarce, and no link has yet been established with used products. Because the motives of supporting an underdog and purchasing second-hand products are similar (Jin & Huang, 2019; McGinnis & Gentry, 2009; Paharia et al., 2011; Schmidt & Steenkamp, 2022), it could be likely that second-hand products are perceived as an underdog and therefore, retaliatory intentions may be attenuated. Hence, the third research question investigates whether used and new products are perceived differently in terms of the underdog effect:

> ***RQ 3:*** *Are second-hand products, compared to new products, more likely to be perceived as an underdog?*

Due to different characteristics of pre-used goods compared to new ones and motivations to purchase them, the existing findings regarding new products in the product failure literature cannot be applied directly to second-hand products. This paper answers calls for future research on second-hand consumption (Crosno & Cui, 2018), consumer brand forgiveness (Tsarenko & Tojib, 2011; Xie & Peng, 2009), and the underdog effect (Paharia et al., 2011). It further proposes a combination of the underdog effect and consumer brand forgiveness to serve as serial mediators in the relationship between transgressing second-hand products and subsequent consumer reactions.

The findings contribute to the extension of the academic literature in two main ways. First, the findings extend the knowledge of the effects of product failures regarding second-hand products. Due to different characteristics of and motivations when buying second-hand products, the existing findings regarding new products in the product failure literature cannot be directly applied to second-hand products. Second, two main constructs are proposed to mediate the relationship between transgressing second-hand products and subsequent consumer reactions, namely consumer brand forgiveness and the underdog effect. Consumer brand forgiveness has not yet been examined regarding different product types (Fetscherin & Sampedro, 2019). Different perceptions of second-hand and new products, that could be explained by the underdog effect, could in consequence justify different levels of forgiveness intentions. These findings thus support the decision-making process for companies when deciding whether to offer their own products on the second-hand market.

Marketing research has mainly looked at the consumption of pre-owned products via P2P platforms (Roux & Guiot, 2020) and neglected the brands' point of view. Due to the increasing interest in the sales of second-hand products by the brands themselves, it needs to be worked out how negatively a failure regarding a second-hand product can directly affect the brands. This study additionally provides initial recommendations for managers for their decision-making process. Managers might profit from realizing how their brands are perceived, especially when broadening their portfolio with new product categories (i.e., second-hand products). The decision-making process regarding whether to offer second-hand products can be simplified by showing multiple consequences that go along with this decision. Managers can adapt their communication strategy accordingly and might therefore be able to attract new customer segments such as eco-conscious audiences or customers that would not be able to afford their products new. Further, marketers might use the underdog effect in order to tell a fitting story around the introduction of second-hand products. By knowing which dimensions of the underdog effect support consumers in forgiving a transgressing product, companies can adapt their recovery strategies accordingly and thereby strengthen the forgiveness process.

In summary, the proposed study will answer several calls for further research on second-hand consumption (Guiot & Roux, 2010; Lo et al., 2019), by examining recommendation and purchase intentions as well as brand attitude toward new and unused products of a brand after consuming second-hand products of the very same brand. For this purpose, the study takes a company perspective and assumes that brands sell their own products pre-owned.

The remainder of this paper is structured as follows: First, the literature on second-hand consumption, consumer brand forgiveness, and the underdog effect is presented, and hypotheses are derived accordingly. The study design, comprising two pretests and one main study is introduced before the findings are analyzed. The paper concludes with a critical discussion of the findings, theoretical and managerial implications, and outlines avenues for further research.

2 Literature Review and Hypotheses

2.1 Status Quo of Research on Second-Hand Consumption

Definition and Delimitation

Second-hand consumption is defined as "the acquisition of used objects" (Roux & Guiot, 2008, p. 66). It can be seen as a form of collaborative consumption (Becker-Leifhold & Iran, 2018) in which products have already been owned and/or used by a third person (Sihvonen & Turunen, 2016). Often, the term second-hand is used synonymously with the term vintage. There is, however, a substantial difference: Vintage products are previously owned, but not necessarily used. Hence, not all second-hand products need to be old, and not all vintage goods are used (Cervellon et al., 2012). Moreover, vintage items represent a specific era, mostly between the 1920s and 1980s and are consequently no longer available on the market (Cervellon et al., 2012). They are mostly bought as collectibles (Sihvonen & Turunen, 2016) which carry a unique historical background. This justifies a higher price for vintage products compared to second-hand products. Therefore, the motives of consuming vintage products, nostalgia and fashion involvement (Cervellon et al., 2012) among others, differ from those of buying second-hand products. The focus of the underlying study is on second-hand products rather than on vintage products.

Motivations to Purchase Second-Hand Products

Popularity of second-hand consumption emerged in the 18th century during the industrialization (van Damme & Vermoesen, 2009). Later, with the introduction of mass production, everyday products were offered at lower prices and were therefore readily available to the general public, thus making second-hand products, that were initially bought due to economic reasons, obsolete (Weinstein, 2014). From there on, second-hand products were increasingly associated with poverty and deprivation (Hamilton, 2009), as a consequence the demand dropped significantly. In the last decades, this point of view has gradually changed. Whereas second-hand consumption for a long time has mainly been seen as a thriftily way to shop, it is now rather driven by recreational, critical, hedonic (Guiot & Roux, 2010; Roux & Guiot, 2008), and fashionability reasons (Ferraro et al., 2016).

One of the main drivers of consuming pre-owned products are *utilitarian* motives. These appear in the form of frugality (Cervellon et al., 2012) as consumers try to preserve money (Bardhi & Arnould, 2005; Guiot & Roux, 2010). By buying second-hand products at cheaper prices, consumers can afford well-known brands without restricting their purchases (Guiot & Roux, 2010).

This could lead to impulsive buying behavior (Parguel et al., 2017). Consumers also purchase second-hand products due to *hedonistic* motives. Some are looking for the thrill of treasure hunting or feelings of nostalgia (Guiot & Roux, 2010); others prefer second-hand products due to their authenticity (Guiot & Roux, 2010) or uniqueness (DeLong et al., 2005; Ferraro et al., 2016). When looking at *biospheric* motives, the link between second-hand consumption and sustainability is not clear. Whereas Cervellon et al. (2012) observed that second-hand consumption is only indirectly driven by environmental awareness through bargain hunting, Ek Styvén and Mariani (2020) discovered a direct link between the attitude toward purchasing second-hand products and their sustainability. The attitude that buying pre-owned products has a positive impact on the environment is positively related to second-hand purchases (Borusiak et al., 2020). In addition, consumers often use sustainability aspects to justify their purchase decision for second-hand products (Silva et al., 2021). Moreover, second-hand consumption provides a way for consumers to disassociate themselves from consumer society (Ek Styvén & Mariani, 2020; Guiot & Roux, 2010).

Consumer Reactions and Second-Hand Consumption

Companies and brands also take advantage of these motives to sell products a second time. To measure their success, various indicators are used to assess consumer responses. One important indicator is (re)purchase intention. Three groups can be made out as more likely to (re)purchase pre-owned products: Satisfied customers (as with sales of new products) (Ashfaq et al., 2019), customers looking for bargains and good deals, and customers holding positive attitudes toward second-hand consumption (Padmavathy et al., 2019). However, pre-owned products run the risk of being perceived as unsanitary. This is especially true when second-hand products are advertised with persons wearing the product. Such advertisement might lead to a decreased purchase intention. These negative outcomes can be circumvented by such simple measures as presenting second-hand products in their original packaging (Bezançon et al., 2019).

Price presentations also play a decisive role in the sale of second-hand products. Purchase intentions are higher when prices for pre-owned products are presented as all-inclusive (i.e., the displayed prices include shipping costs) rather than partitioned pricing (i.e., the additional shipping fee is not yet included in the price). However, this phenomenon is reversed for high quality brands. This effect can be explained by customers using different decision frames for the purchase of new versus second-hand products (Crosno & Cui, 2018).

In addition, indulgent consumption could be stimulated as consumers might justify their purchases through sustainable behavior and buy more goods accordingly (Parguel et al., 2017). Consumers often compare prices of second-hand products to prices of new products. This is especially true with regard to expensive goods (Sihvonen & Turunen, 2016). Consumers purchasing second-hand products perceive higher levels of uncertainty (Fernando et al., 2018). Hence, second-hand products are usually priced lower compared to their unused counterparts (Guiot & Roux, 2010). However, in a consumer-to-consumer (C2C) context, sellers often charge higher prices for their products as they feel emotionally attached to them (Genesove & Mayer, 2001).

Further, information on products being tested for functionality before re-selling them to consumers can increase consumer's trust in second-hand products. This leads to a higher willingness to purchase a second-hand product (Lee & Lee, 2005). However, especially in a C2C-setting, consumers often lack trust with regard to variability, size, and quality of clothing (Armstrong et al., 2016).

Another important factor is WOM intention, as it is recognized as influential when it comes to consumer reactions (Daugherty & Hoffman, 2014). Promoting the hedonic and ethical benefits of second-hand consumption leads to a higher perceived social acceptability of second-hand garment shopping. This further increases willingness to recommend second-hand consumption. For economic benefits, a direct link on willingness to recommend second-hand purchasing was found (Lo et al., 2019).

Activities to Promote Second-Hand Products

The purchase of pre-owned products is still associated with unsanitariness and might elicit feelings of embarrassment (Silva et al., 2021). Hygiene concerns and a desire for new products are the most often mentioned barriers when it comes to purchasing second-hand products. Marketing activities to present pre-owned products in a positive light are creative. They often include efforts that make second-hand products look like new; however, these have proven difficult so far (Bezançon et al., 2019). Nevertheless, second-hand consumption has become more and more popular in recent years. To date, marketing research has mainly looked at sales between consumers (Roux & Guiot, 2020) and disregarded the business perspective. However, more and more companies enter the re-commerce business as they take advantage of the rising second-hand trend: They profit from several benefits, like upselling effects due to loyalty toward a brand (Miller & Brannon, 2022). Further, companies can benefit from attracting customers that try their products second-hand at a cheaper price first (Abbes et al., 2020). So far, most studies have focused on the purchasing behavior regarding second-hand products, however, potential crossover effects onto the brand have been neglected. In addition, no research regarding second-hand products causing a transgression has been conducted so far.

2.2 Product Failures of Second-Hand Products

The aforementioned consumer reactions toward the consumption of second-hand products might differ depending on the experience with the purchased product. Experiences with products can either be positive or negative and may be activated through direct or indirect contact. An example for direct contact would be product usage, while indirect contact might occur via advertisements (Meyer & Schwager, 2007). Brand and product failures cannot be completely avoided (Hassey, 2019) and can have severe consequences for the brand. Reactions toward a transgressing brand or product depend on the severity of the transgression. Product and brand failures can affect the consumer-brand relationship (Aaker et al., 2004). In addition, purchase intentions might decrease (e.g., Smith & Bolton, 1998) and negative WOM intentions could increase (e.g., Grégoire & Fisher, 2008). Research to date has mainly focused on transgressions regarding new products. Whether responses to transgressions regarding second-hand products are the same or whether they differ from transgressions regarding new products remains unclear. Since negative experiences with new products trigger negative consumer reactions, it can be assumed that this holds true for second-hand products as well.

Second-hand products differ from new products in several characteristics and consumers' purchase motivations are different when it comes to these. Therefore, we must consider that these differences between new and pre-owned goods could have an impact on the relationship between experiences with the product and the resulting consumer responses. Pre-owned goods are offered at a lower price (Bardhi & Arnould, 2005; Guiot & Roux, 2010), which may increase the amount purchased. They are further perceived as original and authentic and evoke nostalgic feelings with consumers (Guiot & Roux, 2010). These characteristics might alter consumer reactions following a product failure. It is therefore proposed that in the case of a product failure concerning pre-owned products, the negative effect on WOM or purchase intentions could be reversed or at least mitigated. Consumers might be more likely to exert positive attitudes toward the transgressing brand and display positive crossover effects on new products of the transgressing brand:

H1: When experiencing a product failure with a second-hand product compared to a new product, consumers…

a) *…will be less likely to engage in negative WOM behavior.*
b) *…will be more likely to purchase a new product of the transgressing brand.*
c) *…have a more positive attitude toward the transgressing brand.*

Additionally, in a positive experience with a second-hand product, consumers could encounter a positive disconfirmation (Oliver, 1980). This could be due to the fact that their initial expectations regarding the second-hand product in form of positive experiences with pre-owned products are exceeded. Further, a halo effect might occur (Koschate-Fischer et al., 2019) and experiences with used products might be transferred to other characteristics of the brand. Such a halo effect was not only found to carry over to the brand itself, but also to related products (Ahluwalia & Gürhan-Canli, 2000), to similar attributes of the same product (Ahluwalia et al., 2001), or to brand extensions (Fedorikhin et al., 2008). This might result in an increased purchase intention for new products of the brand. Hence, it is proposed that if consumers have a positive experience with pre-owned products compared to new products, they are more likely to buy new products of the same brand:

H2: When experiencing a positive experience with a second-hand product compared to a new product, consumers…

a) *…will be more likely to engage in positive WOM behavior.*
b) *…will have a more positive attitude toward the brand.*

These hypotheses could be explained by the diverse attitudes, consumers hold toward different categories of products. A further indication could be provided through consumer brand forgiveness. Consumer brand forgiveness might explain why a relationship is weakened less after a product failure regarding second-hand products compared to one regarding new products (Christodoulides et al., 2021).

2.3 The Mediating Role of Consumer Brand Forgiveness

Consumer forgiveness is defined as a *"consumer's willingness to give up retaliation, alienation, and other destructive behaviors, and to respond in constructive ways after an organizational violation of trust and the related recovery efforts"* (Xie & Peng, 2009, p. 578). Similar to interpersonal forgiveness (McCullough et al., 1998), consumer brand forgiveness can be triggered after a transgression by a brand (Christodoulides et al., 2021), as consumers tend to have close relationships with certain products or brands. Forgiveness is seen as an important process in repairing interpersonal relationships (Tsang et al., 2006). It helps consumers regain their trust in the transgressing brand (Xie & Peng, 2009). Christodoulides et al. (2021, p. 1690) define consumer brand forgiveness as "*the consumer's cognitive, affective, and behavioral response to a brand's perceived wrongdoing, with the aim of maintaining a constructive relationship with the brand.*" *Affective forgiveness* is an emotional motive to forgive the culprit (Worthington et al., 2015). It brings along feelings of betrayal and a loss of trust in the brand (Christodoulides et al., 2021). Product failures elicit different kinds of negative emotions, such as hate, fear, or sadness. These might lead to reservations in consumers' willingness to engage with the brand again. Within the *cognitive* dimension, the brand will be reassessed after the transgression (Christodoulides et al., 2021). Consumers therefore first assess the severity of the failure (Fetscherin & Sampedro, 2019). *Behavioral forgiveness* "encompasses consumers' behavioral intentions" (Christodoulides et al., 2021, p. 1690), which often precedes emotional forgiveness (Tsarenko & Tojib, 2012).

Factors influencing forgiveness intentions are ambiguous (Yao et al., 2017). It is, for example, unclear whether forgiveness intentions depend on the type of product (Fetscherin & Sampedro, 2019) or if the product is pre-used. With the social acceptance of second-hand consumption increasing in recent years, consumers tend to talk more about their purchasing experiences with pre-owned goods. Consumers increasingly engage in (negative) WOM behavior (Lo et al., 2019), in which they find a way to vent their anger. Such negative feelings and public exchanges might decrease the willingness to forgive a transgressing brand (Christodoulides et al., 2021; Joireman et al., 2016). Providing consumers with guidelines that impose restrictions on their language of complaints could mitigate negative emotions (Christodoulides et al., 2021). Subsequently, positive emotions, such as empathy (Wei et al., 2020), can increase consumer brand forgiveness (Rahman et al., 2021).

Further, forgiveness intentions depend on company-related aspects. The severity of the misconduct (Fetscherin & Sampedro, 2019) or the brand's perceived control over the transgression (Yagil & Luria, 2016) can influence consumer brand forgiveness. The question of controllability cannot be answered unambiguously for a transgressing second-hand product. This is particularly difficult to assess when second-hand products are purchased via platforms. On platforms, the controllability could be attributed to either the seller or the platform provider (Padmavathy et al., 2019; Suri et al., 2019). But even if the brand itself sells pre-used products, the question regarding responsibility attribution remains unanswered. Consumers tend to be more forgiving toward less controllable transgressions. However, this is only the case when failures can be attributed to an individual provider (e.g., individual entity selling on ebay) compared to when the enabling organization (e.g., ebay) is to be accused. On the other hand, for highly controllable transgressions, no matter who is to blame, consumers tend to be less forgiving (Suri et al., 2019).

The assessment of the severity of misconduct by a second-hand product could differ from that of new products due to a violation of basic needs. A customer's need for health or hygiene might be violated: Previous owners could have transferred their odor to the goods (Roux, 2010). Further, pre-used products are perceived as riskier than new goods (Bezançon et al., 2019) and bring along a fear of contagion or unsanitary perceptions (Gullstrand Edbring et al., 2016). However, consumers who buy second-hand products are usually aware of these facts. Therefore, there is the possibility that consumers weigh transgressions regarding second-hand products less harshly compared to ones regarding new products. Referring to the confirmation-disconfirmation-paradigm (Oliver, 1980), consumers might purchase second-hand products with lower initial expectations compared to when they buy new products. Summarizing, pre-used products, compared to new products, might bring along more risk. However, consumers are aware of these risks at the point of purchase. Therefore, consumers might be more inclined to forgive a transgressing second-hand product compared to a transgressing new product, where consumers' initial expectations are much higher and therefore more likely to be disconfirmed:

> *H3: After a product failure, consumer brand forgiveness intentions toward the transgressing brand will be higher if the delinquency was conducted by a second-hand product compared to a new product.*

As previously outlined, the properties of second-hand products differ from those of new products. One explanation why consumers might perceive pre-owned products differently, is the so-called underdog effect. The underdog effect is also proposed to explain higher forgiveness intentions toward a transgressing brand.

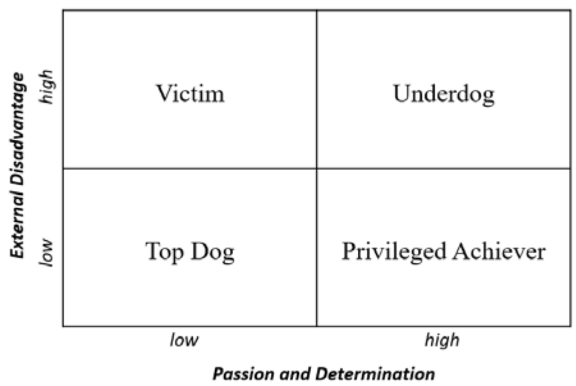
2.4 The Interplay of the Underdog Effect and Second-Hand Products

The term *underdog* itself emerged in the 19th century and originated from dog fights. The losing dog, which usually lay under the winning dog, was declared "the underdog". Later, this term was symbolically transferred to various contexts and referred to those holding a competitive disadvantage (Goldschmied & Vandello, 2012).

The underdog effect is a phenomenon where people hold positive beliefs toward and support weaker entities (Kim et al., 2019) while simultaneously disregarding disadvantages such as a low likelihood of succeeding (Goldschmied & Vandello, 2012). In contrast to an underdog, a topdog enjoys more resources and has a higher possibility of winning (Jin & Huang, 2019). In marketing, brands that are positioned as underdogs are perceived as being highly determined and passionate, while at the same time they possess restricted resources or power (Kim et al., 2019) and a less privileged position in the marketplace (Paharia et al., 2011). When positioning a brand as an underdog, the orientation does not concentrate on specific characteristics. It rather focuses on the unpretentious background and the products' noble intentions (Kim & Park, 2020).

Contrary to underdog brands, topdog brands are often large national or international brands, such as Starbucks, that overshadow regional competitors (Paharia et al., 2011). Topdog brands are often associated with negative actions, such as unethical or unsustainable production or working conditions. However, these large companies also exploit the underdog effect, for example by showing passion and emphasizing their difficult beginnings. Apple, for example, uses the underdog effect by telling the story about the launch of the company from a bedroom or garage (Paharia et al., 2011). Through this targeted storytelling, topdogs try to position themselves as underdogs.

Figure 1. Underdog Disposition Matrix based on Paharia et al. (2011, p. 778)



However, as depicted in Figure 1 it is not sufficient for a company or brand to solely tell the story of a tough beginning. They also have to show passion and determination while outlining their external disadvantages. The two dimensions *external disadvantage* and *passion and determination* jointly stimulate the underdog effect (Paharia et al., 2011).

Consumers' motives to support an underdog are empathy, nostalgia, exerting individuality, freedom of choice, and inspiration (McGinnis & Gentry, 2009). Further, consumers tend to see themselves as underdogs in their daily lives (Kim & Park, 2020) and are hence rather prone to identifying with underdog brands than with topdog brands (Li & Zhao, 2018). This identification could also be attributed to a tendency toward anti-consumption (McGinnis & Gentry, 2009), which is similarly observed in the purchase motivation of second-hand products (Ek Styvén & Mariani, 2020; Guiot & Roux, 2010). Besides the intention to distance oneself from consumer society, second-hand products are purchased out of different motives than new products are. Consumers buy pre-owned goods out of inspiration and nostalgia, which resemble the motives of supporting an underdog (McGinnis & Gentry, 2009). Comparably to an underdog, second-hand products, in contrast to new goods, employ a weaker position in the market and possess fewer resources[2] (Jin & Huang, 2019; Paharia et al., 2011). Moreover, their positioning concentrates on their humble nature (i.e., having been previously used and not being brand-new) and their virtuous intentions (Kim & Park, 2020). In conclusion, as pre-owned goods exert the ability to evoke nostalgic feelings (Guiot & Roux, 2010), are seen as being sustainable (Becker-Leifhold & Iran, 2018), and embody the possibility of getting a second chance, it is proposed that these characteristics could lead to second-hand products being perceived as underdogs:

> *H4: Underdog perception will be higher for second-hand products than for new products.*

The introduced mediators should not be considered isolated from each other, as the underdog might be able to influence a consumer's forgiveness intention. Subsequently, consumer brand forgiveness can cause different consumer responses, which is especially important when dealing with product failures.

2.5 Overall Effect: Second-Hand Consumption, Underdog Effect, and Consumer Brand Forgiveness

Forgiveness intentions depend on the type of failure. Consumers tend to exert more negative emotions toward an underdog when the transgression was of ethical (Kim & Park, 2020) or relational nature (Kim et al., 2019) as the purchase motivation for an underdog brand contains moral or prosocial intentions (Schmidt & Steenkamp, 2022). These intentions are similar to the

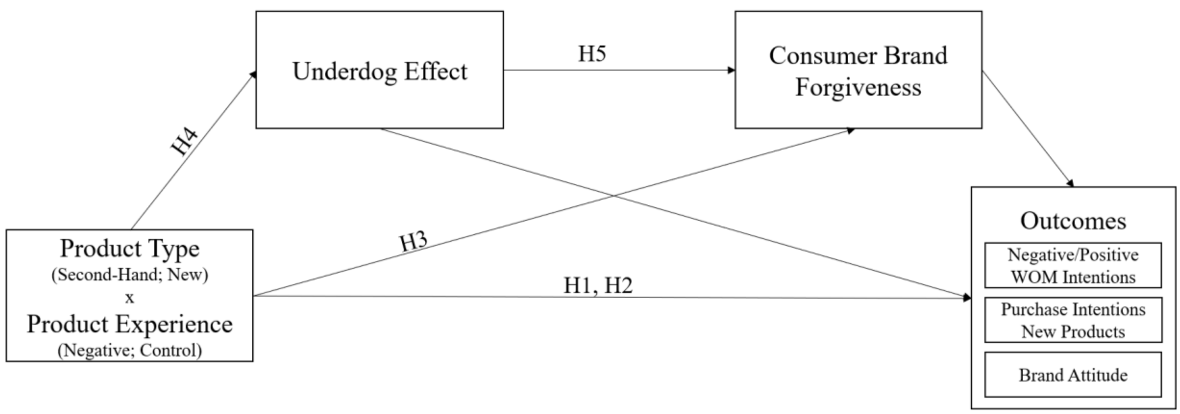
[2] About You (German online retailer for clothing, shoes, and accessories) completely stops selling second-hand products during their discount campaigns due to the lower profit margin.

sustainability motivation of purchasing second-hand products (Guiot & Roux, 2010). Therefore, failures of second-hand products, when compared to failures of new products, could be more likely to prompt negative emotions toward the product or brand, which subsequently affects forgiveness intentions. However, humans tend to exert sympathy for the underprivileged and possess an aversion against unfairness and inequalities (Fehr & Schmidt, 1999), which yields more positive emotions and cognitions toward an underdog (Einwiller et al., 2006). This might subsequently lead to higher forgiveness intentions. As second-hand products resemble underdogs in certain characteristics, it is suggested that their peculiarities could attenuate retaliatory intentions and further affect consumers' repurchase and WOM intentions.

> *H5: When experiencing a product failure with a second-hand product compared to a product failure with a new product, the effect that consumers…*
>
> a) *…will be less likely to engage in negative WOM behavior…*
> b) *…will be more likely to purchase a new product of the transgressing brand…*
> c) *…have a more positive attitude toward the transgressing brand…*
>
> *…is driven by the underdog effect (i.e., external disadvantage and passion and determination) and consumer brand forgiveness.*

Figure 2 summarizes the hypotheses and graphically presents the conceptual framework of this study.

Figure 2. Research Framework



3 Methodology

The overarching goal of this study is to test whether consumers are more forgiving toward a transgressing pre-owned or a transgressing new product and to study whether the underdog effect and consumer brand forgiveness subsequently influence WOM intentions, purchase behavior toward new products, and brand attitude. First, two pretests regarding the failure and transgression experiences are run. Second, in the first step of the main study, the direct relationship between transgressing second-hand products compared to new ones and different consumer reactions (i.e., WOM intention, purchase intention toward new products, and brand attitude) is examined using a scenario experiment. This contains a negative experience and a control condition with a purchased pre-used or new product. In the second step, consumer responses regarding the transgression experience are investigated by comparing the differences between a second-hand product and new product. The emphasis is on examining the underdog effect and consumer brand forgiveness, considering them as serial mediators.

3.1 Pretest

Pretest 1

A first pretest was conducted to identify how the product experience scenarios (i.e., negative and control) were perceived by participants. The pre-tested scenarios can be found in Appendix 1. Participants were recruited from the survey database Surveycircle (SurveyCircle, 2022) (Age: 48% between 26 and 35 years old; 57% females; 41% postgraduates). A total of 84 valid responses could be gathered after the responses were adjusted based on an attention check[3]. A 2 x 2 x 2 between-subjects-design (product type: second-hand versus new; experience: negative versus control; experience type: product-related versus personal) was used.

Participants were presented with one of the eight scenarios resulting from the 2 x 2 x 2 between-subjects design, where they were first told about a fictitious company selling either second-hand products in addition to new products or selling exclusively brand-new fashion products. Subjects were further asked to imagine that they have recently bought either a new or a second-hand jacket. A jacket was chosen as the example product in the scenario because among second-hand garments, jackets and coats are top sellers (Momox, 2022), and they are pieces of apparel which all genders could wear. Further, participants were either exposed to a negative or a control scenario with a product-related experience (i.e., jacket is in a (1) much worse or (2) in the

[3] Participants were asked to select “strongly disagree” (1 = *strongly disagree*, 7 = *strongly agree*) as their answer to show that they are paying attention, adapted from Gruzd et al. (2020).

same good condition a couple of weeks after it was bought) or personal experience (i.e., a friend tells you that the jacket (1) suits you or (2) does not suit you at all) and that you made a (1) good or (2) bad purchase with the jacket. Participants then answered questions on how satisfied they were with the jacket based on the introduced scenario, using three different statements adapted from Allen et al. (2015), on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree): "I am satisfied with this jacket.", "I think that I did the right thing when I selected this jacket.", "I am happy with this jacket." Further, subjects assessed the experience of using the jacket from the scenario on a bipolar seven-point scale with two items adapted from Zhang et al. (2014): "very negative - very positive" and "extremely unsatisfying - extremely satisfying". Following the statement "The experience with the jacket from the scenario caused me …", subjects evaluated the failure severity of the experience with the jacket from the scenario, using a bipolar seven-point scale with three items adapted from Grégoire and Fisher (2008): "minor problems - major problems", "small inconveniences - big inconveniences", and "minor aggravation - major aggravation".

According to the results, subjects were less satisfied with the negative scenario ($M = 3.64$; $SD = 1.73$) than with the control scenario (($M = 5.32$; $SD = 1.13$), $t(72.61)= -5.247$, $p < .001$). Further, participants perceived the experience of using the product from the negative scenario ($M = 3.31$; $SD = 1.36$) as less positive than the control scenario (($M = 5.43$; $SD = 1.07$), $t(82) = -7,88$, $p < .001$). In addition, failure severity was also assessed to be more severe with the negative scenario ($M = 3.37$; $SD = 1.26$) than the control scenario (($M = 2.55$; $SD = 1.42$), $t(82) = 2.80$, $p = 0.006$). Further details can be found in Appendix 2.

Pretest 2

Since there were only about 20 participants in each of the individual types of transgression scenarios (i.e., product-related versus personal), a second pretest was conducted to determine whether the different types of transgressions are perceived differently. Further, an initial test about whether second-hand products compared to new products are rather perceived as underdogs and a manipulation check for the product type (i.e., second-hand product versus new product) was conducted. The wording of the scenarios has been adapted slightly and can be found in the Appendix 3.

Participants were again recruited from the survey database Surveycircle (SurveyCircle, 2022)[4] (Age: 63% between 18 and 25 years old; 70% females; 33% undergraduates). A total of 95

[4] It was ensured that participants from the first pretest did not participate again in the second pretest.

valid responses could be gathered after adjusting the responses based on the same attention check as in the first pretest. For the pretest, a 2 x 2 between-subjects-design (product type: second-hand product versus new product; experience type: product-related versus personal) was used.

After introducing participants to the initial situation (see Appendix 3) about a fictitious company offering both new and second-hand products, subjects were asked to answer questions on the underdog effect. The questions on underdog perception were asked directly after the introduction of the company since one remains an underdog only as long as there has been no victory or defeat (Goldschmied & Vandello, 2012). To assess the single dimensions of the underdog effect, subjects evaluated statements on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree), using seven items for the dimension of *external disadvantage*, for example, “The product starts from a disadvantaged position compared to other products” and six items for the dimension of *passion and determination*, for example, “The product shows more resilience than other products in the face of adversity.”

Participants were further presented with one of the negative scenarios from the first pretest to test which of the failure types (i.e., product-related or personal) was perceived to be more negative. Subjects answered the same questions on satisfaction (adapted from Allen et al., 2015), usage experience (adapted from Zhang et al., 2014), and failure severity (adapted from Grégoire & Fisher, 2008) as in the first pretest.

As a manipulation check, participants stated whether they realized that the product from the scenario was previously owned, assessing three items on a seven-point scale (1 = strongly disagree, 7 = strongly agree): “I realize that this jacket was previously owned.”, “I realize that this jacket is not new.”, and “I realize that this jacket is a second-hand product.” extended and adapted from Bezançon et al. (2019).

To assess the results, both factors of the underdog measure were combined as an overall scale for underdog perception (Paharia et al., 2011). According to the results, participants were more likely to perceive the pre-used product as an underdog ($M = 4.64$; $SD = 1.19$) than the new jacket (($M = 3.28$; $SD = 1.63$), $t(93) = 4.67$, $p < .001$). Details can be found in Appendix 4.

Furthermore, subjects were less satisfied with the product-related failure scenario ($M = 1.94$; $SD = 1.06$) than with the personal failure scenario (($M = 4.45$; $SD= 1.41$), $t(72.61) = -5.247$, $p < .001$). Participants perceived the experience in using the product from the product-related failure scenario as less positive ($M = 2.31$; $SD = 1.39$) than the personal failure scenario

((M = 3.91; SD = 1.19), t(93) = 6.02, p < .001). In addition, failure severity was perceived to be more severe with the product-related failure (M = 4.28; SD = 1.66) than with the personal failure ((M = 3.01; SD = 1.24), t(87.08) = -4.24, p < .001) as well. See Appendix 5 for more information.

When assessing the manipulation check, the scenarios referring to the second-hand jacket were rather perceived as being second-hand (M = 6.16; SD = 1.26) than the new jacket ((M = 3.70; SD = 2.07), t(73.45) = 6.96, p < .001). Further details are displayed in Appendix 6.

In reference to the results of the pretest, the main study used the product-related failure scenario. Further, the wording of single items from some of the measures where adapted.

3.2 Main Study

Subjects and Design

Subjects for the scenario-based experiment were employees and students of LMU Munich interested in research. 202 complete responses were registered. The experiment consisted of a 2 (product type: new product, second-hand product) x 2 (experience: negative experience, control) between-subjects design. Subjects were randomly assigned to one of the four groups. 15 subjects failed an attention check, where they were asked to select “strongly disagree” on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree). These subjects were removed from further analysis, resulting in a total sample size of 187 respondents between the age of 17 and 70 years, with a mean age of 25.04. The plurality of respondents is female (67.9%), has a high-school (42.8%) or Master’s degree (20.3%), is a student (81.3%), and earns an annual gross household income below 50,000 € (55.1%). A complete overview of all assessed demographics can be found in Table 1.

Table 1. Overview of Sample Descriptive Statistics of Main Study

		Mean/Frequency	Percentage	Min	Max
Age		25.04		17	70
Gender					
	Male	57	30.5%		
	Female	127	67.9%		
	Non-binary/third gender	1	0.5%		
	Prefer not to say	2	1.1%		
Education					
	Some school but not degree	4	2.1%		
	High school graduate	80	42.9%		
	Some college but no degree	24	12.8%		
	Bachelor's degree	27	14.4%		
	Master's degree	38	20.3%		
	Professional degree	1	0.5%		
	Doctorate degree	8	4.3%		
	Other	5	2.7%		
Employment					
	Full-time	19	10.2%		
	Part-time	7	3.7%		
	Student	152	81.3%		
	Stay-at-home-parent	2	1.1%		
	Unemployed	1	0.5%		
	Retired	3	1.6%		
	Self-employed	3	1.6%		
Income (voluntary)					
	< 10,000 €	63	33.7%		
	10,001 € – 50,000 €	40	21.4%		
	50,001 € – 90,000 €	22	11.8%		
	90,001 € - 150,000 €	11	5.9%		
	> 150,001 €	4	2.1%		
	No indication	47	25.1%		
		N = 187			

Stimuli and Procedure

Subjects were introduced to the fictitious global clothing company *wearIMM*, which offers second-hand products in addition to brand-new fashion products. Based on the findings from the pretests, the scenarios were adapted slightly. Study participants were randomly assigned to one of the four scenarios. They were told to either imagine that they bought a (1) new or a (2) second-hand jacket from the brand *wearIMM*. Participants were further told that they are invited to the movies and have decided to wear the new or second-hand jacket. When taking the jacket out of the cabinet, the product was either (1) in the same good condition as it was a couple of weeks ago, and the subject felt comfortable wearing it, or (2) in a much worse condition than it was a couple of weeks ago, and the participant does not feel comfortable wearing it anymore. See Appendix 7 for a detailed presentation of the different scenarios. After subjects were presented with the scenario, they were asked to answer several questions. Further, participants were asked background questions.

3.3 Measures

A detailed list of all applied constructs and items is presented in Appendix 8. Unless otherwise noted, all constructs were measured on a seven-point scale (1 = strongly disagree; 7 = strongly agree). Cronbach's alpha is used to assess internal consistency and presented next to the constructs.

Negative WOM Intentions (α = .88): Since the study is mainly interested in the outcome after a product failure has taken place, negative WOM intentions were measured on a three-item scale adapted from Grégoire and Fisher (2006). Participants were asked to indicate how strongly they agree with the statements "I will spread negative word-of-mouth about the brand *wearIMM*.", "I will speak ill of the brand *wearIMM* to my friends", and "When my friends look for similar products, I tell them not to buy from the brand *wearIMM*".

Purchase Intention Toward New Products: Purchase intention can be understood as a doubly concrete construct (Rossiter, 2002) and has been successfully measured as a single-item in top research journals (e.g., Argo et al., 2006; Sundie et al., 2011). Therefore, purchase intention was measured using a single-item measure. As the study is interested in crossover effects of negative experiences with pre-owned products, participants were asked about their purchase intention for new products of the transgressing brand: "Based on the experience from the scenario with the brand *wearIMM*, how likely are you to buy a new product from the brand *wearIMM*?". Participants were asked to state their consent using a single-item rated on a seven-point scale (1 = very unlikely to buy, 7 = very likely to buy) (Argo et al., 2006).

Brand Attitude (α = .95): Brand attitudes often serve as a basis for product choice and purchase behavior (Priester et al., 2004) and are defined as a consumer's assessment of a brand (Keller, 2013). Participants were asked to indicate how well one or the other adjective in the pairs "bad – good", "like – dislike", "unpleasant – pleasant", and "useless – useful" describes their overall feeling toward the brand on a seven-point semantic differential scale (Bergkvist & Rossiter, 2009).

Positive WOM Intentions (α = .97): Positive WOM intentions were measured using four items (Brüggen et al., 2011) on a seven-point Likert-type scale (1 = extremely unlikely, 7 = extremely likely), for example, "I am likely to say positive things about the brand *wearIMM* to other people". The scale was adapted to the situation of the scenario, as the original items were developed to measure WOM intentions with a restaurant. However, the items seem to be applicable to use with a variety of entities (Bruner, 2013).

Consumer Brand Forgiveness (α = .89): Only few studies use multidimensional and multi-item scales to measure forgiveness in the marketing context (Fetscherin & Sampedro, 2019); most studies focus on one or the other of the dimensions introduced. Behavioral measures of forgiveness are especially appropriate for use in experimental settings (Fernández-Capo et al., 2017). Due to the peculiarities of second-hand products, consumer brand forgiveness is measured using a three-dimensional scale comprising cognitive (e.g., "I disapprove of this brand"), affective (e.g., "I feel sympathetic toward this brand"), and behavioral (e.g., "I avoid using this brand") aspects (Christodoulides et al., 2021).

Underdog Effect (α = .86): The underlying research design aims to examine whether pre-owned products are perceived as underdogs. The measure of the underdog was queried before the actual manipulation of the experience with the product (negative versus control) was introduced. This was considered useful because the term underdog is usually attached to an entity before a perceived outcome. After the outcome is known, the entities are either winners or losers (Goldschmied & Vandello, 2012). An underdog brand or product was found to consist of two dimensions: external disadvantage (α = .90), passion and determination (α = .70) (Paharia et al., 2011). To ensure a high construct validity, i.e., that the operational definition actually reflects the construct that it is supposed to measure (Sarstedt & Mooi, 2014), a shortened version of the underdog scale developed by Paharia et al. (2011) was adapted to the context of product types and used to cover these two dimensions. For the dimension of external disadvantage, seven items were used, for which the participants were asked, for example, whether they agreed that "The product starts from a disadvantaged position compared to other products". The dimension of passion and determination uses six items, and here participants were asked, for example, if they agree with the following statement: "The product is more resistant than other products in the face of adversity".

Experience with Second-Hand Clothing (α = .96): As experience with second-hand shopping is found to influence acceptance of pre-owned goods (O'Reilly et al., 1984), it is included as a covariate and measures whether participants are familiar with the consumption of second-hand clothing. Participants will rate their previous experience on a three-item scale adapted from Lo et al. (2019), stating, for example, “I have a great deal of experience in buying second-hand clothes”.

Preference for Second-Hand Clothing (α = .94): As individual product preferences influence consumer reactions, a customer’s preference toward second-hand clothes is included as a control variable. The scale is adapted to the context of the study and measured using three-items (Cheng et al., 2017): “Which product type do you like more in terms of clothing?”, “Which product type are you more favorable toward in terms of clothing?”, “Which product type are you more likely to buy in terms of clothing?”. The scale was recoded before the analysis (1 = new clothing, 7 = second-hand clothing).

Product Involvement (α = .97): Involvement is defined as a consumer’s perceived importance of a product or a product category, based on their values, interests, and needs (Zaichkowsky, 1985). As the involvement with clothing might influence subsequent consumer reactions, the scale is included as a covariate variable, measured on a seven-point Likert-type scale (Brocato et al., 2015). Subjects were asked to declare their agreement on several statements like “In general, I have a strong interest in clothing”.

Demographic Information: Participants were asked several questions about their demographic characteristics at the end of the survey. Data on gender, age, education, employment, and income was gathered. The declaration of income was voluntary.

4 Analysis and Results

The following chapter is divided into two parts. Within the manipulation check, it was tested whether the scenarios and the product type manipulation were perceived as intended. Next, the hypotheses were tested by means of comparing group differences and running a mediation analysis. Table 2 displays the descriptive statistics of the mediator measures used in the main study by product type.

Table 2. Descriptive Statistics of Mediator Measures Used in the Main Study by Product Type

	Mean	N	SD	Min	Max
Underdog (External Disadvantage & Passion and Determination)					
Second-Hand Product	4.3615	98	0.71150	3.00	6.01
New Product	3.5464	89	1.01574	1.00	6.10
Underdog (Dimension: External Disadvantage)					
Second-Hand Product	4.3863	98	1.16892	1.86	7.00
New Product	3.2520	89	1.27703	1.00	6.29
Underdog (Dimension: Passion and Determination)					
Second-Hand Product	4.3367	98	0.79047	1.83	6.50
New Product	3.8408	89	0.99237	1.00	6.33
Scarcity					
Second-Hand Product	4.8197	98	1.45978	1.00	7.00
New Product	3.6255	89	1.40510	1.00	7.00

Table 3 displays the descriptive statistics of the measures used in the main study by product type and type of experience.

Table 3. Descriptive Statistics of Measures Used in the Main Study by Product Type and Type of Experience

		Mean	N	SD	Min	Max
Consumer Brand Forgiveness						
Second-Hand Product	*Negative Experience*	3.6859	52	1.21192	1.00	6.78
	Control Experience	5.3164	46	0.58136	3.78	7.00
New Product	*Negative Experience*	2.8561	44	0.90095	1.00	5.00
	Control Experience	5.1481	45	0.60022	3.67	6.33
Brand Attitude						
Second-Hand Product	*Negative Experience*	3.0288	52	1.22640	1.00	6.25
	Control Experience	5.9130	46	0.72881	4.00	7.00
New Product	*Negative Experience*	2.0682	44	0.86496	1.00	4.00
	Control Experience	6.0111	45	0.77415	3.75	7.00
Negative Word-of-Mouth						
Second-Hand Product	*Negative Experience*	3.8526	52	1.60936	1.00	7.00
	Control Experience	1.6667	46	0.81043	1.00	3.67
New Product	*Negative Experience*	5.1894	44	1.15559	1.67	7.00
	Control Experience	1.8815	45	1.04726	100	5.33
Purchase Intention Toward a New Product						
Second-Hand Product	*Negative Experience*	3.6538	52	1.57037	1.00	7.00
	Control Experience	4.4783	46	1.64302	1.00	7.00
New Product	*Negative Experience*	2.7500	44	1.69986	1.00	7.00
	Control Experience	5.2222	45	1.45990	1.00	7.00

4.1 Manipulation Check

Whether participants perceived the product from the scenario to be either pre-used or new was measured on a three item seven-point Likert scale (1 = strongly disagree, 7 = strongly agree), for example, “This jacket was previously owned” adapted from Bezançon et al. (2019). As expected, the majority of participants in the second-hand product scenario stated that they perceived the product from the scenario to be second-hand ($M = 6.54$, $SD = 0.70$) than the participants from the new product scenario (($M = 3.13$, $SD = 1.84$), $t(185) = 17$, $p < .001$).

To check whether the failure scenario was perceived more negatively than the control setting, participants were asked to rate the experience from their scenario. This was done using two items which were measured on a seven-point scale (1= very negative/extremely unsatisfied, 7 = very positive/extremely satisfied) (Zhang et al., 2014). Further, failure severity was surveyed by asking participants if the experience with the product from their scenario caused them “minor” to “major problems”, “small” to “big inconveniences”, and “minor” to “major troubles” on a seven-point scale (1 = minor/small, 7 = major/big) (Grégoire & Fisher, 2008). As expected, the negative experience scenario was perceived as more negative and unsatisfying

($M = 2.43$, $SD = 1.16$) than the control scenario (($M = 5.97$, $SD= .81$), $t(185) = -24.14$, $p<0.001$). Regarding failure severity, the negative scenario was also perceived to cause participants more major problems ($M = 4.03$, $SD = 1.42$) than the control scenario (($M = 1.92$, $SD = 1.18$), $t(185) = 11.10$, $p<0.001$).

4.2 Hypotheses Testing

First, an Analysis of Covariance (ANCOVA) was conducted to assess group differences in the endogenous variables among the scenarios. Next, a mediation analysis was conducted to test the hypotheses.

4.2.1 Group Differences

A two-way factor ANCOVA is an extension of single factor ANOVA in which one or more variables can be controlled for. The assumptions of measurement independence, a minimum interval-scaled dependent variable, an independent nominal-scaled variable, and minimum interval-scaled covariates, had already been considered in the study design. Homogeneity of regression slopes, presence of outliers in groups, normally distributed residuals, and homoscedasticity were statistically tested for in SPSS (Field, 2011). For the combined underdog scale, only one outlier was found, which was left in the data analysis. Further, for brand attitude, consumer brand forgiveness, and the combined underdog scale, data did not show a homogeneity of variance. However, an ANCOVA is said to be robust against this violation, therefore, the analysis was performed.

The conducted ANCOVAs all controlled for product involvement, experience with second-hand clothing, and preference for second-hand clothing. To compare the four different scenarios, a Bonferroni post hoc test as one of the most conservative post hoc tests (Field, 2011) was used. To compare the different scenarios, a variable called “group” was designed, which comprised the different scenario combinations (e.g., second-hand product and negative experience). Appendix 9 - Appendix 12 display the results of the group comparisons.

Negative WOM Intentions: An ANCOVA revealed a significant main effect of the different groups on negative WOM behavior ($F_{3,180} = 85.42$, $p < .001$). As might be expected, participants were less likely to engage in negative WOM behavior in the control condition than in the negative experience setting. H1a predicted that customers would be less likely to engage in negative WOM behavior when the failure was conducted by a second-hand product compared to a new product. In line with H1a, a significant difference in negative WOM behavior between a second-hand and a new product was found ($M_{second\text{-}hand,\ negative} = 3.86$ versus $M_{new,\ negative} = 5.18$, $p < 0.001$), supporting H1a.

Purchase Intention Toward New Product: A significant main effect of the different groups regarding purchase intention for a new product ($F_{3,180} = 21.72$, $p < 0.001$) was revealed. Customers are more likely to purchase a new product of the same brand if they have had a negative experience with a second-hand product compared to a negative experience with a new product ($M_{second\text{-}hand,\ negative} = 3.65$ versus $M_{new,\ negative} = 2.70$, $p = 0.024$), supporting H1b. When looking at a positive experience (i.e., control scenario) with a second-hand product compared to a new product, no significant differences were found ($M_{second\text{-}hand,\ control} = 4.48$ versus $M_{new,\ control} = 5.28$, $p = .104$). Therefore, H2a is rejected.

Brand Attitude: A significant effect for the group on brand attitude ($F_{3,180} = 207.38$, $p < .001$) was found. As might be expected, a negative experience with a product leads to a lower brand attitude. H1c predicted that customers would have a higher brand attitude toward brands where a transgression was conducted by a second-hand product than a new product ($M_{second\text{-}hand,\ negative} = 3.03$ versus $M_{new,\ negative} = 2.08$, $p < 0.001$). A significant difference was found, supporting H1c. H2b predicted that with a positive experience, brand attitude would also be higher for second-hand products than for new products. However, no significant difference was found ($M_{second\text{-}hand,\ control} = 5.92$ versus $M_{new,\ control} = 6.00$, $p = 1$). Therefore, H2b is rejected.

Consumer Brand Forgiveness: An ANCOVA revealed a significant main effect of the grouping variable for consumer brand forgiveness ($F_{3,180} = 82.20$, $p < .001$). H3 predicted that consumers are more likely to forgive a brand after a product transgression regarding second-hand products than after one regarding new products. In line with this hypothesis, a significant difference was found between the two product types, where customers are more likely to forgive the brand after a product transgression regarding a second-hand product than after a misconduct regarding a new product ($M_{second\text{-}hand,\ negative} = 3.68$ versus $M_{new,\ negative} = 2.86$, $p < .001$).

Underdog Perceptions: A significant main effect of product type on the measure of underdog perceptions ($F_{1,182} = 40.05$, $p < .001$) was revealed. H4 predicted that second-hand products are more likely to be perceived as an underdog than new products. In line with this hypothesis, a significant difference was found between both product types: Pre-used products are more likely to be perceived as an underdog than new products ($M_{second\text{-}hand} = 4.36$ versus $M_{new} = 3.55$, $p < .001$).

Scarcity: Although no effects were hypothesized for scarcity perceptions, the construct was measured within the questionnaire and respondents' evaluations were analyzed. An ANCOVA on scarcity revealed a significant main effect of product type ($F_{1,182} = 32.48$, $p < .001$), where second-hand products are perceived to be scarcer than new products ($M_{second\text{-}hand} = 4.82$ versus $M_{new} = 3.64$, $p < .001$).

4.2.2 Mediation Analysis

The mediation analysis was conducted using IBM SPSS Statistics and PROCESS macro version 4.1 developed by Hayes (2022). Bootstrapping was applied to estimate the indirect effects. As the study is mainly interested in the mediation within a negative product experience scenario, only the group presented with the negative experience was analyzed. The coefficients are interpreted in relation to each other, as the independent variable, i.e., product type[5], is categorical.

A combination of parallel and serial mediation was chosen. In the parallel multiple mediator model the independent variable is modeled to influence the dependent variable both directly and indirectly through two mediators, where no mediator influences the other (Hayes, 2022). This assumption is relaxed for a serial multiple mediator model, where the independent variable causes the first mediator, and this first mediator causes the second mediator (Hayes, 2022). Since the underdog effect has two different dimensions (i.e., external disadvantage and passion and determination) which can have different effects on consumer brand forgiveness, it was decided to regard the dimensions individually. In the presented mediation model (PROCESS model = 80; 5,000 resamples, Davidson and MacKinnon (1993) heteroscedasticity-consistent inference) negative WOM, purchase intention for new products, and brand attitude were used as dependent variables, and product type was used as the independent variable. The dimensions of the underdog scale, external disadvantage and passion and determination, as well as consumer brand forgiveness were used as mediating variables. Product involvement, past experience with second-hand apparel, and preference for second-hand clothing were included as covariates. Covariates are only discussed when they were significant.

The indirect effects can be found in Table 4. Table 5 provides an overview of all tested hypotheses and its results. Further information can also be found in Appendix 15 until Appendix 18.

[5] 0 = new product; 1 = second-hand product

Table 4. Indirect Effects Mediation Models

Variables	Effect	BootSE [a]	LLCI [b]	ULCI [b]
negativeWOM				
product type → external disadvantage → consumer brand forgiveness → negativeWOM	0.1733	0.0854	0.0261	0.3525
product type → passion and determination → consumer brand forgiveness → negativeWOM	-0.2205	0.1230	-0.4776	-0.0039
Brand Attitude				
product type → external disadvantage → consumer brand forgiveness → brand attitude	-0.1259	0.0623	-0.2590	-0.0166
product type → passion and determination → consumer brand forgiveness → brand attitude	0.1601	0.0858	0.0058	0.3386
Purchase Intention Toward New Product				
product type → external disadvantage → consumer brand forgiveness → purchase intention toward new products	-0.1436	0.0717	-0.2972	-0.0155
product type → passion and determination → consumer brand forgiveness → purchase intention toward new product	0.1827	0.1044	0.0037	0.4117

Note: N = 96, [a] Standard errors from the mean result of bootstrapping;
[b] LLCI/ULCI = lower-/upper-level of bias corrected bootstrap 95%-confidence interval.

Underdog Dimensions: Compared to new products, second-hand products were found to exert a positive effect on the underdog dimensions external disadvantage, (B= .7381, t = 2.7980, p = .0063) and passion and determination (B = .5977, t = 3.2364, p = .0017).

Consumer Brand Forgiveness: The underdog dimension of external disadvantage negatively affected consumer brand forgiveness (B= -.2319, t = -2.1860, p = .0314). The dimension of passion and determination positively predicted consumer brand forgiveness (B = .3643, t = 2.0801, p = 0.0404).

Negative WOM Intentions: Consumer brand forgiveness is found to negatively predict negative WOM intentions (B = -1.0124, t = -10.6457, p < .001). On the one hand, it was found that the relationship between negative WOM intentions and product type is partially mediated by the dimension of external disadvantage and consumer brand forgiveness, in the case that it leads to higher negative WOM intentions for a transgressing second-hand product (second-hand product indirect effect = .1733, LLCI = .0261, ULCI = .3525). On the other hand, the dimension of passion and determination also partially mediates this relationship. However, the dimension of passion and determination leads to higher forgiveness intentions (B = .3643, t = 2.0801, p = 0.0404), which further leads to an overall negative indirect effect on negative WOM intention (second-hand product indirect effect = -.2205, LLCI = -.4776, ULCI = -.0039). Therefore, hypothesis H5a is partially supported.

Purchase Intention Toward New Product: Purchase intention for a new product was also positively predicted by consumer brand forgiveness (B = .8392, t = 5.9431, p < .001). The relationship between the intention to buy a new product from the transgressing brand and product type is mediated by the dimension of external disadvantage and consumer brand forgiveness, which leads to a lower purchasing intention for new products (second-hand product indirect effect = -.1436, LLCI = -.2972, ULCI = -.0155). Additionally, the relationship between purchase intention for a new product after a transgression regarding a second-hand product compared to one regarding a new product is mediated by the dimension of passion and determination. It leads to higher purchase intentions for a new product of the very same brand (second-hand product indirect effect = .1827, LLCI = .0037, ULCI = .4117). Therefore, hypothesis H5b is partially supported.

Brand Attitude: Consumer brand forgiveness also positively predicted brand attitude (B = .7354, t = 9.3896, p < .001). The relationship between brand attitude and product type is partially mediated by the dimension of external disadvantage and consumer brand forgiveness, in the case that it leads to lower brand attitude for a transgressing second-hand product than for a transgressing new product (second-hand product indirect effect = -.1259, LLCI = -.2590, ULCI = -.0166). The underdog dimension of passion and determination partially mediates the relationship between a transgression regarding a second-hand product versus a transgression regarding a new product. This leads to higher brand attitude (second-hand product indirect effect = .1601, LLCI = .0022, ULCI = .3386). Hypothesis H5c is partially supported.

Table 5. Overview of Hypotheses Test Results

H1: When experiencing a product failure with a second-hand product compared to a new product, consumers…	
a) …will be less likely to engage in negative WOM behavior.	Supported
b) …will be more likely to purchase a new product of the transgressing brand.	Supported
c) …have a more positive attitude toward the transgressing brand.	Supported
H2: When experiencing a positive experience with a second-hand product compared to a new product, consumers…	
a) …will be more likely to engage in positive WOM behavior.	Not supported
b) …will have a more positive attitude toward the brand.	Not supported
H3: After a product failure, consumer brand forgiveness intentions toward the transgressing brand will be higher if the delinquency was conducted by a second-hand product compared to a new product.	Supported
H4: Underdog perception will be higher for second-hand products than for new products.	Supported
H5: When experiencing a product failure with a second-hand product compared to a product failure with a new product, the effect that consumers…	
a) …will be less likely to engage in negative WOM behavior…	Partially supported
b) …will be more likely to purchase a new product of the transgressing brand…	Partially supported
c) …have a more positive attitude toward the transgressing brand…	Partially supported
…is driven by the underdog effect (i.e., external disadvantage and passion and determination) and consumer brand forgiveness.	

5 Conclusion

Discussion

Second-hand purchases are becoming more and more popular. The majority of pre-owned products are still sold via P2P-platforms. However, the trend has not completely bypassed retailers and some already sell their own used products. In academia, second-hand consumption was already discussed several decades ago, with early papers mainly dealing with flea markets or other offline possibilities of second-hand sales. In recent years, research has increased and focused mainly on the motivators of buying and consuming pre-owned products. A business perspective, i.e., brands selling their own pre-used products, has largely been neglected.

The presented study addresses this research gap and takes on a company perspective. This perspective is examined by investigating the sales of second-hand products by companies and brands. Additionally, acting on a general desire for further research on second-hand consumption, especially however on the shortcomings of second-hand purchases (Guiot & Roux, 2010), the proposed study investigates consumer reactions toward the experience with pre-owned products. Since negative experiences are only a matter of time (Hassey, 2019) and can have detrimental consequences for the brand, the study initially concentrates on the question of how a product failure regarding pre-used products affects consumer reactions (H1). Thus, the study aims to investigate whether a transgression regarding pre-owned products has similar negative effects on the brand as those that have already been investigated for new products. In the case of a positive experience with a pre-used product, it was suggested that a crossover effect should be even stronger than for new products (H2). As one possible explanation for differing consumer outcomes, consumer brand forgiveness (H2) and the underdog effect (H3) are suggested and investigated as parallel and serial mediators (H5).

As initially suggested, when experiencing a product failure with a second-hand product compared to one with a new product, consumers were found to be less likely to engage in negative WOM behavior, more likely to purchase a new product of the transgressing brand and exert a more positive attitude toward the transgressing brand. This is especially interesting as a pre-used good has different starting conditions than the same product when new. A negative experience leads people to be more likely to engage in negative behavior, such as negative WOM, but positive experiences lead consumers to be more inclined to talk positively about their experience with the brand or product (Olson & Ahluwalia, 2021). However, this effect was only observed for negative experiences with second-hand products. When experiencing a positive experience with a pre-used good compared to one with a new product, consumers were not found to engage in more positive WOM behavior or exert a higher brand attitude. One possible explanation could be the *negativity effect*, which explains that negative information is weighted more heavily than positive information (Ahluwalia, 2002). In addition, the *frequency-based attribution* states that positive information is less influential as it is already more prevalent (Chen & Lurie, 2013) in society. This could lead to the fact that consumers are more willing to talk about a negative experience with a product compared to a positive one, which they do not perceive as noteworthy.

Furthermore, a transgression regarding a second-hand product is more likely to be forgiven compared to one by a new product. In line with the initial argumentation, the confirmation-disconfirmation paradigm (Oliver, 1980) could be decisive here, as consumers might exert different expectations toward a second-hand product. The underdog effect was proposed as a possible justification, and it was found that underdog perceptions are higher for second-hand products than for new products. When combining both mediators, i.e., the underdog effect and consumer brand forgiveness, an interesting outcome was observed. When looking at the two dimensions of the underdog effect, a second-hand product yields higher scores for both the dimension of external disadvantage as well as for passion and determination. However, those two dimensions have different effects on consumer brand forgiveness and therefore on customer reactions.

As suggested, a second-hand product triggers a higher score on the underdog dimension of passion and determination, which leads to higher consumer brand forgiveness intentions. These forgiveness intentions then lead to less negative WOM behavior, higher purchase intentions toward new products, and a higher brand attitude toward the transgressing brand. When looking at the dimension of external disadvantage, second-hand products also score higher. However, the dimension of external disadvantage exerts a negative effect on consumer brand forgiveness, meaning that consumers are less likely to forgive a transgressing brand when only focusing on the product's external disadvantage. In this setting, external disadvantage rather functions as a suppressor (Rucker et al., 2011).

One possible explanation for this could be *contamination theory*. Hygiene concerns and a desire for new products are the most frequently mentioned barriers to purchasing second-hand products. In the case of negative contamination (Belk, 1988), people are concerned that a product previously owned by another person might transmit their odors or dirtiness (Roux, 2010). When advertising a product as externally disadvantaged (e.g., as "just like new"), products are perceived as less desirable (Ackerman & Hu, 2017). Therefore, consumers might not want to know about the previous life of the product, which can be seen as an external disadvantage.

Furthermore, *identification theory* (e.g., Kim et al., 2019) is often mentioned as an explanation of why consumers are more supportive toward underdogs than topdogs. Referring to identification theory, consumers have a positive attitude toward brands or products with which they can identify (Einwiller et al., 2006). As consumers might feel like underdogs in their everyday life (Kim & Park, 2020), they are more likely to identify with underdog than with topdog brands (Li & Zhao, 2018). This leads to more positive emotions and cognitions toward the underdog (Einwiller et al., 2006). However, when looking at the different dimensions of the underdog effect, consumers might be less willing to identify with someone that had a rough start and is at a disadvantage, compared to someone who is not and rather focus on the positive side, where passion and determination are at the forefront.

Limitations and Further Research

The findings of this paper extend current research and challenges previously presented papers. With nonrelational transgressions conducted by an underdog brand, it had already been found that consumers show greater forgiveness intentions toward underdog brands than toward topdog brands. However, when facing relational transgressions, the underdog effect might backfire and lead to lower forgiveness intentions (Kim et al., 2019). These findings should be reexamined from a different perspective, taking the individual dimensions of the underdog effect into account rather than looking at the construct as a whole.

The presented study is not without limitations, which could serve as fruitful avenues for further research. For starters, the scale used for measuring the underdog perceptions was adapted from Paharia et al. (2011), where the scale was used in an interpersonal context. For future research on brand and product perceptions, the development of a measurement instrument for these purposes could be helpful. Further, one contributing factor to the different perceptions of passion and determination and external disadvantage regarding second-hand products, may be attributed to the utilization of fictitious stimuli within the present study. This could engender an artificial effect that diverges from perceptions observed in real brand contexts. Future research could test the derived assumptions, using real brands.

The data was gathered using a student sample. While multiple advantages of using student samples have been found in literature (e.g., Darley et al., 2010), it could be interesting to see whether the attitude of consumers differs when using a more mature sample. Furthermore, as attitudes toward second-hand consumption have undergone major changes in the western countries due to an increased awareness of sustainable and ethical consumption, second-hand products might be perceived differently in a different cultural context. Therefore, conducting a similar study with a more diverse sample could provide more generalizable results.

The examined product category (i.e., fashion) could play a decisive role as well. Using different product categories that are also often sold and bought second-hand, such as furniture or cars, could yield interesting results. Further, the study excluded luxury products as they exert different characteristics compared to daily goods. It could be interesting to see whether perceptions of luxury products differ compared to products consumed in everyday lives.

Furthermore, no statistical difference was found for customer outcomes when experiencing a positive situation with a second-hand product compared to one with a new product. On the one hand, this could be due to the already mentioned negativity or frequency-based attribution effects. On the other hand, a second experiment with a more concrete manipulation of the positive experience with such products could shed more light onto this finding.

Implications for Marketers

Since few companies currently offer their own products as second-hand products, the results serve as a basis for decision-making. A brand's overall reputation can benefit from offering second-hand products. Although no significant differences between a pre-used and a new product were found when having a positive experience, this also shows that pre-used products will not harm brand attitude when being introduced to the product portfolio. Moreover, when having a negative experience with a pre-owned product, the perception of the used products as underdogs has a mitigating impact on the brand.

Further, offering second-hand products might increase turnover. It has been shown that the additional sale of second-hand clothing does not cannibalize the company's own sales (Ghose et al., 2006), but that companies can save on advertising costs (Strähle, 2021) and address a new customer segment (Miller & Brannon, 2022). Despite the economic motivations, the purchase of pre-owned clothes is not related to gross household income (Momox, 2021). This is particularly relevant as consumers that are financially well off are able to purchase more second-hand products within their budget than consumers that are financially less well off (Roux & Guiot, 2020). Another advantage is that customer segments with a lower income can be addressed more easily, as these consumers can test products they would not be able to afford new, especially in the high-priced segment (Abbes et al., 2020). Besides the economic reasons, the sustainability motivator plays an important role in the purchase of second-hand products. Therefore, entirely new customer segments could be addressed. For example, customers who prefer sustainable products could now turn to second-hand products from commercial suppliers. Marketers can adapt their communication strategy accordingly to target appropriate audiences.

Lastly, the underdog effect opens doors for suitable storytelling. Marketers can leverage this effect and adapt their story around second-hand products so that these are more aligned with the underdog effect by focusing on the product's nature and background. This effect can, among others, increase brand identification (Delgado-Ballester, 2021), but also help mitigate certain types of failures (Kim et al., 2019). Based on the findings, marketers need to be especially careful which dimension of the underdog they promote. Focusing on the dimension of external disadvantage might lead to less intentions to forgive a transgressing second-hand product. Therefore, a promotion of the product's passion and determination should be targeted, as this leads to higher forgiveness intentions.

Conclusion

In conclusion, second-hand consumption is a rising trend and has mainly been offered on P2P-platforms. However, companies and brands benefit from synergy effects by offering their own products as second-hand products. In everyday use, when there is no product failure, second-hand products are perceived no worse than new products. However, when experiencing a product failure with a pre-used product compared to one with a new product, consumers are more forgiving toward the pre-owned good, which results in less negative WOM, higher purchase intentions for the brand's new products, and a higher brand attitude. However, brands should be especially careful on how to promote their pre-used products to customers. Second-hand products are perceived as underdogs. Marketers should focus on the dimension of passion and determination to promote the product rather than on the dimension of external disadvantage, as it might have negative effects on a consumer's willingness to forgive. As second-hand consumption has only received renewed interest in the academic community in recent years, there remains a potential for further research, especially concerning consumer reactions toward the consumption of second-hand products.

Appendix

Appendix 1. Overview of Initial Situation and Scenarios of Pretest 1

About *wearIMM:* *wearIMM* is a global clothing company with a focus on fashion for men, women, children, and teenagers. NEW: Since its founding in 1997, *wearIMM* offers **brand-new fashion products**. SECOND-HAND: Since its founding in 1997, *wearIMM* has steadily expanded its product portfolio and also offers **second-hand (i.e., previously owned)** in addition to new fashion products. Imagine you have recently purchased a **[new/second-hand] jacket** online from *wearIMM* to wear at some future occasion.		
	Product Experience (negative)	**Product Experience (positive/control)**
Product-Related	A couple of weeks after you have bought your **[new/second-hand] jacket,** you are invited to the movies and decide to wear the **[new/second-hand] jacket.** You take the **[new/second-hand] jacket** out of the cabinet and are disappointed as it is in a **much worse condition** than it was a couple of weeks ago, and you do **not** feel comfortable wearing it anymore.	A couple of weeks after you have bought your **[new/second-hand] jacket**, you are invited to the movies and decide to wear the **[new/second-hand] jacket.** You take the **[new/second-hand] jacket** out of the cabinet and are **happy** as it is in the **same good condition** as it was a couple of weeks ago, and you feel **comfortable** wearing it.
Personal	A couple of weeks after you have bought your **[new/second-hand] jacket**, you are invited to the movies and decide to wear the **[new/second-hand] jacket**. A friend of yours approaches you in front of the movies and directly notices your **[new/second-hand] jacket**. S/he tells you that the **[new/second-hand] jacket** does **not** suit you at all and that you have made a **bad** purchase.	A couple of weeks after you have bought your **[new/second-hand] jacket**, you are invited to the movies and decide to wear the **[new/second-hand] jacket**. A friend of yours approaches you in front of the movies and directly notices your **[new/second-hand] jacket**. S/he tells you that the **[new/second-hand] jacket** does **suit** you and that you have made a **good** purchase.

Appendix 2. Comparison of Failure Scenarios (i.e., Negative versus Control) – Pretest 1

		N	Mean	SD	SE Mean
Satisfaction	*Negative*	43	3.6434	1.73113	0.26399
	Control	41	5.3171	1.12781	0.17613
Experience of Product Usage Based on Scenario	*Negative*	43	3.3140	1.36287	0.20784
	Control	41	5.4268	1.06982	0.16708
Failure Severity	*Negative*	43	3.3721	1.26242	0.19252
	Control	41	2.5528	1.41737	0.22136

Note: Only data used where participants passed the attention check.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% CI of the Difference	
	F	Sig.	t	df	Two-Sided p	Mean Difference	SE Difference	Lower	Upper
Satisfaction									
Equal variances assumed	13.398	0.000	-5.223	82	0.000	-1.67366	0.32046	-2.31117	-1.03616
Equal variances not assumed			-5.274	72.608	0.000	-1.67366	0.31736	-2.30621	-1.04111
Product Usage Experience									
Equal variances assumed	2.914	0.092	-7.878	82	0.000	-2.11288	0.26820	-2.64640	-1.57935
Equal variances not assumed			-7.923	79.127	0.000	-2.11288	0.26667	-2.64365	-1.58210
Failure Severity									
Equal variances assumed	2.594	0.111	2.800	82	0.006	0.81925	0.29255	0.23728	1.40122
Equal variances not assumed			2.793	79.874	0.007	0.81925	0.29336	0.23542	1.40307

Note: Only data used where participants passed the attention check.

Appendix 3. Overview of Initial Situation of Pretest 2

About wearIMM:
wearIMM is a global clothing company with a focus on fashion for men, women, children, and teenagers. Since its founding in 1997, *wearIMM* has steadily expanded its product portfolio and also offers second-hand products (i.e., used and previously owned) in addition to brand-new fashion products.
Imagine you have bought a **[new/second-hand] jacket** from *wearIMM* to wear at some future occasion (e.g., wedding, birthday party).

Appendix 4. Comparison of Product Type (i.e., Second-Hand Products Versus New Products) on Underdog Perceptions – Pretest 2

		N	Mean	SD	SE Mean
Underdog Effect (combined – both dimensions)	*Second-Hand Product*	49	4.6429	1.18805	0.16972
	New Product	46	3.2826	1.62840	0.24009
Underdog Effect (External Disadvantage)	*Second-Hand Product*	49	4.2245	1.21534	0.17362
	New Product	46	3.5875	1.47274	0.21714
Underdog Effect (Passion and Determination)	*Second-Hand Product*	49	4.2898	0.83871	0.11982
	New Product	46	4.1391	1.10342	0.16269

Note: Only data used where participants passed the attention check.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% CI of the Difference	
	F	Sig.	t	df	Two-Sided p	Mean Difference	SE Difference	Lower	Upper
Underdog Effect (combined – both dimensions)									
Equal variances assumed	6.894	0.010	4.672	93	0.000	1.36025	0.29117	0.78203	1.93846
Equal variances not assumed			4.626	82.011	0.000	1.36025	0.29403	0.77534	1.94516
Underdog Effect (External Disadvantage)									
Equal variances assumed	2.190	0.142	2.305	93	0.023	0.63702	0.27634	0.08826	1.18577
Equal variances not assumed			2.291	87.429	0.024	0.63702	0.27802	0.08446	1.18957
Underdog Effect (Passion and Determination)									
Equal variances assumed	2.747	0.101	0.752	93	0.454	0.15067	0.20033	-0.24715	0.54848
Equal variances not assumed			0.746	83.910	0.458	0.15067	0.20205	-0.25114	0.55247

Note: Only data used where participants passed the attention check.

Appendix 5. Comparison of Failures Scenarios (i.e., Product-Related versus Personal Failure Scenario) – Pretest 2

		N	Mean	SD	SE Mean
Satisfaction	*Personal*	47	4.4539	1.41494	0.20639
	Product-Related	48	1.9375	1.05556	0.15236
Product Experience	*Personal*	47	3.9149	1.19017	0.17360
	Product-Related	48	2.3125	1.39385	0.20118
Failure Severity	*Personal*	47	3.0142	1.24131	0.18106
	Product-Related	48	4.2847	1.65741	0.23923

Note: Only data used where participants passed the attention check.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% CI of the Difference	
	F	Sig.	t	df	Two-Sided p	Mean Difference	SE Difference	Lower	Upper
Satisfaction									
Equal variances assumed	7.778	0.006	9.839	93	0.000	2.51640	0.25576	2.00852	3.02428
Equal variances not assumed			9.809	85.070	0.000	2.51640	0.25653	2.00635	3.02645
Product Experience									
Equal variances assumed	0.001	0.979	6.020	93	0.000	1.60239	0.26618	1.07382	2.13097
Equal variances not assumed			6.030	9,320	0.000	1.60239	0.26573	1.07457	2.13021
Failure Severity									
Equal variances assumed	5.264	0.024	-4.222	93	0.000	-1.27054	0.30092	-1.86811	-0.67297
Equal variances not assumed			-4.235	87.076	0.000	-1.27054	0.30002	-1.86686	-0.67422

Note: Only data used where participants passed the attention check.

Appendix 6. Manipulation Check – Second-Hand Versus New Product Scenario – Pretest 2

		N	Mean	SD	SE Mean
Manipulation Check	*Second-Hand Product*	49	6.1633	1.26220	0.18031
	New Product	46	3.6957	2.07254	0.30558

Note: Only data used where participants passed the attention check.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% CI of the Difference	
	F	Sig.	t	df	Two-Sided p	Mean Difference	SE Difference	Lower	Upper
Manipulation Check									
Equal variances assumed	22.663	0.000	7.057	93	0.000	2.46761	0.34965	1.77327	3.16195
Equal variances not assumed			6.955	73.445	0.000	2.46761	0.35481	1.76054	3.17468

Note: Only data used where participants passed the attention check.

Appendix 7. Overview of Initial Situation and Different Scenarios – Main Study

About *wearIMM*: *wearIMM* is a global clothing company with a focus on fashion for men, women, children, and teenagers. Since its founding in 1997, *wearIMM* has steadily expanded its product portfolio and also offers second-hand products (i.e., used and previously owned) in addition to brand-new fashion products.

Imagine you have bought a [**new/second-hand jacket]** from *wearIMM* to wear at some future occasion (e.g., wedding, birthday party).

Now, imagine yourself in the following situation:	
Product experience (control)	A couple of weeks after you have bought your **[new/second-hand] jacket** from the brand *wearIMM*, you are invited to the movies and decide to wear the **[new/second-hand] jacket**. You take the **[new/second-hand] jacket** out of the cabinet and are happy as it is in **the same good condition** as it was a couple of weeks ago, and you feel **comfortable** wearing it.
Product experience (negative)	A couple of weeks after you have bought your **[new/second-hand] jacket** from the brand *wearIMM***,** you are invited to the movies and decide to wear the **[new/second-hand] jacket.** You take the **[new/second-hand] jacket** out of the cabinet and are disappointed as it is in a **much worse condition** than it was a couple of weeks ago, and you do **not** feel comfortable wearing it anymore.

Appendix 8. Overview of Measures

Construct/ Variable	Item/Proxy	Precedents/ Sources
	Dependent Variables	
Negative WOM Intentions *(α = .88)*	The construct is measured on a three-item seven-point Likert-type scale from “strongly disagree” to “strongly agree”. Please indicate how strongly you agree with the following statements. (1) I will spread negative word-of-mouth about the **brand *wearIMM***. (2) I will speak ill of the **brand *wearIMM*** to my friends. (3) When my friends look for similar products, I tell them not to buy from the **brand *wearIMM***.	Adapted from Grégoire and Fisher (2006)
Purchase Intention NEW Products	The construct is measured on a one-item seven-point semantic differential scale. Based on the experience from the **scenario** with the brand *wearIMM* how likely are you to buy a **new** product from the **brand *wearIMM***? (1) Very unlikely to buy – very likely to buy	Adapted from Argo et al. (2006)
Brand Attitude *(α = .95)*	All items are measured on a four-item seven-point semantic differential scale. Below you will find four pairs of adjectives. Indicate how well one or the other adjective in each pair describes your overall feeling of the **brand *wearIMM***. (1) Bad - Good (2) Like -Dislike (reverse) (3) Unpleasant – Pleasant (4) Useless – Useful	Adapted from (Bergkvist & Rossiter, 2009)
Positive WOM Intentions *(α = .97)*	The construct is measured on a four item seven-point Likert-type from “extremely unlikely” to “extremely likely”. Please state how likely you are to engage in the following behavior. (1) I am likely to say positive things about the **brand *wearIMM*** to other people. (2) I am likely to recommend the **brand *wearIMM*** to a friend or colleague. (3) I am likely to say positive things about the **brand *wearIMM*** in general to other people. (4) I am likely to encourage friends and relatives to shop the **brand *wearIMM.***	Adapted from Brüggen et al. (2011) and Zeithaml et al. (1996)

Appendix 8. Overview of Measures (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
	Mediation and Independent Variables	
Consumer Brand Forgiveness *(α = .89)*	The construct is measured on a nine-item seven-point Likert scale from “strongly disagree” to “strongly agree”. Please indicate the extent to which you agree with the following statements regarding the **brand *wearIMM***. ***Cognitive*** (1) I disapprove of this brand. (reverse) (2) I think the brand should get what it deserves. (reverse) (3) I wish that others could see that this brand is unworthy. (reverse) ***Affective*** (4) I feel sympathetic toward this brand. (5) I have compassion for the brand. (6) I feel as if I have faith in this brand. ***Behavioral*** (7) I avoid using this brand. (reverse) (8) I do not consider this brand anymore when evaluating alternatives. (reverse) (9) I am less likely to try this brand again. (reverse)	Christodoulides et al. (2021)
Underdog **(Dimension: External Disadvantage)** *(α = .90)*	The construct is measured on a four-item seven-point Likert scale from “strongly disagree” to “strongly agree”. Please indicate how strongly you agree with the following statements about the **[new/second-hand]** jacket. (1) The product starts from a disadvantaged position compared to other products. (2) There are more barriers in the way of this product succeeding compared to other products. (3) The product struggles more than other products to be successful. (4) The product is in a minority trying to gain acceptance. (5) The odds are against the product in performing well compared to other products. (6) The product has to compete with other products, which have a more resourceful background than this product. (7) The product has to face more discrimination compared to other products.	Adapted from Paharia et al. (2011)
Underdog **(Dimension: Passion and Determination)** *(α = .70)*	The construct is measured on a four-item seven-point Likert scale from “strongly disagree” to “strongly agree”. Please indicate how strongly you agree with the following statements about the **[new/second-hand]** jacket. (1) The product is more resistant than other products in the face of adversity. (2) Compared to other products, the sale of the product requires more passion. (3) When others expect the product to fail, it will remain on the market. (4) Compared to other products, this product is not given up easily. (5) Even when the product failed, the product is not given up. (6) The product has it harder compared to other products to succeed even if there are obstacles.	Adapted from Paharia et al. (2011)

Appendix 8. Overview of Measures (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
Scarcity *(α = .92)*	The construct is measured on a three-item seven-point Likert scale from “not likely at all” to “very likely”. How likely is it that the jacket from the scenario is... (1) … in short supply. (2) … scarce. (3) … not available in sufficient quantities.	Adapted from Mukherjee and Lee (2016)
	Manipulation Check	
Product Type (second-hand versus new) *(α = .96)*	All items are measured on a three-item seven-point scale from “strongly disagree” to “strongly agree”. Please indicate the extent to which you agree with the following statements regarding the **jacket from the scenario**. (1) This jacket was previously owned. (2) This jacket is not new. (3) This jacket is a second-hand product.	Extended and adapted from Bezançon et al. (2019)
Perception of Product Experience *(α = .97)*	The construct measured on a two-item seven-point scale from “very negative” to “very positive” and “extremely unsatisfied” to “extremely satisfied”. How would you rate the experience as stated in the scenario in using the **jacket**? (1) very negative - very positive (2) extremely unsatisfied - extremely satisfied	Adapted from Zhang et al. (2014)
Product Failure Severity *(α = .96)*	The construct is measured on a three-item seven-point bipolar semantic differential scale. The experience with the **jacket** from the scenario caused me… (1) minor problems – major problems (2) small inconveniences – big inconveniences (3) minor troubles – major troubles	Adapted from Grégoire and Fisher (2008)
	Attention Check	
Attention Check	The construct is measured on a one-item seven-point scale from “strongly disagree” to “strongly agree”. (1) Please select "strongly disagree" as your answer.	Adapted from Gruzd et al. (2020)
	Covariates and Controls	
Experience with Second-Hand Clothing *(α = .96)*	The construct is measured on a seven-point Likert-type scale from “strongly disagree” to “strongly agree”. Thinking about your purchasing habits please indicate how strongly you agree with the following statements about **second-hand clothing**. (1) I have a great deal of experience in buying **second-hand clothes.** (2) I frequently shop for **second-hand clothes**. (3) I am very confident in shopping for **second-hand clothes.**	Adapted from Lo et al. (2019)

Appendix 8. Overview of Measures (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
Preference for Second-Hand Clothing *(α = .94)*	The construct is measured on a seven-point bipolar semantic differential scale from "new clothing" to "second-hand clothing". Thinking about your shopping preferences, please answer the following questions: (1) Which product type do you like more in terms of clothing? (2) Which product type are you more favorable toward in terms of clothing? (3) Which product type are you more likely to buy in terms of clothing?	Adapted from Cheng et al. (2017)
Product Involvement *(α = .97)*	The construct is measured on a four-item seven-point Likert-type scale from "strongly disagree" to "strongly agree". Please indicate to what extent you agree with the following statements: (1) In general, I have a strong interest in clothing. (2) Clothing is very important to me. (3) Clothing matters a lot to me. (4) Clothing means a lot to me.	Adapted from Brocato et al. (2015)
	Demographics	
Gender	Please indicate which gender you feel most closely aligned with: • Male • Female • Non-binary/third gender • Prefer not to say	Self-developed
Age	How old are you? (Open question) ___________	Self-developed
Education	What is the highest level of education you have achieved? • Some school but no degree • High school graduate • Some college but no degree • Bachelor's degree • Master's degree • Professional degree • Doctorate degree • Other	Adapted from Lo et al. (2019)
Employment	What is your current employment status? • Full-time • Part-time • Self-employed • Student • Stay-at-home parent • Unemployed • Retired	Adapted from Lo et al. (2019)

Appendix 8. Overview of Measures (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
Income	The declaration of average gross yearly income is voluntary. What is the level of your annual gross household income? (1) < $10,000 (2) $10,001 – $50,000 (3) $50,001 – $90,000 (4) $90,001 - $150,000 (5) > $150,001	Adapted from Lo et al. (2019)

Appendix 9. Comparison of Product Type and Experience Scenario on Negative WOM – Main Study

			95% CI	
	Mean	**SE**	**Lower Bound**	**Upper Bound**
Second-Hand Product, Negative Experience	3.860[a]	0.168	3.528	4.191
Second-Hand Product, Control Experience	1.649[a]	0.179	1.295	2.002
New Product, Negative Experience	5.179[a]	0.183	4.818	5.540
New Product, Control Experience	1.902a	0.181	1.544	2.259

a. Covariates appearing in the model are evaluated at the following values: Product involvement = 4.3797, Experience with second-hand clothing = 3.5027, Preference for second-hand clothing = 2.9857.

					95% CI for Difference[b]	
		Mean Differ-ence (I-J)	**SE**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
Second-Hand Product, Negative Experience	*Second-Hand Product, Control Experience*	2.211*	0.246	0.000	1.554	2.868
	New Product, Negative Experience	-1.319*	0.248	0.000	-1.981	-0.657
	New Product, Control Experience	1.958*	0.247	0.000	1.298	2.617
Second-Hand Product, Control Experience	*Second-Hand Product, Negative Experience*	-2.211*	0.246	0.000	-2.868	-1.554
	New Product, Negative Experience	-3.530*	0.257	0.000	-4.215	-2.846
	New Product, Control Experience	-0.253	0.255	1.000	-0.933	0.427
New Product, Negative Experience	*Second-Hand Product, Negative Experience*	1.319*	0.248	0.000	0.657	1.981
	Second-Hand Product, Control Experience	3.530*	0.257	0.000	2.846	4.215
	New Product, Control Experience	3.277*	0.258	0.000	2.588	3.966
New Product, Control experience	*Second-Hand Product, Negative Experience*	-1.958*	0.247	0.000	-2.617	-1.298
	Second-Hand Product, Control Experience	0.253	0.255	1.000	-0.427	0.933
	New Product, Negative Experience	-3.277*	0.258	0.000	-3.966	-2.588

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 10. Comparison of Product Type and Experience Scenario on Purchase Intention Toward New Products – Main Study

			95% CI	
	Mean	Std. Error	Lower Bound	Upper Bound
Second-Hand Product, Negative Experience	3.646[a]	0.219	3.214	4.078
Second-Hand Product, Control Experience	4.480[a]	0.234	4.019	4.941
New Product, Negative Experience	2.701[a]	0.239	2.230	3.172
New Product, Control Experience	5.278[a]	0.236	4.812	5.744

a. Covariates appearing in the model are evaluated at the following values: Product involvement = 4.3797, Experience with second-hand clothing = 3.5027, Preference for second-hand clothing = 2.9857.

		Mean Difference (I-J)	SE	Sig.[b]	95% CI for Difference[b] Lower Bound	Upper Bound
Second-Hand Product, Negative Experience	*Second-Hand Product, Control Experience*	-0.834	0.321	0.061	-1.691	0.023
	New Product, Negative Experience	.945*	0.324	0.024	0.082	1.808
	New Product, Control Experience	-1.632*	0.322	0.000	-2.492	-0.772
Second-Hand Product, Control Experience	*Second-Hand, Negative Experience*	0.834	0.321	0.061	-0.023	1.691
	New Product, Negative Experience	1.779*	0.334	0.000	0.886	2.671
	New Product, Control Experience	-0.798	0.332	0.104	-1.685	0.088
New Product, Negative Experience	*Second-Hand Product, Negative Experience*	-.945*	0.324	0.024	-1.808	-0.082
	Second-Hand Product, Control Experience	-1.779*	0.334	0.000	-2.671	-0.886
	New Product, Control Experience	-2.577*	0.337	0.000	-3.475	-1.678
New Product, Control experience	*Second-Hand Product, Negative Experience*	1.632*	0.322	0.000	0.772	2.492
	Second-Hand Product, Control Experience	0.798	0.332	0.104	-0.088	1.685
	New Product, Negative Experience	2.577*	0.337	0.000	1.678	3.475

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 11. Comparison of Product Type and Experience Scenario on Brand Attitude – Main Study

	Mean	SE	95% CI Lower Bound	95% CI Upper Bound
Second-Hand Product, Negative Experience	3.028[a]	0.130	2.772	3.284
Second-Hand Product, Control Experience	5.915[a]	0.138	5.642	6.187
New Product, Negative Experience	2.078[a]	0.141	1.800	2.357
New Product, Control Experience	6.000[a]	0.140	5.725	6.276

a. Covariates appearing in the model are evaluated at the following values: Product involvement = 4.3797, Experience with second-hand clothing = 3.5027, Preference for second-hand clothing = 2.9857.

		Mean Difference (I-J)	SE	Sig.[b]	95% CI for Difference[b] Lower Bound	95% CI for Difference[b] Upper Bound
Second-Hand Product, Negative Experience	*Second-Hand Product, Control Experience*	-2.887*	0.190	0.000	-3.394	-2.380
	New Product, Negative Experience	.950*	0.191	0.000	0.439	1.460
	New Product, Control Experience	-2.972*	0.191	0.000	-3.481	-2.464
Second-Hand Product, Control Experience	*Second-Hand Product, Negative Experience*	2.887*	0.190	0.000	2.380	3.394
	New Product, Negative Experience	3.837*	0.198	0.000	3.309	4.364
	New Product, Control Experience	-0.086	0.197	1.000	-0.610	0.439
New Product, Negative Experience	*Second-Hand Product, Negative Experience*	-.950*	0.191	0.000	-1.460	-0.439
	Second-Hand Product, Control Experience	-3.837*	0.198	0.000	-4.364	-3.309
	New Product, Control Experience	-3.922*	0.199	0.000	-4.454	-3.391
New Product, Control Experience	*Second-Hand Product, Negative Experience*	2.972*	0.191	0.000	2.464	3.481
	Second-Hand Product, Control Experience	0.086	0.197	1.000	-0.439	0.610
	New Product, Negative Experience	3.922*	0.199	0.000	3.391	4.454

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 12. Comparison of Product Type and Experience Scenario on Consumer Brand Forgiveness – Main Study

			95% CI	
	Mean	**SE**	**Lower Bound**	**Upper Bound**
Second-Hand Product, Negative Experience	3.678[a]	0.122	3.438	3.918
Second-Hand Product, Control Experience	5.328[a]	0.130	5.072	5.584
New Product, Negative Experience	2.856[a]	0.132	2.594	3.117
New Product, Control Experience	5.146[a]	0.131	4.887	5.404

a. Covariates appearing in the model are evaluated at the following values: Product involvement = 4.3797, Experience with second-hand clothing = 3.5027, Preference for second-hand clothing = 2.9857.

					95% CI for Difference[b]	
		Mean Difference (I-J)	**SE**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
Second-Hand Product, Negative Experience	*Second-Hand Product, Control Experience*	-1.650*	0.178	0.000	-2.126	-1.175
	New Product, Negative Experience	.823*	0.180	0.000	0.343	1.302
	New Product, Control Experience	-1.467*	0.179	0.000	-1.945	-0.990
Second-Hand Product, Control Experience	*Second-Hand Product, Negative Experience*	1.650*	0.178	0.000	1.175	2.126
	New Product, Negative Experience	2.473*	0.186	0.000	1.978	2.968
	New Product, Control Experience	0.183	0.184	1.000	-0.309	0.675
New Product, Negative Experience	*Second-Hand Product, Negative Experience*	-.823*	0.180	0.000	-1.302	-0.343
	Second-Hand Product, Control Experience	-2.473*	0.186	0.000	-2.968	-1.978
	New Product, Control Experience	-2.290*	0.187	0.000	-2.789	-1.791
New Product, Control Experience	*Second-Hand Product, Negative Experience*	1.467*	0.179	0.000	0.990	1.945
	Second-Hand Product, Control Experience	-0.183	0.184	1.000	-0.675	0.309
	New Product, Negative Experience	2.290*	0.187	0.000	1.791	2.789

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 13. Comparison of Product Type on the Underdog Effect – Main Study

			95% CI	
	Mean	SE	Lower Bound	Upper Bound
Second-Hand Product	4.358[a]	0.088	4.185	4.532
New Product	3.550[a]	0.092	3.367	3.732

a. Covariates appearing in the model are evaluated at the following values: Product involvement = 4.3797, Experience with second-hand clothing = 3.5027, Preference for second-hand clothing = 2.9857.

					95% CI for Difference[b]	
		Mean Differ-ence (I-J)	SE	Sig.[b]	Lower Bound	Upper Bound
Second-Hand Product	*New Product*	.809*	0.128	0.000	0.556	1.061
New Product	*Second-Hand Product*	-.809*	0.128	0.000	-1.061	-0.556

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 14. Comparison of Product Type on Scarcity – Main Study

			95% CI	
	Mean	**SE**	**Lower Bound**	**Upper Bound**
Second-Hand Product	4.821[a]	0.145	4.535	5.106
New Product	3.625[a]	0.152	3.325	3.924

a. Covariates appearing in the model are evaluated at the following values: Product involvement = 4.3797, Experience with second-hand clothing = 3.5027, Preference for second-hand clothing = 2.9857.

		Mean Differ-ence (I-J)	**SE**	**Sig.[b]**	**95% CI for Difference[b] Lower Bound**	**Upper Bound**
Second-Hand Product	*New Product*	-1.650*	0.178	0.000	-2.126	-1.175
New Product	*Second-Hand Product*	1.650*	0.178	0.000	1.175	2.126

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 15. Mediation Analysis

Underdog Dimension: External Disadvantage

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.3485	0.1215	1.4965	3.2841	4	91	0.0146

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	4.0956	0.5755	7.1167	0	2.9524	5.2387
Product Type	0.7381	0.2638	2.798	0.0063	0.2141	1.2622
Product Involvement	-0.0443	0.0764	-0.5801	0.5633	-0.1961	0.1075
Experience Second-Hand Clothing	-0.0472	0.0945	-0.4998	0.6185	-0.235	0.1405
Preference for Second-Hand Clothing	-0.105	0.1123	-0.935	0.3523	-0.3282	0.1181

Underdog Dimension: Passion and Determination

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.3892	0.1514	0.7599	4.2293	4	91	0.0035

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	3.222	0.3803	8.4726	0	2.4666	3.9774
Product Type	0.5977	0.1847	3.2364	0.0017	0.2308	0.9645
Product Involvement	0.0679	0.0604	1.1255	0.2633	-0.052	0.1878
Experience Second-Hand Clothing	0.028	0.091	0.3072	0.7594	-0.1528	0.2088
Preference for Second-Hand Clothing	0.0613	0.1063	0.5769	0.5654	-0.1498	0.2725

Consumer Brand Forgiveness

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.4742	0.2248	1.0993	3.6725	6	89	0.0026

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	2.4885	0.7388	3.3683	0.0011	1.0205	3.9565
Product Type	0.7684	0.2334	3.2923	0.0014	0.3047	1.2322
External Disadvantage	-0.2319	0.1061	-2.186	0.0314	-0.4426	-0.0211
Passion and Determination	0.3643	0.1751	2.0801	0.0404	0.0163	0.7123
Product Involvement	-0.0599	0.0717	-0.836	0.4054	-0.2024	0.0825
Experience Second-Hand Clothing	0.0398	0.0805	0.4942	0.6224	-0.1201	0.1996
Preference for Second-Hand Clothing	-0.0306	0.113	-0.271	0.7871	-0.255	0.1938

Appendix 16. Mediation Analysis on Negative WOM

R	R-sq	MSE	F(HC3)	df1	df2	p
0.8278	0.6853	0.8301	27.1813	7	88	0

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	7.1961	0.6106	11.7859	0.0000	5.9827	8.4095
Product Type	-0.6216	0.2380	-2.6122	0.0106	-1.0946	-0.1487
External Disadvantage	0.0699	0.0986	0.7091	0.4801	-0.1260	0.2658
Passion and Determination	0.1562	0.1576	0.9914	0.3242	-0.1569	0.4694
Consumer Brand Forgiveness	-1.0124	0.0951	-10.6457	0.0000	-1.2014	-0.8234
Product Involvement	0.0320	0.0608	0.5255	0.6005	-0.0889	0.1528
Experience Second-Hand Clothing	0.0610	0.0651	0.9373	0.3512	-0.0683	0.1903
Preference for Second-Hand Clothing	-0.1075	0.0889	-1.2091	0.2299	-0.2842	0.0692

Total Effect Model

R	R-sq	MSE	F(HC3)	df1	df2	p
0.455	0.2071	2.0225	6.1256	4	91	0.0002

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	5.2393	0.5493	9.5375	0.0000	4.1481	6.3304
Product Type	-1.3018	0.2906	-4.4791	0.0000	-1.8792	-0.7245
Product Involvement	0.0647	0.0806	0.8025	0.4244	-0.0954	0.2248
Experience Second-Hand Clothing	0.0004	0.0907	0.0043	0.9966	-0.1798	0.1806
Preference for Second-Hand Clothing	-0.1215	0.1307	-0.9300	0.3548	-0.3812	0.1381

Total, Direct, and Indirect Effects of X on Y

Total effect of X on Y

Effect	se(HC3)	t	p	LLCI	ULCI
-1.3018	0.2906	-4.4791	0	-1.8792	-0.7245

Direct effect of X on Y

Effect	se(HC3)	t	p	LLCI	ULCI
-0.6216	0.238	-2.6122	0.0106	-1.0946	-0.1487

Indirect effect(s) of X on Y:

	Effect	BootSE	Boot-LLCI	BootULCI
TOTAL	-0.6802	0.2420	-1.1489	-0.1985
Product Type → External Disadvantage → Negative WOM	0.0516	0.0730	-0.0914	0.2041
Product Type → Passion and Determination → Negative WOM	0.0934	0.0949	-0.0756	0.3021
Product Type → Consumer Brand Forgiveness → Negative WOM	-0.7780	0.2251	-1.2314	-0.3296
Product Type → External Disadvantage → Consumer Brand Forgiveness → Negative WOM	0.1733	0.0854	0.0261	0.3525
Product Type → Passion and Determination → Consumer Brand Forgiveness → Negative WOM	-0.2205	0.1230	-0.4776	-0.0039

Appendix 17. Mediation Analysis on Purchase Intention Toward New Product

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.6356	0.404	1.8252	10.1824	7	88	0

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	-0.1766	0.7930	-0.2227	0.8243	-1.7524	1.3993
Product Type	0.0896	0.3457	0.2590	0.7962	-0.5975	0.7766
External Disadvantage	0.0893	0.1400	0.6378	0.5253	-0.1889	0.3674
Passion and Determination	0.1710	0.2154	0.7938	0.4295	-0.2571	0.5991
Consumer Brand Forgiveness	0.8392	0.1412	5.9431	0.0000	0.5586	1.1198
Product Involvement	0.0173	0.0833	0.2077	0.8359	-0.1483	0.1829
Experience Second-Hand Clothing	0.0420	0.1023	0.4105	0.6824	-0.1612	0.2452
Preference for Second-Hand Clothing	-0.2352	0.1219	-1.9303	0.0568	-0.4774	0.0069

Total Effect Model

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.3146	0.0989	2.6684	2.2736	4	91	0.0673

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	3.016	0.576	5.240	0.000	1.873	4.160
Product Type	0.942	0.351	2.681	0.009	0.244	1.639
Product Involvement	0.004	0.101	0.040	0.968	-0.197	0.205
Experience Second-Hand Clothing	0.094	0.139	0.676	0.501	-0.182	0.369
Preference for Second-Hand Clothing	-0.221	0.155	-1.426	0.157	-0.528	0.087

Total, Direct, and Indirect Effects of X on Y

Total effect of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
	0.9416	0.3513	2.6807	0.0087	0.2439	1.6393

Direct effect of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
	0.0896	0.3457	0.259	0.7962	-0.5975	0.7766

Indirect effect(s) of X on Y:

	Effect	BootSE	Boot-LLCI	BootULCI
TOTAL	0.8521	0.2305	0.4222	1.3128
Product Type → External Disadvantage → Purchase Intention Toward New Product	0.0659	0.1089	-0.1200	0.3229
Product Type → Passion and Determination → Purchase Intention Toward New Product	0.1022	0.1297	-0.1546	0.3784
Product Type → Consumer Brand Forgiveness → Purchase Intention Toward New Product	0.6449	0.2051	0.2566	1.0662
Product Type → External Disadvantage → Consumer Brand Forgiveness → Purchase Intention Toward New Product	-0.1436	0.0717	-0.2972	-0.0155
Product Type → Passion and Determination → Consumer Brand Forgiveness → Purchase Intention Toward New Product	0.1827	0.1044	0.0037	0.4117

Appendix 18. Mediation Analysis on Brand Attitude

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.8186	0.6701	0.4906	20.9516	7	88	0

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	0.2184	0.3078	0.7094	0.48	-0.3934	0.8301
Product Type	0.4159	0.1726	2.4099	0.018	0.0729	0.7588
External Disadvantage	-0.1012	0.0588	-1.7208	0.0888	-0.218	0.0157
Passion and Determination	0.0048	0.1102	0.0438	0.9652	-0.2141	0.2238
Consumer Brand Forgiveness	0.7354	0.0783	9.3896	0.0000	0.5797	0.8910
Product Involvement	-0.0032	0.0456	-0.0695	0.9448	-0.0939	0.0875
Experience Second-Hand Clothing	0.0303	0.0491	0.6181	0.5381	-0.0672	0.1279
Preference for Second-Hand Clothing	-0.0032	0.059	-0.0542	0.9569	-0.1205	0.1141

Total Effect Model

	R	R-sq	MSE	F(HC3)	df1	df2	p
	0.4405	0.194	1.1591	5.0615	4	91	0.001

	coeff	se(HC3)	t	p	LLCI	ULCI
Constant	1.8144	0.3941	4.6035	0.0000	1.0315	2.5973
Product Type	0.9434	0.2201	4.2865	0.0000	0.5062	1.3806
Product Involvement	-0.0167	0.0640	-0.2608	0.7948	-0.1438	0.1104
Experience Second-Hand Clothing	0.0800	0.0779	1.0276	0.3069	-0.0747	0.2348
Preference for Second-Hand Clothing	0.0196	0.1068	0.1831	0.8551	-0.1926	0.2317

Total, Direct, and Indirect Effects of X on Y

Total effect of X on Y

Effect	se(HC3)	t	p	LLCI	ULCI
0.9434	0.2201	4.2865	0	0.5062	1.3806

Direct effect of X on Y

Effect	se(HC3)	t	p	LLCI	ULCI
0.4159	0.1726	2.4099	0.018	0.0729	0.7588

Indirect effect(s) of X on Y:

	Effect	BootSE	Boot-LLCI	BootULCI
TOTAL	0.5275	0.1857	0.1739	0.9096
Product Type → External Disadvantage → Brand Attitude	-0.0747	0.0510	-0.1905	0.0064
Product Type → Passion and Determination → Brand Attitude	0.0029	0.0646	-0.1219	0.1421
Product Type → Consumer Brand Forgiveness → Brand Attitude	0.5651	0.1870	0.2232	0.9429
Product Type → External Disadvantage → Consumer Brand Forgiveness → Brand Attitude	-0.1259	0.0623	-0.2590	-0.0166
Product Type → Passion and Determination → Consumer Brand Forgiveness → Brand Attitude	0.1601	0.0858	0.0022	0.3386

References

Aaker, J., Fournier, S., & Brasel, S. A. (2004). When Good Brands Do Bad. *Journal of Consumer Research*, *31*(1), 1–16. https://doi.org/10.1086/383419

Abbes, I., Hallem, Y., & Taga, N. (2020). Second-Hand Shopping and Brand Loyalty: The Role of Online Collaborative Redistribution Platforms. *Journal of Retailing and Consumer Services*, *52*, 1–11. https://doi.org/10.1016/j.jretconser.2019.101885

Ackerman, D. S., & Hu, J. (2017). Assuring Me That it is as 'Good as New' Just Makes Me Think About How Someone Else Used It. Examining Consumer Reaction Toward Marketer-Provided Information About Secondhand Goods. *Journal of Consumer Behaviour*, *16*(3), 233–241. https://doi.org/10.1002/cb.1631

Ahluwalia, R. (2002). How Prevalent Is the Negativity Effect in Consumer Environments? *Journal of Consumer Research*, *29*(2), 270–279. https://doi.org/10.1086/341576

Ahluwalia, R., & Gürhan-Canli, Z. (2000). The Effects of Extensions on the Family Brand Name: An Accessibility-Diagnosticity Perspective. *Journal of Consumer Research*, *27*(3), 371–381. https://doi.org/10.1086/317591

Ahluwalia, R., Unnava, H. R., & Burnkrant, R. E. (2001). The Moderating Role of Commitment on the Spillover Effect of Marketing Communications. *Journal of Marketing Research*, *38*(4), 458–470. https://doi.org/10.1509/jmkr.38.4.458.18903

Allen, A. M., Brady, M. K., Robinson, S. G., & Voorhees, C. M. (2015). One Firm's Loss is Another's Gain: Capitalizing on Other Firms' Service Failures. *Journal of the Academy of Marketing Science*, *43*(5), 648–662. https://doi.org/10.1007/s11747-014-0413-6

Argo, J. J., Dahl, D. W., & Morales, A. C. (2006). Consumer Contamination: How Consumers React to Products Touched by Others. *Journal of Marketing*, *70*(2), 81–94. https://doi.org/10.1509/jmkg.70.2.081

Armstrong, C. M., Niinimäki, K., Lang, C., & Kujala, S. (2016). A Use-Oriented Clothing Economy? Preliminary Affirmation for Sustainable Clothing Consumption Alternatives. *Sustainable Development*, *24*(1), 18–31. https://doi.org/10.1002/sd.1602

Ashfaq, M., Yun, J., Waheed, A., Khan, M. S., & Farrukh, M. (2019). Customers' Expectation, Satisfaction, and Repurchase Intention of Used Products Online: Empirical Evidence From China. *SAGE Open*, *9*(2), 1-14. https://doi.org/10.1177/2158244019846212

Bardhi, F., & Arnould, E. J. (2005). Thrift Shopping: Combining Utilitarian Thrift and Hedonic Treat Benefits. *Journal of Consumer Behaviour*, *4*(4), 223–233. https://doi.org/10.1002/cb.12

Becker-Leifhold, C., & Iran, S. (2018). Collaborative Fashion Consumption – Drivers, Barriers and Future Pathways. *Journal of Fashion Marketing and Management: An International Journal*, *22*(2), 189–208. https://doi.org/10.1108/JFMM-10-2017-0109

Belk, R. W. (1988). Possessions and the Extended Self. *Journal of Consumer Research*, *15*(2), 139. https://doi.org/10.1086/209154

Bergkvist, L. I., & Rossiter, J. R. (2009). Tailor-Made Single-Item Measures of Doubly Concrete Constructs. *International Journal of Advertising*, *28*(4), 607–621. https://doi.org/10.2501/S0265048709200783

Bezançon, M., Guiot, D., & Le Nagard, E. (2019). The Role of Negative Physical Contagion in the Online Purchase of Second-Hand Products. *Recherche Et Applications En Marketing (English Edition)*, *34*(4), 2–28. https://doi.org/10.1177/2051570719881944

Borusiak, B., Szymkowiak, A., Horska, E., Raszka, N., & Żelichowska, E. (2020). Towards Building Sustainable Consumption: A Study of Second-Hand Buying Intentions. *Sustainability*, *12*(3), 875. https://doi.org/10.3390/su12030875

Brandt, M. (2021, March 16). *Second-hand? Net, spasibo!* Statista Global Consumer Survey. https://de.statista.com/infografik/24421/umfrage-zum-kauf-von-second-hand-produkten/

Brocato, E. D., Baker, J., & Voorhees, C. M. (2015). Creating Consumer Attachment to Retail Service Firms Through Sense of Place. *Journal of the Academy of Marketing Science*, *43*(2), 200–220. https://doi.org/10.1007/s11747-014-0381-x

Brüggen, E. C., Foubert, B., & Gremler, D. D. (2011). Extreme Makeover: Short- and Long-Term Effects of a Remodeled Servicescape. *Journal of Marketing*, *75*(5), 71–87. https://doi.org/10.1509/jmkg.75.5.71

Bruner, G. C. (2013). *Marketing Scales Handbook: Multi-Item Measures for Consumer Insight Research* (Volume 7). GCBII Productions LLC.

Casidy, R., Duhachek, A., Singh, V., & Tamaddoni, A. (2021). Religious Belief, Religious Priming, and Negative Word of Mouth. *Journal of Marketing Research*, *58*(4), 762–781. https://doi.org/10.1177/00222437211011196

Cervellon, M.-C., Carey, L., & Harms, T. (2012). Something Old, Something Used: Determinants of Women's Purchase of Vintage Fashion vs Second-Hand Fashion.

International Journal of Retail & Distribution Management, *40*(12), 956–974. https://doi.org/10.1108/09590551211274946

Chen, Z., & Lurie, N. H. (2013). Temporal Contiguity and Negativity Bias in the Impact of Online Word of Mouth. *Journal of Marketing Research*, *50*(4), 463–476. https://doi.org/10.1509/jmr.12.0063

Cheng, Y., Mukhopadhyay, A., & Schrift, R. Y. (2017). Do Costly Options Lead to Better Outcomes? How the Protestant Work Ethic Influences the Cost–Benefit Heuristic in Goal Pursuit. *Journal of Marketing Research*, *54*(4), 636–649. https://doi.org/10.1509/jmr.15.0105

Christodoulides, G., Gerrath, M. H. E. E., & Siamagka, N. T. (2021). Don't be Rude! The Effect of Content Moderation on Consumer-Brand Forgiveness. *Psychology & Marketing*, *38*(10), 1686–1699. https://doi.org/10.1002/mar.21458

Crosno, J. L., & Cui, A. P. (2018). Something Old, Something New: The Role of Partitioned Pricing in Consumers' Preference for New Versus Used Products. *Journal of Consumer Marketing*, *35*(4), 353–365. https://doi.org/10.1108/JCM-02-2017-2091

Darley, W. K., Blankson, C., & Luethge, D. J. (2010). Toward an Integrated Framework for Online Consumer Behavior and Decision Making Process: A Review. *Psychology & Marketing*, *27*(2), 94–116. https://doi.org/10.1002/mar.20322

Daugherty, T., & Hoffman, E. (2014). eWOM and the Importance of Capturing Consumer Attention Within Social Media. *Journal of Marketing Communications*, *20*(1-2), 82–102. https://doi.org/10.1080/13527266.2013.797764

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press. http://www.loc.gov/catdir/enhancements/fy0604/92012048-d.html

Delgado-Ballester, E. (2021). Effect of Underdog (vs Topdog) Brand Storytelling on Brand Identification: Exploring Multiple Mediation Mechanisms. *Journal of Product & Brand Management*, *30*(4), 626–638. https://doi.org/10.1108/JPBM-11-2019-2639

DeLong, M., Heinemann, B., & Reiley, K. (2005). Hooked on Vintage! *Fashion Theory*, *9*(1), 23–42. https://doi.org/10.2752/136270405778051491

Einwiller, S. A., Fedorikhin, A., Johnson, A. R., & Kamins, M. A. (2006). Enough Is Enough! When Identification No Longer Prevents Negative Corporate Associations. *Journal of the Academy of Marketing Science*, *34*(2), 185–194. https://doi.org/10.1177/0092070305284983

Ek Styvén, M., & Mariani, M. M. (2020). Understanding the Intention to Buy Secondhand Clothing on Sharing Economy Platforms: The Influence of Sustainability, Distance from the Consumption System, and Economic Motivations. *Psychology & Marketing*, *37*(5), 724–739. https://doi.org/10.1002/mar.21334

Fedorikhin, A., Park, C. W., & Thomson, M. (2008). Beyond Fit and Attitude: The Effect of Emotional Attachment on Consumer Responses to Brand Extensions. *Journal of Consumer Psychology*, *18*(4), 281–291. https://doi.org/10.1016/j.jcps.2008.09.006

Fehr, E., & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Fernández-Capo, M., Fernández, S. R., Sanfeliu, M. G., Benito, J. G., & Worthington, E. L. (2017). Measuring Forgiveness. *European Psychologist*, *22*(4), 247–262. https://doi.org/10.1027/1016-9040/a000303

Fernando, A. G., Sivakumaran, B., & Suganthi, L. (2018). Comparison of Perceived Acquisition Value Sought by Online Second-Hand and New Goods Shoppers. *European Journal of Marketing*, *52*(7/8), 1412–1438. https://doi.org/10.1108/EJM-01-2017-0048

Ferraro, C., Sands, S., & Brace-Govan, J. (2016). The Role of Fashionability in Second-Hand Shopping Motivations. *Journal of Retailing and Consumer Services*, *32*, 262–268. https://doi.org/10.1016/j.jretconser.2016.07.006

Fetscherin, M., & Sampedro, A. (2019). Brand Forgiveness. *Journal of Product & Brand Management*, *28*(5), 633–652. https://doi.org/10.1108/JPBM-04-2018-1845

Field, A. (2011). *Discovering Statistics Using SPSS: (and Sex and Drugs and Rock 'n' Roll)* (3. ed., reprinted.). SAGE.

Genesove, D., & Mayer, C. (2001). Loss Aversion and Seller Behavior: Evidence From the Housing Market. *Quarterly Journal of Economics*, *116*(4), 1233–1260.

Ghose, A., Smith, M. D., & Telang, R. (2006). Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact. *Information Systems Research*, *17*(1), 3–19. https://doi.org/10.1287/isre.1050.0072

Goldschmied, N. P., & Vandello, J. A. (2012). The Future is Bright: The Underdog Label, Availability, and Optimism. *Basic and Applied Social Psychology*, *34*(1), 34–43. https://doi.org/10.1080/01973533.2011.637726

Grégoire, Y., & Fisher, R. J. (2006). The Effects of Relationship Quality on Customer Retaliation. *Marketing Letters*, *17*(1), 31–46. https://doi.org/10.1007/s11002-006-3796-4

Grégoire, Y., & Fisher, R. J. (2008). Customer Betrayal and Retaliation: When your Best Customers Become your Worst Wnemies. *Journal of the Academy of Marketing Science*, *36*(2), 247–261. https://doi.org/10.1007/s11747-007-0054-0

Gruzd, A., Jacobson, J., & Dubois, E. (2020). Cybervetting and the Public Life of Social Media Data. *Social Media + Society*, *6*(2). https://doi.org/10.1177/2056305120915618

Guiot, D., & Roux, D. (2010). A Second-Hand Shoppers' Motivation Scale: Antecedents, Consequences, and Implications for Retailers. *Journal of Retailing*, *86*(4), 383–399. https://doi.org/10.1016/j.jretai.2010.08.002

Gullstrand Edbring, E., Lehner, M., & Mont, O. (2016). Exploring Consumer Attitudes to Alternative Models of Consumption: Motivations and Barriers. *Journal of Cleaner Production*, *123*, 5–15. https://doi.org/10.1016/j.jclepro.2015.10.107

Hamilton, K. (2009). Consumer Decision Making in Low-Income Families: The Case of Conflict Avoidance. *Journal of Consumer Behaviour*, *8*(5), 252–267. https://doi.org/10.1002/cb.285

Hassey, R. V. (2019). How Brand Personality and Failure-Type Shape Consumer Forgiveness. *Journal of Product & Brand Management*, *28*(2), 300–315. https://doi.org/10.1108/JPBM-09-2017-1563

Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis : A Regression-Based Approach*. Guilford Publications. http://ebookcentral.proquest.com/lib/ub-lmu/detail.action?docID=6809031

Hegner, S. M., Fenko, A., & Teravest, A. (2017). Using the Theory of Planned Behaviour to Understand Brand Love. *Journal of Product & Brand Management*, *26*(1), 26–41. https://doi.org/10.1108/JPBM-06-2016-1215

Hubert, J. (2022, March 24). *Umsatz ausgewählter Anbieter von Secondhandkleidung in Deutschland im Jahr 2019: (in Millionen Euro)*. https://de.statista.com/statistik/daten/studie/1198643/umfrage/umsatz-ausgewaehlter-anbieter-von-secondhandmode/

Jin, L., & Huang, Y. (2019). How Power States Influence the Persuasiveness of Top-Dog versus Underdog Appeals. *Journal of Consumer Psychology*, *29*(2), 243–261. https://doi.org/10.1002/jcpy.1069

Joireman, J., Grégoire, Y., & Tripp, T. M. (2016). Customer Forgiveness Following Service Failures. *Current Opinion in Psychology*, *10*, 76–82. https://doi.org/10.1016/j.copsyc.2015.11.005

Keller, K. L. (2013). *Strategic Brand Management: Building, Measuring, and Managing Brand Equity* (4. ed., global ed., Pearson international ed.). Pearson. http://www.myilibrary.com?id=523713

Kim, Y., & Park, K. (2020). When the Underdog Positioning Backfires! The Effects of Ethical Transgressions on Attitudes Toward Underdog Brands. *Frontiers in Psychology*, *11*(1988), 1–15. https://doi.org/10.3389/fpsyg.2020.01988

Kim, Y., Park, K., & Stacey Lee, S. (2019). The Underdog Trap: The Moderating Role of Transgression Type in Forgiving Underdog Brands. *Psychology & Marketing*, *36*(1), 28–40. https://doi.org/10.1002/mar.21155

Koschate-Fischer, N., Hoyer, W. D., & Wolframm, C. (2019). What if Something Unexpected Happens to my Brand? Spillover Effects from Positive and Negative Events in a Co-Branding Partnership. *Psychology & Marketing*, *36*(8), 758–772. https://doi.org/10.1002/mar.21210

Lee, S. M., & Lee, S. J. (2005). Consumers' Initial Trust toward Second-Hand Products in the Electronic Market. *Journal of Computer Information Systems*, *46*(2), 85–98. https://doi.org/10.1080/08874417.2006.11645887

Li, Y., & Zhao, M. (2018). Must the Underdog Win? The Moderation Effect of Product Type in the Underdog Effect of Brand Stories. *Asian Journal of Social Psychology*, *21*(4), 237–245. https://doi.org/10.1111/ajsp.12326

Lo, C. J., Tsarenko, Y., & Tojib, D. (2019). To Tell or Not to Tell? The Roles of Perceived Norms and Self-Consciousness in Understanding Consumers' Willingness to Recommend Online Secondhand Apparel Shopping. *Psychology & Marketing*, *36*(4), 287–304. https://doi.org/10.1002/mar.21179

McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Brown, S. W., & Hight, T. L. (1998). Interpersonal Forgiving in Close Relationships: II. Theoretical Elaboration and Measurement. *Journal of Personality and Social Psychology*, *75*(6), 1586–1603. https://doi.org/10.1037/0022-3514.75.6.1586

McGinnis, L. P., & Gentry, J. W. (2009). Underdog Consumption: An Exploration Into Meanings and Motives. *Journal of Business Research*, *62*(2), 191–199. https://doi.org/10.1016/j.jbusres.2008.01.026

Meyer, C., & Schwager, A. (2007). Understanding Customer Experience. *Harvard Business Review*, *85*(2), 116-126.

Miller, C. J., & Brannon, D. C. (2022). Pursuing Premium: Comparing Pre-Owned Versus New Durable Markets. *Journal of Product & Brand Management*, *31*(1), 1–15. https://doi.org/10.1108/JPBM-02-2020-2769

Momox. (2020). *Studie zeigt: Klare Motivation für Kauf von Second Hand Kleidung bleibt Nachhaltigkeit*. https://momox.biz/presse/28-pressemeldungen/2020/1053-studie-zeigt-klare-motivation-fuer-kauf-von-second-hand-kleidung-bleibt-nachhaltigkeit

Momox. (2021). *Second Hand Report 2021 Bücher & Medien*. https://www.momox.de/studien/second-hand-report-2021/

Momox. (2022). *Second Hand Fashion Report 2022*. https://momox.biz/images/Pressemitteilungen/2022/Booklet_Second_Hand_Fashion_Report_momox_fashion_2022.pdf

Mukherjee, A., & Lee, S. Y. (2016). Scarcity Appeals in Advertising: The Moderating Role of Expectation of Scarcity. *Journal of Advertising*, *45*(2), 256–268. https://doi.org/10.1080/00913367.2015.1130666

O'Reilly, L., Rucker, M., Hughes, R., Gorang, M., & Hand, S. (1984). The Relationship of Psychological and Situational Variables to Usage of a Second-Order Marketing System. *Journal of the Academy of Marketing Science*, *12*(3), 53–76. https://doi.org/10.1007/BF02739319

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, *17*(4), 460. https://doi.org/10.2307/3150499

Olson, N. J., & Ahluwalia, R. (2021). When Sharing Isn't Caring: The Influence of Seeking The Best on Sharing Favorable Word of Mouth about Unsatisfactory Purchases. *Journal of Consumer Research*, *47*(6), 1025–1046. https://doi.org/10.1093/jcr/ucaa052

Padmavathy, C., Swapana, M., & Paul, J. (2019). Online Second-Hand Shopping Motivation – Conceptualization, Scale Development, and Validation. *Journal of Retailing and Consumer Services*, *51*, 19–32. https://doi.org/10.1016/j.jretconser.2019.05.014

Paharia, N., Keinan, A., Avery, J., & Schor, J. B. (2011). The Underdog Effect: The Marketing of Disadvantage and Determination through Brand Biography. *Journal of Consumer Research*, *37*(5), 775–790. https://doi.org/10.1086/656219

Parguel, B., Lunardo, R., & Benoit-Moreau, F. (2017). Sustainability of the Sharing Economy in Question: When Second-Hand Peer-to-Peer Platforms Stimulate Indulgent Consumption. *Technological Forecasting and Social Change*, *125*, 48–57. https://doi.org/10.1016/j.techfore.2017.03.029

Priester, J. R., Nayakankuppam, D., Fleming, M. A., & Godek, J. (2004). The A2SC2 Model: The Influence of Attitudes and Attitude Strength on Consideration and Choice. *Journal of Consumer Research*, *30*, 574–587.

Rahman, R., Langner, T., & Temme, D. (2021). Brand Love: Conceptual and Empirical Investigation of a Holistic Causal Model. *Journal of Brand Management*, 1–34. https://doi.org/10.1057/s41262-021-00237-7

Rossiter, J. R. (2002). The C-OAR-SE Procedure for Scale Development in Marketing. *International Journal of Research in Marketing*, *19*(4), 305–335. https://doi.org/10.1016/S0167-8116(02)00097-6

Roux, D. (2010). Identity and Self-Territory in Second Hand Clothing Transfers. In M. C. Campbell, J. Inman, & R. Pieters (Chairs), *Association for Consumer Research*.

Roux, D., & Guiot, D. (2008). Measuring Second-Hand Shopping Motives, Antecedents and Consequences. *Recherche Et Applications En Marketing (English Edition)*, *23*(4), 63–91. https://doi.org/10.1177/205157070802300404

Roux, D., & Guiot, D. (2020). Second-Hand Markets: Alternative Forms of Acquiring, Disposing of, and Recirculating Consumer Goods. In L. M. Visconti, L. Peñaloza, & N. Toulouse (Eds.), *Marketing Management: A Cultural Perspective* (2nd ed., pp. 388–407). Routledge. https://doi.org/10.4324/9780203710807-29

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation Analysis in Social Psychology: Current Practices and New Recommendations. *Social and Personality Psychology Compass*, *5*(6), 359–371. https://doi.org/10.1111/j.1751-9004.2011.00355.x

Sarstedt, M., & Mooi, E. (2014). *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics* (2nd ed.). *Springer texts in business and economics*. Springer.

Schmidt, H. J., & Steenkamp, P. (2022). Beware, an Underdog May Bite: Literature Review and Brand Management Framework in the Context of Underdog Brands. *Journal of Brand Management*, *29*(1), 85–110. https://doi.org/10.1057/s41262-021-00259-1

Schnebelen, S., & Bruhn, M. (2018). An Appraisal Framework of the Determinants and Consequences of Brand Happiness. *Psychology & Marketing*, *35*(2), 101–119. https://doi.org/10.1002/mar.21073

Sihvonen, J., & Turunen, L. L. M. (2016). As Good as New – Valuing Fashion Brands in the Online Second-Hand Markets. *Journal of Product & Brand Management*, *25*(3), 285–295. https://doi.org/10.1108/JPBM-06-2015-0894

Silva, S. C., Santos, A., Duarte, P., & Vlačić, B. (2021). The Role of Social Embarrassment, Sustainability, Familiarity and Perception of Hygiene in Second-Hand Clothing

Purchase Experience. *International Journal of Retail & Distribution Management*, *49*(6), 717–734. https://doi.org/10.1108/IJRDM-09-2020-0356

Smith, A. K., & Bolton, R. N. (1998). An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounters. *Journal of Service Research*, *1*(1), 65–81. https://doi.org/10.1177/109467059800100106

Strähle, J. (2021, August 13). *Secondhand Boom bei Zalando und About You: Nachhaltig oder smartes Business?* YouTube. https://www.youtube.com/watch?v=um-DppT51wPU

Sundie, J. M., Kenrick, D. T., Griskevicius, V., Tybur, J. M., Vohs, K. D., & Beal, D. J. (2011). Peacocks, Porsches, and Thorstein Veblen: Conspicuous Consumption as a Sexual Signaling System. *Journal of Personality and Social Psychology*, *100*(4), 664–680. https://doi.org/10.1037/a0021669

Suri, A., Huang, B., & Sénécal, S. (2019). I Can Forgive You, But I Can't Forgive the Firm: An Examination of Service Failures in the Sharing Economy. *Journal of Marketing Theory and Practice*, *27*(4), 355–370. https://doi.org/10.1080/10696679.2019.1644958

SurveyCircle. (2022). *Forschungswebseite SurveyCircle. Published 2016.* https://www.surveycircle.com

ThredUp. (2022). *Resale Report: 2022*. https://www.thredup.com/resale/static/2022-resaleReport-full-92a77020598ceca50f43227326100cc2.pdf

Tsang, J.-A., McCullough, M. E., & Fincham, F. D. (2006). The Longitudinal Association Between Forgiveness and Relationship Closeness and Commitment. *Journal of Social and Clinical Psychology*, *25*(4), 448–472. https://doi.org/10.1521/jscp.2006.25.4.448

Tsarenko, Y., & Tojib, D. (2011). A Transactional Model of Forgiveness in the Service Failure Context: A Customer-Driven Approach. *Journal of Services Marketing*, *25*(5), 381–392. https://doi.org/10.1108/08876041111149739

Tsarenko, Y., & Tojib, D. (2012). The Role of Personality Characteristics and Service Failure Severity in Consumer Forgiveness and Service Outcomes. *Journal of Marketing Management*, *28*(9-10), 1217–1239. https://doi.org/10.1080/0267257X.2011.619150

Tsarenko, Y., & Tojib, D. (2015). Consumers' Forgiveness After Brand Transgression: The Effect of the Firm's Corporate Social Responsibility and Response. *Journal of Marketing Management*, *31*(17-18), 1851–1877. https://doi.org/10.1080/0267257X.2015.1069373 van Damme, I., & Vermoesen, R. (2009). Second-Hand Consumption as a Way of Life: Public Auctions in the Surroundings of Alost in the Late Eighteenth Century. *Continuity and Change*, *24*(2), 275–305. https://doi.org/10.1017/S0268416009007188

Vandello, J. A., Goldschmied, N. P., & Richards, D. A. R. (2007). The Appeal of the Underdog. *Personality & Social Psychology Bulletin*, *33*(12), 1603–1616. https://doi.org/10.1177/0146167207307488

Wei, C., Liu, M. W., & Keh, H. T. (2020). The Road to Consumer Forgiveness is Paved with Money or Apology? The Roles of Empathy and Power in Service Recovery. *Journal of Business Research*, *118*, 321–334. https://doi.org/10.1016/j.jbusres.2020.06.061

Weidemann, T. (2020, September 29). Mehr als Nachhaltigkeit: Zalando und H&M steigen ins Second-Hand-Geschäft ein. *T3n Magazin*. https://t3n.de/news/mehr-nachhaltigkeit-zalando-hm-second-1324956/

Weinstein, J. F. (2014). *Reframe, Reuse, Re-Style: Deconstructing the Sustainable Second-Hand Consumer* (Doctoral Dissertation). Wesleyan University. https://doi.org/10.14418/wes01.1.1033

Worthington, E. L., Lavelock, C., vanOyen Witvliet, C., Rye, M. S., Tsang, J.-A., & Toussaint, L. (2015). Measures of Forgiveness. In *Measures of Personality and Social Psychological Constructs* (pp. 474–502). Elsevier. https://doi.org/10.1016/B978-0-12-386915-9.00017-6

Xie, Y., & Peng, S. (2009). How to Repair Customer Trust After Negative Publicity: The Roles of Competence, Integrity, Benevolence, and Forgiveness. *Psychology & Marketing*, *26*(7), 572–589. https://doi.org/10.1002/mar.20289

Yagil, D., & Luria, G. (2016). Customer Forgiveness of Unsatisfactory Service: Manifestations and Antecedents. *Service Business*, *10*(3), 557–579. https://doi.org/10.1007/s11628-015-0282-1

Yao, S., Chen, J., Yu, X., & Sang, J. (2017). Mediator Roles of Interpersonal Forgiveness and Self-Forgiveness Between Self-Esteem and Subjective Well-Being. *Current Psychology*, *36*(3), 585–592. https://doi.org/10.1007/s12144-016-9447-x

Zaichkowsky, J. L. (1985). Measuring the Involvement Construct. *Journal of Consumer Research*, *12*(3), 341. https://doi.org/10.1086/208520

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The Behavioral Consequences of Service Quality. *Journal of Marketing*, *60*(2), 31. https://doi.org/10.2307/1251929

Zhang, Y., Feick, L., & Mittal, V. (2014). How Males and Females Differ in Their Likelihood of Transmitting Negative Word of Mouth. *Journal of Consumer Research*, *40*(6), 1097–1108. https://doi.org/10.1086/674211

Sandra Baringhorst | Manfred Schwaiger | Louisa Weritz

II Examination of the Role of Social Media in Health-Care

Abstract

In recent years, interactions on Instagram have grown significantly, especially in health-related discussions. This study investigates whether factors like influencers' personal experiences, medical qualifications, and follower counts impact patients' willingness to visit a doctor. A first descriptive analysis found that nearly 28% of patients change their opinion whether to visit a doctor after being exposed to an Instagram post. In a second step, a mixed ANCOVA and a logistic regression revealed that the advice of an expert on Instagram, as well as the recommendation not to go to the doctor, increases the probability that patients want to postpone the appointment or do not want to visit the doctor at all. This study demonstrates that Instagram postings exert an influence on a patient's opinion formation. Findings provide directions for policymakers regarding awareness campaigns on the reliability of online sources.

1 Motivation

Interactions, whether with friends or strangers, are increasingly taking place online. This trend can be seen, among other aspects, by the fact that in the last ten years, social media usage has increased by 200% (Rabe, 2022). Compared to all social media platforms, Instagram has overtaken other providers in terms of usage times in Germany (Najorka, 2021). In the United States, Instagram is the second most popular social media platform with 61% of adults using it (Ruby, 2023).

Whereas informal interactions dominated in the beginning, followers are now looking for specific content and information. Of growing importance is the demand for and supply of health-related information. Studies show that people are already using social media to look for such information (Brandt, 2020) and that nearly one in three people has already discussed health-related issues on social media (Honigman, 2015). In a study conducted by the Pew Research Center in the United States, it was revealed that 72% of adult internet users have engaged in online searches to acquire information about various health matters. The most commonly sought-after information pertains to specific illnesses and potential remedies. Additionally, 26% of adult internet users have, within the last year, either read or viewed personal accounts of health-related experiences shared by others. Furthermore, 16% of adult internet users in the United States have utilized online platforms in the previous year to connect with individuals who have similar health issues (Fox, 2014). This exchange takes not only place regionally, over the last decade, social media have expanded the communication landscape across geographic boundaries, providing a possibility to tackle public health problems through collaboration with influencers worldwide (McCosker, 2018).

Social media can have a positive impact on diagnosis, prevention, and treatment of disorders and certain conditions (Leist, 2013). Health-related topics that individuals frequently search on social media are diverse and can include topics such as: specific diseases or medical problems, treatments or procedures, weight control, health insurance, food safety or recalls, drug safety or recalls, advertised drugs, medical test results, aging, pregnancy, childbirth, and health-care costs (Fox & Duggan, 2013).

Patients participate in social media not only to seek health-related information but also to find social and emotional support (Zhao & Zhang, 2017). The digital community thus fulfills an important function, namely the recognition that one is not alone. In addition, the content is mostly written in simple terms and available at a low threshold. Instagram posts are easier to understand compared to studies and official diagnoses, and less intimidating than the prospect of talking to friends and family or a professional about one's symptoms and concerns. However, the quality of the content varies and can also depend on who communicates such. It can be alarming when context and factual basis are lacking. What leads to an empowering self-diagnosis for some may trigger an unfounded fear of suffering from an illness in others. This variance in content can be, among other things, due to the fact that there are almost no access restrictions. Further, this multitude of followers and influencers, each with different interests and intentions, makes it difficult for educational work to address all of them with a common communication strategy. Strategies must be individually targeted, but this requires knowledge about the actors' characteristics on the social networks. Engagement and credibility of posts might alter a patient's intention to consult a doctor. To date, there is no uniform definition of the term credibility, but particularly with respect to Web-based search, it is generally accepted that credibility results from a combination of characteristics of the sender, the message, and the recipient (Wathen & Burkell, 2002).

For the underlying study, the focus lies on the senders' characteristics (i.e., the influencer). Information sources are considered as being credible if they are perceived as being trustworthy and proficient (Rieh & Danielson, 2007). Based on a literature review, three distinct factors could be identified that exhibit a high importance: affectedness (i.e., whether influencers are affected themselves, so-called patient influencers), expertise (i.e., degree of expertise on health-related topics), and the popularity of the Instagram account (i.e., number of followers).

We call people that have been or are still suffering from a disease themselves *patient* or *health* influencers. These influencers can provide experiences about certain health patterns and symptoms. As patient influencers suffer from a disease themselves, they may often provide valuable and authentic insights and forms of social support (McCosker, 2018; Tian et al., 2017) to patients and other interested audiences. As such, following their accounts may have beneficial effects. It is already known that followers look for like-minded people on social media channels. However, it is questionable whether authentic insights and social support weigh more than an expert opinion of a medical professional. Therefore, we derive the first research question:

> ***RQ1:*** *Do postings of patient influencers on Instagram alter a patient's willingness to demand a doctor's appointment?*

Considering that any person who uses social media can become an influencer and content creator, background information on the sender's expertise is of great importance with regard to the credibility of information (L. Campbell et al., 2016). Such expertise-bloggers use social media to share health-related information, network with colleagues, disseminate research, market their practice, and engage in health advocacy (Chretien & Kind, 2013). No research to date has looked at the relative benefits of professional versus unprofessional postings on social networking sites (Lefebvre & Bornkessel, 2013). It is unclear whether followers on Instagram prefer to look for expert opinions on social media to avoid negative side effects of a doctor's visit (e.g., long waiting times, making appointments) or whether the focus of the search lies on discourse and opinions with like-minded people. Moreover, investigating whether patients exhibit a higher propensity to adhere to the guidance provided by an expertise-blogger in contrast to a patient influencer, or whether they primarily consider the popularity of a profile, could contribute insights into the role of social media in self-diagnosis. Accordingly, the second research question is as follows:

> ***RQ2:*** *Does the medical expertise of an Instagram influencer alter a patient's willingness to demand a doctor's appointment?*

At first glance, Instagram profiles differ in terms of their number of followers, which serves as a cue to assess their popularity. The more popular an influencer is, the more trustworthy, extravert, and approachable the person is perceived (S.-A. A. Jin & Phua, 2014; Veirman et al., 2017). However, literature also states that micro- and nano-influencers are perceived as more authentic compared to macro- and mega-influencers (C. Campbell & Farrell, 2020). It is questionable whether this popularity is sufficient to change patients' minds and persuade them to (not) see a doctor, especially as trustworthiness is an important factor in health-related topics. In addition, it is interesting to see whether a posting by an influencer with a high number of followers, but low medical expertise preempts a posting by an influencer with a lower number of followers but high medical expertise. Therefore, the third research question is as follows:

> ***RQ3:*** *Does the popularity of an Instagram influencer alter a patient's willingness to demand a doctor's appointment?*

The three identified factors seem to be decisive regarding a post's evaluation. However, these factors have not yet been tested, neither individually nor in their combination, on the degree of influence they have to alter a patient's willingness to demand a doctor's appointment. Therefore, we would like to examine which of those factors (or a particular combination) has a major impact regarding the follower's decision to (not) visit a doctor. Of particular interest is whether social media content could alter the followers' first tendency to (not) visit a doctor due to their symptoms. Therefore, we formulate:

> ***RQ4:*** *Which combination of the factors of the status of (not) being a patient influencer, having expertise, and the popularity of an influencer on Instagram can alter a patient's willingness to demand a doctor's appointment?*

The effect health influencers have on their followers, and thus on potential patients, is still largely unanswered (Sugawara et al., 2012). Patients self-assess that social media make them seek a second opinion, see a particular doctor or ask about a particular medication, and reconsider the treatment of an illness (Lefebvre & Bornkessel, 2013). Nevertheless, no evidence has yet been found that peer to peer online support has an impact on patient health or opinion change (Eysenbach et al., 2004).

Further, this paper answers calls for further research on the effectiveness of social media applications (i.e., Instagram) for health communication (Moorhead et al., 2013) and its impact on health-related behavior change (Moorhead et al., 2013). The research questions test the identified variables with the help of a scenario experiment. The purpose primarily lies in finding whether certain characteristics of influencers and their Instagram posts are decisive in altering the opinion of followers about a visit to the doctor. This is especially helpful for public policymakers, as they can target influencers to address either the public (e.g., for educational purposes), but also specific target groups (e.g., preventive care for depression). Further, doctors can benefit as well with regards to addressing their patients in an unconventional way.

The remainder of this paper is structured as follows. The literature on social media and Instagram usage is introduced and a bow is drawn to health-care. Further, the study design, which comprises a pretest and one main study is presented before analyzing the data. The paper rounds off with a critical discussion, implications, and suggestions for policymakers, before outlining avenues for further research.

2 Literature Review

In recent years, a new phenomenon has emerged in the health-care industry. Physicians complain about patients seeking advice from a medical professional too late and looking for support elsewhere (Frakt, 2016). Timely medical treatments promise a milder course and earlier recovery for many medical conditions. However, with the rise of social media, the ways in which people receive and engage with health information have changed. Whereas physicians were once people's primary and often only source for health advice, today, social media provide new spaces for accessing health-related information from various sources and for sharing and interacting with the information in different ways (Chou et al., 2009). It is safe to assume that patients nowadays spend more time online than at the doctor's office (Hackworth & Kunz, 2011). However, this transformation has brought new challenges, including the spread of non-expert opinions as well as misinformation (Bode & Vraga, 2018; Heiss, 2020).

2.1 Status Quo of Research on Health-Related Topics in Social Media

Relevance of Social Media for Health-Care

Social media are *"Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation and exchange of User Generated Content"* (Kaplan & Haenlein, 2010, p. 61). Such applications include (micro)blogs, social networking sites, virtual worlds, collaborative projects, content community sites, and feedback and review sites (Mangold & Faulds, 2009). The main focus of the underlying study lies on social networks (i.e., Instagram), as these focus on establishing and maintaining social contacts (Meffert et al., 2019). People use social media platforms to *''present themselves in an online profile, accumulate friends who can post comments on each other's pages, and view each other's profiles''* (Ellison et al., 2007, p. 1143). With social media, users have the opportunity to develop parasocial interactions with other followers and influencers as they are regularly exposed to another person's life (Boerman, 2020).

These network sites provide the opportunity to share positive and negative experiences (Hackworth & Kunz, 2011). Further, for most people, social networking sites are already part of their daily routine. Incorporating health-related issues in this routine is very easy and convenient (Benetoli et al., 2017). With the development of Web 2.0, social media are an ever-evolving topic that is increasingly influencing people's health behavior (Zhao & Zhang, 2017). Social media are used by the general public, health professionals, patients, and health service providers to share health-related topics (Gupta et al., 2022). Users are able to access such information from other users or may ask for help and make contributions to others while sharing their own experiences. Social media have changed the way people consume and share information about their well-being and physical condition (Gupta et al., 2022) with peers as they have the advantage of being interactive. Patients are not only able to consume information, they can also ask questions, or share their own stories (Benetoli et al., 2017; Gupta et al., 2022). Further, social media provide access to videos on health-related topics (Gupta et al., 2022). In addition, social media make use of pictures or a combination of text, images, and videos (Zhang et al., 2013). Therefore, social media are a good tool for people with lower literacy (Moorhead et al., 2013).

Patients do not necessarily turn to social media to bypass a doctor's visit; rather, they use this option due to the health-care system's inability to accommodate all individual needs and requests (Smailhodzic et al., 2016). Contrary to the assumption that only affected patients or their relatives enquire about health topics online, people anxious about public health concerns (issues that affect the well-being of populations, e.g., outbreaks of infectious diseases, vaccinations and immunizations) also use online media (Lewis et al., 2005). Accordingly, the number of interested parties is much broader and not limited to patients (Zhao & Zhang, 2017). With social media, people are able to find relevant information more effectively than they have been with the use of traditional media (Zhao & Zhang, 2017). It is possible to obtain information on certain clinical pictures more quickly, but also to find out about certain medications or doctors and to obtain testimonials from former users or patients (Zhao & Zhang, 2017).

Besides all the buzz around social media, they are said to be one of the least reliable sources for health-related information online. A study from the Pew Research Center found that for a considerable time, Americans have displayed a higher inclination toward placing trust in information originating from local and national news outlets compared to information found on social media platforms. This pattern endures in the present, except when it comes to the youngest adults. Individuals under the age of 30 now exhibit nearly equivalent levels of trust in information sourced from social media sites and information coming from national news sources (Liedke & Gottfried, 2022).

Instagram and its Functionalities

The underlying study uses Instagram as the investigated social media platform. Instagram is a platform which allows users to share photos and videos with their followers (Instagram, 2023) and exerts a strong focus on social interactions and engagement. The platform was created in

2010 by Kevin Systrom and Mike Krieger. In 2012, Instagram was acquired by Facebook and is therefore now operated by Meta. Meta possesses four of the largest social media platforms, each boasting more than one billion monthly active users: Facebook (its core platform), WhatsApp, Facebook Messenger, and Instagram (Dixon, 2023a). In May 2021, 57% of Gen Z social media users were active on Instagram. Instagram predominantly appeals to individuals in the 25 to 34-year-old age bracket, and by December 2021, approximately one-third of all users in the United States fell within this age group (Dixon, 2023b).

Instagram has become a powerful social media platform that combines visual storytelling with social interaction, enabling users to share moments from their lives and discover content from others. The core functionality of Instagram revolves around posting visual content, such as photos and videos, either in the form of permanent posts or temporary stories that disappear after 24 hours (Picazo-Sánchez et al., 2022). Users can improve their posts by using filters, editing tools, and captions. They can also tag other users and add location information to their postings and can follow other accounts to see their posts in their feed. Further, the platform features an explore section, where users can discover new influencers and content based on their own interests (Instagram, 2023). Followers can like, comment, repost, and engage with the influencer and other followers for example via direct messaging (Comp et al., 2021; Instagram, 2023).

Benefits and Motivations of Social Media Usage in Health-Care

Instagram has different benefits and consumers therefore use it for different motivations. Within a literature review, we identified three main motivations to use Instagram in health-care: Information search, peer-to-peer (P2P) support, and social and emotional support. In this chapter we will focus on these distinct motivational factors.

Information Search

Social media can be utilized to contact the own physician (Sumayyia et al., 2019), however, in most cases, social media are used to look for information regarding treatments, symptoms, medicines, or about information about the own physician (Gupta et al., 2022; Lefebvre & Bornkessel, 2013; van de Belt et al., 2013; Zhao & Zhang, 2017). Patients encourage each other to scrutinize doctor's advice (Stewart Loane & D'Alessandro, 2014). However, whether patients accept and implement the information they receive via social media depends above all on the quality of the information (J. Jin et al., 2016).

The interaction on social media is mostly done through certain groups or followership of one or many persons (Benetoli et al., 2017). Such a group membership or followership is especially useful as news are directly accessible via the newsfeed (Benetoli et al., 2017) and is therefore up to date. Participants can either search for or provide information themselves (Zhang et al., 2013). However, social media users do not solely search for information on social media for themselves, but also do so for friends and acquaintances (Sumayyia et al., 2019).

Patients that are newly diagnosed need a lot of information about their condition and upcoming treatments (Smailhodzic et al., 2016). Therefore, they are interested in how other patients are treated with the same condition, and their experiences (Benetoli et al., 2017). Depending on the clinical picture, very personal data may also be disclosed, such as blood values (Zhang et al., 2013). However, people often do not feel comfortable sharing such personal information online (Sumayyia et al., 2019).

In addition to sharing information on doctors, treatment methods, or the personal experience, information is also shared about alternative treatment options and non-prescription medications (Benetoli et al., 2017; Stewart Loane & D'Alessandro, 2014). In some cases, based on online discussions, social media could also lead to patients switching their doctors (Smailhodzic et al., 2016). Further, information on how adaptions in the patients' lifestyle might improve the patients' health conditions are also exchanged (Zhang et al., 2013). Additionally, users look for possible side effects of medications, and how other sufferers tolerate or have dealt with them (Benetoli et al., 2017; Greene et al., 2011; van de Belt et al., 2013). Also when being unsure how to exactly take medication, social networking sites might help (Scanfeld et al., 2010). In some cases, the information gathered through social media is shared and discussed later with a patient's health-care professional (Bhaskaran et al., 2017; Stewart Loane & D'Alessandro, 2014).

Peer-to-Peer Support

We define P2P support as a rather active communication among several individuals with the help of social networking sites. Social networking sites can be a good tool to build a sense of a community of like-minded people with similar conditions (Hyun Jung Oh et al., 2013) where patients can get in touch with other patients (Antheunis et al., 2013). What is valued most about online platforms is the exchange in the community (Stewart Loane & D'Alessandro, 2014). Patients are able to talk about their own experiences online and share it with a larger community (Benetoli et al., 2017). Some users even openly provoke emotional support from peers, often in combination with information gathering (Zhang et al., 2013). Encouraging messages and providing support to peers are often offered, especially if someone describes him or herself in a difficult situation. These messages do not have to be addressed to a specific person, rather, they are addressed to a broader audience, i.e., the community of like-minded people. Further, patients are able to express their story and emotions openly without being concerned about the emotions and feelings of their loved ones (Smailhodzic et al., 2016). From negative emotions such as anger, down or sarcastic feelings to positive messages such as upbeat and cheerful emotions (Zhang et al., 2013). Additionally, people are no longer restricted to a regional support group, using social networking sites, they can overcome language barriers and talk to people from all over the world with a similar condition (Zhang et al., 2013).

Social and Emotional Support

A significant majority of social media users engaged in health-related topics tend to play a passive role. Due to privacy reasons most of those social media users keep their identity to themselves (Benetoli et al., 2017) and do not share personal information (Antheunis et al., 2013). Still, patients want the benefits of social and emotional support. We recognize social and emotional support as a rather passive motive to engage in social networking activities such as following a blog or influencer without actively engaging in the communication.

Social media are often used due to their capacity to provide social and emotional reinforcement, engendering a sense of companionship and shared experience among individuals confronting such issues (Hyun Jung Oh et al., 2013; Zhao & Zhang, 2017). This inclination is underscored by the perception that engagement with these platforms imparts a sensation of solidarity (Gupta et al., 2022; Smailhodzic et al., 2016). Emotional support can be defined as *"communication that meets an individual's emotional or affective needs"* (Smailhodzic et al., 2016). Participants seek a channel to express their emotions, look for emotional support, or provide support to their peers (Zhang et al., 2013).

Patients possess the capability to engage in comparative assessments with their peers, thereby discerning the severity of their condition and the efficaciousness of diverse therapeutic interventions (Smailhodzic et al., 2016). Such support cannot be provided by either the health-care professionals or family and friends, as neither of them are affected themselves (Benetoli et al., 2017).

Looking for health-related social support online also improves one's own health self-efficacy (Hyun Jung Oh et al., 2013), which refers to *"individual's beliefs about their ability to manage* *their health"* (Lee et al., 2008, p. 362). Patients who believe that they are able to achieve certain health goals, are more likely to seek health-related information compared to those that doubt their ability to improve one's health (Lee et al., 2008). Especially emotional support was found to be an important predictor for health self-efficacy (Hyun Jung Oh et al., 2013).

2.2 Health-Care Influencers

Besides the motives, why patients engage in health-related communication on social media, within literature two types of health information, namely expertise-based information which is produced by medical professionals or experience-based information, which is based on a lay-persons experience with health-related topics (Song et al., 2016) have been identified. We will look at both of them in the next chapters. Another factor that is repeatedly discussed in the literature but also in practice is the number of followers. The follower count is always paid attention to regardless of the industry, therefore, we will also investigate the followership in more detail.

2.2.1 Patient Influencers

Health communicators or influencers are social media influencers that are actively engaging in content creation of a specific health-related field or niche such as fitness (Durau et al., 2022), nutrition (Jenkins, Ilicic, Barklamb, & McCaffrey, 2020; Rogers et al., 2022), or health conditions (Heiss & Rudolph, 2022).

Health influencers could prompt a shift in a follower's attitude, enhance knowledge about specific health conditions, and consequently drive behavioral changes (Heiss & Rudolph, 2022). These so-called *patient influencers* are a group of health influencers that have been or are still suffering from mental or physical conditions (Heiss & Rudolph, 2022). They are defined as *"health influencers who have been patients of long-term physical or mental conditions that are difficult to cope with, including various noncommunicable (e.g. cancer and diabetes) and communicable (e.g. HIV) diseases"* (Heiss & Rudolph, 2022, p. 2). These influencers often have personal experiences with chronic illnesses, rare diseases, disabilities, mental health conditions, or other health-related issues and are therefore often perceived as knowledgeable and experts within the field of their disease (Willis & Delbaere, 2022; Willis et al., 2023).

Patient influencers serve as a bridge between the health-care community and patients, providing valuable firsthand insights into the realities of living with a specific condition (McCosker, 2018; Tian et al., 2017). Their authentic and relatable content helps to educate, inspire, and empower others who may be going through similar health challenges. Through their social media presence, they share their journeys, including symptoms, treatments, coping strategies, and daily life experiences. Further, patient influencers share their knowledge about medications and potential side effects, and also talk about their personal success stories (Willis et al., 2023). Due to similar medical conditions and experiences, patients may identify with patient influencers, therefore patients might actively ask their physician for a certain treatment (Willis et al., 2023).

Their communication is articulated on a personalized basis, which is undermined by pictures and videos of themselves. Often, health influencers post daily and disclose personal details about themselves (Willis et al., 2023). Thereby, they exude a sense of intimacy making them more tangible and relatable, and exerting stronger persuasive power. This leads to patient influencers being perceived as authentic "superpeers" (Janssen et al., 2022) leading to a stronger emotional relationship between influencer and follower compared to the typical follower-influencer bond (Willis et al., 2023).

Patient influencers are often intrinsically motivated to assist fellow patients, as they might have experienced past struggles with their condition and a lack of social and emotional support. Their primary intention is to utilize their personal experiences to offer emotional and social support and raise awareness about their condition. This serves as a coping mechanism, addressing problems and emotions, which can alleviate stress and enhance the overall well-being of their followers (Heiss & Rudolph, 2022; Willis et al., 2023).

Patient influencers also provide valuable and authentic insights and forms of social support (McCosker, 2018; Tian et al., 2017) to patients and other interested audiences. This emotional and social support is one of the main reasons why affected followers turn to social media (Heiss & Rudolph, 2022). Within a supportive community, patient influencers often connect individuals who share common health concerns. They foster a sense of belonging and create safe spaces for discussion, where people can openly share their experiences and find encouragement from others who understand their struggles.

Patient influencers do not solely help their fellow patients. The public and indirectly affected individuals, such as family members and close friends also engage with patient influencers to gather further information (Heiss & Rudolph, 2022).

Even if there are existing doubts about the content's credibility, there is also the possibility that social media may act as a deterrent from seeking out health professionals (Kim, 2009). The rise of patient influencers has been fueled by the increasing use of social media platforms and the desire for authentic, relatable stories, and experiences. As such, following their accounts may have beneficial effects. It is already known that followers look for like-minded people on social media channels. However, information provided by patient influencers might be generalized for a broader audience or oversimplified. It was found that information provided by patient influencers without medical background is misleading in 50% of the postings (Yeung et al., 2022). Still, most patients feel confident to distinguish trustworthy from doubtful statements (Stewart Loane & D'Alessandro, 2014). However, patient influencers are often far away from being proficient in the topics they are talking about. They might be influenced by third parties with commercial interests, posing challenges to the health-care system (Byrne et al., 2017; Heiss & Rudolph, 2022).

2.2.2 Expertise

The main recurring limitations of social media use for health communication are quality concerns (Adams, 2010; Moen et al., 2009; Orizio et al., 2010) and the lack of reliability of the health information (Adams, 2010; Kukreja et al., 2011; Moen et al., 2009). It also seems to be more difficult for individuals to discern the reliability of information found online (Adams, 2010). If patients do not trust the online health service, they are less likely to look for information there (Li & Wang, 2018).

Considering that any person who uses social media can become an influencer and content creator, background information on the sender's expertise is of great importance with regard to the credibility of information (L. Campbell et al., 2016). In the underlying study, influencers with expertise are defined as those who are trained professionals in the health-care field. This can include, for example, physicians (e.g., Rothfischer, 2021), nurses (e.g., Kerr et al., 2020), nutritionists (e.g., Jenkins, Ilicic, Molenaar, et al., 2020), or natural health professionals. Those groups can also be termed as "doctorfluencers" (Vassallo et al., 2022). Usually, doctorfluencers are active in the health-related domain where they have been trained at (Rothfischer, 2021). However, patients are often not aware of the different stages of professions (e.g., pre-licensed and licensed therapist) and therefore tend to rely on the existence of a professional title in general (Triplett et al., 2022; White & Hanley, 2023).

Doctorfluencers disclose their professional background either in their profile description (Jenkins, Ilicic, Molenaar, et al., 2020) or via a link, which leads to a website with additional information (Triplett et al., 2022). Another common option is to request a verified 'public interest account' on social media platforms like Instagram, identified by a blue checkmark (DiSilvestro, 2021; Sabbagh et al., 2020). When doctorfluencers disclose their affiliations, credentials, or qualifications, it triggers the expertise heuristic among their followers, leading to increased source credibility (Borah & Xiao, 2018; Meinert & Krämer, 2022). Consequently, doctorfluencers are automatically perceived as more credible, authentic, and trustworthy by their followers due to their professional background.

Nevertheless, it is essential to note that doctorfluencers may still provide inaccurate information. Yeung et al. (2022) conducted a study in which they analyzed ADHD-related (Attention Deficit Hyperactivity Disorder) posts to assess the accuracy of the information presented. They categorized the videos as either misleading, based on personal experience, or useful. Videos shared by health-care providers had significantly higher quality and usefulness, indicating accurate information. However, it was found that nearly 30% of the posts by health-care professionals were classified as misleading. This is particularly concerning due to the expertise heuristic, which may lead people to rely mainly on the professional credentials of a post's author to determine its truthfulness (Meinert & Krämer, 2022). Above all, there are instances where doctorfluencers venture outside their primary area of expertise, lacking formal education and qualifications (White & Hanley, 2023).

Still, these doctorfluencers or expertise-bloggers use social media to share health-related information, network with colleagues, disseminate research, market their practice, and engage in health advocacy (Chretien & Kind, 2013). A key motivation for doctorfluencers to participate in social media platforms is raising awareness about specific diseases. They achieve this by presenting evidence-based information in a clear and accessible manner to a wide audience, thereby combating the spread of misinformation (Vassallo et al., 2022).

Doctorfluencers have experienced an interesting added benefit from their online presence: they are perceived as skilled and competent by their younger audience (White & Hanley, 2023). Similarly, a study conducted by Kolmes and Taube (2016) found that 68.2% of participants had a more positive view of therapists who were active on social media platforms, as it signaled a greater understanding of their clients.

2.2.3 Popularity

The impact of health influencers on social media is linked to both the size of their follower base and their ability to influence their opinions, even though there is no specific follower count threshold to identify them (Veirman et al., 2017). The number of followers serves as a cue to assess an influencer's popularity. The more popular an influencer is, the more trustworthy, extravert, and approachable the person is perceived (S.-A. A. Jin & Phua, 2014; Veirman et al., 2017). Therefore, the count of followers is frequently utilized for categorization in both academic research and marketing applications (Haenlein et al., 2020). Further, the level of online popularity significantly influences the perceived source credibility of social media users (S.-A. A. Jin & Phua, 2014).

Social media influencers are being categorized based on their followership, with some researchers proposing the division into micro-influencers (small followership) and macro-influencers (large followership) (Kay et al., 2020). Influencers with a rather small audience (up to 10,000 followers) are called nano-influencers, whereas influencers with a follower base up to 100,000 are termed micro-influencers. Those with a larger audience, such as macro influencers, present a followership with up to 1,000,000 followers or more in the case of mega influencers (C. Campbell & Farrell, 2020).

Having a larger follower count significantly boosts the influencer's likeability. This increase can be attributed mainly to the perception of higher popularity among the audience. Additionally, a smaller portion of this effect is due to people attributing more opinion leadership to the influencer based on their perceived popularity. Consequently, a substantial number of followers can lead to heightened perceptions of popularity and, consequently, greater likeability. However, it should be noted that this does not automatically imply that the influencer will be perceived as an opinion leader (Veirman et al., 2017).

However, an ever-increasing number of followers may not always lead to an ideal influencer-follower relationship (Qutteina et al., 2019). In fact, it could result in the influencer being perceived as a celebrity, which may diminish their credibility and authenticity in the eyes of their followers.

Often, micro- and nano-influencers are considered more intimate and authentic compared to macro- and mega-influencers , which enhances their persuasive impact on followers (C. Campbell & Farrell, 2020). Micro-influencers also tend to have a dedicated and loyal follower base, often targeted based on geographical location (C. Campbell & Farrell, 2020). Further, nano-influencers are often considered to be in the early stages of their potential career, with their relatively small audience comprising primarily friends and family. As a result, they tend to exhibit the highest engagement rates compared to other influencer categories (C. Campbell & Farrell, 2020).

2.2.4 Patient Influencers, Expertise, and Popularity

It is already known that followers look for like-minded people on social media channels. However, it is questionable whether authentic insights and social support weigh more than an expert opinion of a medical professional. Patient influencers are likely to foster stronger emotional connections with their followers due to the shared experience of enduring the same chronic disease.

Further, it is unclear whether followers on Instagram prefer to look for expert opinions on social media to avoid negative side effects of a doctor's visit (e.g., long waiting times, making appointments) or whether the focus of the search lies on discourse and opinions with like-minded people. The examination of whether patients are more likely to follow the advice of an expertise-blogger compared to a patient influencer or might solely look at the popularity of a profile could shed light on social media's use for self-diagnosis.

Follower count reflects an influencer's popularity, impacting their perceived trustworthiness, extraversion, and approachability (S.-A. A. Jin & Phua, 2014; Veirman et al., 2017). Mega- and macro-influencers are seen as experts and opinion leaders, gaining more trust from their followers (C. Campbell & Farrell, 2020). However, research on health influencers also suggests that as they gain more followers, the emotional and intimate influencer-follower relationship may be compromised. This is not applicable to nano- or micro-influencers operating in specific health-related domains, where their perceived expertise stems from. Having a smaller audience has its advantages, as influencers can respond more promptly and effectively to their followers' requests compared to those with a larger followership. Further, health influencers often cover only a niche due to their expert knowledge (Rothfischer, 2021).

It is questionable whether popularity is sufficient to change patients' minds and persuade them (not) to see a doctor, especially as trustworthiness is an important factor in health-related topics. In addition, it is interesting to see whether a posting by an influencer with a high number of followers, but low medical expertise preempts a posting by an influencer with a lower number of followers but high medical expertise.

The three identified factors seem to be decisive regarding a post's evaluation. Herefore, we would like to examine which of those factors (or a particular combination) has a major influence regarding the follower's decision to (not) visit a doctor. Of particular interest is whether social media content could alter the followers' first tendency to (not) visit a doctor due to their symptoms.

3 Study Design

The goal of this study is to test whether a patient's willingness to visit a doctor, changes after patients are being exposed to an Instagram posting. This study uses a 2 x 2 x 2 x 2 within-between subjects design in an online experiment with different scenarios (i.e., postings and profiles). The main differences in the postings where the type of influencer (i.e., was the influencer affected him- or herself with the disease s/he is talking about), expertise (i.e., influencer earned a medical degree but is currently not practicing versus someone with average knowledge about diseases and health-care (i.e., someone like you and me)), and popularity (i.e., macro influencers with 251,712 followers versus nano influencers with 238 followers). Further, the influencer either suggested to visit a doctor or mentioned that it is not necessary to visit a doctor. To make sure that the manipulations work as intended, a pretest was run. After data was gathered within the main study, in the first step, the data was analyzed descriptively. It was examined if and under which conditions patients change their predisposition to visit a doctor. In the second step, a mixed between-within ANCOVA and a logistic regression was run to investigate which characteristics of an Instagram posting is most likely to change a patient's willingness to visit a doctor.

3.1 Pretest

A pretest was conducted to test how the manipulations where perceived by the participants. Participants were recruited from the database Prolific. A total of 205 valid responses were gathered. First, a screener question was used to assess whether participants are using Instagram daily. Further, participants were presented to one of the 16 influencer profiles and postings from the 2 x 2 x 2 x 2 (patient influencer: yes, no; expertise: yes, no; popularity: nano, macro; recommendation: yes, no) between-subjects design. Recommendation was included as a further variable in order to be able to examine whether participants follow the influencer's recommendation. To check whether participants perceive influencers with a larger number of followers as more popular, several questions on the size of the community ("Please indicate how many followers the influencer profile from the initial scenario had." (1) large community; (2) small community) and the influencer's popularity ("The person from whom I read the Instagram posting is" (1) popular, (2) quite accepted, etc.) adapted from Scott and Judge (2009) were asked. To assess how the influencer's expertise is perceived, subjects were asked whether the influencer has a medical degree (yes or no) and how they would rate the influencer's medical expertise relative to others (1 = one of the least knowledgeable, 7 = one of the most knowledgeable) adapted from Yoo (2014). Further, the influencer's degree of affectedness was measured by querying whether the influencer suffered from the disease her/himself (yes, no) and the influencer's call to action was surveyed by asking whether the influencer recommends visiting a doctor (yes, no).

Patient Influencer: Subjects that were exposed to postings and influencers from people that do also suffer from the disease they are talking about, also perceived those influencers to be affected themselves ($M = 1.03$, $SD = .171$) (1 = yes suffering from hemorrhoids themselves, 2 = not suffering from hemorrhoids themselves) compared to those that are not affected themselves (($M = 1.86$, $SD = .353$), $t(149.59) = 21.43$, $p < .001$).

Expertise: Those influencers mentioning that they possess a medical degree where also perceived by the subjects to own a medical degree ($M = 1.12$, $SD = .325$) (1 = own medical degree, 2 = no medical degree) compared to those that mentioned that they are people like you and me (($M = 1.94$, $SD = .234$), $t(181.47) = 20.75$, $p < .001$). Similar, when looking at the degree of expertise of the influencer, influencers and postings that mentioned that they do possess a medical degree but are currently not practicing are perceived to be more knowledgeable regarding their medical expertise compared to others ($M = 4.74$, $SD = 1.324$) than those that have an average knowledge about diseases and health-care (i.e., someone like you and me) (($M = 3.48$, $SD = 1.307$), $t(203) = -6.866$, $p < .001$). Expertise was measured on a one-item scale from 1 = one of the least knowledgeable to 7 = one of the most knowledgeable.

Number of Followers and Popularity: According to the results, profiles and postings with 251,712 followers were perceived to have a larger audience ($M = 1.24$; $SD = 0.426$) (1 = large community, 2 = small community) than profiles with only 238 followers (($M = 1.9$, $SD = 0.298$), $t(180.39) = 12.991$, $p < .001$). Further, the profiles with a larger number of followers were also perceived to be more popular ($M = 4.9$, $SD = 0.93$) compared to those with a smaller audience (($M = 3.9$, $SD = 1.04$), $t(203) = -7.019$, $p < .001$). Popularity was measured on an eight-item seven-point Likert type scale from 1 = strongly disagree to 7 = strongly agree.

Recommendation Visiting a Doctor: The posting where influencers recommended visiting a doctor where also perceived as such ($M = 1.24$, $SD = .426$) (1 = recommendation yes, 2 = recommendation no) compared to those where influencers mention that their followers do not need to see a doctor (($M = 1.84$, $SD = .364$), $t(197.54) = 11.00$, $p < .001$).

3.2 Main Study

Subjects and Design

Participants for the scenario-based experiment were recruited via Prolific. 909 participants took part in the study. The experiment consisted of a 2 (patient influencer: yes, no) x 2 (expertise: yes, no) x 2 (popularity: nano, macro) x 2 (recommendation: yes, no) within-between subjects design. Subjects were randomly assigned to one of the 16 groups. Subjects first validated a pre-defined screener question via Prolific: "Which of the following social media sites do you use on a regular basis (at least once a month)?". As we were only interested in those that are using Instagram on a regular basis, they were able to choose between "Instagram" and "Others but no Instagram". Ten users do not use Instagram on a regular basis and were therefore dismissed from the survey. Further, 14 subjects failed the attention check, where they had to select "strongly disagree" on a seven-point scale (1 = strongly disagree, 7 = strongly agree) adapted from Gruzd et al. (2020). Therefore, these participants were removed from the analysis, leaving us with a sample of 885. The average participant is 36 years old. Further, the plurality of respondents is female (49.5%), has a Bachelor's degree (43.2%), is working full-time (56.9%), and has an annual gross household income of between 10,000 and 50,000 USD (45.1%). All details on the demographics can be found in Table 1.

Table 1. Sample Descriptive Statistics of Main Study

		Mean/Frequency	Percentage	Min	Max
Age		35.73		18	74
Gender					
	Male	434	49.0%		
	Female	438	49.5%		
	Non-binary/third gender	13	1.5%		
	Prefer not to say	0	0%		
Education					
	Some school but no degree	8	0.9%		
	High school graduate	133	15.0%		
	Some college but no degree	180	20.3%		
	Bachelor's degree	382	43.2%		
	Master's degree	127	14.4%		
	Professional degree	37	4.2%		
	Doctorate degree	13	1.5%		
	Other	5	0.5%		
Employment					
	full-time	504	56.9%		
	part-time	120	13.6%		
	self-employed	85	9.6%		
	student	62	7.0%		
	stay-at-home parent	32	3.6%		
	unemployed	61	6.9%		
	retired	21	2.4%		
Income					
	<$10,000	49	5.5%		
	$10,000 - $50,000	399	45.1%		
	$50,001 - $ 90,000	274	31.0%		
	$90,001 - $150,000	102	11.5%		
	>$150,001	22	2.5%		
	prefer not to say	39	4.4%		
		N = 885			

Stimuli and Procedure

Subjects were introduced to an initial situation. They were asked to imagine that they woke up one day and experienced health-related issues, such as itching, burning, and oozing at the anus or pain during defecation. Further, they were told that their pain level reaches a six (1 = very low, 10 = very high) and that they came up with the possible diagnosis of hemorrhoids after an online search. The disease pattern of hemorrhoids was chosen due to pre-selected criteria (i.e., assignable symptoms, illness has to be applicable to males and females, mitigating home remedies, it has to be clearly diagnosable, treatment costs should tend to increase the longer the patient does not have it treated medically), which were elaborated with different urologists. Participants were further informed that they can have the symptoms treated medically, which leads to less pain and faster recovery, however, that they do not have to have it treated necessarily, as symptoms will go away on their own. The full initial situation can be found in Appendix 1.

After being exposed to the initial situation, participants were asked, for the first time, whether they would like to visit a doctor. Then, participants were randomly assigned to one of the 16 scenarios, which were slightly adapted based on the results from the pretest. We used *Alex* as influencer, as this name represents a gender-neutral name. The profile gave an insight into Alex's number of followers, if s/he is suffering from hemorrhoids her/himself, and whether Alex completed a medical degree or is someone like you and me. Within the posting, Alex referred to the already mentioned characteristics of her/his profile and further described a disease pattern which matched the one of hemorrhoids. S/he further gives some advice how to deal with hemorrhoids at home. In the end, Alex either suggested to visit a doctor or mentioned that her/his followers do not need to see a doctor. A detailed presentation of the scenarios can be found in Appendix 2. After subjects were presented with the different scenarios, they were asked several questions regarding the situation and the scenario, among others including a second question on their willingness to visit a doctor. At the end of the survey, participants answered several demographic questions.

3.3 Measures

A list of all constructs and items can be found in Appendix 3. All constructs were measured on a seven-point scale (1 = strongly disagree; 7 = strongly agree), unless otherwise mentioned. To assess internal consistency, Cronbach's alpha is used, which is displayed right next to the constructs.

Likelihood to Visit a Doctor: Since the study is interested in the change of willingness to visit a doctor, the question for a patient's likelihood to visit a doctor was measured twice. First, it was gathered after subjects were introduced to the initial situation, however, before the manipulation took place. Second, participants were asked the exact same self-developed questions after being exposed to the manipulated Instagram post. Subjects were asked to indicate how likely they are to consult a doctor on a seven-point scale (1 = extremely unlikely; 7 = extremely likely) and whether they would consult a doctor (1 = I would not consult a doctor; 2 = I would try to make an appointment in the next few weeks; 3 = I would try to make an appointment in the next few days; 4 = I would try to make an appointment immediately). The second scale was recoded before the analysis.

Instagram Affinity (α = .76): Since affinity with Instagram could also influence the results, we included Instagram usage as a covariate. Participants were asked to state their agreement with statements such as “I use Instagram more often than other people do”; “I am interested in Instagram”, “I am experienced in using Instagram”, and “In general, Instagram is important for me”. The scale was adapted from Schumann et al. (2014).

Hypochondriasis (α = .93): The willingness to see a doctor also depends strongly on the personal tendency to hypochondriasis. For this purpose, we adapted the Whiteley Index (Hiller et al., 2002). Participants were asked to indicate their agreement with the statements “I often worry about the possibility that I have got a serious illness.” or “I am afraid of illness.”.

Risk Aversity (α = .63): Furthermore, risk aversion plays an important role when it comes to one's own health. Participants indicated their agreement on a three-item scale adapted from Donthu and Gilliland (1996): “I would rather be safe than sorry.”, “I want to be sure before I take over-the-counter medications.”, and “I avoid risky things.”.

Trust in Health-Care System (α = .95): Overall trust in the health-care system also plays a role in the decision to see a doctor. The original scale from Egede and Ellis (2008) provides three dimensions for trust in the health-care system: Trust in health-care providers, in health-care payers, and in health-care institutions. For our purpose only the first and third subscale were appropriate, the items were queried accordingly. The first dimension used ten items, for example “My health care provider is usually considerate of my needs and puts them first.” and the third subscale three items, for example “Healthcare institutions provide the highest quality in medical care.”.

Demographic Information: Within the demographic question section at the end of the survey data on gender, age, employment, and gross annual income was gathered. Further, education was queried as the level of education seems to be associated with social media usage for information searching (Sumayyia et al., 2019).

4 Analysis and Results

The following chapter is divided into three parts. First, a manipulation check is conducted to test whether all manipulations within the experiment worked as intended. Second, the descriptive analyses are introduced to highlight the changes in willingness to visit a doctor after being exposed to one of the influencers' posting. Third, the research questions are tested using a mixed between-within ANCOVA and a logistic regression.

4.1 Manipulation Check

Patient Influencer: Subjects were asked to rate whether the influencer suffers from hemorrhoids herself/himself (1 = yes, 2 = no). In the scenario with the influencer suffering from hemorrhoids him/herself, the influencer was more likely to be perceived to be affected him/herself ($M = 1.03$, $SD = 0.175$) than in the scenario with the influencer that was "healthy and well" (($M = 1.77$, $SD = 0.422$), $t(588.40) = 33.96$, $p < 0.001$).

Expertise: To measure the level of medical expertise of the influencer, participants were asked whether the influencer from the initial scenario has a medical degree (1 = yes, 2 = no). In the scenario where the influencer was said to have a medical degree, the influencer was also more likely to be perceived to hold a medical degree ($M = 1.12$, $SD = .32$) compared to the scenario where the influencer stated that s/he was just like anyone else ("someone like you") (($M = 1.93$, $SD = .255$), $t(838.70) = 41.84$, $p < 0.001$). A second manipulation check verified these findings. Subjects were further asked how they would rate the influencer's medical expertise relative to others (1 = one of the least knowledgeable, 7 = one of the most knowledgeable). The influencer in the scenario where s/he mentioned that s/he completed a medical degree but is currently not practicing was perceived to have more medical expertise ($M = 4.67$, $SD = 1.10$) compared to the scenario where the influencer only mentioned that s/he is "someone like you" (($M = 3.46$, $SD = 1.20$), $t(877.41) = -15.65$, $p < 0.001$).

Number of Followers and Popularity: Subjects were asked to indicate how many followers the influencer profile from the initial scenario had (1 = large community, 2 = small community). In the scenario where the influencer displayed a large community (i.e., 251,712 followers), the influencer was also perceived to have a larger community ($M = 1.19$, $SD = 0.396$) compared to the influencer in the scenario which only displayed a small community (i.e., 238 followers) (($M = 1.84$, $SD = .364$), $t(877.77) = 25.44$, $p < 0.001$). Further, we measured popularity using an eight-item seven-point Likert type scale, i.e., "The person from whom I read the Instagram posting" e.g., (1) "is popular" or (2) "is quite accepted" (1 = strongly disagree; 7 = strongly agree). The same results were found on the popularity scale, the influencer displaying a larger community was perceived to be more popular (M = 5.12, SD = .86) compared to the influencer only displaying a smaller community ((M = 4.08, SD = .89), t(883) = -17.78, p < 0.001).

Recommendation Visiting a Doctor: Participants were also asked whether the influencer recommended visiting a doctor in her/his posting (1 = strongly disagree, 7 = strongly agree). The scenario in which the influencer recommended seeing a doctor ("you should see a doctor") was perceived to be more likely to recommend visiting a doctor (M = 5.58, SD = 1.67) compared to the group where the influencer did not recommend visiting a doctor ("you do not need to see a doctor") ((M = 2.29, SD = 1.51), t(871.71) = -30.51, p < 0.001).

4.2 Descriptive Analysis

Within this subchapter, the change in willingness to visit a doctor is examined descriptively. Table 2 and Appendix 4 present an overview of the change in willingness to visit a doctor from T0 (initial situation) to T1 (after being exposed to the Instagram profile and posting). Further, we looked at the different influencer profile and posting characteristics. These are only discussed in more detail if one characteristic exceeds the other in terms of percentage. The distribution of demographics was similar to that in the overall sample.

Table 2. Overview of Change in Willingness to Visit a Doctor

			T1: Willingness to Visit a Doctor…				
			1	2	3	4	**Total**
T0: Willingness to Visit a Doctor…	1	n	200	10	12	0	**222**
		%	90.1%	4.5%	5.4%	0.0%	**100.0%**
	2	n	18	129	23	2	**172**
		%	10.4%	75.0%	13.4%	1.2%	**100.0%**
	3	n	19	40	199	18	**276**
		%	6.9%	14.5%	72.1%	6.5%	**100.0%**
	4	n	5	16	24	170	**215**
		%	2.3%	7.4%	11.2%	79.1%	**100.0%**

1 = not consult a doctor; 2 = appointment in the next couple of weeks; 3 = appointment in the next couple of days; 4 = appointment immediately // T0 = initial situation; T1 = after being exposed to the Instagram profile and posting

T0 Willingness to Visit a Doctor: Not Consult a Doctor

90.1% who decided not to see a doctor after reading the initial situation, did not change their mind after reading the Instagram post and still do not want to visit a doctor. The remaining 9.9% changed their initial opinion into visiting a doctor. Those subjects were exposed to an

Instagram post recommending visiting a doctor (86%) and/or a posting from a not-affected influencer (63.6%).

T0 Willingness to Visit a Doctor: Appointment in the Next Couple of Weeks

When looking at those who wanted to visit a doctor in the next few weeks, 75% did not change their opinion. On the one hand, 10.4% decided not to go to a doctor at all after being exposed to a posting of a nano influencer (55.6%) and/or a posting not recommending visiting a doctor (72.2%). On the other hand, 14.6% of participants who initially decided to make an appointment in the next few weeks changed their mind into seeing a doctor earlier. Those who have decided to go to the doctor earlier in T1 have seen postings from not-affected influencers (56%) and/or postings recommending visiting a doctor (72%).

T0 Willingness to Visit a Doctor: Appointment in the Next Couple of Days

Of those participants that initially decided to make an appointment in the next few days 27.9% changed their mind after being exposed to the Instagram post. 21.4% of participants switched their opinion to not go to the doctor or to making an appointment later. Here, the majority who changed their disposition into consulting a doctor later or never were exposed to an Instagram post not recommending visiting a doctor (59.2%) and/or a posting of an expert influencer (57.3%). The majority who changed their opinion into consulting a doctor immediately were exposed to an Instagram post recommending visiting a doctor (70%) and/or a posting of a non-expert influencer (60%) and/or someone who has a small community (60%).

T0 Willingness to Visit a Doctor: Appointment Immediately

Of the subjects who initially decided to make an appointment immediately, 20.9% of participants changed their mind to seeing a doctor later or never. Those were exposed to an Instagram post not recommending visiting a doctor (73.3%).

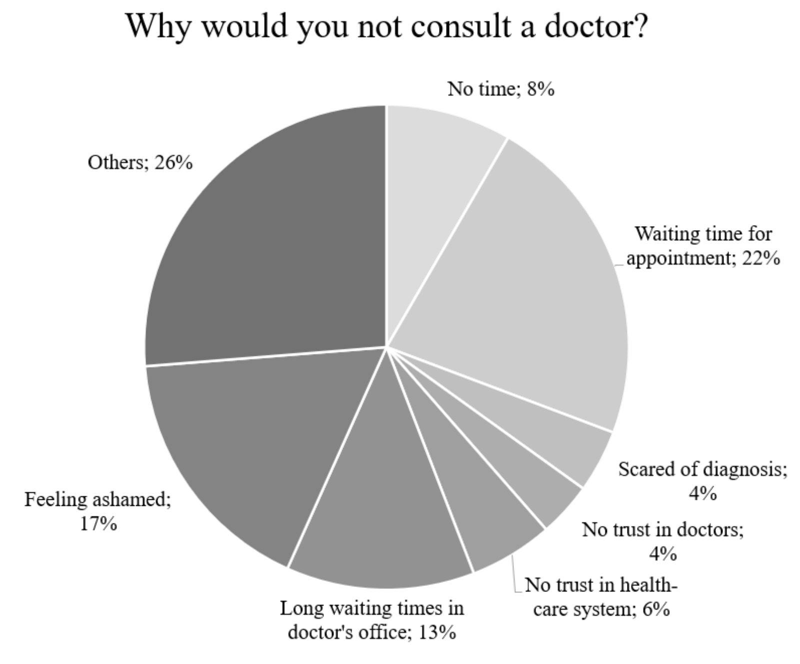
In summary, the majority did not change their initial disposition after being exposed to the Instagram postings or profiles. Still, between 9.9% and 27.2% did change their opinion either into seeing a doctor earlier or never/later. We will have a closer look at these groups.

4.3 Qualitative Analysis

Additionally, we examined those that decided to not visit a doctor in T1. 27.3% decided to not visit a doctor regardless of the Instagram posting and profile the participants they were exposed to. We asked those subjects to further state the reasons why they decided not to visit a doctor.

Participants had the chance to state several reasons. An overview can be found in Figure 1. Participants most likely would not like to consult a doctor as they have to wait quite long for an appointment (22%), are feeling ashamed (17%), or do not want to spend much time waiting in the doctor's office (13%). Other reasons are that they do have no time (8%), have no trust in the health-care system (6%), are scared of the diagnosis (4%), or have no trust in doctors (4%). These results are also reflected in the open-ended responses. We investigated the open answers of the remaining 26% that stated other reasons. The given statements show that most people do not consider the problem serious enough to see a doctor. They prefer to use home remedies or over-the-counter medications. Some do not want to burden the health service or feel uncomfortable taking up the doctor's time. Most believe they can treat the problem themselves and see no need for medical advice. Financial constraints and confidence in self-treatment also play a role.

Figure 1. Overview of Reasons why Patients Refrain from Visiting a Doctor



4.4 Mixed Between-Within ANCOVA

We conducted a mixed between-within ANCOVA to examine the impact of the various characteristics of Instagram profiles and posts on participants' inclination to visit a doctor. This method was chosen as it combines within-subject and between-subject designs. In the mixed ANCOVA, there is at least one variable as the within-subject factor (i.e., likelihood to visit a doctor measured before (T0) and after (T1) the treatment). The within-subject factor is therefore *time* since we measured the dependent variable twice in the same person during the experiment. Further, there is at least one between-subject factor (i.e., Instagram profiles and postings: number of followers and popularity, expertise, patient influencer, recommendation visiting a doctor). By employing a mixed ANCOVA, we can assess the impact of different Instagram profiles and types on the alteration of the inclination to visit a doctor.

Since our primary focus is on the shift in opinions, we initially performed a median split based on the willingness to visit a doctor before exposure to the manipulation (i.e., subjects were asked to indicate how likely they are to consult a doctor on a seven-point scale). All participants stating a four or higher were categorized into *"willing to visit a doctor T0"*, the others were *"not willing to visit a doctor T0"*. Further, based on the descriptive statistics the type of recommendation seems to be the most important criteria for a change in opinion. Therefore, "recommendation" was used to split the dataset before analyzing the data further.

Assumptions of Mixed Between-Within ANCOVA

In total, there are eight requirements that must be met to calculate a mixed ANCOVA. The requirements are first introduced and later discussed in detail. Three requirements were considered when setting up the research design. The dependent variable (i.e., likelihood to visit a doctor) must be at least interval-scaled. In the experiment, likelihood to visit a doctor was measured on a Likert-type scale from extremely unlikely to extremely likely. Further, the between-subjects factor should be independent and nominally scaled. Based on the different scenarios of the experiment, participants were randomly divided into different groups. The within-subject factor should also be independent and nominally scaled, here, the two different points of time in measurements of the dependent variable are used as within-subject factor. The residuals of the dependent variable should be approximately normally distributed for each group. However, this assumption is considered the least important and mixed ANCOVAs are considered to be robust against the violation of this assumption, especially for moderate violations or when the sample size is appropriately large ($n > 30$). For analyses with more than 30 subjects it can be assumed that, according to the central limit theorem, the sampling distribution will be approximately normally distributed (Bortz & Schuster, 2010). The same applies for outliers. Sphericity, which depicts the equality of variances between the individual groups, should be given. It is one of the most important assumptions of mixed ANCOVAs with more than two levels. However, in our case, the within-subject factor *time* has only two levels, therefore the Mauchly test cannot be calculated. In such a case, sphericity is given (O'Brien & Kaiser, 1985). Furthermore, the variances should be homoscedastic.

T0: Willing to Visit a Doctor

We first had a look at the group that was initially willing to visit a doctor, which was assessed by the means of a median split. After further splitting the dataset based on the group of recommendation (recommendation to visit a doctor: yes versus no), homogeneity of the error variances, as assessed by Levene's test ($p > .05$) was found for both groups "recommendation yes" and "recommendation no" for both measures during T0 and T1. Further, the box test for equality of covariance matrices, is relevant due to the mixed model design. As the error term in a mixed design averages the error terms for each level of the between-subjects factor, the interaction should not only be equal from one pair to another pair within levels, however also from a pair in one group with the same pair in other groups (Cohen, 2008). For the group "recommendation no", homogeneity of covariances, as assessed by Box's test ($p = .709$) was found, however, for the group "recommendation yes", no homogeneity of covariances was found ($p = .027$). In scientific practice, however, the results are often interpreted anyway, and we also follow this procedure in our paper.

T0: Not Willing to Visit a Doctor

Subsequently, we examined the requirements for individuals in the group who initially expressed reluctance to visit a doctor. Following the division of the dataset into those advised to visit a doctor and those advised not to, we examined the homogeneity of error variances through the Levene's test. For both groups during T0 and T1, we found homogeneity of error variances ($p > .05$). Further, by looking at the Box's test of equality of covariance matrices, for the group "recommendation no", no homogeneity of covariances, as assessed by the Box's test ($p = .048$) was found, however, for the group "recommendation yes", homogeneity of covariances was found ($p = .178$). We also follow the scientific approach in interpreting these results.

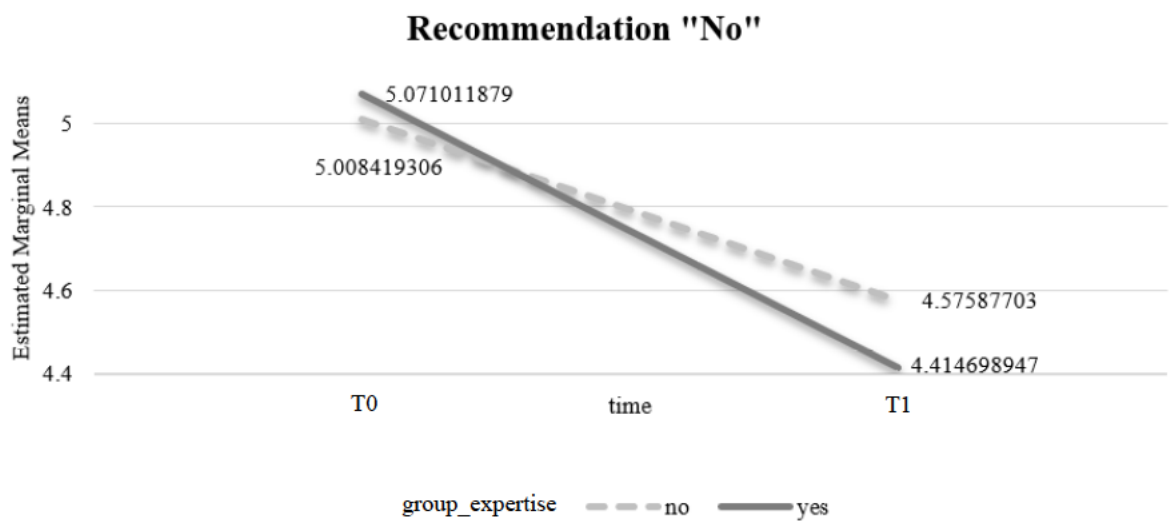
Interpretation of Results

When interpreting the results, we are mainly interested in the interaction between time and the different characteristics of the Instagram posting.

T0: Willing to Visit a Doctor

Since the study predominantly concentrated on patients who altered their opinions, we examined participants initially willing to visit a doctor but were advised against it. We identified only one statistically significant interaction between time and group_expertise, $F(1, 287) = 4.65$, $p = 0.032$, partial $\eta^2 = 0.016$, which equals a rather low effect size of 0.13. All other interactions were insignificant (see Appendix 5 for more details). Further, there were no significant main effects on time or any of the other groups, see also Appendix 5 for more details. Examining the graphical distinctions among expert groups, Figure 2 illustrates that participants exposed to a post by an expert influencer recommending against visiting a doctor are more inclined to shift their opinion toward not visiting a doctor. However, when conducting an independent t-test, no significant difference was found ($t(297) = -.570$, $p = .569$). This finding will be further investigated using a logistic regression (Chapter 4.5).

Figure 2. Graphical Depiction of the Difference in Expert Groups With Recommendation "No"



T0: Not Willing to Visit a Doctor

Additionally, we investigated the group of patients initially unwilling to visit a doctor who were subsequently advised to do so. One significant interaction was found between time and group_patient_influencer $F(1, 138) = 4.29$, $p = 0.040$, partial $\eta^2 = 0.030$, which also represents a rather low effect size of 0.18. When conducting an independent t-test, no significant difference between the group being affected and the one not being affected on likelihood to visit a doctor ($t(148) = -1.076$, $p = .284$) was found. See Appendix 6 for further details.

4.5 Logistic Regression Analysis

Assumptions

In order to verify the results from 4.4 and to investigate which characteristics lead to a change in opinion, a logistic regression was used in the next step. It is examined whether participants are willing to change their disposition from making an appointment (1) in the next couple of days to delaying the appointment or to not consult a doctor at all and whether patients are willing to change their opinion from (2) making an appointment immediately to delaying the appointment or to not consult a doctor at all.

First, the assumptions of a logistic regression are tested. The dependent variable is nominally scaled with exactly two values (dichotomous). For case (1), the values of the dependent variable are 0 for no change from T0 to T1 (i.e., subjects stick to their opinion of making an appointment in the next few days) and 1 for a change in opinion from T0 to T1 (i.e., subjects decide to postpone the appointment or to not consult a doctor at all). For case (2), the values of the dependent variable are coded similarly: 0 for no change from T0 to T1 (i.e., subjects stick to their opinion of making an appointment immediately) and 1 for a change in opinion from T0 to T1 (i.e., subjects decide to postpone the appointment or to not consult a doctor at all).

The independent variables are either nominally scaled or at least interval scaled. Furthermore, the independence of observations is important. There is no relationship between the observations in each category of the dependent variables or the observations in each category of the independent variable. No outliers were found for (1), for (2) one outlier was found which was left in the dataset. Linearity was tested assessed using the Box-Tidwell (Box & Tidwell, 1962) procedure. All variables were found to follow a linear relationship. Correlations between predictor variables were low - (1) $r < .30$; (2) $r < .40$ - indicating that multicollinearity was not a confounding factor in the analysis.

Analysis (1) – T0 = Appointment in the Next Couple of Days

The binomial logistic regression model was statistically significant, $\chi^2(13) = 37.560$, $p < .001$, resulting in an acceptable amount of explained variance (Backhaus et al., 2003), as shown by Nagelkerke's $R^2 = .206$. Further, Goodness-of-fit was assessed using the Hosmer-Lemeshow-Test, indicating a good model fit, $\chi^2(8) = 10.293$, $p > .05$. Overall percentage of accuracy in classification was 78.3%, with a sensitivity of 16.9% and a specificity of 96.5%. The results show that for every additional unit of perceived medical expertise, the probability of postponing a doctor's appointment increases by a factor of 1.333 ($p < 0.05$). Furthermore, the results show that the odds of postponing the appointment or not going to the doctor at all are 5.525 ($p < 0.001$) times higher if the influencer recommends not going to the doctor.

To conclude, the advice of an expert on Instagram, as well as the recommendation not to go to the doctor, increases the probability that patients who originally wanted to make an appointment in the next few days either do not go to the physician at all or do so at a later date.

Table 3. Logistic Regression for T0 = Appointment in the Next Couple of Days

	B	SE	Wald	Sig.	Odds Ratio	95% CI for Odds Ratio	
						Lower	Upper
Gender	-0.209	0.336	0.387	0.534	0.812	0.420	1.566
Age	-0.003	0.016	0.040	0.842	0.997	0.965	1.029
Education	-0.215	0.162	1.773	0.183	0.806	0.588	1.107
Employment	-0.022	0.106	0.045	0.832	0.978	0.794	1.204
Income	-0.104	0.165	0.400	0.527	0.901	0.652	1.245
Group "Affected"	-0.347	0.326	1.132	0.287	0.707	0.373	1.339
Group "Recommendation"	-1.711	0.379	20.340	0.000	0.181	0.086	0.380
Expertise	0.287	0.142	4.111	0.043	1.333	1.010	1.760
Popularity	0.316	0.166	3.614	0.057	1.372	0.990	1.901
Instagram Usage	0.189	0.196	0.931	0.335	1.208	0.823	1.773
Hypochondriasis	-0.128	0.156	0.673	0.412	0.880	0.648	1.195
Risk Aversion	0.184	0.182	1.021	0.312	1.202	0.841	1.719
Overall Trust in Health-Care System	-0.106	0.192	0.308	0.579	0.899	0.618	1.309
Constant	-2.518	1.873	1.807	0.179	0.081		

a. Variable(s) entered on step 1: Gender, age, education, employment, income, group_affected, group_recommendation, expertise, popularity, Instagram usage, hypochondriasis, risk aversion, overall trust in health-care system.

Note. Degrees of freedom were 1 for all Wald statistics

Analysis (2) – T0 = Appointment Immediately

The binomial logistic regression model was statistically significant, $\chi^2(13) = 32.588$, $p = .002$. This results in an acceptable amount of explained variance (Backhaus et al., 2003), as presented by Nagelkerke’s $R^2 = .219$. Goodness-of-fit was assessed using the Hosmer-Lemeshow-Test, indicating a good model fit, $\chi^2(8) = 11.281$, $p > .05$. The overall percentage of accuracy in classification was 81.4%, with a sensitivity of 20.0% and a specificity of 97.6%. The findings show that for every additional unit of perceived medical expertise, the probability of postponing a doctor's appointment increases by a factor of 1.306 ($p < 0.1$). Furthermore, the results demonstrate that the odds of postponing the appointment or not going to the doctor at all are 3.717 ($p < 0.001$) times higher if the influencer recommends not going to the doctor.

To conclude, the advice of an expert on Instagram, as well as the recommendation not to go to the doctor, increases the probability that patients who originally wanted to make an appointment immediately either do not visit a doctor at all or postpone the appointment.

Table 4. Logistic Regression for T0 = Appointment Immediately

	B	SE	Wald	Sig.	Odds Ratio	95% CI for Odds Ratio	
						Lower	Upper
Gender	-0.325	0.364	0.796	0.372	0.723	0.354	1.475
Age	-0.016	0.021	0.584	0.445	0.984	0.945	1.025
Education	0.006	0.147	0.002	0.969	1.006	0.754	1.341
Employment	-0.020	0.120	0.027	0.869	0.980	0.775	1.240
Income	-0.064	0.187	0.118	0.731	0.938	0.649	1.354
Group "Affected"	-0.314	0.380	0.686	0.408	0.730	0.347	1.537
Group "Recommendation"	-1.313	0.403	10.637	0.001	0.269	0.122	0.592
Expertise	0.267	0.151	3.131	0.077	1.306	0.972	1.754
Popularity	-0.078	0.181	0.188	0.665	0.925	0.649	1.318
Instagram Usage	0.528	0.227	5.381	0.020	1.695	1.085	2.647
Hypochondriasis	0.283	0.162	3.057	0.080	1.327	0.966	1.821
Risk Aversion	-0.230	0.198	1.350	0.245	0.795	0.539	1.171
Overall Trust in Health-Care System	0.024	0.178	0.019	0.892	1.025	0.723	1.452
Constant	-2.632	2.010	1.715	0.190	0.072		

a. Variable(s) entered on step 1: Gender, age, education, employment, income, group_affected, group_recommendation, expertise, popularity, Instagram usage, hypochondriasis, risk aversion, overall trust in health-care system.

Note. Degrees of freedom were 1 for all Wald statistics

For *T0 = not consult a doctor* only the group "recommendation" led to a significant change in willingness to visit a doctor. As this characteristic was also significant in all other settings, except for *T0 = appointment in the next couple of weeks*, where no significant difference was found, it seems that a call to action is an important tool to change a followers' initial decision.

5 Discussion, Limitations, and Future Research

Discussion

In the recent years, interactions with friends and strangers are increasingly taking place online with Instagram being the second most popular social media platform in the United States. Interactions are shifting from mere general communication to information gathering, such as health-related issues. Those discussed health-related topics are diverse and can include subjects such as specific diseases or medical problems, medical treatments or procedures (Fox & Duggan, 2013). In addition to acquiring health information, the provision of social and emotional support also plays a pivotal role in patients' involvement in social media (Zhao & Zhang, 2017). Influencers can differ based on their account's characteristics. A literature review identified three distinct factors: patient influencers (i.e., was the influencer affected him- or herself with the disease s/he is talking about), expertise (i.e., influencer has completed medical degree but is currently not practicing versus someone like you and me), and the popularity of the Instagram account (i.e., macro influencers with 251,712 followers versus nano influencers with 238 followers). These three factors seem to be decisive regarding a post's evaluation. This study thus examined whether those factors, individually or in their combination, exert an influence on the patient's initial willingness to demand a doctor's appointment.

The data was first examined descriptively. We analyzed whether participants change their initial disposition in visiting a doctor after being exposed to the Instagram posting. Between 9.9% and 27.2% did change their opinion either into seeing a doctor earlier or never/later. A qualitative examination of individuals who chose not to visit a doctor after reading the post uncovered that participants made this decision due to extended waiting times for appointments, feelings of shame, or a reluctance to spend time in the doctor's waiting room. These descriptive findings were further tested whether they exert a significant influence on the change of willingness to visit a doctor.

Subsequently, the data underwent analysis through a mixed between-within ANCOVA. Our focus was on individuals who were originally willing to visit a doctor but received a recommendation not to do so. Individuals who encounter a post from an expert influencer advising against visiting a doctor are more inclined to shift their opinion toward postponing the doctor's visit or not seeking medical attention at all. Following the conduct of a logistic regression in a second step, a further statistically significant result emerged. Our second research question whether the medical expertise of an Instagram influencer alters a patient's willingness to demand a doctor was answered in the affirmative. Trustworthy information and the word of an expert seem to weigh more on social media than the word of a like-minded person.

The study contributes to the investigation of social media's role in health-care. It extends academic literature in the way that it examines different influencer characteristics. Medical expertise seems to be the only criterion that has an influence on the change of willingness to see a doctor. The other criteria, such as popularity or affectedness, appear to be negligible. Influencers who take on an expert role (i.e., previous medical expertise rather than being a person like you and me) have an impact on changing patients' minds. Patients are therefore more inclined to postpone a visit to the doctor or not to see a doctor at all, when being exposed to a posting of an expert. In addition, a "call to action" is crucial. Followers are more likely not to see a doctor or to postpone the appointment if they have been advised not to do so in advance. The study therefore shows that patients are more likely to be influenced by a post and a recommendation from an expert than by a post from a person who is similar to them.

Given that the current study was conducted in the health-care sector, where highly sensitive data can be rapidly accessed, it is not surprising that the results, coupled with research on trust in social media, align with expectations. The media keep telling us how important social media have become in recent years. People are following dangerous health trends (Landwehr, 2023), such as drinking Borax which is a substance often used in laundry detergents to decrease inflammation and joint pain (Bendix & Yang, 2023). Teens are also using social media to self-diagnose themselves (Murphy Kelly, 2023). Especially in health-care, the question of the trustworthiness of the data arises.

In scientific literature, the main recurring limitations being discussed are quality concerns (Adams, 2010; Moen et al., 2009; Orizio et al., 2010) and the lack of reliability of health information (Adams, 2010; Kukreja et al., 2011; Moen et al., 2009). Furthermore, it also seems to be more difficult for individuals to discern the reliability of information found online (Adams, 2010). However, even if there are existing doubts about the content's credibility, there is also the possibility that social media may act as a deterrent from seeking out health professionals (Kim, 2009). In a study conducted with a Dutch sample it was found that health-related information found on social media was perceived to be least reliable (van de Belt et al., 2013). If we look at practice, a similar picture emerges. A study conducted in Germany on general trust in social online networks shows that in winter 2021/2022 around 70% of respondents said they tended not to trust such networks (Europäische Kommission, 2022). In the United States, however, younger adults under the age of 30 are now almost as likely to trust information on social media as compared to information from national news outlets (Liedke & Gottfried, 2022). The question of how important and influential social media are in health-care cannot be fully answered and provides fruitful avenues for future research.

Practical Implications for Health-Care

Policymakers should ensure that there are regulations in place to deal with the dissemination of information on social networks. As the study found that especially recommendations from medical professionals are of importance in willingness to demand a doctor's visit, practitioners should focus on the social media savvy group and inform them about misinformation on social media. This could be done by running campaigns to educate people about 1) the truthfulness of information and 2) offering workshops to learn how to distinguish trustworthy sources from untrustworthy ones.

An alternative perspective could be derived from the field of marketing. Instagram and other channels could introduce an additional button, such as the "paid advertising" button, or could have policymakers in different countries require influencers to first verify that they are from a medical background or have been trained to make certain medical statements. Alternatively, they might be required to include an addendum to a post explicitly indicating that it represents a personal opinion or is based on subjective grounds and advising followers to promptly consult a physician. YouTube serves as a pioneer here: YouTube aims at addressing medical misinformation which might be present on its platform. With this initiative, YouTube will eliminate content that contradicts the guidance provided by health authorities regarding the prevention and transmission of various medical conditions (Suter, 2023). Additionally, it will take down content that opposes recommended treatments, including videos that promote unproven remedies as an alternative to seeking proper medical care, as well as content that denies the existence of specific conditions, such as COVID-19. The platform specified that these new policies will be enforced in cases where content contradicts the recommendations of local health authorities or the World Health Organization (WHO) in relation to specific health conditions, treatments, and substances (Suter, 2023).

Limitations and Future Research

In a lot of different studies on the effects of social media in health-related issues, chronic diseases such as diabetes (Zhang et al., 2013) or inflammatory bowel disease (Stewart Loane & D'Alessandro, 2014) were chosen to be the subject of interest. An illness that is only temporary may not be taken as seriously. A disease that is chronic and that patients have to deal with throughout their lives has already become part of their lives. Patients are therefore much more involved, which can lead to a desire for alternative treatments or new medications and testimonials from others if their doctor's initial advice did not help them. If the physician is unable to offer assistance or only provides information that is already known, this target group might be more inclined to switch to alternative channels. Therefore, another starting point for future research would be to conduct a similar experiment in a setting with chronically ill patients. In the underlying setting, participants had to imagine themselves in a fictitious scenario dealing with a health problem. If people are affected by a disease themselves, they have a different relationship to it and would, especially if conventional medicine can no longer help, possibly look for alternatives through other channels. One possibility would therefore be to carry out a field experiment, which would deal with chronically ill patients and first track down the usage behavior on social media in qualitative or descriptive studies.

Furthermore, demographic aspects could also be decisive. The sample was limited to English-speaking participants from the United States and the United Kingdom and displayed and average age of 36 years. However, Turkey, Chile, and Argentina have the highest share of Instagram users in the population aged 13 and over, worldwide (Lohmeier, 2023). In 2022, around 24.5% of global Instagram users were male and belonged to the 25-34 age group. Aged 25 and 34 and female were around 22.3% of Instagram users worldwide (Lohmeier, 2023). During aging, the younger generation will also suffer from health problems in due course. Hence, the utilization of social media for health-related issues, might become pertinent for the target group in the coming years. Subsequent research could commence preliminary studies on usage behavior with the primary target group, facilitating the observation of trends in the coming years.

Conclusion

There is no one answer to whether social media are a curse or a blessing for the health-care industry. When using social media in health-care, the main focus should be on clarifying whether the influencer has prior medical knowledge, as this is the crucial factor for a patient's change in opinion. The health-care system should address this substantial patient volume to ensure timely and adequate treatments, thereby minimizing subsequent costs. Still, social media should be used wisely, especially in the health-care industry.

Appendix

Appendix 1. Overview of Initial Situation

Imagine you woke up and experience the following symptoms:

- Itching, burning, and oozing at the anus.
- Foreign body or pressure sensation in the anal region.
- Sitting becomes increasingly uncomfortable.
- Pain during defecation.

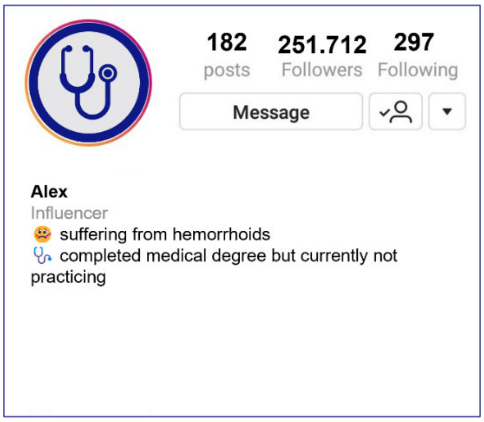
On a scale from 1 (very low) to 10 (very high), your pain level reaches a 6. You do some research on the Internet and come across the following possible diagnosis: **Hemorrhoids**.

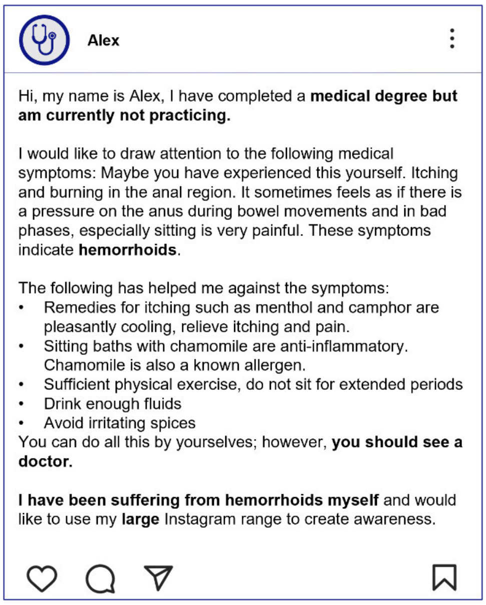
Hemorrhoids can be treated medically which leads to faster recovery and less pain. However, you do not necessarily have to have it treated professionally, it will take a little longer, but the symptoms might go away on their own.

Appendix 2. Overview of Initial Situation

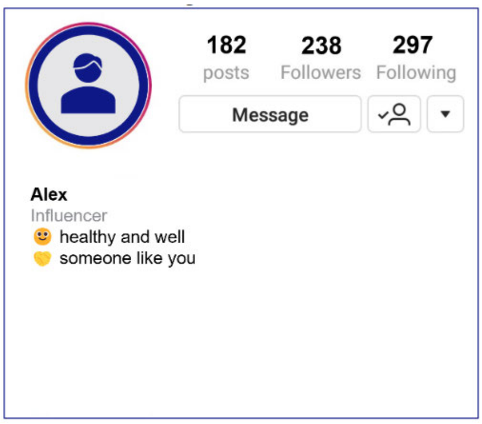
Please read the following scenario.

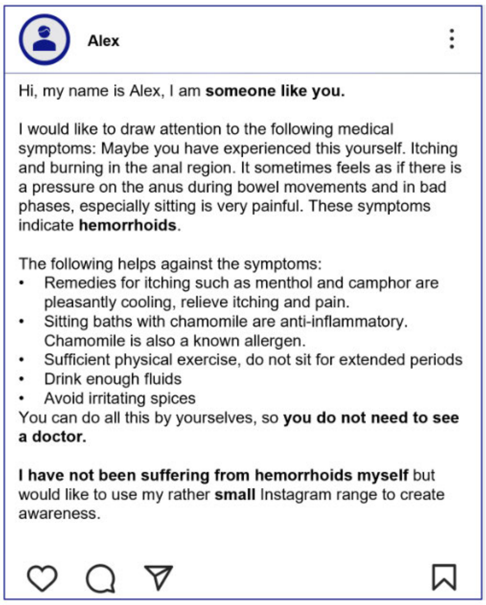
Recommendation " yes"





Recommendation " no"





Key

- **Followers:** nano = 238 followers vs. macro = 251,712 followers
- **(Not) affected:** suffering from hemorrhoids vs. healthy and well
- **Expertise:** completed medical degree but currently not practicing vs. someone like you

- **Followers:** small vs. large Instagram range
- **(Not) affected:** I have been suffering from hemorrhoids myself vs. I have not been suffering from hemorrhoids myself
- **Expertise:** I have completed a medical degree but am currently not practicing vs. I am someone like you

Appendix 3. Overview of Measures and Stimuli

Construct/ Variable	Item/Proxy	Precedents/ Sources
	Dependent Variables	
Likelihood to Visit a Doctor Ia	The construct is measured a one-item six-point Likert-type scale from "extremely unlikely" to "extremely likely". How likely are you to consult a doctor? (1) Extremely unlikely – extremely likely	Self-developed
Likelihood to Visit a Doctor Ib	The construct was measured using a one-item six-point scale. Would you consult a doctor? (1) I would **not** consult a doctor. (2) I would try to make an appointment in the next few **days**. (3) I would try to make an appointment in the next few **weeks**. (4) I would try to make an appointment **immediately**.	Self-developed
Likelihood to Visit a Doctor IIa	The construct is measured a one-item six-point Likert-type scale from "extremely unlikely" to "extremely likely". Now that you have read the scenario, how likely are you to consult a doctor? (1) Extremely unlikely – extremely likely	Self-developed
Likelihood to Visit a Doctor IIb	The construct was measured using a one-item six-point scale. Please answer the following questions: (1) I would **not** consult a doctor. (2) I would try to make an appointment in the next few **days**. (3) I would try to make an appointment in the next few **weeks**. (4) I would try to make an appointment **immediately**.	Self-developed
Reasons to Not Visit a Doctor	This question featured seven reasons, which were identified during a literature review, and one open answer field. Why would you not consult a doctor? (1) No Time (2) Waiting time for appointment (3) Scared of diagnosis (4) No trust in doctors (5) No trust in health-care system (6) Long waiting times in doctor's office (7) Feeling ashamed (8) Others: ________________	Self-developed
	Manipulation Check	
Community Influencer	The construct was measured using two items. Please indicate how many followers the influencer profile from the initial scenario had. (1) Large community (2) Small community	Self-developed

Appendix 3. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
Popularity *(α = .93)*	The construct is measured on an eight-item seven-point Likert type scale from “strongly disagree” to “strongly agree”. The person from whom I read the Instagram posting… (1) is popular. (2) is quite accepted. (3) is well-known. (4) is generally admired. (5) is liked. (6) is socially visible. (7) is viewed fondly. (8) is not popular.	Adapted from Scott and Judge (2009)
Professional-ism	The manipulation was measured on a one-item scale. The influencer from the initial scenario has a medical degree. (1) Yes (2) No	Self-developed
Expertise	The construct is measured on a one-item semantic differential scale. How would you rate the influencer’s medical expertise relative to others? (1) One of the least knowledgeable - One of the most knowledgeable	Adapted from Yoo (2014)
Affected Influencer	The manipulation was measured on a one-item scale. The influencer from the initial scenario suffers from hemorrhoids him/herself. (1) Yes (2) No	Self-developed
Recommen-dation	The construct is measured on a one-item seven-point Likert type scale from “strongly disagree” to “strongly agree”. (1) In the posting, the influencer recommended to visit a doctor.	Self-developed
	Attention Check	
Attention Check	The construct is measured on a one-item seven-point scale from “strongly disagree” to “strongly agree”. (1) Please select "strongly disagree" as your answer.	Adapted from Gruzd et al. (2020)

Appendix 3. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
	Covariates and Controls	
Instagram Affinity *(α = .76)*	The construct measured on a four-item seven-point Likert type scale from “strongly disagree” to “strongly agree”. How much do you agree with the following statements? (1) I use Instagram more often than other people do. (2) I am interested in Instagram. (3) I am experienced in using Instagram. (4) In general, Instagram is important for me.	Schumann et al. (2014)
Hypochon-driasis *(α = .93)*	The construct is measured on a 14-item seven-point Likert type scale from “strongly disagree” to “strongly agree”. How much do you agree with the following statements? (1) I often worry about the possibility that I have got a serious illness. (2) I am bothered by many aches and pains. (3) I am often aware of various things happening in my body. (4) I worry a lot about my health. (5) I often have symptoms of very serious illnesses. (6) If a disease is brought to my attention (through the radio, television, newspapers, or someone I know) I worry about getting it myself. (7) If I feel ill and someone tells me that I am looking better, I become annoyed. (8) I find that I am bothered by many different symptoms. (9) I find it difficult to forget about myself and think about all sorts of other things. (10) I find it hard to believe the doctor when he/she tells me there is nothing for me to worry about. (11) I get the feeling that people are not taking my illness seriously enough. (12) I think that I worry about my health more than most people. (13) I think there is something seriously wrong with my body. (14) I am afraid of illness.	Adapted from Hiller et al. (2002)
Risk Aversity *(α = .63)*	The construct is measured on a three-item seven-point Likert type scale from “strongly disagree” to “strongly agree”. How much do you agree with the following statements? (1) I would rather be safe than sorry. (2) I want to be sure before I take over-the-counter medications. (3) I avoid risky things.	Adapted from Donthu and Gilliland (1996)

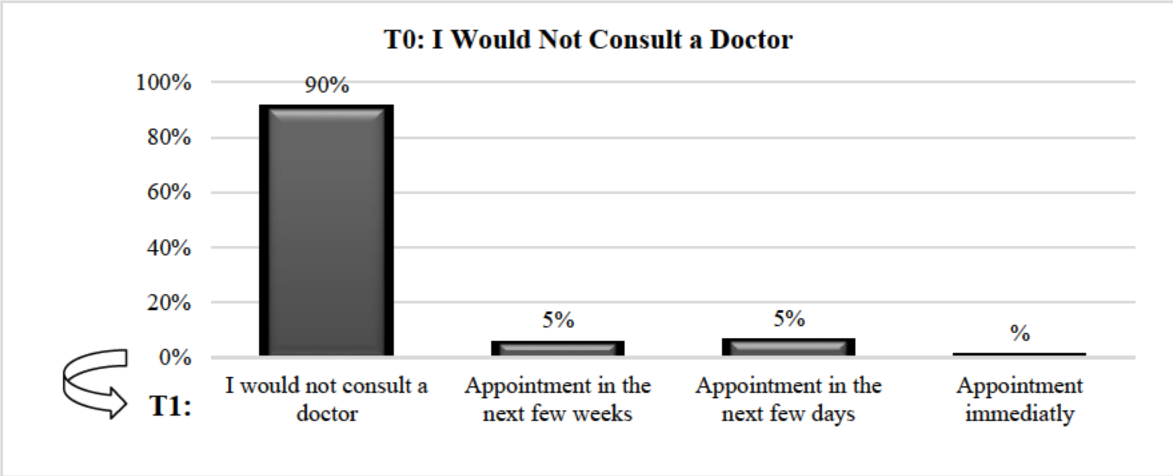
Appendix 3. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
Trust in Health Care System *(α = .95)*	The construct is measured on a 13-item seven-point Likert type scale from “strongly disagree” to “strongly agree”. How much do you agree with the following statements? (1) My health care provider is usually considerate of my needs and puts them first. (2) I have so much trust in my health care provider that I always try to follow his/her advice. (3) I trust my health care provider so much that whatever he/she tells me, it must be true. (4) Sometimes, I do not trust my health care provider’s opinion and therefore I feel I need a second one. (5) I can trust my health care provider´s judgments concerning my medical care. (6) My health care provider will do whatever it takes to give me the medical care that I need. (7) Because my health care provider is an expert, he/she is able to treat medical problems like mine. (8) I can trust my health care provider’s decisions on which medical treatments are best for me. (9) My health care provider offers me the highest quality in medical care. (10) All things considered, I completely trust my health care provider. (11) Health care institutions only care about keeping medical costs down, and not what is needed for my health. (12) Healthcare institutions provide the highest quality in medical care. (13) When treating my medical problems, health care institutions put my medical needs above all other considerations, including costs.	Adapted from Egede and Ellis (2008)
	Demographics	
Gender	Please indicate which gender you feel most closely aligned with: • Male • Female • Non-binary/third gender • Prefer not to say	Self-developed
Age	How old are you? ______________________	Self-developed
Education	What is the highest level of education you have achieved? • Some school but no degree • High school graduate • Some college but no degree • Bachelor’s degree • Master’s degree • Professional degree • Doctorate degree • Other	Adapted from Lo et al., 2019

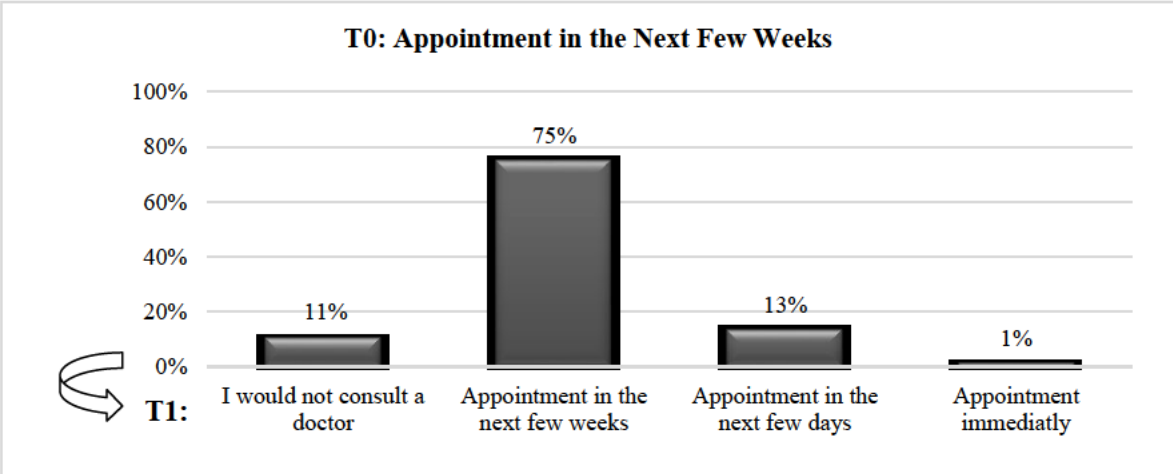
Appendix 3. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources
Employ-ment	What is your current employment status? • Full-time • Part-time • Self-employed • Student • Stay-at-home parent • Unemployed • Retired	Adapted from Lo et al., 2019
Income	What is the level of your annual gross household income? • < $10,000 • $10,000 – $50,000 • $50,001 – $90,000 • $90,001 - $150,000 • >$150,001 • Prefer not to say	Adapted from Lo et al., 2019

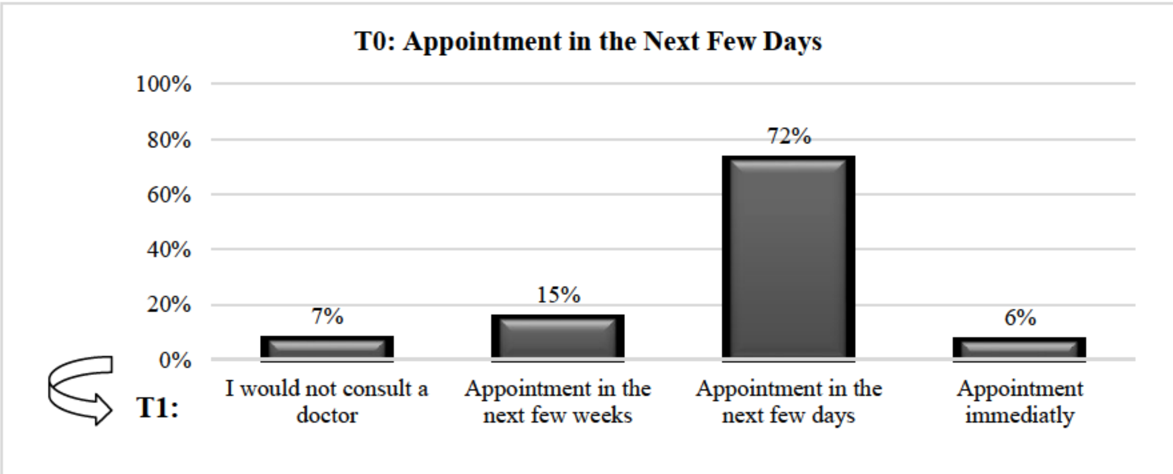
Appendix 4. Change in Willingness to Visit a Doctor From T0 to T1 in Percentage



T0: I Would Not Consult a Doctor

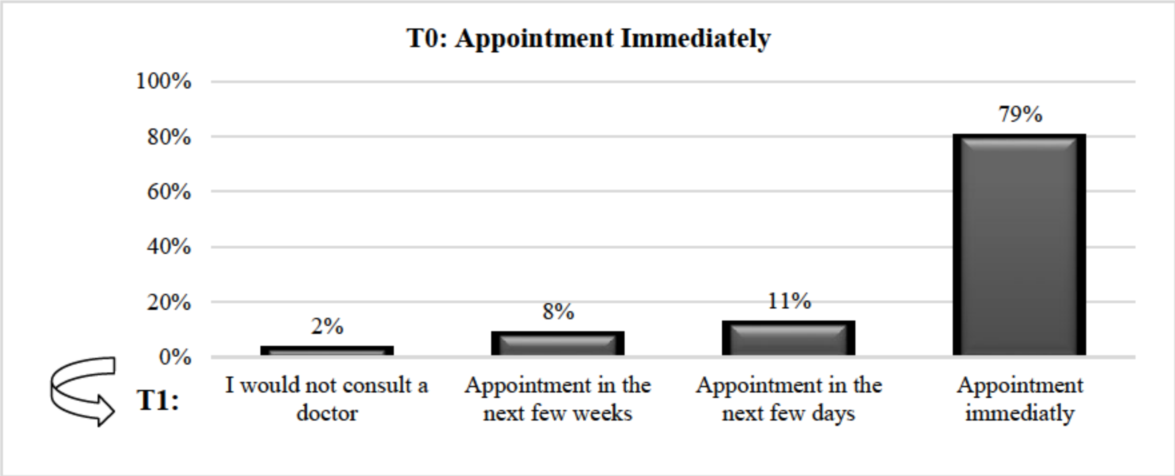


T0: Appointment in the Next Few Weeks



T0: Appointment in the Next Few Days

Appendix 4. Change in Willingness to Visit a Doctor From T0 to T1 in Percentage (continued)



Appendix 5. Test of Within-Subjects Effects for the group that was initially willing to visit a doctor and "Recommendation no"; Measure: Likelihood of Visiting a Doctor

		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Sphericity Assumed	0.227	1	0.227	0.570	0.451	0.002
	Greenhouse-Geisser	0.227	1.000	0.227	0.570	0.451	0.002
	Huynh-Feldt	0.227	1.000	0.227	0.570	0.451	0.002
	Lower-bound	0.227	1.000	0.227	0.570	0.451	0.002
*time * SCALE_Instagram_Usage*	Sphericity Assumed	0.924	1	0.924	2.322	0.129	0.008
	Greenhouse-Geisser	0.924	1.000	0.924	2.322	0.129	0.008
	Huynh-Feldt	0.924	1.000	0.924	2.322	0.129	0.008
	Lower-bound	0.924	1.000	0.924	2.322	0.129	0.008
*time * SCALE_Hypochondriasis*	Sphericity Assumed	0.942	1	0.942	2.367	0.125	0.008
	Greenhouse-Geisser	0.942	1.000	0.942	2.367	0.125	0.008
	Huynh-Feldt	0.942	1.000	0.942	2.367	0.125	0.008
	Lower-bound	0.942	1.000	0.942	2.367	0.125	0.008
*time * SCALE_risk_aversion*	Sphericity Assumed	0.267	1	0.267	0.672	0.413	0.002
	Greenhouse-Geisser	0.267	1.000	0.267	0.672	0.413	0.002
	Huynh-Feldt	0.267	1.000	0.267	0.672	0.413	0.002
	Lower-bound	0.267	1.000	0.267	0.672	0.413	0.002
*time * SCALE_overall_trust_HealthCare System*	Sphericity Assumed	1.188	1	1.188	2.987	0.085	0.010
	Greenhouse-Geisser	1.188	1.000	1.188	2.987	0.085	0.010
	Huynh-Feldt	1.188	1.000	1.188	2.987	0.085	0.010
	Lower-bound	1.188	1.000	1.188	2.987	0.085	0.010
*time * group_patient_influencer*	Sphericity Assumed	0.043	1	0.043	0.109	0.742	0.000
	Greenhouse-Geisser	0.043	1.000	0.043	0.109	0.742	0.000
	Huynh-Feldt	0.043	1.000	0.043	0.109	0.742	0.000
	Lower-bound	0.043	1.000	0.043	0.109	0.742	0.000
*time * group_expertise*	Sphericity Assumed	1.850	1	1.850	4.652	0.032	0.016
	Greenhouse-Geisser	1.850	1.000	1.850	4.652	0.032	0.016
	Huynh-Feldt	1.850	1.000	1.850	4.652	0.032	0.016
	Lower-bound	1.850	1.000	1.850	4.652	0.032	0.016
*time * group_popularity*	Sphericity Assumed	0.505	1	0.505	1.269	0.261	0.004
	Greenhouse-Geisser	0.505	1.000	0.505	1.269	0.261	0.004
	Huynh-Feldt	0.505	1.000	0.505	1.269	0.261	0.004
	Lower-bound	0.505	1.000	0.505	1.269	0.261	0.004

Appendix 5. Test of Within-Subjects Effects for the group that was initially willing to visit a doctor and "Recommendation no"; Measure: Likelihood of Visiting a Doctor (continued)

		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
*time * group_patient_influencer * group_expertise*	Sphericity Assumed	1.006	1	1.006	2.530	0.113	0.009
	Greenhouse-Geisser	1.006	1.000	1.006	2.530	0.113	0.009
	Huynh-Feldt	1.006	1.000	1.006	2.530	0.113	0.009
	Lower-bound	1.006	1.000	1.006	2.530	0.113	0.009
*time * group_ patient_influencer * group_popularity*	Sphericity Assumed	0.040	1	0.040	0.099	0.753	0.000
	Greenhouse-Geisser	0.040	1.000	0.040	0.099	0.753	0.000
	Huynh-Feldt	0.040	1.000	0.040	0.099	0.753	0.000
	Lower-bound	0.040	1.000	0.040	0.099	0.753	0.000
*time * group_expertise * group_popularity*	Sphericity Assumed	0.297	1	0.297	0.746	0.388	0.003
	Greenhouse-Geisser	0.297	1.000	0.297	0.746	0.388	0.003
	Huynh-Feldt	0.297	1.000	0.297	0.746	0.388	0.003
	Lower-bound	0.297	1.000	0.297	0.746	0.388	0.003
*time * group_patient_influencer * group_expertise * group_popularity*	Sphericity Assumed	0.013	1	0.013	0.034	0.854	0.000
	Greenhouse-Geisser	0.013	1.000	0.013	0.034	0.854	0.000
	Huynh-Feldt	0.013	1.000	0.013	0.034	0.854	0.000
	Lower-bound	0.013	1.000	0.013	0.034	0.854	0.000
Error(time)	Sphericity Assumed	114.162	287	0.398			
	Greenhouse-Geisser	114.162	287.000	0.398			
	Huynh-Feldt	114.162	287.000	0.398			
	Lower-bound	114.162	287.000	0.398			

Appendix 6. Test of Within-Subjects Effects for group initially not willing to visit a doctor and "Recommendation yes"; Measure: Likelihood of Visiting a Doctor

		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Sphericity Assumed	0.086	1	0.086	0.234	0.630	0.002
	Greenhouse-Geisser	0.086	1.000	0.086	0.234	0.630	0.002
	Huynh-Feldt	0.086	1.000	0.086	0.234	0.630	0.002
	Lower-bound	0.086	1.000	0.086	0.234	0.630	0.002
*time * SCALE_Instagram_Usage*	Sphericity Assumed	0.239	1	0.239	0.649	0.422	0.005
	Greenhouse-Geisser	0.239	1.000	0.239	0.649	0.422	0.005
	Huynh-Feldt	0.239	1.000	0.239	0.649	0.422	0.005
	Lower-bound	0.239	1.000	0.239	0.649	0.422	0.005
*time * SCALE_Hypochondriasis*	Sphericity Assumed	0.054	1	0.054	0.147	0.702	0.001
	Greenhouse-Geisser	0.054	1.000	0.054	0.147	0.702	0.001
	Huynh-Feldt	0.054	1.000	0.054	0.147	0.702	0.001
	Lower-bound	0.054	1.000	0.054	0.147	0.702	0.001
*time * SCALE_risk_aversion*	Sphericity Assumed	0.008	1	0.008	0.021	0.884	0.000
	Greenhouse-Geisser	0.008	1.000	0.008	0.021	0.884	0.000
	Huynh-Feldt	0.008	1.000	0.008	0.021	0.884	0.000
	Lower-bound	0.008	1.000	0.008	0.021	0.884	0.000
*time * SCALE_overall_trust_HealthCareSystem*	Sphericity Assumed	0.272	1	0.272	0.738	0.392	0.005
	Greenhouse-Geisser	0.272	1.000	0.272	0.738	0.392	0.005
	Huynh-Feldt	0.272	1.000	0.272	0.738	0.392	0.005
	Lower-bound	0.272	1.000	0.272	0.738	0.392	0.005
*time * group_patient_influencer*	Sphericity Assumed	1.583	1	1.583	4.293	0.040	0.030
	Greenhouse-Geisser	1.583	1.000	1.583	4.293	0.040	0.030
	Huynh-Feldt	1.583	1.000	1.583	4.293	0.040	0.030
	Lower-bound	1.583	1.000	1.583	4.293	0.040	0.030
*time * group_expertise*	Sphericity Assumed	0.547	1	0.547	1.485	0.225	0.011
	Greenhouse-Geisser	0.547	1.000	0.547	1.485	0.225	0.011
	Huynh-Feldt	0.547	1.000	0.547	1.485	0.225	0.011
	Lower-bound	0.547	1.000	0.547	1.485	0.225	0.011

Appendix 6. Test of Within-Subjects Effects for group initially not willing to visit a doctor and "Recommendation yes"; Measure: Likelihood of Visiting a Doctor (continued)

		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
*time * group_popularity*	Sphericity Assumed	0.037	1	0.037	0.101	0.751	0.001
	Greenhouse-Geisser	0.037	1.000	0.037	0.101	0.751	0.001
	Huynh-Feldt	0.037	1.000	0.037	0.101	0.751	0.001
	Lower bound	0.037	1.000	0.037	0.101	0.751	0.001
*time * group pa-tient_influencer * group_expertise*	Sphericity Assumed	0.008	1	0.008	0.022	0.883	0.000
	Greenhouse-Geisser	0.008	1.000	0.008	0.022	0.883	0.000
	Huynh-Feldt	0.008	1.000	0.008	0.022	0.883	0.000
	Lower-bound	0.008	1.000	0.008	0.022	0.883	0.000
*time * group_pa-tient influencer * group_popularity*	Sphericity Assumed	0.710	1	0.710	1.925	0.167	0.014
	Greenhouse-Geisser	0.710	1.000	0.710	1.925	0.167	0.014
	Huynh-Feldt	0.710	1.000	0.710	1.925	0.167	0.014
	Lower-bound	0.710	1.000	0.710	1.925	0.167	0.014
*time * group_expertise * group_popularity*	Sphericity Assumed	0.011	1	0.011	0.031	0.860	0.000
	Greenhouse-Geisser	0.011	1.000	0.011	0.031	0.860	0.000
	Huynh-Feldt	0.011	1.000	0.011	0.031	0.860	0.000
	Lower-bound	0.011	1.000	0.011	0.031	0.860	0.000
*time * group pa-tient_influencer * group expertise * group_popularity*	Sphericity Assumed	0.069	1	0.069	0.186	0.667	0.001
	Greenhouse-Geisser	0.069	1.000	0.069	0.186	0.667	0.001
	Huynh-Feldt	0.069	1.000	0.069	0.186	0.667	0.001
	Lower-bound	0.069	1.000	0.069	0.186	0.667	0.001
Error(time)	Sphericity Assumed	50.872	138	0.369			
	Greenhouse-Geisser	50.872	138.00	0.369			
	Huynh-Feldt	50.872	138.00	0.369			
	Lower-bound	50.872	138.00	0.369			

References

Adams, S. A. (2010). Revisiting the Online Health Information Reliability Debate in the Wake of "Web 2.0": An Inter-Disciplinary Literature and Website Review. *International Journal of Medical Informatics*, *79*(6), 391–400. https://doi.org/10.1016/j.ijmedinf.2010.01.006

Antheunis, M. L., Tates, K., & Nieboer, T. E. (2013). Patients' and Health Professionals' Use of Social Media in Health Care: Motives, Barriers and Expectations. *Patient Education and Counseling*, *92*(3), 426–431. https://doi.org/10.1016/j.pec.2013.06.020

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2003). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Springer.

Bendix, A., & Yang, A. (2023, July 22). *Drinking Borax is the Latest Social Media Trend Medical Authorities are Working Feverishly to Debunk*. https://www.nbcnews.com/health/health-news/drinking-borax-tiktok-trend-medical-authorities-debunk-rcna95526

Benetoli, A., Chen, T. F., & Aslani, P. (2017). Consumer Health-Related Activities on Social Media: Exploratory Study. *Journal of Medical Internet Research*, *19*(10), e352. https://doi.org/10.2196/jmir.7656

Bhaskaran, N., Kumar, M., & Janodia, M. D. (2017). Use of Social Media for Seeking Health Related Information – An Exploratory Study. *Journal of Young Pharmacists*, *9*(2), 267–271. https://doi.org/10.5530/jyp.2017.9.52

Bode, L., & Vraga, E. K. (2018). See Something, Say Something: Correction of Global Health Misinformation on Social Media. *Health Communication*, *33*(9), 1131–1140. https://doi.org/10.1080/10410236.2017.1331312

Boerman, S. C. (2020). The Effects of the Standardized Instagram Disclosure for Micro- and Meso-Influencers. *Computers in Human Behavior*, *103*, 199–207. https://doi.org/10.1016/j.chb.2019.09.015

Borah, P., & Xiao, X. (2018). The Importance of 'Likes': The Interplay of Message Framing, Source, and Social Endorsement on Credibility Perceptions of Health Information on Facebook. *Journal of Health Communication*, *23*(4), 399–411. https://doi.org/10.1080/10810730.2018.1455770

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., vollständig überarbeitete und erweiterte Auflage). *Springer-Lehrbuch*. Springer-Verlag Berlin Heidelberg. http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10448295 https://doi.org/10.1007/978-3-642-12770-0

Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics*, 4(4), 531–550. https://doi.org/10.1080/00401706.1962.10490038

Brandt, M. (2020, March 17). *Sprechstunde bei Doktor Google*. https://de.statista.com/infografik/8497/nutzung-des-internets-fuer-gesundheits-informationen/

Byrne, E., Kearney, J., & MacEvilly, C. (2017). The Role of Influencer Marketing and Social Influencers in Public Health. *Proceedings of the Nutrition Society*, *76*(OCE3). https://doi.org/10.1017/S0029665117001768

Campbell, C., & Farrell, J. R. (2020). More Than Meets the Eye: The Functional Components Underlying Influencer Marketing. *Business Horizons*, *63*, 469–479. https://api.semanticscholar.org/CorpusID:216190412

Campbell, L., Evans, Y., Pumper, M., & Moreno, M. A. (2016). Social Media Use by Physicians: A Qualitative Study of the New Frontier of Medicine. *BMC Medical Informatics and Decision Making*, *16*, 91. https://doi.org/10.1186/s12911-016-0327-y

Chou, W. S., Hunt, Y. M., Beckjord, E. B., Moser, R. P., & Hesse, B. W. (2009). Social Media Use in the United States: Implications for Health Communication. *Journal of Medical Internet Research*, *11*(4), e48. https://doi.org/10.2196/jmir.1249

Chretien, K. C., & Kind, T. (2013). Social Media and Clinical Care: Ethical, Professional, and Social Implications. *Circulation*, *127*(13), 1413–1421. https://doi.org/10.1161/CIRCULATIONAHA.112.128017

Cohen, B. H. (2008). *Explaining Psychological Statistics* (3. ed.). Wiley.

Comp, G., Dyer, S., & Gottlieb, M. (2021). Is TikTok The Next Social Media Frontier for Medicine? *AEM Education and Training*, *5*(3). https://doi.org/10.1002/aet2.10532

DiSilvestro, A. (2021, September 23). *How to Get Your Facebook Page or Account Verified*. https://www.searchenginejournal.com/how-to-verify-facebook/417649/#close

Dixon, S. J. (2023a, February 14). *Global Social Networks Ranked by Number of Users 2023*. https://www-statista-com.emedien.ub.uni-muenchen.de/statistics/272014/global-social-networks-ranked-by-number-of-users/

Dixon, S. J. (2023b, August 29). *Leading Countries Based on Instagram Audience Size as of January 2023.* Statista in Coperation with Kepios. https://www.statista.com/statistics/578364/countries-with-most-instagram-users/

Donthu, N., & Gilliland, D. (1996). The Infomercial Shopper. *Journal of Advertising Research,* 36(2), 69-77.

Durau, J., Diehl, S., & Terlutter, R. (2022). Motivate Ne to Exercise With You: The Effects of Social Media Fitness Influencers on Users' Intentions to Engage in Physical Activity and the Role of User Gender. *Digital Health*, *8*, 20552076221102769. https://doi.org/10.1177/20552076221102769

Egede, L. E., & Ellis, C. (2008). Development and Testing of the Multidimensional Trust in Health Care Systems Scale. *Journal of General Internal Medicine*, *23*(6), 808–815. https://doi.org/10.1007/s11606-008-0613-1

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Europäische Kommission. (2022, April 4). *Wie sehr vertrauen Sie sozialen Netzwerken im Internet?* https://de.statista.com/statistik/daten/studie/377228/umfrage/umfrage-in-deutschland-zum-vertrauen-in-soziale-netzwerke-im-internet/

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., & Stern, A. (2004). Health Related Virtual Communities and Electronic Support Groups: Systematic Review of the Effects of Online Peer to Peer Interactions. *BMJ (Clinical Research Ed.)*, *328*(7449), 1166. https://doi.org/10.1136/bmj.328.7449.1166

Fox, S. (2014, January 15). *The Social Life of Health Information.* Pew Research Center. https://www.pewresearch.org/short-reads/2014/01/15/the-social-life-of-health-information/

Fox, S., & Duggan, M. (2013, January). *Health Online 2013.* https://www.pewresearch.org/internet/2013/01/15/health-online-2013/

Frakt, A. (2016, July 11). *Using the Web or an App Instead of Seeing a Doctor? Caution Is Advised*. https://www.nytimes.com/2016/07/12/upshot/using-the-web-or-an-app-before-seeing-a-doctor-caution-is-advised.html

Greene, J. A., Choudhry, N. K., Kilabuk, E., & Shrank, W. H. (2011). Online Social Networking by Patients With Diabetes: A Qualitative Evaluation of Communication with Facebook. *Journal of General Internal Medicine*, *26*(3), 287–292. https://doi.org/10.1007/s11606-010-1526-3

Gruzd, A., Jacobson, J., & Dubois, E. (2020). Cybervetting and the Public Life of Social Media Data. *Social Media + Society*, *6*(2), 205630512091561. https://doi.org/10.1177/2056305120915618

Gupta, P., Khan, A., & Kumar, A. (2022). Social Media Use by Patients in Health Care: A Scoping Review. *International Journal of Healthcare Management*, *15*(2), 121–131. https://doi.org/10.1080/20479700.2020.1860563

Hackworth, B. A., & Kunz, M. B. (2011). Health Care and Social Media: Building Relationships via Social Networks. *Academy of Health Care Management Journal*, *7*(2), 1–14. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=64873311&site=ehost-live

Haenlein, M., Anadol, E., Farnsworth, T., Hugo, H., Hunichen, J., & Welte, D. (2020). Navigating the New Era of Influencer Marketing: How to be Successful on Instagram, TikTok, & Co. *California Management Review*, *63*(1), 5–25. https://doi.org/10.1177/0008125620958166

Heiss, R. (2020). Fighting Health Infodemics: The Role of Citizen Empowerment. *Eurohealth*, *26*(3), 23–25.

Heiss, R., & Rudolph, L. (2022). Patients as Health Influencers: Motivations and Consequences of Following Cancer Patients on Instagram. *Behaviour & Information Technology*, 1–10. https://doi.org/10.1080/0144929X.2022.2045358

Hiller, W., Rief, W., & Fichter, M. M. (2002). Dimensional and Categorical Approaches to Hypochondriasis. *Psychological Medicine*, *32*(4), 707–718. https://doi.org/10.1017/s0033291702005524

Honigman, B. (2015). *24 Stats on How Social Media has Impacted the Healthcare Industry*. https://getreferralmd.com/2013/09/healthcare-social-media-statistics/

Hyun Jung Oh, Carolyn Lauckner, Boehmer, J., Ryan Fewins-Bliss, & Kang Li (2013). Facebooking for Health: An Examination Into the Solicitation and Effects of Health-Related Social Support on Social Networking Sites. *Computers in Human Behavior*, *29*(5), 2072–2080. https://doi.org/10.1016/j.chb.2013.04.017

Instagram. (2023). *Instagram: Give People the Power to Build Community*. https://about.instagram.com/

Janssen, L., Schouten, A. P., & Croes, E. A. J. (2022). Influencer Advertising on Instagram: Product-Influencer Fit and Number of Followers Affect Advertising Outcomes and Influencer Evaluations via Credibility and Identification. *International Journal of Advertising*, *41*(1), 101–127. https://doi.org/10.1080/02650487.2021.1994205

Jenkins, E. L., Ilicic, J., Barklamb, A. M., & McCaffrey, T. A. (2020). Assessing the Credibility and Authenticity of Social Media Content for Applications in Health Communication: Scoping Review. *Journal of Medical Internet Research*, *22*(7), e17296. https://doi.org/10.2196/17296

Jenkins, E. L., Ilicic, J., Molenaar, A., Chin, S., & McCaffrey, T. A. (2020). Strategies to Improve Health Communication: Can Health Professionals be Heroes? *Nutrients*, *12*(6). https://doi.org/10.3390/nu12061861

Jin, J., Yan, X., Li, Y [Yijun], & Li, Y [Yumei] (2016). How Users Adopt Healthcare Information: An Empirical Study of an Online Q&A Community. *International Journal of Medical Informatics*, *86*, 91–103. https://doi.org/10.1016/j.ijmedinf.2015.11.002

Jin, S.-A. A., & Phua, J. (2014). Following Celebrities' Tweets About Brands: The Impact of Twitter-Based Electronic Word-of-Mouth on Consumers' Source Credibility Perception, Buying Intention, and Social Identification With Celebrities. *Journal of Advertising*, *43*(2), 181–195. https://doi.org/10.1080/00913367.2013.827606

Kaplan, A. M., & Haenlein, M. (2010). Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Kay, S., Mulcahy, R., & Parkinson, J. (2020). When Less is More: The Impact of Macro and Micro Social Media Influencers' Disclosure. *Journal of Marketing Management*, *36*(3-4), 248–278. https://doi.org/10.1080/0267257X.2020.1718740

Kerr, H., Booth, R., & Jackson, K. (2020). Exploring the Characteristics and Behaviors of Nurses Who Have Attained Microcelebrity Status on Instagram: Content Analysis. *Journal of Medical Internet Research*, *22*(5), e16540. https://doi.org/10.2196/16540

Kim, S. (2009). Content Analysis of Cancer Blog Posts. *Journal of the Medical Library Association*, *97*(4), 260–266. https://doi.org/10.3163/1536-5050.97.4.009

Kolmes, K., & Taube, D. O. (2016). Client Discovery of Psychotherapist Personal Information Online. *Professional Psychology: Research and Practice*, *47*(2), 147–154. https://doi.org/10.1037/pro0000065

Kukreja, P., Heck Sheehan, A., & Riggins, J. (2011). Use of Social Media by Pharmacy Preceptors. *American Journal of Pharmaceutical Education*, *75*(9). https://doi.org/10.5688/ajpe759176

Landwehr, J. (2023, August 3). *Why Are People Drinking Borax? Experts Warn Against Trying Dangerous Social Media Trend.* Health. https://www.health.com/ingesting-borax-trend-not-safe-7565551

Lee, S. Y., Hwang, H., Hawkins, R., & Pingree, S. (2008). Interplay of Negative Emotion and Health Self-Efficacy on the Use of Health Information and Its Outcomes. *Communication Research*, *35*(3), 358–381. https://doi.org/10.1177/0093650208315962

Lefebvre, R. C., & Bornkessel, A. S. (2013). Digital Social Networks and Health. *Circulation*, *127*(17), 1829–1836. https://doi.org/10.1161/CIRCULATIONAHA.112.000897

Leist, A. K. (2013). Social Media Use of Older Adults: A Mini-Review. *Gerontology*, *59*(4), 378–384. https://doi.org/10.1159/000346818

Lewis, D., Eysenbach, G., Kukafka, R., Stavri, P. Z., & Jimison, H. B. (2005). *Consumer Health Informatics*. Springer New York. https://doi.org/10.1007/0-387-27652-1

Li, Y [Yibai], & Wang, X. (2018). Seeking Health Information on Social Media. *Journal of Organizational and End User Computing*, *30*(1), 1–22. https://doi.org/10.4018/JOEUC.2018010101

Liedke, J., & Gottfried, J. (2022, October 27). *U.S. Adults Under 30 Now Trust Information From Social Media Almost as Much as From National News Outlets.* Pew Research Center. https://www.pewresearch.org/short-reads/2022/10/27/u-s-adults-under-30-now-trust-information-from-social-media-almost-as-much-as-from-national-news-outlets/

Lo, C. J., Tsarenko, Y., & Tojib, D. (2019). To Tell or Not to Tell? The Roles of Perceived Norms and Self-Consciousness in Understanding Consumers' Willingness to Recommend Online Secondhand Apparel Shopping. *Psychology & Marketing*, *36*(4), 287–304. https://doi.org/10.1002/mar.21179

Lohmeier, L. (2023, May 26). *Unternehmen & Produkte: Instagram 2023.* Statista. https://de-statista-com.emedien.ub.uni-muenchen.de/statistik/studie/id/32325/dokument/instagram-statista-dossier/

Mangold, W. G., & Faulds, D. J. (2009). Social Media: The New Hybrid Element of the Promotion Mix. *Business Horizons*, *52*(4), 357–365. https://doi.org/10.1016/j.bushor.2009.03.002

McCosker, A. (2018). Engaging Mental Health Online: Insights From Beyondblue 's Forum Influencers. *New Media & Society*, *20*(12), 4748–4764. https://doi.org/10.1177/1461444818784303

Meffert, H., Burmann, C., Kirchgeorg, M., & Eisenbeiß, M. (2019). *Marketing: Grundlagen marktorientierter Unternehmensführung Konzepte - Instrumente - Praxisbeispiele* (13., überarbeitete und erweiterte Auflage). *Lehrbuch*. Springer Gabler. https://doi.org/10.1007/978-3-658-21196-7

Meinert, J., & Krämer, N. C. (2022). How the Expertise Heuristic Accelerates Decision-Making and Credibility Judgments in Social Media by Means of Effort Reduction. *PLoS One*, *17*(3), e0264428. https://doi.org/10.1371/journal.pone.0264428

Moen, A., Smørdal, O., & Sem, I. (2009). Web-Based Resources for Peer Support - Opportunities and Challenges. *Studies in Health Technology and Informatics*, *150*, 302–306.

Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A New Dimension of Health Care: Systematic Review of the Uses, Benefits, and Limitations of Social Media for Health Communication. *Journal of Medical Internet Research*, *15*(4), e85. https://doi.org/10.2196/jmir.1933

Murphy Kelly, S. (2023, July 20). *Teens are Using Social Media to Diagnose Themselves with ADHD, Autism and More. Parents are Alarmed*. https://edition.cnn.com/2023/07/20/tech/tiktok-self-diagnosis-mental-health-wellness/index.html

Najorka, K. (2021). *Social Media Nutzung in Deutschland 2021*. https://www.heise-regioconcept.de/social-media/social-media-nutzung-deutschland-unternehmenssicht

O'Brien, R. G., & Kaiser, M. K. (1985). Manova Method for Analyzing Repeated Measures Designs: An Extensive Primer. *Psychological Bulletin*, *97*(2), 316–333. https://doi.org/10.1037/0033-2909.97.2.316

Orizio, G., Schulz, P., Gasparotti, C., Caimi, L., & Gelatti, U. (2010). The World of E-Patients: A Content Analysis of Online Social Networks Focusing on Diseases. *Telemedicine Journal and E-Health : The Official Journal of the American Telemedicine Association*, *16*(10), 1060–1066. https://doi.org/10.1089/tmj.2010.0085

Picazo-Sánchez, L., Domínguez-Martín, R., & García-Marín, D. (2022). Health Promotion on Instagram: Descriptive-Correlational Study and Predictive Factors of Influencers' Content. *International Journal of Environmental Research and Public Health*, *19*(23). https://doi.org/10.3390/ijerph192315817

Qutteina, Y., Hallez, L., Mennes, N., Backer, C. de, & Smits, T. (2019). What Do Adolescents See on Social Media? A Diary Study of Food Marketing Images on Social Media. *Frontiers in Psychology*, *10*, 2637. https://doi.org/10.3389/fpsyg.2019.02637

Rabe, L. (2022, January 31). *Anzahl der Social-Media-Nutzer weltweit in den Jahren 2012 bis 2022: (in Milliarden)*. https://de.statista.com/statistik/daten/studie/739881/umfrage/monatlich-aktive-social-media-nutzer-weltweit/#:~:text=Im%20Januar%202022%20lag%20die,bei%20rund%201%2C48%20Milliarden.

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A Multidisciplinary Framework. *Annual Review of Information Science and Technology*, *41*(1), 307–364. https://doi.org/10.1002/aris.2007.1440410114

Rogers, A., Wilkinson, S., Downie, O., & Truby, H. (2022). Communication of Nutrition Information by Influencers on Social Media: A Scoping Review. *Health Promotion*

Journal of Australia: Official Journal of Australian Association of Health Promotion Professionals, *33*(3), 657–676. https://doi.org/10.1002/hpja.563

Rothfischer, K. (2021). Social Media - Key Opinion Leaders of the Future? *Journal of European CME*, *10*(1), 2014094. https://doi.org/10.1080/21614083.2021.2014094

Ruby, D. (2023, June 7). *Social Media Users 2023 — (Global Demographics)*. https://www.demandsage.com/social-media-users/#:~:text=The%20USA%20has%20302.35%20million,74.2%25%20of%20adults%20using%20it.

Sabbagh, C., Boyland, E., Hankey, C., & Parrett, A. (2020). Analysing Credibility of UK Social Media Influencers' Weight-Management Blogs: A Pilot Study. *International Journal of Environmental Research and Public Health*, *17*(23). https://doi.org/10.3390/ijerph17239022

Scanfeld, D., Scanfeld, V., & Larson, E. L. (2010). Dissemination of Health Information Through Social Networks: Twitter and Antibiotics. *American Journal of Infection Control*, *38*(3), 182–188. https://doi.org/10.1016/j.ajic.2009.11.004

Schumann, J. H., Wangenheim, F. von, & Groene, N. (2014). Targeted Online Advertising: Using Reciprocity Appeals to Increase Acceptance among Users of Free Web Services. *Journal of Marketing*, *78*(1), 59–75. https://doi.org/10.1509/jm.11.0316

Scott, B. A., & Judge, T. A. (2009). The Popularity Contest at Work: Who Wins, Why, and What do They Receive? *The Journal of Applied Psychology*, *94*(1), 20–33. https://doi.org/10.1037/a0012951

Smailhodzic, E., Hooijsma, W., Boonstra, A., & Langley, D. J. (2016). Social Media Use in Healthcare: A Systematic Review of Effects on Patients and on Their Relationship With Healthcare Professionals. *BMC Health Services Research*, *16*(1), 442. https://doi.org/10.1186/s12913-016-1691-0

Song, H., Omori, K., Kim, J., Tenzek, K. E., Morey Hawkins, J., Lin, W.-Y., Kim, Y.-C., & Jung, J.-Y. (2016). Trusting Social Media as a Source of Health Information: Online Surveys Comparing the United States, Korea, and Hong Kong. *Journal of Medical Internet Research*, *18*(3), e25. https://doi.org/10.2196/jmir.4193

Stewart Loane, S., & D'Alessandro, S. (2014). Empowered and Knowledgeable Health Consumers: The Impact of Online Support Groups on the Doctor–Patient Relationship. *Australasian Marketing Journal*, *22*(3), 238–245. https://doi.org/10.1016/j.ausmj.2014.08.007

Sugawara, Y., Narimatsu, H., Hozawa, A., Shao, L., Otani, K., & Fukao, A. (2012). Cancer Patients on Twitter: A Novel Patient Community on Social Media. *BMC Research Notes*, *5*(1), 699. https://doi.org/10.1186/1756-0500-5-699

Sumayyia, M. D., Al-Madaney, M. M., & Almousawi, F. H. (2019). Health Information on Social Media. Perceptions, Attitudes, and Practices of Patients and Their Companions. *Saudi Medical Journal*, *40*(12), 1294–1298. https://doi.org/10.15537/smj.2019.12.24682

Suter, T. (2023, August 15). *YouTube Announces New Policies to Target Medical Misinformation*. https://thehill.com/policy/technology/4153540-youtube-announces-new-policies-to-target-medical-misinformation/

Tian, L., An, R., Zhang, J [J.], Sun, Y., Zhao, R., & Liu, M. (2017). Effect of the Patient-to-Patient Communication Model on Dysphagia Caused by Total Laryngectomy. *The Journal of Laryngology and Otology*, *131*(3), 253–258. https://doi.org/10.1017/S0022215116009968

Triplett, N. T., Kingzette, A., Slivinski, L., & Niu, T. (2022). Ethics for Mental Health Influencers: MFTs as Public Social Media Personalities. *Contemporary Family Therapy*, *44*(2), 125–135. https://doi.org/10.1007/s10591-021-09632-3

van de Belt, T. H., Engelen, L. J. L. P. G., Berben, S. A. A., Teerenstra, S., Samsom, M., & Schoonhoven, L. (2013). Internet and Social Media for Health-Related Information and Communication in Health Care: Preferences of the Dutch General Population. *Journal of Medical Internet Research*, *15*(10), e220. https://doi.org/10.2196/jmir.2607

Vassallo, A., Jones, A., & Freeman, B. (2022). Social Media: Frenemy of Public Health? *Public Health Nutrition*, *25*(1), 61–64. https://doi.org/10.1017/S136898002100269X

Veirman, M. de, Cauberghe, V., & Hudders, L. (2017). Marketing Through Instagram Influencers: The Impact of Number of Followers and Product Divergence on Brand Attitude. *International Journal of Advertising*, *36*(5), 798–828. https://doi.org/10.1080/02650487.2017.1348035

Wathen, C. N., & Burkell, J. (2002). Believe it or Not: Factors Influencing Credibility on the Web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134–144. https://doi.org/10.1002/asi.10016

White, E., & Hanley, T. (2023). Therapist + Social Media = Mental Health Influencer? Considering the Research Focusing Upon Key Ethical Issues Around the Use of Social Media by Therapists. *Counselling and Psychotherapy Research*, *23*(1), 1–5. https://doi.org/10.1002/capr.12577

Willis, E., & Delbaere, M. (2022). Patient Influencers: The Next Frontier in Direct-to-Consumer Pharmaceutical Marketing. *Journal of Medical Internet Research*, *24*(3), e29422. https://doi.org/10.2196/29422

Willis, E., Friedel, K., Heisten, M., Pickett, M., & Bhowmick, A. (2023). Communicating Health Literacy on Prescription Medications on Social Media: In-depth Interviews With "Patient Influencers". *Journal of Medical Internet Research*, *25*, e41867. https://doi.org/10.2196/41867

Yeung, A., Ng, E., & Abi-Jaoude, E. (2022). Tiktok and Attention-Deficit/Hyperactivity Disorder: A Cross-Sectional Study of Social Media Content Quality. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, *67*(12), 899–906. https://doi.org/10.1177/07067437221082854

Yoo, C. Y. (2014). Branding Potentials of Keyword Search Ads: The Effects of Ad Rankings on Brand Recognition and Evaluations. *Journal of Advertising*, *43*(1), 85–99. http://www.jstor.org/stable/24749872

Zhang, Y., He, D., & Sang, Y. (2013). Facebook as a Platform for Health Information and Communication: A Case Study of a Diabetes Group. *Journal of Medical Systems*, *37*(3), 9942. https://doi.org/10.1007/s10916-013-9942-7

Zhao, Y., & Zhang, J [Jin] (2017). Consumer Health Information Seeking in Social Media: A Literature Review. *Health Information and Libraries Journal*, *34*(4), 268–283. https://doi.org/10.1111/hir.12192

Sandra Baringhorst | Manfred Schwaiger | Louisa Weritz

III Investigating the Suitability of Customized Coupons for Personalized Pricing

Abstract

This study explores personalized pricing, a promising yet controversial strategy for revenue increase. While it promises financial benefits, concerns about negative customer reactions have deterred many companies. The study specifically investigates the impact of personalized coupons on different consumer reactions. Study 1 is based on a self-assessment and uses an innovative approach with the help of the van Westendorp method. The main aim is to find out whether there is a tolerable range for customers in which coupon value differences do not lead to negative consumer reactions. However, Study 2 challenges the results of Study 1, revealing negative reactions even within the defined tolerated range. Interestingly, despite dissatisfaction, customers rarely voice complaints publicly, and loyalty remains unaffected for most scenarios. The study suggests that if long-term loyalty persists, the perceived loss of intangible assets may be less significant than anticipated. In conclusion, personalized coupons as the only framing method used in this paper cannot offset the negative effects of personalized pricing. Accordingly, it can be assumed that perceived fairness can predominantly be achieved if every customer is offered an identical price for the same product. Any personal deviations from the standard price trigger perceived unfairness.

1 Motivation

Personalization is one of the current top-notch topics in marketing. Standing out but still being part of the crowd seems to be a common shopping mantra for customers (Chandra et al., 2022). Hence, an important marketing practice is customer targeting for differential promotional activities (Rossi et al., 1996). According to the Cambridge Dictionary personalization is "the process of making something suitable for the needs of a particular person". In this respect, personalization is a way to acknowledge the uniqueness of each customer by satisfying them with products, services, content or prices tailored to their preferences (Liang et al., 2006; Shiller, 2014; Surprenant & Solomon, 1987; Yun & Hanson, 2020).

Muji, an internationally well-known Japanese lifestyle brand, implemented personalized promotions successfully by collecting data on customers' online browsing and in-store purchase history. In addition to using in-app push notifications, Muji issued personalized coupons and achieved a 100% increase in coupon redemption as well as a 46% increase in in-store sales over a two-year period (Treasure Data, 2021). Success stories like these, sound promising but are still scarce. However, they give reason to investigate whether customized coupons are a potential successful implementation method for personalized pricing.

The focus of this study is therefore on the personalization of prices[1]. In practice, this means that a minimum of two customers are offered the same product at the same time by the same supplier at different prices. The underlying concept behind personalized pricing is to skim off the customer's maximum willingness to pay (WTP), which is often referred to as first-degree price discrimination (Pigou, 1920). The use of customized pricing is still a controversial topic due to two conflicting strands of literature. On the one hand, proponents have addressed the need for, influence, and benefits of personalized pricing (Choe et al., 2022; Elmaghraby & Keskinocak, 2003; Haws & Bearden, 2006; Kung et al., 2002). From a business perspective, the application of personalized pricing is promising as it can contribute to significant profit increases and revenue maximization. Various studies confirm that personalized pricing through the use of detailed customer data can potentially increase profits by between 10% to 50% (Dong et al., 2009; Dubb & Misra, 2017; Kung et al., 2002; Sahay, 2007; Shiller, 2020; Smith et al., 2023). On the other hand, behavioral researchers have expressed doubts about the predicted success of customized prices because negative consumer reactions such as lower repurchase intentions, satisfaction, loyalty, perceived fairness as well as lower trust in retailers represent negative

[1] Personalized, customized, or tailored pricing are used equivalently.

consequences of personalized pricing (Garbarino & Lee, 2003; Grewal et al., 2004; Haws & Bearden, 2006; Hufnagel et al., 2022). Organizations are thus faced with the dilemma between foregoing potential profit growth or accepting negative customer reactions. Because of this dichotomy, potential mitigating effects should be explored that temper or eliminate negative customer reactions to customized prices so that the benefit of increased profit can still be realized.

One potential approach would be to present prices differently. As Krishna et al. (2002) have already shown, different price presentations (e.g., free gift framing led to the greatest impact on customers' perceived savings) can influence the customer's evaluation of the offer. In line with this, early research findings show that promotions of the same magnitude that are designed differently (e.g., display the same discount value either as a percentage-off or as a dollar-off coupon) are also perceived individually in terms of gains or reduced losses (Diamond & Campbell, 1989). In this context, and particularly important for this study, we assume that price promotions (such as coupons or vouchers[2]) are more likely to elicit positive perceptions of price differences compared to uniform transactions. Weisstein et al. (2013) are the first ones who addressed this issue by examining tactical ways for e-retailers to mitigate consumers' negative reactions to personalized pricing. They show that consumer reactions are enhanced as the level of perceived transaction dissimilarity increases through the use of different price framing strategies compared to no framing. However, one of their findings reveals that providing two customers with the same framing format (customer 1 gets 10% discount and customer 2 gets 20% discount) does not lead to a decrease in negative customer reactions. As it is nearly impossible to provide every customer a different framing format, we would like to further elaborate on this finding to analyze under which circumstances one and the same framing format can lead to positive or more neutral customer reactions.

In this context, it is important to note that coupons seem to be a particularly suitable promotion method because they are generally perceived positively (Chiou-Wei & Inman, 2008; Diamond & Sanyal, 1990; Park & Gomez, 2004) and constitute a gift. Thus, we assume that they have a positive effect on customers a priori. The results of Inman et al. (1990) support this assumption by demonstrating that the mere signal of a price reduction enhances consumer responses without requiring an actual price reduction.

So even if literature is scarce on results concerning personalized coupons and varying percentage values, we can find the two just presented and opposing results. Inman et al. (1990) predict

[2] Coupons and vouchers are used equivalently.

that any coupon that is perceived as a present will lead to positive consumer reactions. These findings counter Weisstein et al.'s (2013) results, which demonstrate that a particular difference of 10% and 20% coupons leads to perceived unfairness and has no attenuating effect on negative consumer reactions. Presumably, fairness is an absolute concept, which in our case is only achieved if there are no price deviations for the same product. In other words, the voucher values should also be identical. However, what is interesting and conceivable with regard to the contrasting results of the two studies is that there might be a so-called indifference interval. This interval or tolerated range describes the difference in percentage points between my voucher and that of another customer (i.e., a friend), which according to self-assessment is perceived as somewhat fair. Outside this tolerated range, vouchers would again be perceived as unfair, despite their gift character, and thus tend to have an adverse impact on customers. The van Westendorp method has so far been used to determine the optimum price for a product from the customer's point of view by indirectly trying to determine the respective WTP. The ultimate result of this method calculates a price range within which a minimum and maximum price is obtained and an optimum price is also calculated (van Westendorp, 1976). After modifying the methodological questions to our case, the van Westendorp method is used to determine whether there is a tolerated range for coupon differences and thus answer the first research question for this paper:

> ***RQ1:*** *Is there a tolerated range between voucher values that is perceived as somewhat fair by the customer?*

While the first study attempts to define a tolerated range descriptively - based on self-assessments - the second study aims to capture how consumers behave within and outside this range in a less theoretical and more practice-oriented scenario. Different constructs from psychology (i.e., information processing as proxy for cognitive effort and need for cognition) play a crucial role. Both contribute to this study by explaining how consumers process information during price promotions in more detail. Of particular interest here is how a perceived "need to think" (Shugan, 1980) or a possible relaxation through the use of heuristics (DelVecchio, 2005) can change the corresponding customer reactions to personalized coupons within and outside the originally self-assessed tolerated range.

Referring back to our first research question, consumers would be expected to determine a particular range that is somewhat fair to them (Study 1), which contains a percentage point range in which the deviation of coupon values is perceived tolerated. Study 2, on the other hand, is intended to examine whether the theoretically determined tolerated coupon value range is valid in practice, i.e., whether or not there is a discrepancy between theoretical assessment and applied practice.

The need to test this range in a more practical scenario stems from the assumption that if customers suddenly find themselves in a real situation, heuristic approaches are often used to assess price fairness and price differences (Grewal et al., 1996). Customers thus want to minimize the cognitive effort required to draw conclusions. In their study Ozanne et al. (1992) found that information discrepancies between products (e.g., price discrepancies) influence customers' willingness to process information. According to their results, information processing is greatest when the discrepancy is moderate. In other words, if the discrepancy is either very small or large, the consumer feels no need to further analyze the reasons for this discrepancy, so the cognitive effort is kept to a minimum and the difference is accepted.

Transferring these findings to price promotions, it can be assumed that percentage coupon differences within the tolerated percentage range determined in Study 1 tend to lead to greater cognitive effort and perceived unfairness in reality because the customer wants to understand the price difference. It is therefore assumed that the percentage point difference perceived as somewhat fair in Study 1 would no longer be perceived as fair in a realistic scenario because the customer would be prompted to think about it. While in Study 1, the focus is on the percentage difference, and consumers may have assessed themselves more generously, in Study 2 consumers now find themselves in an experiment where the focus is less on the percentage difference and more on the product purchase. In such a situation, the percentage differences appear to have a different effect on the consumer. The second research question of our study is, therefore, as follows:

> ***RQ2:*** *How do customers perceive voucher differences within and outside the previously determined tolerated coupon range?*

With this paper, we would like to follow the call for research and investigate potential benefits of personalized coupons (Keller et al., 2022) as a possible strategy to implement personalized pricing (Weisstein et al., 2013). Specifically, we are interested in identifying whether there is a range between two coupon percentages that consumers perceive as more positive (or negative) and thus affect customer reactions to personalized pricing via tailored coupons. The resulting implication for management would be a guideline and a range in which personalized coupons should be allocated at best and which differences should be avoided. Theoretically, we would like to emphasize that theoretical and self-assessed considerations often do not coincide with experienced and realistic situations.

In the following, we would like to address the two research questions by first providing a theoretical introduction to the applied constructs. Second, we address Studies 1 and 2 by explaining the respective methodology, analysis, and results. Third, we discuss the findings, including implications for practice and theory. And finally, the limitations of the study are identified and suggestions for future research are derived.

2 Literature Review

Our research is primarily concerned with personalized marketing and draws on three areas of literature: personalized pricing, personalized coupons, and individual cognitive effort. Below, we discuss each of these areas and highlight our respective contributions.

2.1 Price Discrimination and Customer Reactions

In principle, there are two overarching options for setting prices. One approach is to set prices statically, so that every customer pays the same price for the same product. Another approach is to set prices dynamically, so that not every customer is offered the same price for the same product due to different types of price discrimination. In theory and according to Pigou (1920), there are three classic forms of price discrimination:

Third-degree price discrimination means that customers are divided into different segments based on certain characteristics and therefore cannot decide for themselves which segment they want to belong to. Each segment pays different prices for one and the same product. However, a constant price is paid for each product unit. Examples of this are prices for students, pensioners, etc.

Second-degree price discrimination or nonlinear pricing occurs when companies allow customers to self-select. Among others, unit prices could either change by the number of products (volume discounts as an example), by product bundling, or by offering different product qualities. So unlike in third-degree price discrimination, where prices differ across segments (i.e., across customers), here the prices differ across units but not across customers.

First-degree price discrimination is also referred to as perfect price discrimination and specifies the case when each customer is offered an individual price for the same product that ideally corresponds to the customer's maximum individual WTP.

The main reason for price discrimination is the potential increase in profits by segmenting customers according to their demand sensitivity and a corresponding price adjustment. The basis for any form of segmentation is customer heterogeneity. Heterogeneity in the pricing context refers primarily to differences in the individual WTP. In some cases, heterogeneity can be observed directly and a company can base its pricing on contractually defined consumer characteristics (e.g., students, retirees) (Varian, 1989). In other cases, heterogeneity is not directly noticeable, but may be indirectly induced by offering product and price menus, allowing consumers to self-select. In both cases, the company aims to set the price of its goods in line with the underlying demand elasticity of individual consumers in order to generate more customer surplus and increase sales from more elastic customers (Stole, 2007). Consumer surplus is usually defined for an individual customer as the difference between one's WTP and the price charged by the company (N. Chen & Gallego, 2019; Varian, 2014) and in general, customers only purchase a product if their WTP is greater than or equal to the product's price (Dhebar & Oren, 1985). According to Stigler (1950) a company discriminates through price if the ratio of prices differs from the ratio of marginal costs for two goods offered by a company. In other words, when the difference in price for the same good cannot be entirely explained by the variations in marginal costs price discrimination is applied.

While second- and third-degree price discrimination has been common practice for many years, perfect price discrimination has been more of a theoretical ideal taught in textbooks. However, major advances in data collection and analysis have made first-degree price discrimination a viable prospect (Baik et al., 2023; Esteves & Resende, 2019). More specifically, perfect price discrimination in the online context is about algorithmic pricing. Now various types of analysis and data collection make it possible to change prices in real-time. Although second- and third-degree price discrimination are already the first forms of dynamic pricing, "modern" dynamic pricing involves not only the static setting of segment or unit prices under certain conditions, but also the inclusion of, for example, supply and demand to discriminate consumers on the basis of their WTP. In this paper, we refer to a specific form of dynamic pricing, namely personalized pricing (Seele et al., 2021). The meaning and differences of these two forms of pricing strategies are explained in the following.

Modern Dynamic Pricing

In a first step moving away from uniform pricing, prices were automatically adjusted to supply and demand or time constraint (M. Chen & Chen, 2015). This method is also known as dynamic pricing or yield management (Seele et al., 2021) and is familiar to us primarily from the tourism industry (N. Chen & Gallego, 2019; McAfee & Te Velde, 2006). Here, the airfare or hotel price changes in particular due to time constraints and respective demand uncertainties (N. Chen & Gallego, 2019). The shorter the booking period before the vacation, the more expensive the price becomes. But also the increasing demand and/or small number of remaining seats or hotel rooms leads to higher prices (N. Chen & Gallego, 2019). However, it is important that the price at a particular time is the same for two customers under the same conditions[3] and that an individual price is not offered to each customer.

There are two prerequisites for the successful implementation of dynamic pricing, which are not mandatory but provide immediate benefits. First, the product expires at a specific point in time, such as hotel rooms, flights, or time-limited products ("sell before"). Second, capacity is predetermined and can only be expanded at relatively high marginal cost. These requirements mean that the opportunity costs of a sale can widely vary, as the opportunity cost of a sale represents a potential foregone opportunity to sell at a later date (McAfee & Te Velde, 2006).

From a technical perspective dynamic pricing depicts the first generation of algorithmic pricing requiring rather adaptive algorithms (Calvano et al., 2019). Those are less complex algorithms considering for example simple if-then procedures (e.g., if only 20% of the hotel rooms are left, price goes up by 15%).

Personalized Pricing

Personalized pricing is a further development of the so-called first generation and thus a specific type of dynamic pricing (Calvano et al., 2019; Seele et al., 2021). First of all, to illustrate the technical differences, personalized pricing belongs to the second generation of algorithmic pricing, and requires the implementation of learning algorithms (Calvano et al., 2019). Those are algorithms high in complexity using artificial intelligence (AI) and machine learning (Calvano et al., 2019; Elmachtoub et al., 2021). By generating input, these algorithms evolve over time and become more accurate in their price determination (Calvano et al., 2019; F. Xia et al., 2019). Specifically for personalized pricing, these algorithms process input data about markets and players, taking into account numerous factors such as competitor prices, consumer demand, personal demand, geographic location, device used, individual purchase behavior, and characteristics to predict an individual customer's WTP and set the output price in relation to the highest (i.e., profit-maximizing) revenue achievable (Cohen, 2018; Fisher et al., 2018; Keskin & Zeevi, 2014; F. Xia et al., 2019; Juanjuan Zhang, 2011).

The major difference to dynamic pricing is that in personalized pricing, two or more customers are offered the same product at the same time by the same supplier at different prices because a large number of the individual factors just mentioned are analyzed and used to set the individual price. Based on this method, companies can ideally skim the entire consumer surplus

[3] Conditions: same supplier, same product, same time

(Esteves & Shuai, 2022), depending on the accuracy of the algorithm (Wang et al., 2023). Thus, personalized pricing leads to profit maximization and revenue increase (N. Chen & Gallego, 2019; Kallus & Zhou, 2021; Wang et al., 2023) and is therefore considered a promising and profitable marketing strategy.

However, from a corporate perspective, this promising pricing strategy is challenged by several factors. Implementing any form of price discrimination, especially perfect price discrimination, is costly and difficult (Elmachtoub et al., 2021). Information systems for data storage and analysis must be provided, algorithms must be built, and the relevant customer variables for each company must be identified (Arora et al., 2008). However, the resulting costs can be considered an industry problem which may be offset and outdated by the advances in technology (Rossi et al., 1996) and potential increases in profits.

The more severe disadvantage of customized pricing is the negative impact on customers and the associated reactions and perceptions. Considering that price discrimination hurts social welfare unless the total output increases after price discrimination (Varian, 1985), and that personalized pricing is the ultimate form of price discrimination, it is reasonable that consumers react with resentment. Price fairness seems to play a central role in the evaluation of customized prices, as it functions as a driver for the acceptance of dynamic and personalized prices and can cause immediate changes in consumer behavior (Dickson & Kalapurakal, 1994).

Perceived price fairness is the customer's *"assessment and associated feelings about whether the difference (or lack of difference) between a seller's price and another comparator's price is reasonable, acceptable, or justified"* (L. Xia et al., 2004, p. 3). Therefore, the reference price is, by definition, an important criterion for assessing price fairness and may include comparisons with prices previously paid (Campbell, 2007), with competitors' prices (Gourville & Moon, 2004), and/or with other customers' prices (Khandeparkar et al., 2020).

Although the literature on perceived price fairness can be divided into drivers (L. E. Bolton et al., 2003; Darke & Dahl, 2003) and outcomes of perceived price (un-)fairness , we focus on the outcomes (Malc et al., 2016; L. Xia et al., 2004). Regarding the drivers, for this study it is only important to note that there is a negative effect of personalized prices on perceived price fairness for both preferred and disadvantaged customers (Hufnagel et al., 2022). The perceived unfairness of customized prices leads, among other things, to a lower repurchase intention, less benevolence trust, reduced customer satisfaction, as well as an increased willingness to spread negative word-of-mouth (WOM) privately and publicly and to increasingly search for potential purchase alternatives (Garbarino & Maxwell, 2010; Haws & Bearden, 2006; Lii & Sy, 2009). In summary, it becomes apparent that customers, whether positively or negatively affected by personalized pricing, react negatively in all respects (Hufnagel et al., 2022). To avoid potential losses from dissatisfied customers and to enable companies to use personalized pricing profitably, some researchers have already looked at potential mitigating or reversing effects.

A recent study investigated whether personalized pricing has a different effect on consumer responses depending on the type of product purchased (hedonic versus functional products). The results showed that functional products (e.g., toolbox, kitchen aid machines, etc.) tended to increase the negative effect and were therefore not suitable for personalized pricing. A slight positive effect on perceived fairness was visible for hedonic products, but the result was not significant (Weritz, 2023). Furthermore multiple studies considered the customer characteristics (i.e., loyal versus new customers), to identify whether those make a particular difference and whether personalized pricing is rather accepted or rejected (L. E. Bolton et al., 2010; Darke & Dahl, 2003; Maxwell & Garbarino, 2010). However, since the characteristics of consumers cannot be influenced by the company but are at best decisive for the selection of a certain consumer group, this study takes a more managerial view.

Studies that have predominantly looked at transaction dissimilarity (i.e., differences between purchasing processes) to alleviate the negative effects on customers show rather promising findings, which can also be directly influenced and controlled by the organization. One promising approach is to enhance transaction dissimilarity through (dynamic) product bundling as it reduces the consumers' perceived unfairness (W. Li et al., 2018). Furthermore, product differentiation in terms of product line depicts a second opportunity to create dissimilarity and shows that personalized pricing with competing firms can lead to an overall increase in consumer welfare (Choudhary et al., 2005). However, both approaches are only suitable for companies that either have the possibility to offer product bundles or to create quality differences and therefore, neither approach is generally applicable. Weisstein et al. (2013) nonetheless, have specifically looked at different personalized price framings and the results reveal that depending on the framing type, the negative reactions of disadvantaged customers were attenuated. Similar to the aforementioned results, it can be seen that the perception of trust, price fairness, and repurchase intentions (of disadvantaged customers) increases with the degree of transaction dissimilarity. Furthermore, a recent study revealed how the mere display of price discounts as a dynamic pricing strategy diminishes negative reactions such as customers' repurchase intention and perceived price (un-)fairness (Keller et al., 2022).

Table 1. Overview of Different Papers on Transaction Dissimilarity[4]

Author(s)		**Choudhary et al. (2005)**	**Weisstein et al. (2013)**	**W. Li et al. (2018)**	**Keller et al. (2022)**	**This paper**
Dynamic Pricing	**Between customers**	✓	✓	✓		✓
	Over time				✓	
Framing Type		Quality differences	%-off, $-off coupons, free gift and gift card	Product bundling	Daily price adjustments vs. static pricing, and %-off	%-off coupons
Model Variables	**Independent Variable**	Personalized pricing implementation	Price framing types (%-off, $-off coupons, free gift, and gift card)	Dynamic bundling vs. dynamic pricing	Price adjustment (dynamic vs. static)	Voucher ranges
	Dependent Variable	Firm profits, consumer welfare	Repurchase intention	Price fairness perception	Purchase intention	Satisfaction, repurchase intention, complaint intentions, loyalty
	Mediator		Perceived transaction dissimilarity perceived fairness, trust	Perceived transaction dissimilarity, comparison intentions	Pricing transparency, price fairness, perceived value	Information processing, perceived fairness
	Moderator	Product quality differentiation (high quality vs. low quality)	Product price level, customer segments, framing formats		Pricing discovery, price display	Need for cognition
Main Result(s)		The model demonstrates that consumers would benefit if higher quality firms adopt personalized pricing In the event that all firms adopt personalized pricing, consumers would benefit the most.	Price-framing tactics can make price-disadvantaged consumers view similar transactions differently. As perceived transaction dissimilarity grows, positive customer reactions increase.	Bundling not only mitigates the negative impact of dynamic pricing on perceived fairness, but also results in fairness perceptions similar to those aroused by fixed pricing.	If retailers display prices as sufficiently high discounts, they can mitigate these negative reactions. Discount displays equal to or greater than 10% are most effective.	Via self-assessment, consumers can define a fairness range for different coupon values. However, in practice neither within nor out of this range coupon personalization is perceived as fair.
Main Limitation		Only a single product by each firm considered, whereas in practice firms often offer multiple products.	It is impossible to provide each customer with a unique framing format. Only one example of the same framing format was tested, without considering a range.	Bundling is not practicable for every company. Only lab-based experiments were used to assess the effectiveness of dynamic bundling.	Their main focus lies on dynamic pricing, i.e., rather on timely price changes than on personalized prices.	Using vignettes instead of a field experiment only allows assumptions about intentions but no conclusion on actual behavior.

[4] Only the most relevant results for the present study are summarized.

Table 1 presents an overview of the different papers. Each of the four studies just mentioned that looked at transaction dissimilarity identified different framing options to reduce consumers' negative reactions to personalized pricing. In this study, we would like to address the content of these studies and pursue the call for research expressed by several researchers to examine how a particular pricing method (in this case, customized coupons) used to implement personalized pricing affects customer responses (Dubé & Misra, 2023; Keller et al., 2022; Neubert, 2022; Weisstein et al., 2013). We are thus complementing previous research with further insights into potential strategies for implementing personalized prices that can be steered by the company and, in the best case, have a positive influence on customer reactions. Theoretical background to existing research on framing and in particular personalized vouchers is provided in the following chapter.

2.2 Couponing as a Potential Framing Method

In the marketing literature it has been known for many years that price plays a decisive role when it comes to customer choice (Dodds et al., 1991). Using different price presentations, also called price framing, is a suitable method for companies to influence the transaction value as perceived by the customer (Manning & Sprott, 2009). Thereby, companies are able to also influence customer choices (Kahneman & Tversky, 1984). The concept of price framing was first brought to attention by Kahneman and Tversky's (1979, 1984) *Prospect Theory*. Broadly speaking, the two scholars found that framing affects people's judgments and thus influences their choices and reactions. Depending on whether offers are framed as a win or a loss, the customer's willingness to take risks and the overall perception of the deal will change. More specifically, in terms of general price offers and this study's content, it can be summarized that the correct framing of price promotions could have a positive influence on customer perceptions regarding deal savings, purchase decisions and transaction value (S. S. Chen et al., 1998; Darke & Chung, 2005).

Price framing plays an important role in personalized marketing, among other strategies through the implementation of personalized coupons (NCH Marketing Services, 2019). The reason why price framing could be an interesting strategy for customized pricing is because the framing effect occurs when the displayed price information causes people to neglect the item's base value (H. Chen et al., 2012). Research has shown that customers tend to ignore the actual base value (e.g., the product's price) to which the percentage applies and focus only on the displayed percentage (M. Li & Chapman, 2013). Thus, the attention could be diverted from the actual price and the underlying price strategy and accordingly, the negatively associated customer reactions might be avoided.

Along with the development and spread of customized vouchers in practice, for example through loyalty programs in brick-and-mortar stores (J. Zhang & Wedel, 2009), personalized coupons have also aroused growing interest in the marketing literature (Rossi et al., 1996; Shaffer & Zhang, 1995; J. Zhang & Krishnamurthi, 2004). The literature on personalized coupons can be divided into three strands, which are briefly summarized below with their most relevant results for this study. The three strands refer to model creation for personalized coupons, research that has taken place in brick-and-mortar retail, or research that focuses on online retail. In some cases, these strands overlap by first creating a model and then testing it using data from either brick-and-mortar or online retail, or by comparing the online and offline markets.

Development of Coupon Optimization Models

Many scholars have focused on coupon optimization and model design to measure the benefits of personalized vouchers. Rossi et al. (1996), were the first to empirically quantify the benefits of adopting personalized coupons by developing a brand choice model. They particularly emphasized the importance of information on the purchase histories of individual households for optimizing coupons. Later, the same method was validated and a temporal dimension was added to the personalization of coupons (Johnson et al., 2013). Furthermore, focusing on the profit potential of customized promotions, a joint model of purchase incidence, choice, and quantity was analyzed by investigating different optimization procedures in online and offline stores. The results demonstrate, that optimization procedures enhance the organizational profit and customized promotions are especially suitable for promotion sensitive product categories (J. Zhang & Wedel, 2009). Recently a very modern approach has been the development of a product choice model that predicts the influence of personalized coupons on the purchase behavior and purchase probability of customers of a large retailer (Gabel & Timoshenko, 2022). The model predicts how customer-specific purchase probabilities change in response to a retailer's marketing efforts, providing input for model-based recommender systems. In an empirical study using experimental data from a leading German grocery retailer, they were able to confirm the effectiveness of their model by demonstrating that coupon optimization methods achieve significantly higher revenue gains.

Coupon Research Focusing on Brick-and-Mortar Stores

In addition to technical components and optimization problems, another group of researchers looked at personalized coupons in the context of brick-and-mortar retailing, focusing on practical implications in this regard. There are different approaches to how retailers implement customized promotional programs in practice. Some do so with the help of in-house experts, while others rely on consumer marketing companies to execute the promotions. Catalina Marketing, for example, is one of the industry leaders in targeted marketing services for retailers and is most commonly established in the grocery industry (Venkatesan & Leliveld, 2009; J. Zhang & Wedel, 2009). The grocery industry was the first to implement personalized vouchers on large scale, which is why much of the research regarding brick-and-mortar stores has used the grocery industry as a data source and context. Using a game-theoretical framework and working with panel data on household purchase behavior Shaffer and Zhang (1995) examined the impact of customized vouchers in a perfectly competitive environment. They highlighted that if competing companies target their coupon promotions to brand switchers, firms necessarily lose profit because regular prices do not increase. Thus, brand switchers should not be targeted through personalized coupons. Further, the grocery industry has been used to demonstrate how information on household purchase history can be used to offer targeted coupons enhancing organizational profits (Rossi et al., 1996). In particular, the importance of purchase history in estimating targeted coupons has been emphasized, as considering just one purchase history observation per customer increases net coupon revenue by 50% more than an untargeted coupon strategy. Smith et al. (2023) were able to generalize the results from Rossi et al. (1996) using supermarket scanner data. They show that demographic data is less useful compared to data on customers' past purchasing behavior when it comes to generating effective input for profitable pricing strategies. In addition, data from a group of regional grocery retailers was analyzed in terms of profit and campaign returns, and the findings of a quasi-experiment indicated that both the exposure and redemption of personalized coupons have a positive impact on customer purchases (Venkatesan & Farris, 2012). Surprisingly, mere exposure to personalized coupons contributed more to campaign returns than coupon redemption. Thus, the mere presentation of personalized coupons could have a positive effect on consumer behavior. In addition, personalized coupons were found to be more effective the higher the discount, the more unexpected (from the customer's perspective), and if they are framed in such a way that they are specifically selected for the individual customers and tailored to their preferences. J. Zhang and Wedel (2009) were then among the first to compare the effectiveness of personalized promotions at different levels of granularity (i.e., mass market, segment specific, and individual) between online and offline stores. Among other things, they show that tailored promotions at all levels are more profitable in online stores than undifferentiated promotions, which is not the case in offline stores. One reason for this could be the particularly low redemption rate, which hinders the success of tailored promotions in offline stores. It can be concluded that personalized promotions are more successful online, which is why e-retailers in particular are recommended to practice personalized couponing.

E-Retail (Online) Coupons

The latter finding leads to the last part of the literature, namely personalized coupons in the online context, which often deals with recommendations for e-retailers. Analyzing consumer behavior, the literature indicates that gains from personalized coupons should not be measured solely in terms of their redemption (Sahni et al., 2017). Using several experiments and data from an online ticket resale platform, they found that the offer of targeted coupons increased consumers' spending behavior by 37.2%. However, this increase initiated by the discounts offered is not only caused by the consumers who made use of the targeted coupons as about 90% of the increase is independent of coupon redemption. A so-called spill-over effect sets in, influencing the purchase behavior even after the coupon has expired by reminding customers of other products, in this case tickets, which would be available for purchase. From an organizational perspective, sales increases can therefore be expected not only from the redemption of vouchers, but also from the customers’ purchasing behavior stimulated by the voucher as a reminder. Another study of consumer behavior focuses on how non-recipients respond to tailored promotions offered to other customers (Feinberg et al., 2002). The betrayal effect shows that consumers have a lower preference for their favored company when the latter offers a tailored promotion to switchers. In addition, the jealousy effect shows that customers prefer their favorite company less when another company offers price reductions to its loyal customers. Overall, they propose that consumers' preference for a particular firm is influenced not only by the price they receive themselves, but also by the prices available to others. This paper is one of the first to move from a conventional customer rationality approach to a more behaviorist approach with respect to personalized promotions. A complementary approach is that of Barone and Roy (2010), who investigate whether the response of recipients of personalized promotions depends on their perception of the exclusivity of the offer (i.e., is the offer available only to me or also to others). They find that particularly male customers and those with an independent self-view (i.e., people who see themselves as independent of others) are the ones favoring personalized promotions over universal promotions. By contrast, female customers and those with interdependent construal (i.e., belonging to a group) prefer universal rather than customized offers. In summary, Barone and Roy (2010) encourage marketers to respond to customer characteristics and to personalize or refrain from personalizing offers accordingly. In this case, the success of personalized coupons is tied to customer characteristics and the marketer has little leeway to create differences in consumer response, e.g., through the design or type of personalized promotion.

Relevant Insights Into the Impact of Coupons in a Dynamic (Personalized) Pricing Strategy

Two studies that have particularly focused on price framing methods in the online dynamic (personalized) pricing context are the studies by Weisstein et al. (2013) (personalized pricing) and Keller et al. (2022) (dynamic pricing). Both studies deal with customer reactions to different framing methods, which can be determined by the company. Weisstein et al. (2013) conducted a very thorough study on different framing methods (i.e., dollar off, percentage off, free gift, and gift card). They focus mainly on the presentation and examination of framing versus no framing effects. In the first two studies, the participant is always disadvantaged (i.e., receiving a lower discount than a friend) and receives a framed prize that should be compared to a non-framed prize of a friend. Here the results are consistent and show that despite the fact that the participant is disadvantaged the degree of perceived transaction dissimilarity mediates the positive effect of price framing (versus no framing) on perceived price fairness, trust, and repurchase intentions. Furthermore, moderation effects of product price level (high versus low price), customer segment (loyal versus new customers) and framing formats (i.e., dollar off, percentage off, free gift, and gift card) were confirmed. Similar to earlier findings by S. S. Chen et al. (1998) and Gendall et al. (2006), who showed that there is a difference in how dollar-off and percentage-off framing formats influence customer perceptions, Weisstein et al.'s (2013) results show that for low-priced products, the percentage-off framing strategy was more effective, while for high-priced products, the dollar-off framing format was more effective. Moreover, it was also found that for loyal customers, the percentage-off framing format was more effective in increasing perceived fairness and repurchase intention than the dollar-off method. Only in the last study do the disadvantaged participants compare their individual offer either with a friend's offer in the same price format ($1050 + 10% off versus $1050 + 20% off) or with a different format ($1050 + 10% off versus $1200 + $360 off). When the two transactions were in the same framing formats, there was no discernible weakening of negative consumer reactions (Weisstein et al., 2013). The fact that no mitigating effects were found in this case may be due, among other things, to the fact that it was previously established that percentage-off coupons are better suited to low-priced products and the study was tested here on high-priced products, and possibly also to the fact that only one percentage point range was tested, namely that between 10% - 20%.

The second study that analyzed a specific price framing method (i.e., percentage-off) was conducted by Keller et al. (2022). It measures consumer reactions to the display of discounts compared to a non-discounted reference price without the customers having to buy the respective product. First, they show that even in the absence of an explicit reference transaction dynamic pricing evokes negative customer reactions. Second, their results reveal that if companies display prices as sufficiently high discounts (higher or equal to 10%), they can alleviate these negative reactions such as repurchase intention and perceived price (un-)fairness. The effect is mediated by transparency as various information were displayed to the participant (discount value, discounted price, original price, and the information whether prices change over time or are maintained static). However, the study by Keller et al. (2022) was conducted in the context of dynamic pricing, so that discounts do not change individually but generally over time. Although the present study focuses on customized promotions, Keller et al. (2022), create a broader understanding of pricing strategies in the context of dynamic pricing, of which personalized pricing is a part. Moreover, they find that merely displaying a discount provides additional price information that is perceived positively tempering negative customer reactions. Their results support the idea that percentage-off framing formats for personalized coupons could have a positive effect on customers' perception of tailored prices.

Implications from Existing Literature for the Underlying Study

However, the positive effect of a voucher (or promotion in general) assumed at the outset with regard to its gift character need not apply universally. If a personalized pricing strategy is adopted by the company, it is almost impossible to offer each customer an individualized price and to ensure that customers do not have a peer in their environment who receives the same framing format as an offer. Thus, it would be possible that customers compare different discount sizes of the same framing type (e.g., only percentage-off sizes) with each other. The risk here, as equity frameworks (Adams, 1965; G. E. Bolton & Ockenfels, 2000; Greenberg, 1986) suggest, is that people make interpersonal comparisons that do not only take the outcomes they receive (nonsocial utility) into account but also how those outcomes compare to those of others (social utility). This would mean that some consumers might be offended if they receive a smaller discount than others (Feinberg et al., 2002).

In contrast, the coupon valuation of customers may also be different if one considers strategic customers in particular, as many scholars do, such as Su (2007), Y. Chen et al. (2019), Aviv and Pazgal (2008), and Y. Chen and Farias (2018), as well as the references therein. In these papers, consumers are generally assumed to be surplus maximizers, and the focus is on identifying optimal pricing policies rather than analyzing welfare gains through dynamic pricing. Hence, when consumers respond to marketing offers with the goal of maximizing their personal welfare (i.e., they look out for themselves), receiving an exclusive offer leads to a beneficial inequality that improves the valuation of the targeted discount among other recipients (Greenberg, 1987; Loewenstein et al., 1989). Since contradictory results can be found in the literature, it cannot be predicted unequivocally whether customers perceive personalized prices in the form of individualized coupons positively and thus whether the negative customer reactions to general personalized prices can be mitigated.

Looking back at the two studies by Weisstein et al. (2013) and Keller et al. (2022) we see two different percentage figures. In Weisstein et al. (2013) the comparison of a 10% coupon from the participant to a 20% coupon from a friend led to negative results. In contrast, Keller et al. (2022) suggest, based on their results, that discounts greater than or equal to 10% already attenuate negative customer reactions. Two different but overlapping percentages leading to different results, which may also be due to different contexts (personalized pricing with reference price versus dynamic pricing without reference price).

Taking a look at the pricing literature, one realizes that prices in the marketplace differ across products, brands, stores, and by time (Inman et al., 1997; Monroe, 1990; Rao & Sieben, 1992). However, the decisive factor is how these price differences are evaluated, because the assessment of price acceptance is an essential link between psychological processes and the organism, i.e., an overt reaction such as the customer's willingness to buy (Jacoby & Olson, 1977; Monroe, 1990). Because of apparent price differences, it is assumed that customers compare the price they are offered with a range of prices that are either implicitly remembered or explicitly seen in the marketplace before making a purchase (Lichtenstein et al., 1988; Monroe, 1990). Such a range is also called an acceptable price range and its lower and upper ends determine the lowest and highest amount to be paid for the product. It is important for marketers to calculate such a price range, as it has a direct influence on the price strategies to be implemented (Monroe, 1990). If these results and assumptions are transferred to the coupon literature, the value of the coupon is also decisive for the willingness to buy, and personalized coupons are evaluated by individual reference values and comparisons. To our knowledge, however, no scholar has yet taken on the task of investigating a range that indicates which percentage point differences between one's own coupon and reference coupons are considered acceptable. From *Social Judgment Theory* and assimilation contrast effects (Monroe, 1990; Sherif, 1963), we can deduce that acceptable price ranges have upper and lower limits. We therefore assume that this also applies to the percentage value ranges of vouchers. Ideally, such a range could be used to personalize vouchers without causing direct negative customer reactions. Accordingly, the following research question characterizes our first aim of this paper:

> ***RQ1:*** *Is there a tolerated range between voucher values that is perceived as somewhat fair by the customer?*

Besides a suitable price percentage range and a self-reported assessment of coupon value discrepancies, it is necessary to also find moderators and mechanism influencing the evaluation of personalized coupons and corresponding customer reactions in a behavioral context.

2.3 Mediators and Moderators

The aim of our second study is to implement the theoretical insights gained in Study 1 and to place the suitable percentage range in a more behavioral context. To this end, we will test a model involving two mediators and a moderator, which are explained below.

2.3.1 First Mediator: Information Processing as a Type of Cognitive Effort

Percentage discounts require an arithmetic calculation by the customer to calculate the final price (H. Chen & Rao, 2007). What we know from consumer behavior studies is that customers who receive a percentage discount either do not calculate the final price at all (Suri et al., 2013) or do so but inaccurately when performing mental arithmetic calculations (H. Chen & Rao, 2007). Instead, consumers, whom we conceptualize as strategic decision simplifiers, often rely on the presence of a deal as a decision heuristic to minimize their cognitive effort (DelVecchio, 2005; Weisstein et al., 2013). Especially when making decisions about product prices and promotions, consumers often use simplifying heuristics to form a general judgment about the offer (Diamond & Sanyal, 1990; Morwitz et al., 1998). Accordingly, the minimization or presence of cognitive effort appears to be an important element in the relationship between price promotions, their corresponding consumer perception, and the resulting customer behavior.

Cognitive effort can be defined as the total amount of cognitive resources, such as perceptions, memory, and assessments, required to perform tasks and make decisions (Cooper-Martin, 1994; Russo & Dosher, 1983). In other words, the perceived cognitive effort can also be defined as the "cost of thinking" as described by Shugan (1980). Accordingly, making decisions requires effort from the customer and the extent of this effort depends on the available alternatives, information, constraining time pressure, and the customer's limited information processing capacity (Shugan, 1980). Since information processing is a key component determining the level of cognitive effort (Tyler et al., 1979), we consider it as a synonym for cognitive effort in the following.

Gotlieb and Swan (1990) were among the first to examine the effect of discount size on the extent of information processing in the context of price promotions. They suggested that displayed price promotions increase consumer involvement and thus also increase the extent of information processing. These findings have been supported indirectly by showing that information discrepancy between products (i.e., different types of cars) influences customers' motivation to process information (Ozanne et al., 1992). Interestingly, the results suggest that the relationship between consumers' willingness to process information and information discrepancy can be represented in the form of an inverted U-curve. According to this, the extent and depth of information processing is greatest at moderate discrepancy values. At high and low discrepancy levels, on the other hand, consumers derive little benefit from extensive information processing, which is why they use simpler heuristics to evaluate an offer in these cases. The study by Ozanne et al. (1992) was conducted in a consumer information search behavior context and the experiment was based on differences in cars. Transferring the results of their study to our research project on personalized pricing is thus a risky approach but substantiated in literature. Referring to the work of Ozanne et al. (1992), Grewal et al. (1996) were already able to show in their study that consumers process most information when they see advertisements with moderately large discounts (compared to low and high discounts). The inverted U-shape relationship was thus confirmed by Grewal et al. (1996) in the context of discount use in advertisements. Thus, the importance of information processing evoked by the discount size (e.g., in absolute values) on the effectiveness of an offer has been supported. Based on these results, we venture to apply the findings of Ozanne et al. (1992) to our personalized coupons context, where their findings could predict a similar relationship between the discrepancy between two price offers (i.e., discount sizes) and the degree to which customers process this information. For our model, we therefore assume that information processing serves as a (first) mediator influencing the relationship between a personalized coupon offer and consumers' perceptions (i.e., perceived fairness) and reactions such as WOM intentions, repurchase intention, satisfaction, and loyalty (see Figure 1).

With regard to the direction of influence, it can be stated that the valence of cognitive effort or increased information processing is perceived as negative by default (C. W. Yoo et al., 2017) and in some cases also has negative effects on customer reactions such as customer loyalty (H. Zhang et al., 2018). Accordingly, we also assume that enhanced information processing (with a medium discrepancy in information) is likely to have a negative effect on customer perceptions and reactions.

As already indicated, information processing can be considered the first mediator, followed by perceived fairness, which will be explained in the following.

2.3.2 Second Mediator: Perceived Fairness

In chapter 2.1 we have already defined that perceived price fairness is about the customer evaluation comparing whether the difference (or lack of difference) between two prices is reasonable, acceptable or justifiable (L. Xia et al., 2004). In the following we will focus on the reason why perceived fairness is part of the serial mediation and is thus influenced by information processing.

Theoretically, a distinction can be made between three types of fairness, namely distributive, procedural, and interactive fairness. In principle, however, the entire purchasing process is always evaluated in terms of perceived price fairness (Haws & Bearden, 2006). It is particularly important that the assessment of fairness always takes place in the eye of the beholder and thus reflects a subjective perception (Greenberg et al., 1991).

Looking at the pricing literature, some researchers have already shown that perceived fairness is the underlying mechanism between personalized prices and different consumer responses (Weritz, 2023), regardless of whether buyers were advantaged or disadvantaged (Hufnagel et al., 2022). In terms of information processing, if we look at the existing fairness literature, there is already much literature on the antecedents of perceived fairness, and some scholars have already investigated the influence of thought processes on perceived fairness. Barclay et al. (2017) are the first to present a motivated cognition approach as a formation for fairness perception. Their approach is based on the idea that every individual has different motivations, such as desires, wishes, and preferences with regard to a given reasoning task (Kunda, 1990). Such motivations then stimulate cognitive thought processes and information processing that influence perceived fairness. In addition, in the pricing context, newer findings explicitly indicate that different thinking styles influence perceived fairness differently (Shaw et al., 2022). Holistic thinkers perceive price increases as fairer compared to analytical thinkers. Based on the above-stated findings, it is reasonable to assume that the information processing of customers triggered by personalized coupons also influences perceived fairness. Therefore, a serial mediation is tested in this study, with information processing as the first and perceived fairness as the second mediator (see Figure 1).

After first determining an appropriate price percentage range in Study 1, we hypothesize the following relationship between coupon value discrepancy, information processing, customer fairness perceptions and reactions:

> *H1: Coupon value differences outside (both below and above) the suitable discount percentage range defined through Study 1 do not increase customers' information processing, and thus mitigate the negative effect on customers' fairness perception. This results in less negative customer reactions such as decreased 1) WOM intention, 2) repurchase intention, 3) satisfaction, and 4) loyalty.*

> *H2: Coupon value differences that are within the suitable discount percentage range defined through Study 1 increase customers' information processing and thus have a negative influence on customers' fairness perception and resulting customer reactions, such as 1) WOM intention, 2) repurchase intention, 3) satisfaction, and 4) loyalty.*

2.3.3 Moderator: Need for Cognition

It is already known that the effects of price framing on customer evaluation are moderated by cognitive processes, among others (Chandon et al., 2000). This can be explained by *Mental Calculation Theory*, as people avoid cognitive effort and opt for mental heuristics that lead them to ignore the underlying base value (Tripathi & Pandey, 2017). In addition to various ways of price framing that can be influenced by the company and could reduce cognitive effort, individual dispositions for cognitive effort can also change the perception of a price. Need for cognition is regarded as a decisive individual difference variable determining the motivation of the respective individual to process information (Haugtvedt et al., 1992).

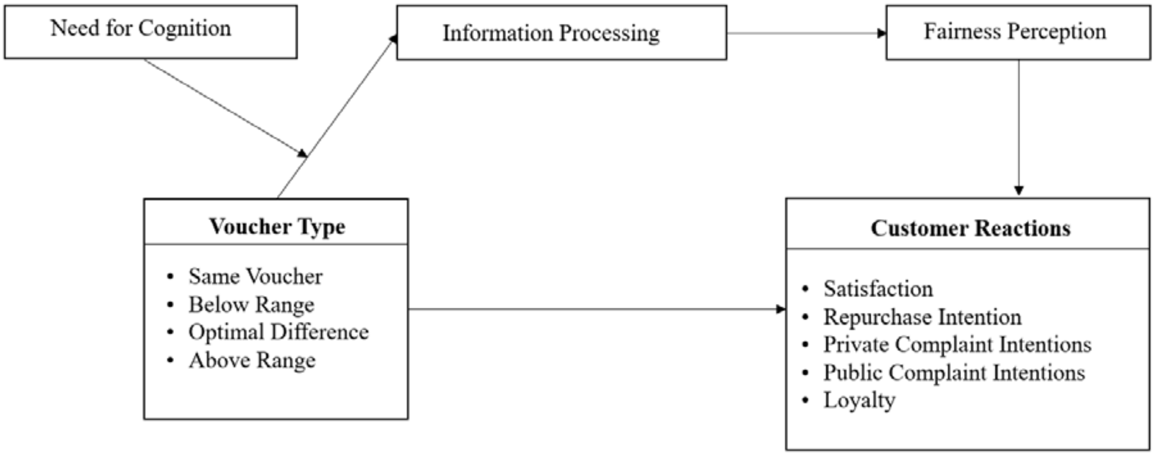
Need for cognition can thus be defined as an individual's stable disposition to actively and extensively engage in cognitive activities, influencing consumers' willingness to process information (Cacioppo et al., 1996). In the context of consumption, consumers with a high need for cognition are intrinsically motivated to obtain information about products and their characteristics (such as price in our case) and to analyze and compare them in detail. In contrast, consumers with a low need for cognition prefer tasks and choices that require little cognitive effort and tend to process information heuristically (Kim & Kramer, 2006). Although the price information of a product is one of the most important decision criteria for a purchase, many customers lack the motivation to process this information accurately (Morwitz et al., 1998; Stiving & Winer, 1997). For these consumers, there is also a lack of willingness to draw comparisons between their own discount and that of a friend and to search for explanations for this difference (Kim & Kramer, 2006).

The research of Inman et al. (1990; 1997) has shown that need for cognition moderates the effectiveness of price promotions by influencing customer reactions to price discounts. Whereas individuals with low need for cognition can be influenced by mere discount signals regardless of the actual price discount, individuals with high need for cognition are only influenced by actual price discounts (Inman et al., 1990).

Within the underlying study, we assume that need for cognition has a moderating influence on the direct effect of personalized coupons on information processing. Based on the results of other studies mentioned above, we presume that for customers with a low need for cognition, the mere display of a voucher is perceived as something rather positive, as no further questioning takes place. These customers would therefore put less effort into information processing. For customers with a high need for cognition, however, personalized coupons are more likely to trigger a higher willingness to make comparisons between different coupon values in order to find reasons for coupon differences. In this case, the willingness to process information would be increased.

Although the need for cognition is not a variable that can be influenced by the company, it is one that can be determined by existing customer data prior to purchase and is therefore decisive in evaluating whether a personalized coupon should be offered or not. The introduced conceptual model can be found in Figure 1.

Figure 1. Conceptual Model of Underlying Study



Covariates: Purchase Experience; Product Involvement; Atttitude Toward Personalization; Time Spent Online; E-Purchase Frequency

3 Study 1

3.1 Study Design

The goal of the first study is to investigate how consumers react toward personalized pricing in the realm of tailored coupons. Participants were asked to indicate how much the value of a coupon given to a close friend could differ from their own coupon and still be considered fair. Expensive and rather moderately priced concert tickets were used in the scenario. The goal of this first study is to gather a percentage range that is tolerated in relation to one's own personalized voucher. To make sure consumers understand the adapted measures and scenarios, a pretest was run. After the data was gathered, first, it was analyzed descriptively. Second, it was analyzed whether the two different products are perceived to be different in their initial price (expensive versus moderate). Further, the tolerated range was computed using an adapted van Westendorp price sensitivity meter. The so gathered range was further investigated in a second study.

Subjects and Design

Participants for the scenario-based survey were employees and students recruited via a large German university The survey consisted of a between-subjects design with two different groups: high priced product (i.e., concert tickets of an international band) and low-priced product (i.e., concert tickets of a regional band). Participants were randomly assigned to one of the two groups. First, participants were given a definition of the concept of personalized pricing and asked a verification question (see Appendix 1), whether they understood the concept correctly. Four participants did not understand the concept correctly, leaving us with a sample of 116 persons. In a second attention check, participants were asked to select "strongly disagree" on a scale from "strongly disagree" to "strongly agree", which was adapted from Gruzd et al. (2020). Ten subjects failed the attention check, leaving us with a sample of 106 participants. For the van Westendorp price sensitivity meter (van Westendorp, 1976), transitivity of the data is a prerequisite, therefore, participants were excluded from the sample if their answer pattern did not show transitive percentages, leaving us with a final sample of 47 participants: 25 in the high-price group and 22 in the low-price group.

The average participant is 28 years old. Further, the plurality of respondents is female (61.7%), has a Master's degree (53.2%), and is studying (55.3%). All details on the demographics can be found in Table 2.

Table 2. Sample Descriptive Statistics of Study 1

		Mean/Frequency	Percentage	Min	Max
Age		28.23		19	70
Gender					
	Male	18	38.3%		
	Female	29	61.7%		
	Non-binary/third gender	0	0.00%		
	Prefer not to say	0	0.00%		
Education					
	High school graduate	9	19.1%		
	Bachelor's degree	12	25.5%		
	Master's degree	25	53.2%		
	Professional degree	0	0.00%		
	Doctorate degree	1	2.1%		
Employment					
	Employed full-time	10	21.3%		
	Employed part-time	7	14.9%		
	Self-employed	1	2.1%		
	Homemaker	0	0.00%		
	Student	26	55.3%		
	Retired	1	2.1%		
	Unemployed	2	4.3%		
		N = 47			

Stimuli and Procedure

Subjects were introduced to an initial situation, depending on whether they were in the high-price or low-price group. They were asked to imagine that their favorite international (low-price: regional) band "The Rolling IMMs" was in town in a couple of weeks. The band played at the Olympic Stadium (low-price: one of the bars) and concert tickets are 175€ per person (low-price: 35 € per person). Participants were further told that they received an email with an attached customized voucher. Based on their previous purchasing behavior, they received a personalized voucher for 10% off the concert tickets. Discount displays equal to or exceeding 10% prove to be effective (Keller et al., 2022). However, very low-price discounts (below 10%) counteract the positive impact of additional price information for dynamic pricing and should be avoided. Additionally, consumers generally perceive discounts of approximately 5% as very low, 10% as low, 20% as medium, 30% as high, and 50% as very high (Keller et al., 2022). Consequently, we opted to use 10% as our baseline case. Further information on the initial situation can be found in Appendix 2. After being exposed to the initial situation, participants were told that their best friend calls and tells them that s/he has also heard about the concert and wants to go as well. Participants were told that their best friend has also received a personalized voucher. Subsequently, participants had to answer questions which were adapted from the van Westendorp price sensitivity meter (van Westendorp, 1976) toward the fairness situation (see Appendix 3). Before running the first main study, a pretest with 18 subjects was conducted to test whether the adapted questions were comprehensible for the participants. Subjects were international students at a German university. Regarding the van Westendorp method, participants had to rate the percentage size of the discount voucher of their friends based on their fairness perceptions, e.g., *"In comparison to your coupon, at which discount size of the voucher that your friend received for the concert ticket would you consider the promotion too unfair and you would not buy the tickets yourself?"* or *"In comparison to your coupon, which discount size of the voucher that your friend received for the concert ticket would you consider too unfair but you would still buy the tickets yourself?".* Participants were able to select discount sizes from 10% to 60%. After subjects were presented with the scenario, they were asked several questions regarding the situation, among others including manipulation checks. At the end of the survey, participants answered several demographic questions.

3.2 Measures

A list of all measures can be found in Appendix 4. Unless otherwise mentioned, all constructs were measured on a seven-point Likert scale (1=strongly disagree; 7=strongly agree). Cronbach's alpha is used to assess internal consistency and displayed right next to the construct name.

Van Westendorp Method: The regular questions based on the van Westendorp method were adapted to the fairness concept and pretested. In the end, the following questions were used: "In comparison to your coupon, at which discount size of the voucher that your friend received for the concert ticket would you consider the promotion too unfair, and you would not buy the tickets yourself?" (= too unfair and not buy); "In comparison to your coupon, which discount size of the voucher that your friend received for the concert ticket would you consider too unfair, but you would still buy the tickets yourself?" (= too unfair but buy); "In comparison to your coupon, which discount size of the voucher that your friend received for the concert ticket would you still consider fair, but you would still be upset that you didn't get such a high discount." (= fair but doubts); "In comparison to your coupon, which discount size of the voucher that your friend received for the concert ticket would you consider fair without any jealous feelings toward your friend?" (= fair).

Attention and Manipulation Checks (also partly Study 2): To determine whether participants read the questionnaire carefully, we included an attention check asking participants to select "strongly disagree" adapted from Gruzd et al. (2020). Additionally, we included one more attention check as a single choice question by asking participants, whether the initial ticket price was 35€ or 175€ (1 = yes, or 2 = no). Coupon manipulation was checked by recalling whether their friend had received a higher or equal discount.

Demographic Information: The last part of the survey addressed demographic data such as gender, age, employment status, or education level.

3.3 Analysis and Results

First, a manipulation check is conducted to test whether all manipulations within the survey worked as they were supposed to. Second, the data is analyzed descriptively. Next, for the first study, the fairness range is determined using the van Westendorp price sensitivity meter (van Westendorp, 1976).

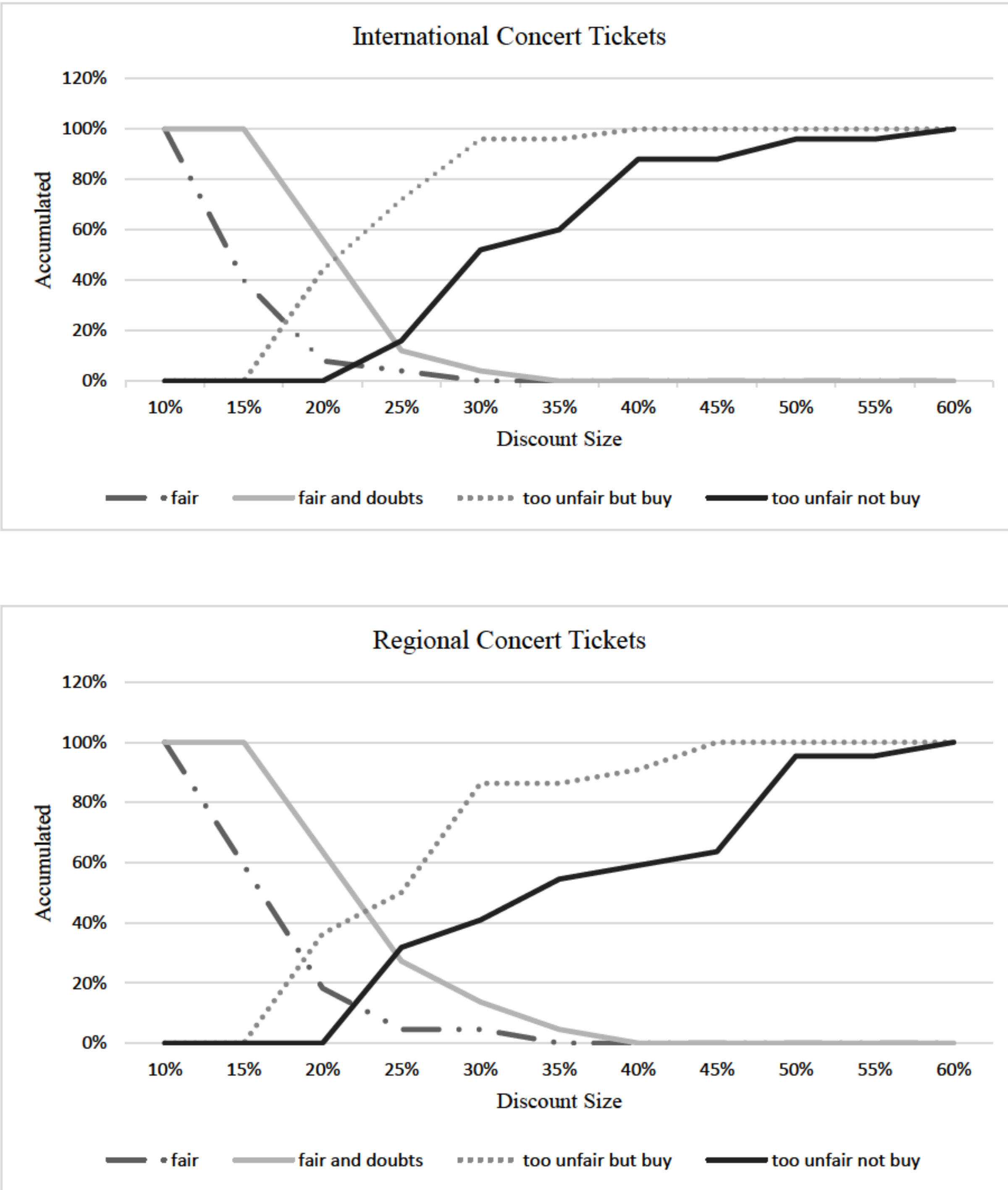
3.3.1 Manipulation Check

Subjects were asked to rate whether the price of the concert ticket (i.e., 35€ or 175€) is a rather low or high price. In comparison to normal concert tickets, the price of the international concert tickets of 175€ was perceived to be of high price ($M = 6.16$, $SD = .99$), whereas the price of the regional concert tickets of 35€ was perceived to be of a rather low price ($M = 1.77$, $SD = 0.87$), $t(45)=16.21$, $p < 0.001$).

3.3.2 Determination of Fairness Range

Within this subchapter, fairness perceptions regarding the discount size of the friend’s voucher are analyzed graphically using the van Westendorp method. Figure 2 presents the graphical depiction using the van Westendorp method for the international and regional concert tickets.

Figure 2. Van Westendorp for Fairness Perception



Further, Python was used to calculate the intersection points where participants are indifferent in their fairness perceptions and the optimal fairness range. We used Python version: 3.11.4 64-bit with Spyder version: 5.4.3 (conda). The exemplary Python code for the high-priced concert tickets can be found in the Appendix 5. To calculate the lower fairness value, one needs the intersection between the fairness assessment "fair" and "too unfair but buy". Further, some assumptions were made. To calculate the function of both curves, it was assumed that there is a linear relationship between the measured fairness points (e.g., fair: points two and three). This assumption was made for the entire function, so that the slope m and the y-axis intercept t could be calculated, yielding the respective function. The two functions were equated and solved using the equation application in Python.

If we take a closer look at the concert tickets of the international band (i.e., higher-priced), we see that the optimal fairness point is 22%, the lower bound fairness point is 17.63%, the upper bound fairness is 24.67%, and the indifferent fairness point is 20.83 %. The results for the regional band concert tickets (i.e., lower-priced) yield similar results: The optimal fairness point is 22%, lower bound fairness 18.83%, upper bound fairness 24.64%, and indifferent fairness point is 22.75%. An overview can be found in Table 3. These results of fairness perception laid the ground for Study 2.

Table 3. Overview of Fairness Perceptions

	Optimal Fairness Point	**Lower Bound Fairness Point**	**Upper Bound Fairness Point**	**Indifferent Fairness Point**
International concert tickets (higher-priced)	x: 22.00%	x: 17.63%	x: 24.67%	x: 20.83%
Regional concert tickets (lower-priced)	x: 22.00%	x: 18.83 %	x: 24.64%	x: 22.75%

4 Study 2

4.1 Study Design

The range gathered in Study 1 was further investigated in a second study using an experimental design. Based on the first study, four different discount rates (i.e., same voucher, optimal difference, below range, above range) are selected and tested on how consumer reactions differ. Further, a serial mediation (i.e., information processing and fairness perceptions) is investigated. In addition, a moderator (i.e., need for cognition) is introduced.

Subjects and Design

Participants for the second study were also recruited via a large German university and through different survey sharing platforms (e.g., SurveyCircle, 2022). 307 participants took part in the study. As recommended by Weisstein et al. (2013), we chose a rather low price for the ticket (i.e., 35€), as it is assumed that percentage discount coupons work better for low-priced products. Additionally, there was nearly no difference between the low- and the high-priced product in the first study.

The experiment consists of a between-subjects design with four different groups: same voucher, optimal difference, below range, above range. Participants were randomly assigned to one of the four groups. As in the first study, participants were given a definition of the concept of personalized pricing and asked a verification question. After being introduced to the scenario, participants were asked whether they would like to purchase the event tickets. 57[5] participants did not understand the concept correctly or decided to not purchase the product. Participants were asked to answer the same attention check as in Study 1 ("Please select "strongly disagree" as your answer") and an additional one, where they had to select whether the initial price of the event ticket is 35€. 54[6] participants did not answer both questions correctly, leaving us with a final sample of 196 participants. Within the final sample of Study 2, the average participant is 26 years old. The plurality of respondents is female (67.3%), has a Bachelor's degree (36.7%), is studying (71.4%), and has an annual gross household income of between 10,001 and 50,000€. All details on the demographics can be found in Table 4.

[5] 19 Participants did not understand the concept of personalized pricing correctly. 38 subjects decided not to purchase the event ticket.
[6] 26 Subjects did not select „strongly disagree" and 28 were not able to recall that the initial price was 35€.

Table 4. Sample Descriptive Statistics of Study 2

		Mean/Frequency	Percentage	Min	Max
Age		25.65		17	72
Gender					
	Male	59	30.1%		
	Female	132	67.4%		
	Non-binary/third gender	4	2.0%		
	Prefer not to say	1	0.5%		
Education					
	High school graduate	53	27.0%		
	Bachelor's degree	72	36.7%		
	Master's degree	59	30.1%		
	Professional degree	5	2.6%		
	Doctorate degree	7	3.6%		
Employment					
	Employed full-time	28	14.3%		
	Employed part-time	17	8.7%		
	Self-employed	7	3.6%		
	Homemaker	0	0.0%		
	Student	140	71.4%		
	Retired	2	1.0%		
	Unemployed	1	0.5%		
	Other	1	0.5%		
Annual Gross Household Income					
	< €10,000	66	33.7%		
	€10,001 – €50,000	59	30.1%		
	€50,001 – €90,000	18	9.2%		
	€90,001 - €150,000	13	6.6%		
	> €150,001	5	2.5%		
	Prefer not to say	35	17.9%		
		N = 196			

Stimuli and Procedure

After the first attention check (definition of personalized pricing), subjects were introduced to a fictitious scenario. They were told that they are looking online for a ticket for a specific event and that tickets are 35€. They were further told that based on their purchase history and external information, such as social media, their favorite online event ticket seller offers them a personalized voucher for 10% off the regular price and that the online ticket seller is known to be using personalized pricing. Participants were then asked to state whether they would like to purchase the event ticket (see Appendix 6). Afterwards, subjects were told that their best friend, who is also a frequent customer of the same online ticket seller, was also offered a voucher on the same day by the same online ticket seller. Based on the findings from the first study, four different scenarios were created about the height of the personalized voucher that the friend receives: same voucher (i.e., friend's voucher is 10%), optimal difference (i.e., friend's voucher is 22%), below range (i.e., friend's voucher is 12%), and above range (i.e., friend's voucher is 32%). Subjects were randomly assigned to one of the four scenarios. See Appendix 7 for further details on the scenario description. After participants were presented with the scenario, they were asked to answer several questions.

4.2 Measures

All measures from Study 2 can also be found in Appendix 4. Again, unless otherwise mentioned, all constructs were measured on a seven-point Likert scale (1=strongly disagree; 7=strongly agree). Cronbach's alpha is used to assess internal consistency and displayed right next to the construct name.

Satisfaction (α = .96): Participants were asked to indicate their satisfaction with the e-ticket seller's promotion strategy evaluating it on a sematic differential scale adapted from Darke and Dahl (2003) and Haws and Bearden (2006) containing four bipolar adjectives (i.e., dissatisfied/satisfied, unhappy/happy, disappointed/delighted, and displeased/pleased).

Repurchase Intention (α = .95): Four items slightly adapted but based on those used by Garbarino and Maxwell (2010) addressed the respondent's willingness to purchase from the e-ticket seller from the scenario now or in the future.

Complaint Intentions: In addition to customer satisfaction and purchase intention we measured complaint intentions (i.e., WOM) as one of the potential customer reactions. Generally complaint intentions can either be expressed privately or publicly (Singh, 1988). Hence, the construct was measured by means of four items for private complaint intentions (α = .76) and five items for public complaint intentions (α = .83) on a seven-point Likert scale (1=very unlikely; 7=very likely) (Garbarino & Maxwell, 2010; Singh, 1988).

Loyalty (α = .89): Brand loyalty measures the degree of customer's favorable attitude toward the e-ticket seller. All items were adapted from B. Yoo and Donthu (2002) and Lai et al. (2010). The likelihood of whether participants would consider buying from this e-ticket seller again, whether they consider it their first choice, and whether they do not buy tickets elsewhere were also tested on a seven-point Likert scale (1=very unlikely; 7=very likely).

Information Processing (α = .89): To measure the first potential mediator of our research model, we used the Subjective Information Processing Awareness (SIPA) scale. It describes the extent to what the interaction with an intelligent system like a personalized pricing algorithm enable customers to experience 1) transparency, 2) understanding, and 3) predictability of the e-ticket seller's information processing for the price setting (Schrills et al., 2022). We adapted the original six items slightly to the context of our study.

Perceived Price Fairness (α = .90): In the serial mediation of our model, the second mediator is perceived price fairness. We measure perceived price fairness following the definition of L. Xia et al. (2004) namely that price fairness is a consumer's assessment of whether a price is reasonable, acceptable, or justifiable relative to a reference. For its measurement we therefore relied on a three item well established scale from Grewal et al. (2004) which was also used in Weisstein et al.'s (2013) study.

Need for Cognition (α = .86): As a potential moderator, we measure the extent to which the participant enjoys thinking or exerting cognitive effort in some way. We used Cacioppo et al.'s (1984) 18-item scale, a self-report measure designed to determine whether the participant has a rather high or low need for cognitive effort.

Confounds and Covariates: To account for individual heterogeneity and increase internal validity, several covariates were included. First, we measured the extent to which participants had previous experience with purchasing event tickets. Participants had to answer three questions regarding the *purchasing experience (α = .89)* (e.g., "I have a great deal of experience in buying event tickets"). The scale was adapted from Lo et al. (2019) The participants *attitude toward personalization (α = .85)* was additionally included as a covariate. The scale was adapted from Taylor and Todd (1995) and comprised four different items, such as "For the e-ticket seller, using the personalized coupons is a good idea.". Comparable to McQuarrie and Munson (1992) consumers' *product involvement (α = .90)* was assessed by using eight items out of the original 20 item semantic differential scale proposed by Zaichkowsky (1985). In addition, we wanted to examine how familiar consumers are with online purchases and make sure there were no significant group differences. We therefore assessed how much time they spend online for their own purposes and how often they have made online purchases in the last twelve months (Doolin et al., 2005; Malhotra et al., 2004).

4.3 Analysis and Results

The following chapter first presents a descriptive overview of the used measures and further tests whether the scenarios are perceived as intended. Next, hypotheses are tested by investigating group differences and running a mediation analysis. Table 5 and Table 6 demonstrate the descriptive statistics of the dependent variables and mediators used in the second study.

Table 5. Descriptive Statistics of Dependent Variable Measures Used in Study 2

	N	Min	Max	Mean	SD
Satisfaction					
Same voucher	49	2.50	7.00	5.1735	1.43088
Below range	48	1.00	6.00	3.2135	1.40666
Optimal range	50	1.00	5.00	2.5250	1.27800
Above range	49	1.00	7.00	2.5714	1.33658
Repurchase Intentions					
Same voucher	49	2.25	7.00	5.2806	1.14636
Below range	48	1.00	7.00	4.2240	1.63346
Optimal range	50	1.00	6.25	3.8200	1.47084
Above range	49	1.00	6.25	3.5612	1.35163
Private Complaint Intentions					
Same voucher	49	1.25	6.00	3.0918	1.11884
Below range	48	2.00	6.50	4.1042	1.22348
Optimal range	50	2.75	7.00	4.6050	1.02779
Above range	49	1.25	6.25	4.4796	1.24774
Public Complaint Intentions					
Same voucher	49	1.00	5.60	1.6571	0.87750
Below range	48	1.00	5.00	2.0875	0.98750
Optimal range	50	1.00	4.80	2.0680	0.99354
Above range	49	1.00	5.00	2.2245	1.05486
Loyalty					
Same voucher	49	1.00	6.50	4.0408	1.34566
Below range	48	1.00	5.25	3.1563	1.31645
Optimal range	50	1.00	5.50	3.1950	1.13759
Above range	49	1.00	6.00	2.7398	1.20000

same voucher = 10%; below range = 12%; optimal range = 22%; above range = 32%

Table 6. Descriptive Statistics of Mediator Measures Used in Study 2

	N	Min	Max	Mean	SD
Fairness					
Same voucher	49	1.67	7.00	5.4082	1.20428
Below range	48	1.00	6.00	3.8056	1.40891
Optimal range	50	1.00	6.00	3.2800	1.23509
Above range	49	1.00	6.67	3.2925	1.51937
Information Processing					
Same voucher	49	1.00	6.50	3.1224	1.37638
Below range	48	1.00	4.50	2.3611	1.00197
Optimal range	50	1.00	5.67	2.5033	1.09549
Above range	49	1.00	5.50	2.2789	1.14937

same voucher = 10%; below range = 12%; optimal range = 22%; above range = 32%

4.3.1 Manipulation Check

Within the manipulation check participants had to decide whether their friend received a higher discount voucher or the same discount voucher (reverse coded) on a seven-point Likert scale (1=strongly disagree, 7=strongly agree). The perception of voucher discounts was statistically different for the different voucher discounts, Welch's $F(3, 103.06) = 226.38$, $p < .001$. A Games-Howell post-hoc analysis revealed a significant difference ($p < .001$) between the height of the voucher of the groups *same voucher* and *below range* (-5.09, 95% CI [-5.62; -4.55], *same voucher* and *optimal range* (-4.92, 95 % CI [-5.61; -4.23], and *same voucher* and *above range* (-5.16, 95% CI [-5.72; -4.61]. Compared to the *same voucher*-scenario, the friend's coupons with discounts *below, optimal*, and *above* range were perceived to receive a higher voucher by the participants.

Based on the first study, 35€ for event tickets were perceived to be rather low-priced. To check, whether this is also the case in Study 2, participants had to state their perception regarding the statement "Compared to normal event tickets, 35€ per event ticket is a rather…" (1=low price, 7=high price). A single-item measure was used as single item scales seem to perform equally well or better than multi-item scales (see Study IV). A median split was conducted, and it was found that 65.8% perceived the 35€ of the event ticket as rather low-priced.

4.3.2 Hypotheses Testing

An Analysis of Covariance (ANCOVA) was conducted to evaluate group differences among the different scenarios. Further, a moderated mediation analysis was conducted.

Group Differences

The requirements of an ANCOVA have been considered when designing the study or tested before running the main analysis. All requirements are fulfilled, in some cases no normality distribution is given, however, ANCOVAs just like ANOVAs are said to be robust against this violation, and therefore, the analysis was conducted.

The ANCOVAs controlled for purchase experience, product involvement, attitude toward personalization, the time subjects spent online, and e-purchase frequency. To compare the different scenarios, a Bonferroni corrected post hoc test was conducted. Further, a variable called *group* was established, which comprises the four different voucher scenarios (i.e., *same voucher* (10% and 10%); *below range* (10% and 12%); *optimal difference* (10% and 22%), and *above range* (10% and 32%)). Appendix 8 - Appendix 14 present the findings of the group comparisons.

Satisfaction: After adjusting for the above mentioned covariates, the level of satisfaction differed statistically significant for the different types of coupons, $F(3, 187) = 33.83$, $p < .001$, partial $\eta^2 = .352$. A Bonferroni-corrected post-hoc analysis revealed several significant differences between satisfaction and the different types of vouchers. Significant differences between the group *same* voucher and voucher *below range* ($M_{same} = 4.94$ versus $M_{below} = 3.31$, $p < 0.001$), the group *same* vouchers and voucher with *optimal range* ($M_{same} = 4.94$ versus $M_{optimal} = 2.63$, $p < 0.001$), and *same voucher* and vouchers *above range* ($M_{same} = 4.94$ versus $M_{above} = 2.61$, $p < 0.001$) were found. Therefore, participants are more satisfied with the scenario where both parties received the *same* discount voucher compared to those where the subject's friend received a higher voucher. Further, significant differences were found between the voucher *below range* and *optimal range* ($M_{below} = 3.31$ versus $M_{optimal} = 2.63$, $p = 0.045$) and *below* and *above range* ($M_{below} = 3.31$ versus $M_{above} = 2.61$, $p = 0.041$).

Repurchase Intention: When looking at repurchase intention, an ANCOVA revealed a statistically significant difference for the different types of vouchers, $F(3, 187) = 10.315$, $p < .001$, partial $\eta^2 = .142$. For the groups *same* voucher and *optimal range* ($M_{same} = 4.98$ versus $M_{optimal} = 3.94$, $p < 0.001$) as well as *same* voucher and *above range* ($M_{same} = 4.98$ versus $M_{above} = 3.60$, $p < 0.001$) a statistical difference is found. The groups with the *same* vouchers are more likely to repurchase from that ticket seller compared to the groups where the friend received either a voucher within the *optimal range* or *above the range*. No significant difference between the groups with the *same* voucher and the voucher *below rang*e was found ($M_{same} = 4.98$ versus

M_{below} = 4.36, p = 0.126). In addition, subjects are also more likely to repurchase from that company again if their friend received a voucher *below range* compared to the group where the friend received a voucher *above range* (M_{below} = 4.36 versus M_{above} = 3.60, p = 0.21).

Private Complaint Intentions: After adjusting for the covariates, private complaint intentions were found to differ statistically significant for the different types of vouchers, F(3, 187) = 11.769, p < .001, partial η^2 = .159. Subjects from the *same* voucher scenario were less likely to engage in private complaint intentions compared to those presented to a friend with a *below range* (M_{same} = 3.32 versus M_{below} = 4.00, p = 0.022), *optimal difference* (M_{same} = 3.32 versus $M_{optimal}$ = 4.51, p < 0.001) or *above range* (M_{same} = 3.32 versus M_{above} = 4.45, p < 0.001) voucher.

Public Complaint Intentions: Other than with private complaint intentions, no significant effect for the different types of coupons on public complaint intentions was found, F(3, 187) = 1.581, p = .195, partial η^2 = .025. Subjects do not display significantly different intentions to complain publicly when being offered the *same* vouchers compared to those whose friends receive coupons *below range* (M_{same} = 1.78 versus M_{below} = 2.05, p = 1), within the *optimal difference* (M_{same} = 1.78 versus $M_{optimal}$ = 2.00, p = 1), or *above the range* (M_{same} = 1.78 versus M_{above} = 2.22, p = 0.192).

Loyalty: When looking at loyalty, an ANCOVA revealed that loyalty differs statistically significant for the different types of vouchers, F(3, 187) = 6.882, p < .001, partial η^2 = .099. In this case, only those that received the *same* vouchers compared to those that received coupons *above* the range differed significantly (M_{same} = 3.82 versus M_{above} = 2.76, p < 0.001). No difference for those being offered the *same* vouchers compared those whose friend was being offered a voucher *below* range (M_{same} = 3.82 versus M_{below} = 3.29, p = 0.174) or within the *optimal difference* (M_{same} = 3.82 versus $M_{optimal}$ = 3.26, p = 0.125) was found.

Price Fairness: Further, also for fairness perceptions, after adjusting for the above mentioned covariates, a statistical difference for the different types of vouchers, F(3, 187) = 22.616, p < .001, partial η^2 = .266, was found. Subjects perceived the *same* voucher scenario to be fairer compared to the *below range* (M_{same} = 5.17 versus M_{below} = 3.95, p < 0.001), *optimal difference* (M_{same} = 5.17 versus $M_{optimal}$ = 3.35, p < 0.001), and *above range scenario* (M_{same} = 5.17 versus M_{above} = 3.32, p < 0.001).

Information Processing: For information processing, an ANCOVA revealed a statistically different result for the different types of vouchers, F(3, 187) = 3.313, p = .021, partial η^2 = .050.

Information processing for those participants in the group with the *same* vouchers was higher compared to those in the *above range* group (M_{same} = 2.98 versus M_{above} = 2.29, p = 0.015). As mentioned above, the scale of information processing as used in the scale rather shows how transparent and accessible the information was for the participants, therefore participants had a rather low effort to process information in the *same* compared the *above* range scenario. For all other groups, information processing did not differ significantly, for example (M_{same} = 2.98 versus M_{below} = 2.46, p = 0,165) or (M_{same} = 2.98 versus $M_{optimal}$ = 2.54, p = 0.341).

Moderated Mediation Analysis

A moderated mediation analysis was performed using IBM SPSS Statistics and PROCESS macro version 4.2 developed by Hayes (2022). Bootstrapping was employed to gauge the indirect impacts. As the independent variable is categorical, the coefficients are interpreted to the baseline scenario *same voucher*. As we were interested in a serial mediation, first model 6 was run. We are also interested in a moderating effect, therefore, model 83 was chosen additionally.

In the presented mediation model (PROCESS model = 6 or 83; 5,000 resamples, Davidson and MacKinnon (1993) heteroscedasticity-consistent inference) satisfaction, repurchase intention, private and public complaint, and loyalty were used as dependent variables, and group (i.e., same voucher, below range, optimal difference, above range) was used as the independent variable. Information processing was used as first mediating variable and fairness for the serial mediation. Purchase experience, product involvement, attitude toward personalization, time spent online, and e-purchase frequency were included as covariates. Further information can be found in the Appendix 15 - Appendix 19.

Satisfaction: Information processing (B = 0.2253, t = 2.5854, p = 0.0105) and fairness perceptions (B = 0.3832, t = 5.2090, p < 0.001) are found to positively predict satisfaction. When looking at the mediation effects information processing is found to negatively affect satisfaction for vouchers *below range* (indirect effect = -.1175, LLCI = -.2862, ULCI = -.0017) *and above range* (indirect effect = -.1576, LLCI = -.3383, ULCI = -.0171) relative to the same voucher scenario. Further, a partial negative mediation was found via fairness perception on satisfaction for coupons *below* range (indirect effect = -.4441, LLCI = -.7520, ULCI = -.2059), with *optimal difference* (indirect effect = -.6777, LLCI = -1.0499, ULCI = -.3750), and *above* range (indirect effect = -.6814, LLCI = -1.0496, ULCI = -.3647) relative to *same* coupons. No indirect effect via information processing and fairness perceptions on satisfaction was found for any of the coupon types. When taking need for cognition as moderator into the model, no moderated mediation effect was found.

Repurchase Intention: Fairness perceptions are found to positively predict repurchase intentions ($B = .3763$, $t = 4.6339$, $p < 0.001$); however, no direct effect of information processing on repurchase intentions was found ($B = .1103$, $t = 1.3212$, $p = .1881$). When looking at the different voucher types, the effect on repurchase intentions was negatively and fully mediated by fairness perceptions for the vouchers *below range* (indirect effect = -.4361, LLCI = -.7615, ULCI = -.1826) and with *optimal difference* (indirect effect = -.6655, LLCI = -1.0658, ULCI = -.3324) and partially mediated for the *above range* voucher (indirect effect = -.6691, LLCI = -1.0653, ULCI = -.3393). There is no serial mediation via information processing and fairness perceptions on repurchase intentions. When including need for cognition as moderator, no effect was found.

Private Complaint Intentions: Fairness perceptions are found to negatively predict private complaint intentions ($B = -.2755$, $t = -4.1120$, $p < 0.001$). Further, fairness perceptions were found to fully negatively mediate the relationship between coupons *below range* (indirect effect = .3193, LLCI = .1308, ULCI = .5626), with *optimal difference* (indirect effect = .4873, LLCI = .2344, ULCI = .7930), and *above range* (indirect effect = .4899, LLCI = .2358, ULCI = .7909) relative to the baseline coupon with *same* vouchers. A moderated mediation with need for cognition on information processing and fairness did not yield any significant results.

Public Complaint Intentions: Fairness perceptions are found to negatively predict public complaint intentions ($B = -.1575$, $t = -2.8489$, $p = .0049$). Fairness further fully mediates public complaint intentions positively for the vouchers *below range* (indirect effect = .1833, LLCI = .0541, ULCI = .3440), with *optimal difference* (indirect effect = .2797, LLCI = .0946, ULCI = .4793), and *above range* (indirect effect = .2812, LLCI = .0889, ULCI = .4949) respective to *same* vouchers. A moderated mediation model with need for cognition as moderator on the serial mediation on information processing and fairness did not yield any significant results.

Loyalty: Only fairness perceptions were found to predict loyalty significantly ($B = .3189$, $t = 4.6518$, $p < 0.001$). When looking at the indirect effects, the effects of all three different voucher types, namely *below range* (indirect effect = -.3696, LLCI = -.6289, ULCI = -.1613*), optimal difference* (indirect effect = -.5639, LLCI = -.9093, ULCI = -.2895), and *above range* (indirect effect = -.5670, LLCI = -.9044, ULCI = -.2852) were found to be negatively mediated by fairness in relation to *same* vouchers. No significant effects of need for cognition as moderator on the serial mediation was found.

In summary, hypotheses 1 and 2 have to be rejected. Information processing was perceived to be rather untransparent in the above voucher scenario. Further the influence on satisfaction, repurchase intention, and private WOM was negative. A mediation effect was only found for fairness perception; however, information processing had no effect on either fairness perceptions or the customer response.

5 Discussion, Limitations, and Future Research

Discussion

In the underlying study we discussed customized prices which is one of the focal topics regarding personalization. Technological advances paved the way for first degree price discrimination, like personalized pricing (Priester et al., 2020). Financially, this pricing strategy is promising for companies as it can contribute to significant increases in revenue and profit (Shiller, 2020; Smith et al., 2023). However, subsequent negative customer reactions and perceptions (Garbarino & Lee, 2003; Grewal et al., 2004) have so far deterred many companies due to the fear that the monetary gain will not exceed the loss of intangible assets, e.g., corporate reputation (Eberl & Schwaiger, 2005).

Current research is increasingly concerned with the search for ways to implement personalized pricing and temper negative customer reactions. One implementation method being discussed is price framing, which has already successfully led to the reversal of negative customer reactions as introduced by Weisstein et al. (2013). Personalized coupons as one particular framing method is already used in practice but poorly researched and could be a possible remedy. Muji an international operating Japanese lifestyle brand highlights this by showing that the introduction of personalized coupons can lead to positive results (Treasure Data, 2021).

In this study, we first looked at whether there is a certain range between different coupon values depicting a kind of indifference interval. Hence, in Study 1, a tolerated range was defined through a self-assessment by the study participants. Using the van Westendorp method, the evaluation showed that the tolerated range is almost identical regardless of the product price (low- versus high-priced product). Accordingly, customers perceive personalized coupons as fair as long as a friend receives a voucher of not over 25% for high-priced products (same for low-priced products) compared to their own 10% personalized voucher. Since this range was determined by means of a self-assessment, it was assumed that coupon differences within this range would be perceived as fair and thus not lead to negative customer reactions despite the personalized nature of the coupons. Yet, self-assessment does not serve as an indicator of perceived fairness. This is evident in our Study 2 experiment, where customers displayed negative reactions to this pricing strategy, even when individual coupon differences fell within the pre-established tolerated range.

Our second study was thus designed to test the fairness range assessed in Study 1 using an experimental design. Based on the current literature, the following hypotheses were derived:

Within the tolerated range, i.e., with medium-sized differences in coupon values, information processing would most strongly be stimulated and a greater consideration of the pricing strategy and background to different coupons would trigger unfairness. This, in turn, would lead to negative customer reactions. It was further assumed that particularly small and large differences between the coupons require significantly less information processing from the customer and therefore have a less negative influence on the perceived fairness and customer reactions. However, the experiment did not yield any significant results. Nevertheless, our data revealed interesting results regarding direct effects from personalized coupons to customer reactions, which we would like to discuss in the following.

Managerial Implications

If we look at the direct effect of individualized coupons on repurchase intention, we can see that the greater the spread between two coupons, the less likely a customer is to buy from the same retailer again. Accordingly, either coupons of the same value or only with minimal (below the range compared to same vouchers) differences trigger a higher repurchase intention. With regard to this result, companies should only implement personalized coupons with minimal differences if introducing such.

However, the question remains open as to whether minimal differences in coupons would lead to worthwhile increases in entrepreneurial profits. This is reminiscent of the previously mentioned overarching question regarding personalized prices and whether the monetary gains exceed the intangible losses. With regard to the risk of losing intangible assets, two interesting results have emerged.

First, the data shows that respondents would communicate their discomfort with the personalized voucher experience in a private setting. However, customers do not harbor any intentions of public complaint intentions. This reduces the reach of negative experience reports. Companies therefore do not have to expect any public outcry such as negative WOM. As WOM is a decisive antecedent for the development of intangible assets, such as reputation (Castellano & Dutot, 2017; Mahon, 2002; Shamma, 2012), trust (Hajli et al., 2014; Jalilvand et al., 2017) and loyalty (Manyanga et al., 2022), no major loss is to be expected in this respect.

In line with this, the data shows that customers with same vouchers, or vouchers below and within the range in reference to the baseline scenario remain loyal to the company. Only customers with a reference voucher that is above the fairness range feel less loyal to the company after the purchase. Therefore, it can be summarized that even if customers are rather dissatisfied with the purchasing experience due to personalized coupons, there is no effect on public complaints and in most cases also no effect on customer loyalty. One question that arises is whether this effect is due to a one-time experience or whether this effect could also be observed in the long term.

If the existing customer loyalty and the lack of public complaints continue in the long term, it can be assumed that the loss of intangible assets may be less than assumed. In this case, companies could be advised to use personalized coupons, as long as these lead to profit increases and the negative effects on long-term customer behavior are limited.

Theoretical Implications

This study was able to confirm the negative influence of personalized pricing strategies (regardless of the coupon framing format used) on perceived customer fairness previously reported in other studies (Hufnagel et al., 2022). The present results show that any difference to one's own coupon has a direct negative influence on perceived fairness. Consistent with the equity and distributive justice theory (Adams, 1965), consumers evaluate the ratio of their inputs and outputs in a transaction compared to the ratio of inputs and outputs of another person's transaction. When these ratios are equal, consumers tend to perceive a state of justice, i.e., they perceive fairness. If, on the other hand, the two ratios differ, consumers perceive inequality and therefore unfairness (Adams, 1965). Social comparison theory comes to the same conclusion (Festinger, 1954). When comparing prices, references to and comparisons with other customers are more pronounced and therefore lead to the highest perceptions of unfairness if the customer is disadvantaged by a price (or in this case a lower voucher and therefore a higher price) (Ashworth & McShane, 2012; Major & Testa, 1989).

In addition to the concrete result on customer perception, the twofold nature of this study shows that theoretically defined reference values (here the defined tolerance range), which were determined by the participants' self-assessment, lose their effect in practical situations. Although a clear tolerance range could be defined through Study 1, no perceived fairness within this range was observed in the experiment in Study 2. Accordingly, we recommend that results from self-reported studies should be validated through field experiments. Otherwise, there is a greater risk of response bias, as self-reporting is prone to hypothetical bias, social desirability, satisficing and other cognitive biases, among others, which could affect the validity of the survey (Hainmueller et al., 2015; Krosnick et al., 2014; Schwarz, 1999).

Limitations and Future Research

Following on from the last point of the theoretical implications, we have used a scenario experiment (or vignette study) in this paper. Although this increases internal validity (Dülmer, 2016) it only captures intentional behavior (Hainmueller et al., 2015) and not actual behavior, therefore lacking external validity. As can be seen from the results, there is a difference between self-assessed behavior and intentional behavior. The aforementioned biases can influence the participants and their behavior in hypothetical scenarios, as there are no costs or consequences attached to them (Bertrand & Mullainathan, 2001; Neill et al., 1994). Therefore, vignette studies only allow inferences to be drawn about the potential behavior of customers, but not about their real reactions (Hainmueller et al., 2015). Future research would be important and needed to observe whether the results prove similar in a field experiment or whether differences can be recognized.

Apart from the method chosen, which may have influenced the results, we were unfortunately unable to find any support for our hypotheses. This indicates that the transfer of the results of Ozanne et al. (1992), previously described as "risky", was not readily possible.

Moreover, in our study we pursue a one-dimensional approach and only consider disadvantaged customers. However, considering the results of Hufnagel et al., (2022) it is evident that both disadvantaged and advantaged customers perceive personalized pricing as unfair and react negatively to it. It would be interesting to see whether this result can also be reproduced with customized coupons, or whether the "gift character" prevails for advantaged customers.

Furthermore, for the experiment we recruited participants via a large German university. More than half of the participants were students (in Study 1: 55% and in Study 2: 71 %). Hence, our results are highly influenced by students as one particular societal group. Student samples are sometimes considered as convenience samples as they are cheaper than regular samples and are criticized as students represent a pool of individuals with specific characteristics that cannot always be generalized (Carter & Irons, 1991; Marwell & Ames, 1981). Especially regarding pricing, students might be more inclined to use vouchers as their financial situation is not comparable to that of professionals. Therefore, students might be influenced differently by vouchers. We suggest conducting a similar study with a regular sample representing the general population of interest.

In the underlying study, we further focused on a single framing method. We used percentage-off coupons for a low-priced product, as suggested in the literature (Weisstein et al., 2013).

However, another suitable framing strategy that works best for high-priced products is dollar-off coupons. Therefore, a similar study should be conducted using the dollar-off framing method to evaluate whether the results are similar or different to percentage-off coupons.

Lastly, and regarding the managerial implications, the results raise the question whether the monetary profit exceeds the loss of intangible assets, as some intangible assets such as customer loyalty do not seem to be damaged. Therefore, we see great potential in the measurement of average profit gains of companies introducing personalized pricing and the simultaneous measurement of medium to long-term customer reactions. Generalizable results from such a study should make it easier for organizations to decide whether to take the "risk" of personalized pricing.

Conclusion

The fairness range derived trough Study 1 could not be verified in Study 2. In addition to perceived fairness as a mediator, neither information processing as an underlying mechanism nor a moderating effect of need for cognition could be confirmed. It can be assumed that regardless of the method used (price framing vs. no framing) to apply personalized prices, perceived fairness is the decisive underlying mechanism. In the case of prices, we presume this means that fair is “equal” and unfair is “unequal” and therefore even personalized coupons with a "gift character" cannot turn the customer's perception into a positive one. Nevertheless, the study provides helpful approaches for further research and, above all, for practice, in which personalized coupons are often common practice.

Appendix

Appendix 1. Screener Question "Personalized Pricing Definition" - Study 1 and Study 2

Understanding **personalized pricing** is essential for the following scenario. Therefore, we would like to facilitate your comprehension by providing a definition: With **personalized pricing**, each customer receives an **individualized price for a standardized product**. The personalized price is determined with the help of new technologies incorporating various types of information (e.g., geographical location, purchasing behavior, search behavior, etc.). In this setting, **personalized discount coupons** are used to offer an **individual price to each customer.** A good example would be you and your best friend searching for the same product on the same website at the same time, **but both getting a different discount.**

Please indicate which of the following situations represents a company offering **personalized pricing.**

(1) You and your best friend are offered a different price at the same time, from the same retailer, for the same product.
(2) You and your best friend are offered the same price at the same time, from the same retailer, for the same product.

Appendix 2. Overview of Initial Scenarios - Study 1

Concert international

Please read the following scenario and questions carefully.

Imagine the following situation: Your favorite **international band** "The Rolling IMMs" is in town in a couple of weeks. They play at the Olympic Stadium in the city center and concert tickets cost **175 € per person**. Since you haven't been to a concert of theirs for a long time, you really want to go. The event organizer has sent you an email and attached a customized voucher for you. Based on your previous purchasing behavior you get a **personalized voucher for 10% off the concert tickets** as the event organizer of "The Rolling IMMs" is known to be using **personalized pricing** for a couple of years now.

Concert regional

Please read the following scenario and questions carefully.

Imagine the following situation: Your favorite **regional band** "The Rolling IMMs" is intown in a couple of weeks. They play at one of the bars in the city center and concert tickets cost **35 € per person**. Since you haven't been to a concert of theirs for a longtime, you really want to go. The event organizer has sent you an email and attached a customized voucher for you. Based on your previous purchasing behavior you get a **personalized voucher for 10% off the concert tickets** as the event organizer of "The Rolling IMMs" is known to be using **personalized pricing** for a couple of years now.

Appendix 3. Overview Van Westendorp Questions - Study 1

Now, your best friend calls you and tells you that s/he has heard about the concert as well and also wants to go. **Your best friend has also received a personalized voucher.**

Too Unfair Not Buy	In comparison to your coupon, at which **discount size** of the voucher **that your friend received** for the concert ticket would you consider the promotion **too unfair** and you would **not** buy the tickets yourself?	• 10% • 15% • 20% • 25% • 30% • 35% • 40% • 45% • 50% • 55% • 60%
Too Unfair But Buy	In comparison to your coupon, which **discount size** of the voucher **that your friend received** for the concert ticket would you consider **too unfair** but you would **still** buy the tickets yourself?	
Fair and Upset	In comparison to your coupon, which **discount size** of the voucher **that your friend received** for the concert ticket would you still consider **fair** but **you would still be upset that you didn't get such a high discount.**	
Fair Not Jealous	In comparison to your coupon, which **discount size** of the voucher **that your friend received** for the concert ticket would you consider **fair without any jealous feelings toward your friend?**	

Appendix 4. Overview of Measures and Stimuli

Construct/ Variable	Item/Proxy	Precedents/ Sources	Study
	Dependent Variables		
Satisfaction *(α = .96)*	This construct is measured a four-item differential scale. With the promotion strategy of the e-ticket seller, I am … (1) dissatisfied - satisfied (2) unhappy -happy (3) disappointed - delighted (4) displeased - pleased	Adapted from Darke and Dahl (2003) and Haws and Bearden (2006)	Study 2
Repurchase Intention *(α = .95)*	This construct is measured a four-item seven-point Likert-type scale from "very unlikely" to "very likely". Please remember the scenario from the beginning. If this situation happened to you, how likely would you be to ... (1) buy something from this e-ticket seller? (2) purchase event tickets from this e-ticket seller if you were looking for event tickets? (3) shop at this e-ticket seller in the future? (4) return to this e-ticket seller?	Adapted from Garbarino and Maxwell (2010)	Study 2
Private Complaint Intentions *(α = .76)*	This construct is measured a four-item seven-point Likert-type scale from "very unlikely" to "very likely". Please remember the scenario from the beginning. If this situation happened to you, how likely would you be to ... (1) forget about the incident? (2) decide not to use this e-ticket seller again? (3) speak to your friends and relatives in a negative way about your experience? (4) convince your friends and relatives not to use this e-ticket seller?	Adapted from Garbarino and Maxwell (2010) and Singh (1988)	Study 2
Public Complaint Intentions *(α = .83)*	This construct is measured a five-item seven-point Likert-type scale from "very unlikely" to "very likely". Please remember the scenario from the beginning. If this situation happened to you, how likely would you be to ... (1) complain to the e-ticket seller? (2) report the experience to a consumer agency? (3) complain to a consumer agency and ask them to make the firm address your concern? (4) write a letter to the local newspaper about your experience? (5) take some legal action against the online shop?	Adapted from Garbarino and Maxwell (2010) and Singh (1988)	Study 2

Appendix 4. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources	Study
Loyalty *(α = .89)*	This construct is measured a four-item seven-point Likert-type scale from "very unlikely" to "very likely". Please remember the scenario from the beginning. If this situation happened to you, how likely would you be to ... (1) consider yourself loyal to the e-ticket seller? (2) consider the same e-ticket seller to be your first choice for your next purchase? (3) not buy elsewhere if this e-ticket seller is available to you? (4) intend to continue to shop with the same e-ticket seller?	Adapted from B. Yoo and Donthu (2002) and (Lai et al., 2010)	Study 2
Information Processing *(α = .89)*	This construct is measured a six-item seven-point Likert-type scale from "strongly disagree" to "strongly agree". Please indicate the degree to which you agree/disagree with the following statements. (1) It was transparent to me which information was collected by the ticket seller. (2) The information that the ticket seller could acquire was observable for me. (3) It was understandable to me how the collected information led to the discount. (4) The ticket seller´s information processing was comprehensible to me. With the information accessible for me, the result was foreseeable for me.	Adapted from Schrills et al. (2022)	Study 2
Perceived Price Fairness *(α = .90)*	This construct is measured a three-item seven-point Likert-type scale from "strongly disagree" to "strongly agree". Compared to the value of the voucher I received, the **value of my friend's voucher** at the online event ticket seller is... (1) fair. (2) acceptable. (3) reasonable.	Adapted from Grewal et al. (2004), Weisstein et al. (2013), and L. Xia et al. (2004)	Study 2

Appendix 4. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources	Study
Need for Cognition *(α = .86)*	This construct is measured an eighteen-item seven-point Likert-type scale from "strongly disagree" to "strongly agree". Please indicate the degree to which you agree/disagree with the following statement. (1) I would prefer complex to simple problems. (2) I like to have the responsibility of handling a situation that requires a lot of thinking. (3) Thinking is not my idea of fun. (4) I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. (5) I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something. (6) I find satisfaction in deliberating hard and for long hours. (7) I only think as hard as I have to. (8) I prefer to think about small, daily projects rather than long-term ones. (9) I like tasks that require little thought once I've learned them. (10) The idea of relying on thought to make my way to the top appeals to me. (11) I really enjoy a task that involves coming up with new solutions to problems. (12) Learning new ways to think doesn't excite me very much. (13) I prefer my life to be filled with puzzles that I must solve. (14) The notion of thinking abstractly is appealing to me. (15) I would prefer a task that is intellectual, difficult and important to one that is somewhat important but does not require much thought. (16) I feel relief rather than satisfaction after completing a task that required a lot of mental effort. (17) It's enough for me that something gets the job done; I don't care how or why it works. (18) I usually end up deliberating about issues even when they do not affect me personally.	Adapted from Cacioppo et al. (1984)	Study 2

Appendix 4. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources	Study
	Manipulation Check		
High versus low price	The construct was measured on a one-item seven-point Likert type scale from "low price" to "high price". Compared to normal concert tickets, **175 € [35 €]** per concert ticket is a rather... (1) … low price. (2) … high price.	Self-developed	Study 1
Height Discount Voucher	The construct was measured on a two-item seven-point Likert type scale from "strongly disagree" to "strongly agree". Please indicate whether you agree with the following statement regarding the scenario. (1) My friend received a **higher discount voucher** than I did. (2) My friend received the **same discount voucher** as I did.	Self-developed	Study 1 Study 2
High or Low Price	The construct was measured on a one-item seven-point Likert type scale from "low price" to "high price". Please answer the following questions: Compared to normal event tickets, **35 €** per event ticket is a rather…. (1) … low price/…high price.	Self-developed	Study 2
	Attention Check		
Attention Check I	The construct is measured on a one-item seven-point scale from "strongly disagree" to "strongly agree". (1) Please select "strongly disagree" as your answer.	Adapted from Gruzd et al. (2020)	Study 1 Study 2
Attention Check II	The attention check was measured on a one-item scale, participants could either select "yes" or "no". Please indicate whether you agree with the following statements regarding **the scenario.** The initial price of the event tickets is **35 €.** (1) Yes (2) No	Self-developed	Study 2

Appendix 4. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources	Study
	Covariates and Controls		
Purchase Experience *(α = .89)*	The construct is measured on a three-item seven-point Likert type scale from "strongly disagree" to "strongly agree". Thinking about your purchasing habits, please indicate how strongly you agree with the following statements about **event tickets.** (1) I have a great deal of experience in buying **event tickets.** (2) I frequently shop for **event tickets.** (3) I am very confident in shopping for **event tickets.**	Adapted from Lo et al. (2019)	Study 2
Product Involvement *(α = .90)*	This construct is measured a seven-item differential scale. To me buying an **event ticket** is… (1) unimportant / important (2) boring / interesting (3) irrelevant / relevant (4) unexciting / exciting (5) unappealing / appealing (6) banal / fascinating (7) worthless / valuable	Adapted from Zaichkows-ky (1985)	Study 2
Attitude Personalization *(α = .85)*	The construct is measured on a four-item seven-point Likert type scale from "strongly disagree" to "strongly agree". Please state the extent to which you agree with the following statement? (1) For the e-ticket seller, using the personalized coupons is a good idea. (2) Using the personalized coupons is a wise idea for the e-ticket seller. (3) I like the idea of the e-ticket seller using the personalized coupons. (4) The e-ticket seller using the personalized coupons would be pleas-ant.	Adapted from Taylor and Todd (1995)	Study 2
Time Spent Online	The construct is measured on a one-item scale with five answer options. How much time do you spend online for private reasons? (1) Less than an hour a day (2) 1-2 hours a day (3) 2-4 hours a day (4) 4-6 hours a day more than 6 hours a day (5) More than 6 hours a day	Adapted from Doolin et al. (2005) and Malhotra et al. (2004)	Study 2
E-Purchase Frequency	The construct is measured on a one-item scale with five answer options. Please indicate the frequency of your online purchases in the last 12 months: (1) Never (2) Once or twice (3) 3-4 times (4) Monthly (5) Weekly	Adapted from Doolin et al. (2005) and Malhotra et al. (2004)	Study 2

Appendix 4. Overview of Measures and Stimuli (continued)

Construct/ Variable	Item/Proxy	Precedents/ Sources	Study
	Demographics		
Gender	Please indicate which gender you feel most closely aligned with: • Male • Female • Non-binary/third gender • Prefer not to say	Self-developed	Study 1 Study 2
Age	How old are you? (open question) ____________	Self-developed	Study 1 Study 2
Education	Please indicate your education level: • High school graduate • Bachelor's degree • Master's degree • Professional degree • Doctorate degree	Self-developed	Study 1 Study 2
Employment	Please indicate your employment status: • Employed full-time • Employed part-time • Self-employed • Homemaker • Student • Retired • Unemployed • Other	Self-developed	Study 1 Study 2
Income	What is the level of your annual gross household income? • < 10,000 € • 10,000 € – 50,000 € • 50,001 € – 90,000 € • 90,001 € - 150,000 € • >150,001 € • Prefer not to say	Adapted from Lo et al. (2019)	Study 1 Study 2

Appendix 5. Exemplary Python Code for Calculating the Range of Fairness Perceptions of the High-Priced Concert Tickets

```
import math
import sympy
from sympy.solvers import solve
from sympy import Symbol
from sympy import symbols, Eq, solve


#Daten
#fair
x1_fair = 10
y1_fair = 100

x2_fair = 15
y2_fair = 40

x3_fair = 20
y3_fair = 8

x4_fair = 25
y4_fair = 4

#fair and doubts
x1_fairanddoubts = 10
y1_fairanddoubts = 100

x2_fairanddoubts = 15
y2_fairanddoubts = 100

x3_fairanddoubts = 20
y3_fairanddoubts = 56

x4_fairanddoubts = 25
y4_fairanddoubts = 12

#too unfair BUT buy
x1_toounfairbutbuy = 10
y1_toounfairbutbuy = 0

x2_toounfairbutbuy = 15
y2_toounfairbutbuy = 0

x3_toounfairbutbuy = 20
y3_toounfairbutbuy = 44

x4_toounfairbutbuy = 25
y4_toounfairbutbuy = 72

#too unfair NOT buy
x1_toounfairnotbuy = 10
y1_toounfairnotbuy = 0

x2_toounfairnotbuy = 15
y2_toounfairnotbuy = 0

x3_toounfairnotbuy = 20
y3_toounfairnotbuy = 0

x4_toounfairnotbuy = 25
y4_toounfairnotbuy = 16
```

Appendix 5. Exemplary Python Code for Calculating the Range of Fairness Perceptions of the High-Priced Concert Tickets (continued)

```
#Intersection - Lower Bound Fairness - "fair" and "too unfair but buy"
##FAIR: Linear function fair between point 2 and 3
##Steigung m = ((y2-y1)/(x2-x1))
m_fair_1 = ((y3_fair-y2_fair)/(x3_fair-x2_fair))
##Y-Achsenabschnitt t = y - mx
t_fair_1 = y2_fair - m_fair_1 * x2_fair

##TOO UNFAIR BUT BUY: Linear function fair between point 2 and 3
##Steigung m = ((y2-y1)/(x2-x1))
m_toounfairbutbuy_1 = ((y3_toounfairbutbuy-y2_toounfairbutbuy)/(x3_toounfairbutbuy-x2_toounfairbutbuy))
##Y-Achsenabschnitt t = y - mx
t_toounfairbutbuy_1 = y2_toounfairbutbuy - m_toounfairbutbuy_1 * x2_toounfairbutbuy

#Intersection - Lower Bound Fairness - "fair" and "too unfair but buy"
x, y = symbols ("x y")
equation_1_fair = Eq((m_fair_1*x+t_fair_1),y)
equation_1_toounfairbutbuy = Eq((m_toounfairbutbuy_1*x+t_toounfairbutbuy_1),y)
print("Equation 1:", equation_1_fair)
print("Equation 2:", equation_1_toounfairbutbuy)
solution_lower_bound_fairness = solve ((equation_1_fair, equation_1_toounfairbutbuy), (x, y))

#Intersection - Upper Bound Fairness - "fair and doubts" and "too unfair not buy"
##FAIR AND DOUBTS: Linear function fair between point 3 and 4
##Steigung m = ((y2-y1)/(x2-x1))
m_fairanddoubts_1 = ((y4_fairanddoubts-y3_fairanddoubts)/(x4_fairanddoubts-x3_fairanddoubts))
##Y-Achsenabschnitt t = y - mx
t_fairanddoubts_1 = y4_fairanddoubts - m_fairanddoubts_1 * x4_fairanddoubts

##TOO UNFAIR NOT BUY: Linear function fair between point 3 and 4
##Steigung m = ((y2-y1)/(x2-x1))
m_toounfairnotbuy_1 = ((y4_toounfairnotbuy-y3_toounfairnotbuy)/(x4_toounfairnotbuy-x3_toounfairnotbuy))
##Y-Achsenabschnitt t = y - mx
t_toounfairnotbuy_1 = y3_toounfairnotbuy - m_toounfairnotbuy_1 * x3_toounfairnotbuy

#Intersection - Upper Bound Fairness - "fair and doubts" and "too unfair not buy"
x, y = symbols ("x y")
equation_1_fairanddoubts_1 = Eq((m_fairanddoubts_1*x+t_fairanddoubts_1),y)
equation_1_toounfairnotbuy_1 = Eq((m_toounfairnotbuy_1*x+t_toounfairnotbuy_1),y)
print("Equation 1:", equation_1_fairanddoubts_1)
print("Equation 2:", equation_1_toounfairnotbuy_1)
solution_upper_bound_fairness = solve ((equation_1_fairanddoubts_1, equation_1_toounfairnotbuy_1), (x, y))

#Intersection - Indifferent Fairness Point - "fair and doubts" and "too unfair but buy"
##FAIR AND DOUBTS: Linear function fair between point 3 and 4
##Steigung m = ((y2-y1)/(x2-x1))
#m_fairanddoubts_1 = ((y4_fairanddoubts-y3_fairanddoubts)/(x4_fairanddoubts-x3_fairanddoubts))
##Y-Achsenabschnitt t = y - mx
#t_fairanddoubts_1 = y4_fairanddoubts - m_fairanddoubts_1 * x4_fairanddoubts

##TOO UNFAIR BUT BUY: Linear function fair between point 3 and 4
##Steigung m = ((y2-y1)/(x2-x1))
m_toounfairbutbuy_1 = ((y4_toounfairbutbuy-y3_toounfairbutbuy)/(x4_toounfairbutbuy-x3_toounfairbutbuy))
##Y-Achsenabschnitt t = y - mx
t_toounfairbutbuy_1 = y4_toounfairbutbuy - m_toounfairbutbuy_1 * x4_toounfairbutbuy

#Intersection - Indifferent Fairness Point - "fair and doubts" and "too unfair but buy"
x, y = symbols ("x y")
equation_1_fairanddoubts_1 = Eq((m_fairanddoubts_1*x+t_fairanddoubts_1),y)
equation_1_toounfairbutbuy_1 = Eq((m_toounfairbutbuy_1*x+t_toounfairbutbuy_1),y)
print("Equation 1:", equation_1_fairanddoubts_1)
print("Equation 2:", equation_1_toounfairbutbuy_1)
solution_indifferent_fairness_point = solve ((equation_1_fairanddoubts_1, equation_1_toounfairbutbuy_1), (x, y))
```

Appendix 5. Exemplary Python Code for Calculating the Range of Fairness Perceptions of the High-Priced Concert Tickets (continued)

```
#Intersection - Optimal Fairness Point - "fair" and "too unfair not buy"
##FAIR: Linear function fair between point 3 and 4
##Steigung m = ((y2-y1)/(x2-x1))
m_fair_2 = ((y4_fair-y3_fair)/(x4_fair-x3_fair))
##Y-Achsenabschnitt t = y - mx
t_fair_2 = y4_fair - m_fair_2 * x4_fair

##TOO UNFAIR NOT BUY: Linear function fair between point 3 and 4


#Intersection - Optimal Fairness Point - "fair" and "too unfair not buy"
x, y = symbols ("x y")
equation_1_fair_2 = Eq((m_fair_2*x+t_fair_2),y)
equation_1_toounfairnotbuy = Eq((m_toounfairnotbuy_1*x+t_toounfairnotbuy_1),y)
print("Equation 1:", equation_1_fair_2)
print("Equation 2:", equation_1_toounfairnotbuy)
solution_optimal_fairness_point = solve ((equation_1_fair_2, equation_1_toounfairnotbuy), (x, y))

print("Solution Lower Bound Fairness:", solution_lower_bound_fairness)
print("Solution Upper Bound Fairness:", solution_upper_bound_fairness)
print("Solution Indifferent Fairness Point:", solution_indifferent_fairness_point)
print("Solution Optimal Fairness Point:", solution_optimal_fairness_point)
```

Appendix 6. Screener Question "Purchase of Tickets" - Study 2

Please read the following scenario and questions carefully.

Imagine the following situation: You are looking online for a ticket for a specific event. The regular **price is 35€.** Based on your purchase history as well as external information, such as from your social media activity, your **favorite online event ticket** seller offers you **a personalized voucher for 10% off the regular price**. The online event ticket seller is known to be **using personalized pricing.**

Based on the just read information, would you buy the event ticket using your 10% voucher?

(1) Yes, I would buy the ticket.
(2) No, I would not buy the ticket.

Appendix 7. Overview of Scenarios – Study 2

Same Voucher	As you decided **to buy the event ticket using the 10% voucher** you are now looking forward to the upcoming event. Shortly after, you talk to your **best friend** about the purchase, as s/he is **also a frequent customer** of the same online ticket seller. It turns out that your friend was **also offered a voucher on the same day by the same online ticket seller for the same event**. Your friend's voucher was **10%**. Therefore, your friend was offered a voucher in the **same amount.**
Optimal Difference Voucher	As you decided **to buy the event ticket using the 10% voucher** you are now looking forward to the upcoming event. Shortly after, you talk to your **best friend** about the purchase, as s/he is **also a frequent customer** of the same online ticket seller. It turns out that your friend was **also offered a voucher on the same day by the same online ticket seller for the same event**. Your friend's voucher was **22%**. Therefore, your friend was offered **a higher value** voucher.
Below Range Voucher	As you decided **to buy the event ticket using the 10% voucher** you are now looking forward to the upcoming event. Shortly after, you talk to your **best friend** about the purchase, as s/he is **also a frequent customer** of the same online ticket seller. It turns out that your friend was **also offered a voucher on the same day by the same online ticket seller for the same event**. Your friend's voucher was **12%**. Therefore, your friend was offered **a higher value** voucher.
Above Range Voucher	As you decided **to buy the event ticket using the 10% voucher** you are now looking forward to the upcoming event. Shortly after, you talk to your **best friend** about the purchase, as s/he is **also a frequent customer** of the same online ticket seller. It turns out that your friend was **also offered a voucher on the same day by the same online ticket seller for the same event**. Your friend's voucher was **32%**. Therefore, your friend was offered **a higher value** voucher.

Appendix 8. Comparison of Voucher Type on Satisfaction – Study 2

			95% CI	
	Mean	SE	Lower Bound	Upper Bound
same voucher	4.936[a]	0.187	4.567	5.305
below range	3.310[a]	0.182	2.951	3.669
optimal range	2.626[a]	0.179	2.273	2.978
above range	2.612[a]	0.178	2.261	2.963

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

					95% CI for Difference[b]	
		Mean Difference (I-J)	SE	Sig.[b]	Lower Bound	Upper Bound
same voucher	*below range*	1.626*	0.267	0.000	0.915	2.337
	optimal range	2.310*	0.266	0.000	1.601	3.020
	above range	2.324*	0.259	0.000	1.634	3.014
below range	*same voucher*	-1.626*	0.267	0.000	-2.337	-0.915
	optimal range	.684*	0.253	0.045	0.010	1.359
	above range	.698*	0.255	0.041	0.017	1.378
optimal range	*same voucher*	-2.310*	0.266	0.000	-3.020	-1.601
	below range	-.684*	0.253	0.045	-1.359	-0.010
	above range	0.013	0.252	1.000	-0.657	0.684
above range	*same voucher*	-2.324*	0.259	0.000	-3.014	-1.634
	below range	-.698*	0.255	0.041	-1.378	-0.017
	optimal range	-0.013	0.252	1.000	-0.684	0.657

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 9. Comparison of Voucher Type on Repurchase Intention – Study 2

			95% CI	
	Mean	SE	Lower Bound	Upper Bound
same voucher	4.983[a]	0.188	4.612	5.354
below range	4.361[a]	0.183	4.000	4.722
optimal range	3.941[a]	0.180	3.586	4.295
above range	3.601[a]	0.179	3.249	3.954

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

					95% CI for Difference[b]	
		Mean Difference (I-J)	SE	Sig.[b]	Lower Bound	Upper Bound
same voucher	*below range*	0.622	0.268	0.128	-0.092	1.337
	optimal range	1.043*	0.267	0.001	0.329	1.756
	above range	1.382*	0.260	0.000	0.689	2.075
below range	*same voucher*	-0.622	0.268	0.128	-1.337	0.092
	optimal range	0.420	0.254	0.602	-0.258	1.098
	above range	.759*	0.257	0.021	0.075	1.443
optimal range	*same voucher*	-1.043*	0.267	0.001	-1.756	-0.329
	below range	-0.420	0.254	0.602	-1.098	0.258
	above range	0.339	0.253	1.000	-0.335	1.014
above range	*same voucher*	-1.382*	0.260	0.000	-2.075	-0.689
	below range	-.759*	0.257	0.021	-1.443	-0.075
	optimal range	-0.339	0.253	1.000	-1.014	0.335

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 10. Comparison of Voucher Type on Private Complaint Intentions – Study 2

	Mean	SE	95% CI Lower Bound	95% CI Upper Bound
same voucher	3.321[a]	0.160	3.005	3.637
below range	3.995[a]	0.156	3.686	4.303
optimal range	4.511[a]	0.153	4.208	4.814
above range	4.454[a]	0.152	4.153	4.754

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

		Mean Diffe-rence (I-J)	SE	Sig.[b]	95% CI for Difference[b] Lower Bound	95% CI for Difference[b] Upper Bound
same voucher	*below range*	-.674*	0.229	0.022	-1.283	-0.064
	optimal range	-1.190*	0.228	0.000	-1.799	-0.582
	above range	-1.133*	0.222	0.000	-1.724	-0.541
below range	*same voucher*	.674*	0.229	0.022	0.064	1.283
	optimal range	-0.516	0.217	0.110	-1.095	0.062
	above range	-0.459	0.219	0.224	-1.043	0.125
optimal range	*same voucher*	1.190*	0.228	0.000	0.582	1.799
	below range	0.516	0.217	0.110	-0.062	1.095
	above range	0.058	0.216	1.000	-0.518	0.633
above range	*same voucher*	1.133*	0.222	0.000	0.541	1.724
	below range	0.459	0.219	0.224	-0.125	1.043
	optimal range	-0.058	0.216	1.000	-0.633	0.518

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 11. Comparison of Voucher Type on Public Complaint Intentions – Study 2

			95% CI	
	Mean	SE	Lower Bound	Upper Bound
same voucher	1.775[a]	0.147	1.485	2.066
below range	2.048[a]	0.143	1.766	2.331
optimal range	1.999[a]	0.141	1.721	2.277
above range	2.215[a]	0.140	1.939	2.491

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

					95% CI for Difference[a]	
		Mean Difference (I-J)	SE	Sig.[a]	Lower Bound	Upper Bound
same voucher	*below range*	-0.273	0.210	1.000	-0.832	0.286
	optimal range	-0.224	0.209	1.000	-0.782	0.334
	above range	-0.439	0.203	0.192	-0.982	0.103
below range	*same voucher*	0.273	0.210	1.000	-0.286	0.832
	optimal range	0.049	0.199	1.000	-0.481	0.580
	above range	-0.166	0.201	1.000	-0.702	0.369
optimal range	*same voucher*	0.224	0.209	1.000	-0.334	0.782
	below range	-0.049	0.199	1.000	-0.580	0.481
	above range	-0.216	0.198	1.000	-0.744	0.312
above range	*same voucher*	0.439	0.203	0.192	-0.103	0.982
	below range	0.166	0.201	1.000	-0.369	0.702
	optimal range	0.216	0.198	1.000	-0.312	0.744

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Appendix 12. Comparison of Voucher Type on Loyalty – Study 2

	Mean	SE	95% CI Lower Bound	95% CI Upper Bound
same voucher	3.822[a]	0.169	3.488	4.156
below range	3.291[a]	0.165	2.966	3.616
optimal range	3.260[a]	0.162	2.941	3.580
above range	2.760[a]	0.161	2.442	3.077

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

		Mean Difference (I-J)	SE	Sig.[b]	95% CI for Difference[b] Lower Bound	95% CI for Difference[b] Upper Bound
same voucher	*below range*	0.531	0.241	0.174	-0.113	1.175
	optimal range	0.562	0.241	0.125	-0.081	1.204
	above range	1.062*	0.234	0.000	0.438	1.687
below range	*same voucher*	-0.531	0.241	0.174	-1.175	0.113
	optimal range	0.031	0.229	1.000	-0.580	0.642
	above range	0.531	0.231	0.135	-0.085	1.147
optimal range	*same voucher*	-0.562	0.241	0.125	-1.204	0.081
	below range	-0.031	0.229	1.000	-0.642	0.580
	above range	0,501	0.228	0.175	-0.107	1.108
above range	*same voucher*	-1.062*	0.234	0.000	-1.687	-0.438
	below range	-0.531	0.231	0.135	-1.147	0.085
	optimal range	-0.501	0.228	0.175	-1.108	0.107

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 13. Comparison of Voucher Type on Fairness Perceptions – Study 2

	Mean	SE	95% CI Lower Bound	95% CI Upper Bound
same voucher	5.168[a]	0.181	4.810	5.526
below range	3.953[a]	0.177	3.605	4.302
optimal range	3.352[a]	0.174	3.009	3.694
above range	3.315[a]	0.172	2.975	3.655

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

		Mean Difference (I-J)	SE	Sig.[b]	95% CI for Difference[b] Lower Bound	95% CI for Difference[b] Upper Bound
same voucher	*below range*	1.215*	0.259	0.000	0.525	1.904
	optimal range	1.817*	0.258	0.000	1.128	2.505
	above range	1.853*	0.251	0.000	1.184	2.522
below range	*same voucher*	-1.215*	0.259	0.000	-1.904	-0.525
	optimal range	0.602	0.245	0.091	-0.053	1.256
	above range	0.638	0.248	0.064	-0.022	1.298
optimal range	*same voucher*	-1.817*	0.258	0.000	-2.505	-1.128
	below range	-0.602	0.245	0.091	-1.256	0.053
	above range	0.036	0.244	1.000	-0.615	0.687
above range	*same voucher*	-1.853*	0.251	0.000	-2.522	-1.184
	below range	-0.638	0.248	0.064	-1.298	0.022
	optimal range	-0.036	0.244	1.000	-0.687	0.615

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 14. Comparison of Voucher Type on Information Processing – Study 2

	Mean	SE	95% CI Lower Bound	95% CI Upper Bound
same voucher	2.984[a]	0.165	2.660	3.309
below range	2.463[a]	0.160	2.146	2.779
optimal range	2.535[a]	0.157	2.225	2.846
above range	2.285[a]	0.157	1.976	2.594

a. Covariates appearing in the model are evaluated at the following values: SCALE_Experience_Purchase = 3.5357, SCALE_Product_Involvement_2 = 4.2041, SCALE_Attitude_Personalization = 4.1543, SCALE_Time_Spent_Online = 3.08, SCALE_E_Purchase_Frequency = 3.86.

		Mean Difference (I-J)	SE	Sig.[b]	95% CI for Difference[b] Lower Bound	95% CI for Difference[b] Upper Bound
same voucher	*below range*	0.522	0.235	0.165	-0.104	1,147
	optimal range	0.449	0.234	0.341	-0.176	1.074
	above range	.700*	0.228	0.015	0.092	1.307
below range	*same voucher*	-0.522	0.235	0.165	-1.147	0.104
	optimal range	-0.073	0.223	1.000	-0.667	0.521
	above range	0.178	0.225	1.000	-0.421	0.777
optimal range	*same voucher*	-0.449	0.234	0.341	-1.074	0.176
	below range	0.073	0.223	1.000	-0.521	0.667
	above range	0.251	0.222	1.000	-0.340	0.842
above range	*same voucher*	-.700*	0.228	0.015	-1.307	-0.092
	below range	-0.178	0.225	1.000	-0.777	0.421
	optimal range	-0.251	0.222	1.000	-0.842	0.340

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Appendix 15. Mediation Analysis on Satisfaction

Model Summary for Outcome Variable Satisfaction

R	R-sq	MSE	F(HC3)	df1	df2	p
0.7759	0.602	1.2575	30.2384	10	185	0

	coeff	se(HC3)	t	p	LLCI	ULCI
constant	1.4845	0.67	2.2155	0.0279	0.1626	2.8064
X1	-1.043	0.2603	-4.0069	0.0001	-1.5566	-0.5295
X2	-1.5132	0.2826	-5.3546	0	-2.0707	-0.9557
X3	-1.4563	0.3079	-4.7304	0	-2.0636	-0.8489
Information Processing	0.2253	0.0871	2.5854	0.0105	0.0534	0.3972
Fairness	0.3832	0.0736	5.209	0	0.2381	0.5284
Purchase Experience	-0.1316	0.0627	-2.0979	0.0373	-0.2553	-0.0078
Product Involvement	0.0279	0.0788	0.3537	0.724	-0.1277	0.1834
Attitude Toward Personalization	0.1955	0.0783	2.4963	0.0134	0.041	0.35
Time Spent Online	0.1369	0.0751	1.8224	0.07	-0.0113	0.2852
E-Purchase Frequency	-0.0225	0.1003	-0.2245	0.8226	-0.2204	0.1754

Direct and Indirect Effects of X on Y

Relative direct effects of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
X1	-1.043	0.2603	-4.0069	0.0001	-1.5566	-0.5295
X2	-1.5132	0.2826	-5.3546	0	-2.0707	-0.9557
X3	-1.4563	0.3079	-4.7304	0	-2.0636	-0.8489

Relative indirect effects of X on Y

	Effect	BootSE	BootLLCI	BootULCI
Group → Information Processing → Satisfaction				
X1	-0.1175	0.0742	-0.2862	-0.0017
X2	-0.1011	0.0746	-0.279	0.0084
X3	-0.1576	0.0836	-0.3383	-0.0171
Group → Fairness Perceptions → Satisfaction				
X1	-0.4441	0.1398	-0.752	-0.2059
X2	-0.6777	0.1737	-1.0499	-0.375
X3	-0.6814	0.1763	-1.0496	-0.3647
Group → Information Processing → Fairness Perceptions → Satisfaction				
X1	-0.0214	0.0218	-0.0774	0.0071
X2	-0.0184	0.0207	-0.0722	0.0073
X3	-0.0287	0.0259	-0.0908	0.0095

Appendix 16. Mediation Analysis on Repurchase Intention

Model Summary for Outcome Variable Intention

R	R-sq	MSE	F(HC3)	df1	df2	p
0.6846	0.4687	1.34	22.5533	10	185	0

	coeff	se(HC3)	t	p	LLCI	ULCI
constant	1.5747	0.5817	2.7069	0.0074	0.427	2.7223
X1	-0.1078	0.2405	-0.4482	0.6545	-0.5823	0.3667
X2	-0.3094	0.3059	-1.0116	0.313	-0.9128	0.294
X3	-0.6073	0.2907	-2.0893	0.0381	-1.1809	-0.0338
Information Processing	0.1103	0.0835	1.3212	0.1881	-0.0544	0.275
Fairness	0.3763	0.0812	4.6339	0	0.2161	0.5365
Purchase Experience	-0.0825	0.0607	-1.3603	0.1754	-0.2022	0.0372
Product Involvement	0.0082	0.0849	0.0969	0.9229	-0.1593	0.1758
Attitude Toward Personalization	0.3223	0.0833	3.8699	0.0002	0.158	0.4866
Time Spent Online	0.097	0.0846	1.147	0.2529	-0.0699	0.264
E-Purchase Frequency	-0.0637	0.1029	-0.6189	0.5367	-0.2668	0.1394

Direct and Indirect Effects of X on Y

Relative direct effects of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
X1	-.1078	0.2405	-0.4482	0.6545	-0.5823	0.3667
X2	-.3094	0.3059	-1.0116	0.313	-0.9128	0.294
X3	-.6073	0.2907	-2.0893	0.0381	-1.1809	-0.0338

Relative indirect effects of X on Y

	Effect	BootSE	BootLLCI	BootULCI
Group → Information Processing → Repurchase Intention				
X1	-0.0575	0.0547	-0.1856	0.0263
X2	-0.0495	0.0525	-0.1842	0.0216
X3	-0.0772	0.0658	-0.2258	0.035
Group → Fairness Perceptions → Repurchase Intention				
X1	-0.4361	0.1482	-0.7615	-0.1826
X2	-0.6655	0.1882	-1.0658	-0.3324
X3	-0.6691	0.1853	-1.0653	-0.3393
Group → Information Processing → Fairness Perceptions → Repurchase Intention				
X1	-0.021	0.0216	-0.0751	0.0067
X2	-0.0181	0.0204	-0.0703	0.0072
X3	-0.0282	0.0256	-0.0898	0.0088

Appendix 17. Mediation Analysis on Private Complaint Intentions

Model Summary for Outcome Private Complaint Intentions

R	R-sq	MSE	F(HC3)	df1	df2	p
0.6441	0.4149	1.0323	15.0902	10	185	0

	coeff	se(HC3)	t	p	LLCI	ULCI
constant	5.6831	0.6109	9.3023	0	4.4778	6.8884
X1	0.3296	0.2388	1.3799	0.1693	-0.1416	0.8008
X2	0.6816	0.2583	2.6392	0.009	0.1721	1.1911
X3	0.6095	0.2738	2.2262	0.0272	0.0693	1.1496
Information Processing	-0.0181	0.0789	-0.2291	0.8191	-0.1738	0.1376
Fairness	-0.2755	0.067	-4.112	0.0001	-0.4077	-0.1433
Purchase Experience	0.0519	0.0525	0.9894	0.3237	-0.0516	0.1555
Product Involvement	0.0087	0.0611	0.1431	0.8864	-0.1117	0.1292
Attitude Toward Personalization	-0.2474	0.0743	-3.3294	0.0011	-0.3941	-0.1008
Time Spent Online	-0.0255	0.0791	-0.3224	0.7475	-0.1815	0.1305
E-Purchase Frequency	0.0005	0.09	0.0051	0.9959	-0.1772	0,1781

Direct and Indirect Effects of X on Y

Relative direct effects of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
X1	0.3296	0.2388	1.3799	0.1693	-0.1416	0.8008
X2	0.6816	0.2583	2.6392	0.009	0.1721	1.1911
X3	0.6095	0.2738	2.2262	0.0272	0.0693	1.1496

Relative indirect effects of X on Y

	Effect	BootSE	BootLLCI	BootULCI
Group → Information Processing → Private Complaint Intentions				
X1	0.0094	0.0432	-0.0798	0.1009
X2	0.0081	0.0398	-0.0714	0.0969
X3	0.0126	0.0561	-0.1077	0.1199
Group → Fairness Perceptions → Private Complaint Intentions				
X1	0.3193	0.1091	0.1308	0.5626
X2	0.4873	0.143	0.2344	0.793
X3	0.4899	0.1439	0.2358	0.7909
Group → Information Processing → Fairness Perceptions → Private Complaint Intentions				
X1	0.0154	0.0165	-0.0052	0.0579
X2	0.0132	0.0157	-0.0054	0.0547
X3	0.0206	0.0194	-0.0073	0.0675

Appendix 18. Mediation Analysis on Public Complaint Intentions

Model Summary for Outcome Public Complaint Intentions

R	R-sq	MSE	F(HC3)	df1	df2	p
0.3448	0.1189	0.9211	2.6308	10	185	0.0051

	coeff	se(HC3)	t	p	LLCI	ULCI
constant	1.9939	0.6607	3.0178	0.0029	0.6904	3.2974
X1	0.1268	0.2202	0.5757	0.5655	-0.3077	0.5613
X2	-0.0237	0.2313	-0.1024	0.9186	-0.4801	0.4327
X3	0.208	0.2311	0.9004	0.3691	-0.2478	0.6639
Information Processing	0.0865	0.0623	1.3892	0.1664	-0.0363	0.2093
Fairness	-0.1575	0.0553	-2.8489	0.0049	-0.2665	-0.0484
Purchase Experience	0.0438	0.0534	0.8212	0.4126	-0.0614	0.1491
Product Involvement	0.0909	0.0708	1.2828	0.2012	-0.0489	0.2306
Attitude Toward Personalization	-0.0584	0.0634	-0.9217	0.3579	-0.1834	0.0666
Time Spent Online	0.0612	0.0847	0.7227	0.4708	-0.1059	0.2283
E-Purchase Frequency	-0.0376	0.0866	-0.4345	0.6644	-0.2085	0.1332

Direct and Indirect Effects of X on Y

Relative direct effects of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
X1	0.1268	0.2202	0.5757	0.5655	-0.3077	0.5613
X2	-0.0237	0.2313	-0.1024	0.9186	-0.4801	0.4327
X3	0.208	0.2311	0.9004	0.3691	-0.2478	0.6639

Relative indirect effects of X on Y

	Effect	BootSE	BootLLCI	BootULCI
Group → Information Processing → Public Complaint Intentions				
X1	-0.0451	0.0382	-0.1341	0.0154
X2	-0.0388	0.0356	-0.1194	0.0186
X3	-0.0605	0.0468	-0.1644	0.0173
Group → Fairness Perceptions → Public Complaint Intentions				
X1	0.1825	0.0756	0.0515	0.348
X2	0.2785	0.1028	0.0881	0.4873
X3	0.28	0.1069	0.0857	0.5072
Group → Information Processing → Fairness Perceptions → Public Complaint Intentions				
X1	0.0088	0.0101	-0.0025	0.0363
X2	0.0076	0.0097	-0.0027	0.0336
X3	0.0118	0.0122	-0.0035	0.0432

Appendix 19. Mediation Analysis on Loyalty

Model Summary for Outcome Loyalty

R	R-sq	MSE	F(HC3)	df1	df2	p
0.6397	0.4092	0.1006	16.6815	10	185	0

	coeff	se(HC3)	t	p	LLCI	ULCI
constant	0.8515	0.6511	1.3077	0.1926	-0.4331	2.1361
X1	-0.0825	0.2563	-0.322	0.7478	-0.5881	0.4231
X2	0.0702	0.2659	0.2639	0.7921	-0.4544	0.5948
X3	-0.3895	0.2726	-1.4289	0.1547	-0.9274	0.1483
Information Processing	0.1171	0.0834	1.4035	0.1621	-0.0475	0.2817
Fairness	0.3189	0.0686	4.6518	0	0.1836	0.4541
Purchase Experience	0.0274	0.0591	0.4642	0.643	-0.0891	0.144
Product Involvement	0.0306	0.0729	0.4201	0.6749	-0.1132	0.1744
Attitude Toward Personalization	0.2436	0.074	3.2931	0.0012	0.0977	0.3895
Time Spent Online	0.0031	0.0829	0.0373	0.9703	-0.1605	0.1667
E-Purchase Frequency	-0.0711	0.1079	-0.6587	0.5109	-0.284	0.1418

Direct and Indirect Effects of X on Y

Relative direct effects of X on Y

	Effect	se(HC3)	t	p	LLCI	ULCI
X1	-.0825	0.2563	-0.322	0.7478	-0.5881	0.4231
X2	.0702	0.2659	0.2639	0.7921	-0.4544	0.5948
X3	-.3895	0.2726	-1.4289	0.1547	-0.9274	0.1483

Relative indirect effects of X on Y

	Effect	BootSE	BootLLCI	BootULCI
Group → Information Processing → Loyalty				
X1	-0.0611	0.0576	-0.2015	0.0203
X2	-0.0526	0.0555	-0.1916	0.0199
X3	-0.0819	0.0674	-0.235	0.0297
Group → Fairness Perceptions → Loyalty				
X1	-0.3696	0.1201	-0.6289	-0.1613
X2	-0.5639	0.1571	-0.9093	-0.2895
X3	-0.567	0.159	-0.9044	-0.2852
Group → Information Processing → Fairness Perceptions → Loyalty				
X1	-0.0178	0.0187	-0.0667	0.0057
X2	-0.0153	0.0179	-0.0612	0.0068
X3	-0.0239	0.0224	-0.0794	0.0081

References

Adams, J. S. (1965). Inequity In Social Exchange. *Advances in Experimental Social Psychology*, *2*, 267–299. https://doi.org/10.1016/S0065-2601(08)60108-2

Arora, N., Dreze, X., Ghose, A., Hess, J. D., Iyengar, R., Jing, B., Joshi, Y., Kumar, V., Lurie, N., Neslin, S., Sajeesh, S., Su, M., Syam, N., Thomas, J., & Zhang, Z. J. (2008). Putting One-to-One Marketing to Work: Personalization, Customization, and Choice. *Marketing Letters*, *19*(3-4), 305–321. https://doi.org/10.1007/s11002-008-9056-z

Ashworth, L., & McShane, L. (2012). Why Do We Care What Others Pay? The Effect of Other Consumers' Prices on Inferences of Seller (Dis)Respect and Perceptions of Deservingness Violation. *Journal of Retailing*, *88*(1), 145–155. https://doi.org/10.1016/j.jretai.2011.10.004

Aviv, Y., & Pazgal, A. (2008). Optimal Pricing of Seasonal Products in the Presence of Forward-Looking Consumers. *Manufacturing & Service Operations Management*, *10*(3), 339–359. https://doi.org/10.1287/msom.1070.0183

Baik, S., Anderson, A., & Larson, N. (2023). Price Discrimination in the Information Age: Prices, Poaching, and Privacy with Personalized Targeted Discounts. *The Review of Economic Studies*, *90*(5), 2085–2115. https://doi.org/10.1093/restud/rdac073

Barclay, L. J., Bashshur, M. R., & Fortin, M. (2017). Motivated Cognition and Fairness: Insights, Integration, and Creating a Path Forward. *Journal of Applied Psychology*, *102*(6), 867–889. https://doi.org/10.1037/apl0000204

Barone, M. J., & Roy, T. (2010). Does Exclusivity Always Pay Off? Exclusive Price Promotions and Consumer Response. *Journal of Marketing*, *74*(2), 121–132. https://doi.org/10.1509/jm.74.2.121

Bertrand, M., & Mullainathan, S. (2001). Do People Mean What They Say? Implications for Subjective Survey Data. *American Economic Review*, *91*(2), 67–72. https://doi.org/10.1257/aer.91.2.67

Bolton, G. E., & Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, *90*(1), 166–193. https://doi.org/10.1257/aer.90.1.166

Bolton, L. E., Keh, H. T., & Alba, J. W. (2010). How Do Price Fairness Perceptions Differ Across Culture? *Journal of Marketing Research*, *47*(3), 564–576. https://doi.org/10.1509/jmkr.47.3.564

Bolton, L. E., Warlop, L., & Alba, J. W. (2003). Consumer Perceptions of Price (Un)Fairness. *Journal of Consumer Research*, *29*(4), 474–491. https://doi.org/10.1086/346244

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition. *Psychological Bulletin*, *119*(2), 197–253. https://doi.org/10.1037/0033-2909.119.2.197

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Calvano, E., Calzolari, G., Denicolò, V., & Pastorello, S. (2019). Algorithmic Pricing What Implications for Competition Policy? *Review of Industrial Organization*, *55*(1), 155–171. https://doi.org/10.1007/s11151-019-09689-3

Campbell, M. C. (2007). "Says Who?!" How the Source of Price Information and Affect Influence Perceived Price (Un)fairness. *Journal of Marketing Research*, *44*(2), 261–271. https://doi.org/10.1509/jmkr.44.2.261

Carter, J. R., & Irons, M. D. (1991). Are Economists Different, and If So, Why? *Journal of Economic Perspectives*, *5*(2), 171–177. https://doi.org/10.1257/jep.5.2.171

Castellano, S., & Dutot, V. (2017). Investigating the Influence of E-Word-of-Mouth on E-Reputation. *International Studies of Management & Organization*, *47*(1), 42–60. https://doi.org/10.1080/00208825.2017.1241088

Chandon, P., Wansink, B., & Laurent, G. (2000). A Benefit Congruency Framework of Sales Promotion Effectiveness. *Journal of Marketing*, *64*(4), 65–81. https://doi.org/10.1509/jmkg.64.4.65.18071

Chandra, S., Verma, S., Lim, W. M., Kumar, S., & Donthu, N. (2022). Personalization in Personalized Marketing: Trends and Ways Forward. *Psychology & Marketing*, *39*(8), 1529–1562. https://doi.org/10.1002/mar.21670

Chen, H., Marmorstein, H., Tsiros, M., & Rao, A. R. (2012). When More is Less: The Impact of Base Value Neglect on Consumer Preferences for Bonus Packs over Price Discounts. *Journal of Marketing*, *76*(4), 64–77. https://doi.org/10.1509/jm.10.0443

Chen, H., & Rao, A. R. (2007). When Two Plus Two Is Not Equal to Four: Errors in Processing Multiple Percentage Changes. *Journal of Consumer Research*, *34*(3), 327–340. https://doi.org/10.1086/518531

Chen, M., & Chen, Z.-L. (2015). Recent Developments in Dynamic Pricing Research: Multiple Products, Competition, and Limited Demand Information. *Production and Operations Management*, *24*(5), 704–731. https://doi.org/10.1111/poms.12295

Chen, N., & Gallego, G. (2019). Welfare Analysis of Dynamic Pricing. *Management Science*, *65*(1), 139–151. https://doi.org/10.1287/mnsc.2017.2943

Chen, S. S., Monroe, K. B., & Lou, Y.-C. (1998). The Effects of Framing Price Promotion Messages on Consumers' Perceptions and Purchase Intentions. *Journal of Retailing*, *74*(3), 353–372. https://doi.org/10.1016/S0022-4359(99)80100-6

Chen, Y., & Farias, V. F. (2018). Robust Dynamic Pricing with Strategic Customers. *Mathematics of Operations Research*, *43*(4), 1119–1142. https://doi.org/10.1287/moor.2017.0897

Chen, Y., Farias, V. F., & Trichakis, N. (2019). On the Efficacy of Static Prices for Revenue Management in the Face of Strategic Customers. *Management Science*, *65*(12), 5535–5555. https://doi.org/10.1287/mnsc.2018.3203

Chiou-Wei, S.-Z., & Inman, J. J. (2008). Do Shoppers Like Electronic Coupons? *Journal of Retailing*, *84*(3), 297–307. https://doi.org/10.1016/j.jretai.2008.07.003

Choe, C., Matsushima, N., & Tremblay, M. J. (2022). Behavior-Based Personalized Pricing: When Firms can Share Customer Information. *International Journal of Industrial Organization*, *82*, 102846. https://doi.org/10.1016/j.ijindorg.2022.102846

Choudhary, V., Ghose, A., Mukhopadhyay, T., & Rajan, U. (2005). Personalized Pricing and Quality Differentiation. *Management Science*, *51*(7), 1120–1130. https://doi.org/10.1287/mnsc.1050.0383

Cohen, M. C. (2018). Big Data and Service Operations. *Production and Operations Management*, *27*(9), 1709–1723. https://doi.org/10.1111/poms.12832

Cooper-Martin, E. (1994). Measures of Cognitive Effort. *Marketing Letters*, *5*(1), 43–56. https://doi.org/10.1007/BF00993957

Darke, P. R., & Chung, C. M. (2005). Effects of Pricing and Promotion on Consumer Perceptions: It Depends on How you Frame it. *Journal of Retailing*, *81*(1), 35–47. https://doi.org/10.1016/j.jretai.2005.01.002

Darke, P. R., & Dahl, D. W. (2003). Fairness and Discounts: The Subjective Value of a Bargain. *Journal of Consumer Psychology*, *13*(3), 328–338. https://doi.org/10.1207/S15327663JCP1303_13

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press. http://www.loc.gov/catdir/enhancements/fy0604/92012048-d.html

DelVecchio, D. (2005). Deal-Prone Consumers' Response to Promotion: The Effects of Relative and Absolute Promotion Value. *Psychology & Marketing*, *22*(5), 373–391. https://doi.org/10.1002/mar.20064

Dhebar, A., & Oren, S. S. (1985). Optimal Dynamic Pricing For Expanding Networks. *Marketing Science*, *4*(4), 336–351. https://doi.org/10.1287/mksc.4.4.336

Diamond, W. D., & Campbell, L. (1989). The Framing of Sales Promotions: Effects on Reference Price Change. *Advances in Consumer Research* (16), Article ed. Thomas Srull, 241–247.

Diamond, W. D., & Sanyal, A. (1990). The Effect of Framing on the Choice of Supermarket Coupons. *Advances in Consumer Research* (17), 488–493.

Dickson, P. R., & Kalapurakal, R. (1994). The Use and Perceived Fairness of Price-Setting Rules in the Bulk Electricity Market. *Journal of Economic Psychology*, *15*(3), 427–448. https://doi.org/10.1016/0167-4870(94)90023-X

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research*, *28*(3), 307. https://doi.org/10.2307/3172866

Dong, X., Manchanda, P., & Chintagunta, P. K. (2009). Quantifying the Benefits of Individual-Level Targeting in the Presence of Firm Strategic Behavior. *Journal of Marketing Research*, *46*(2), 207–221. https://doi.org/10.1509/jmkr.46.2.207

Doolin, B., Dillon, S., Thompson, F., & Corner, J. L. (2005). Perceived Risk, the Internet Shopping Experience and Online Purchasing Behavior. *Journal of Global Information Management*, *13*(2), 66–88. https://doi.org/10.4018/jgim.2005040104

Dubb, J.-P., & Misra, S. (2017). Scalable Price Targeting. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2992257

Dubé, J.-P., & Misra, S. (2023). Personalized Pricing and Consumer Welfare. *Journal of Political Economy*, *131*(1), 131–189. https://doi.org/10.1086/720793

Dülmer, H. (2016). The Factorial Survey. *Sociological Methods & Research*, *45*(2), 304–347. https://doi.org/10.1177/0049124115582269

Eberl, M., & Schwaiger, M. (2005). Corporate Reputation: Disentangling the Effects on Financial Performance. *European Journal of Marketing*, *39*(7/8), 838–854. https://doi.org/10.1108/03090560510601798

Elmachtoub, A. N., Gupta, V., & Hamilton, M. L. (2021). The Value of Personalized Pricing. *Management Science*, *67*(10), 6055–6070. https://doi.org/10.1287/mnsc.2020.3821

Elmaghraby, W., & Keskinocak, P. (2003). Dynamic Pricing in the Presence of Inventory Considerations: Research Overview, Current Practices, and Future Directions. *Management Science*, *49*(10), 1287–1309. https://doi.org/10.1287/mnsc.49.10.1287.17315

Esteves, R.-B., & Resende, J. (2019). Personalized Pricing and Advertising: Who are the Winners? *International Journal of Industrial Organization*, *63*, 239–282. https://doi.org/10.1016/j.ijindorg.2018.11.003

Esteves, R.-B., & Shuai, J. (2022). Personalized Pricing with a Price Sensitive Demand. *Economics Letters*, *213*, 110396. https://doi.org/10.1016/j.econlet.2022.110396

Feinberg, F. M., Krishna, A., & Zhang, Z. J. (2002). Do We Care What Others Get? A Behaviorist Approach to Targeted Promotions. *Journal of Marketing Research*, *39*(3), 277–291. https://doi.org/10.1509/jmkr.39.3.277.19108

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Fisher, M., Gallino, S., & Li, J. (2018). Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments. *Management Science*, *64*(6), 2496–2514. https://doi.org/10.1287/mnsc.2017.2753

Gabel, S., & Timoshenko, A. (2022). Product Choice with Large Assortments: A Scalable Deep-Learning Model. *Management Science*, *68*(3), 1808–1827. https://doi.org/10.1287/mnsc.2021.3969

Garbarino, E., & Lee, O. F. (2003). Dynamic Pricing in Internet Retail: Effects on Consumer Trust. *Psychology & Marketing*, *20*(6), 495–513. https://doi.org/10.1002/mar.10084

Garbarino, E., & Maxwell, S. (2010). Consumer Response to Norm-Breaking Pricing Events in E-Commerce. *Journal of Business Research*, *63*(9-10), 1066–1072. https://doi.org/10.1016/j.jbusres.2008.12.010

Gendall, P., Hoek, J., Pope, T., & Young, K. (2006). Message Framing Effects on Price Discounting. *Journal of Product & Brand Management*, *15*(7), 458–465. https://doi.org/10.1108/10610420610712847

Gotlieb, J. B., & Swan, J. E. (1990). An Application of the Elaboration Likelihood Model. *Journal of the Academy of Marketing Science*, *18*(3), 221–228. https://doi.org/10.1177/009207039001800304

Gourville, J. T., & Moon, Y. (2004). Managing Price Expectations Through Product Overlap. *Journal of Retailing*, *80*(1), 23–35. https://doi.org/10.1016/j.jretai.2004.01.002

Greenberg, J. (1986). Determinants of Perceived Fairness of Performance Evaluations. *Journal of Applied Psychology*, *71*(2), 340–342. https://doi.org/10.1037/0021-9010.71.2.340

Greenberg, J. (1987). A Taxonomy of Organizational Justice Theories. *Academy of Management Review*, *12*(1), 9–22. https://doi.org/10.5465/amr.1987.4306437

Greenberg, J., Bies, R. J., & Eskew, D. E. (Eds.). (1991). *Applied impression management: How image-making affects managerial decisions*. *Establishing Fairness in the Eye of the Beholder: Managing Impressions of Organizational Justice* (R. A. Giacalone & P. Rosenfeld). Thousand Oaks, CA: Sage. https://psycnet.apa.org/record/1991-98424-006

Grewal, D., Hardesty, D. M., & Iyer, G. R. (2004). The Effects of Buyer Identification and Purchase Timing on Consumers' Perceptions of Trust, Price Fairness, and Repurchase Intentions. *Journal of Interactive Marketing*, *18*(4), 87–100. https://doi.org/10.1002/dir.20024

Grewal, D., Marmorstein, H., & Sharma, A. (1996). Communicating Price Information through Semantic Cues: The Moderating Effects of Situation and Discount Size. *Journal of Consumer Research*, *23*(2), 148. https://doi.org/10.1086/209473

Gruzd, A., Jacobson, J., & Dubois, E. (2020). Cybervetting and the Public Life of Social Media Data. *Social Media + Society*, *6*(2). https://doi.org/10.1177/2056305120915618

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating Vignette and Conjoint Survey Experiments Against Real-World Behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2395–2400. https://doi.org/10.1073/pnas.1416587112

Hajli, N., Lin, X., Featherman, M., & Wang, Y. (2014). Social Word of Mouth: How Trust Develops in the Market. *International Journal of Market Research*, *56*(5), 673–689. https://doi.org/10.2501/IJMR-2014-045

Haugtvedt, C. P., Petty, R. E., & Cacioppo, J. T. (1992). Need for Cognition and Advertising: Understanding the Role of Personality Variables in Consumer Behavior. *Journal of Consumer Psychology*, *1*(3), 239–260. https://doi.org/10.1016/S1057-7408(08)80038-1

Haws, K. L., & Bearden, W. O. (2006). Dynamic Pricing and Consumer Fairness Perceptions. *Journal of Consumer Research*, *33*(3), 304–311. https://doi.org/10.1086/508435

Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis : A Regression-Based Approach*. Guilford Publications. http://ebookcentral.proquest.com/lib/ub-lmu/detail.action?docID=6809031

Hufnagel, G., Schwaiger, M., & Weritz, L. (2022). Seeking the Perfect Price: Consumer Responses to Personalized Price Discrimination in E-Commerce. *Journal of Business Research*, *143*, 346–365. https://doi.org/10.1016/j.jbusres.2021.10.002

Inman, J. J., McAlister, L., & Hoyer, W. D. (1990). Promotion Signal: Proxy for a Price Cut? *Journal of Consumer Research*, *17*(1), 74. https://doi.org/10.1086/208538

Inman, J. J., Peter, A. C., & Raghubir, P. (1997). Framing the Deal: The Role of Restrictions in Accentuating Deal Value. *Journal of Consumer Research*, *24*(1), 68–79. https://doi.org/10.1086/209494

Jacoby, J., & Olson, J. C. (1977). *Consumer Response to Price : An Attitudinal, Information Processing Perspective* (in Y. Wind and M.G. Greenberg). *Moving Ahead with Attitude Research*. American Marketing Association. https://cir.nii.ac.jp/crid/1574231874658987392

Jalilvand, M. R., Nasrolahi Vosta, L., Kazemi Mahyari, H., & Khazaei Pool, J. (2017). Social Responsibility Influence on Customer Trust in Hotels: Mediating Effects of Reputation and Word-of-Mouth. *Tourism Review*, *72*(1), 1–14. https://doi.org/10.1108/TR-09-2016-0037

Johnson, J., Tellis, G. J., & Ip, E. H. (2013). To Whom, When, and How Much to Discount? A Constrained Optimization of Customized Temporal Discounts. *Journal of Retailing*, *89*(4), 361–373. https://doi.org/10.1016/j.jretai.2013.08.002

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Kahneman, D., & Tversky, A. (1984). Choices, Values, and Frames. *American Psychologist*, *39*(4), 341–350. https://doi.org/10.1037/0003-066X.39.4.341

Kallus, N., & Zhou, A. (2021). Fairness, Welfare, and Equity in Personalized Pricing. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.* ACM. https://doi.org/10.1145/3442188.3445895

Keller, A., Vogelsang, M., & Totzek, D. (2022). How Displaying Price Discounts can Mitigate Negative Customer Reactions to Dynamic Pricing. *Journal of Business Research*, *148*, 277–291. https://doi.org/10.1016/j.jbusres.2022.04.027

Keskin, N. B., & Zeevi, A. (2014). Dynamic Pricing with an Unknown Demand Model: Asymptotically Optimal Semi-Myopic Policies. *Operations Research*, *62*(5), 1142–1167. https://doi.org/10.1287/opre.2014.1294

Khandeparkar, K., Maheshwari, B., & Motiani, M. (2020). Why Should I Pay More? Testing the Impact of Contextual Cues on Perception of Price Unfairness for the Price-Disadvantaged Segment in Dual Pricing. *Tourism Management*, *78*, 104075. https://doi.org/10.1016/j.tourman.2020.104075

Kim, H. M., & Kramer, T. (2006). The Moderating Effects of Need for Cognition and Cognitive Effort on Responses to Multi-Dimensional Prices. *Marketing Letters*, *17*(3), 193–203. https://doi.org/10.1007/s11002-006-5928-2

Krishna, A., Briesch, R., Lehmann, D. R., & Yuan, H. (2002). A Meta-Analysis of the Impact of Price Presentation on Perceived Savings. *Journal of Retailing*, *78*(2), 101–118. https://doi.org/10.1016/S0022-4359(02)00072-6

Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2014). *The Handbook of Attitudes, eds Albarracin D, Johnson BT, Zanna MP* (eds Albarracin D, Johnson BT, Zanna MP).

Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kung, M., Monroe, K. B., & Cox, J. L. (2002). Pricing on the Internet. *Journal of Product & Brand Management*, *11*(5), 274–288. https://doi.org/10.1108/10610420210442201

Lai, C.-S., Chiu, C.-J., Yang, C.-F., & Pai, D.-C. (2010). The Effects of Corporate Social Responsibility on Brand Performance: The Mediating Effect of Industrial Brand Equity and Corporate Reputation. *Journal of Business Ethics*, *95*(3), 457–469. https://doi.org/10.1007/s10551-010-0433-1

Li, M., & Chapman, G. B. (2013). A Big Fish or a Small Pond? Framing Effects in Percentages. *Organizational Behavior and Human Decision Processes*, *122*(2), 190–199. https://doi.org/10.1016/j.obhdp.2013.07.003

Li, W., Hardesty, D. M., & Craig, A. W. (2018). The Impact of Dynamic Bundling on Price Fairness Perceptions. *Journal of Retailing and Consumer Services*, *40*, 204–212. https://doi.org/10.1016/j.jretconser.2017.10.011

Liang, T.-P., Lai, H.-J., & Ku, Y.-C. (2006). Personalized Content Recommendation and User Satisfaction: Theoretical Synthesis and Empirical Findings. *Journal of Management Information Systems*, *23*(3), 45–70. https://doi.org/10.2753/MIS0742-1222230303

Lichtenstein, D. R., Bloch, P. H., & Black, W. C. (1988). Correlates of Price Acceptability. *Journal of Consumer Research*, *15*(2), 243. https://doi.org/10.1086/209161

Lii, Y., & Sy, E. (2009). Internet Differential Pricing: Effects on Consumer Price Perception, Emotions, and Behavioral Responses. *Computers in Human Behavior*, *25*(3), 770–777. https://doi.org/10.1016/j.chb.2009.02.005

Lo, C. J., Tsarenko, Y., & Tojib, D. (2019). To Tell or Not to Tell? The Roles of Perceived Norms and Self-Consciousness in Understanding Consumers' Willingness to Recommend Online Secondhand Apparel Shopping. *Psychology & Marketing*, *36*(4), 287–304. https://doi.org/10.1002/mar.21179

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social Utility and Decision Making in Interpersonal Contexts. *Journal of Personality and Social Psychology*, *57*(3), 426–441. https://doi.org/10.1037/0022-3514.57.3.426

Mahon, J. F. (2002). Corporate Reputation. *Business & Society*, *41*(4), 415–445. https://doi.org/10.1177/0007650302238776

Major, B., & Testa, M. (1989). Social Comparison Processes and Judgments of Entitlement and Satisfaction. *Journal of Experimental Social Psychology*, *25*(2), 101–120. https://doi.org/10.1016/0022-1031(89)90007-3

Malc, D., Mumel, D., & Pisnik, A. (2016). Exploring Price Fairness Perceptions and Their Influence on Consumer Behavior. *Journal of Business Research*, *69*(9), 3693–3697. https://doi.org/10.1016/j.jbusres.2016.03.031

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355. https://doi.org/10.1287/isre.1040.0032

Manning, K. C., & Sprott, D. E. (2009). Price Endings, Left-Digit Effects, and Choice. *Journal of Consumer Research*, *36*(2), 328–335. https://doi.org/10.1086/597215

Manyanga, W., Makanyeza, C., & Muranda, Z. (2022). The Effect of Customer Experience, Customer Satisfaction and Word of Mouth Intention on Customer Loyalty: The Moderating Role of Consumer Demographics. *Cogent Business & Management*, *9*(1), Article 2082015. https://doi.org/10.1080/23311975.2022.2082015

Marwell, G., & Ames, R. E. (1981). Economists Free Ride, Does Anyone Else? *Journal of Public Economics*, *15*(3), 295–310. https://doi.org/10.1016/0047-2727(81)90013-X

Maxwell, S., & Garbarino, E. (2010). The Identification of Social Norms of Price Discrimination on the Internet. *Journal of Product & Brand Management*, *19*(3), 218–224. https://doi.org/10.1108/10610421011046193

McAfee, R. P., & Te Velde, V. (Eds.). (2006). *Dynamic Pricing in the Airline Industry*. Handbook on economics and information systems.

McQuarrie, E. F., & Munson, J. M. (1992). *A Revised Product Involvement Inventory: Improved Usability and Validity*. ACR North American Advances. https://scholar.google.de/citations?user=pidfybiaaaaj&hl=de&oi=sra

Monroe, K. B. (Ed.). (1990). *Pricing: Making Profitable Decisions*. McGrawHill.

Morwitz, V. G., Greenleaf, E. A., & Johnson, E. J. (1998). Divide and Prosper: Consumers' Reactions to Partitioned Prices. *Journal of Marketing Research*, *35*(4), 453–463. https://doi.org/10.1177/002224379803500404

NCH Marketing Services. (2019). *2018 Year-End Coupon Facts at a Glance*. https://www.nchmarketing.com/2018-year-end-coupon-facts-at-a-glance.aspx.

Neill, H. R., Cummings, R. G., Ganderton, P. T., Harrison, G. W., & McGuckin, T. (1994). Hypothetical Surveys and Real Economic Commitments. *Land Economics*, *70*(2), 145. https://doi.org/10.2307/3146318

Neubert, M. (2022). *A Systematic Literature Review of Dynamic Pricing Strategies.*

Ozanne, J. L., Brucks, M., & Grewal, D. (1992). A Study of Information Search Behavior During the Categorization of New Products. *Journal of Consumer Research*, *18*(4), 452. https://doi.org/10.1086/209273

Park, K., & Gomez, M. I. (2004). *The Coupon Report: A Study of Coupon Discount Methods*. Charles H. Dyson School of Applied Economics and Management, Cornell University. https://ecommons.cornell.edu/handle/1813/65053

Pigou, A. (1920). *The Economics of Welfare*. Macmillan.

Priester, A., Robbert, T., & Roth, S. (2020). A Special Price Just For You: Effects of Personalized Dynamic Pricing on Consumer Fairness Perceptions. *Journal of Revenue and Pricing Management*, *19*(2), 99–112. https://doi.org/10.1057/s41272-019-00224-3

Rao, A. R., & Sieben, W. A. (1992). The Effect of Prior Knowledge on Price Acceptability and the Type of Information Examined. *Journal of Consumer Research*, *19*(2), 256. https://doi.org/10.1086/209300

Rossi, P. E., McCulloch, R. E., & Allenby, G. M. (1996). The Value of Purchase History Data in Target Marketing. *Marketing Science*, *15*(4), 321–340. https://doi.org/10.1287/mksc.15.4.321

Russo, J. E., & Dosher, B. A. (1983). Strategies for Multiattribute Binary Choice. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *9*(4), 676–696. https://doi.org/10.1037/0278-7393.9.4.676

Sahay, A. (2007). How to Reap Higher Profits With Dynamic Pricing. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/how-to-reap-higher-profits-with-dynamic-pricing/?switch_view=pdf

Sahni, N. S., Zou, D., & Chintagunta, P. K. (2017). Do Targeted Discount Offers Serve as Advertising? Evidence from 70 Field Experiments. *Management Science*, *63*(8), 2688–2705. https://doi.org/10.1287/mnsc.2016.2450

Schrills, T. P. P., Kargl, S., Bickel, M., & Franke, T. (2022). *Perceive, Understand & Predict - Empirical Indication for Facets in Subjective Information Processing Awareness.* https://doi.org/10.31234/osf.io/3n95u

Schwarz, N. (1999). Self-Reports: How the Questions Shape the Answers. *American Psychologist*, *54*(2), 93–105. https://doi.org/10.1037/0003-066x.54.2.93

Seele, P., Dierksmeier, C., Hofstetter, R., & Schultz, M. D. (2021). Mapping the Ethicality of Algorithmic Pricing: A Review of Dynamic and Personalized Pricing. *Journal of Business Ethics*, *170*(4), 697–719. https://doi.org/10.1007/s10551-019-04371-w

Shaffer, G., & Zhang, Z. J. (1995). Competitive Coupon Targeting. *Marketing Science*, *14*(4), 395–416. https://doi.org/10.1287/mksc.14.4.395

Shamma, H. M. (2012). Toward a Comprehensive Understanding of Corporate Reputation: Concept, Measurement and Implications. *International Journal of Business and Management*, *7*(16). https://doi.org/10.5539/ijbm.v7n16p151

Shaw, S., Chovancová, M., & Bejtkovský, J. (2022). Managing Price Changes: Role of Consumer Thinking Styles on Perceived Price Fairness and Purchase Intention. *Innovative Marketing*, *18*(2), 212–223. https://doi.org/10.21511/im.18(2).2022.18

Sherif, C. W. (1963). Social Categorization as a Function of Latitude of Acceptance and Series Range. *The Journal of Abnormal and Social Psychology*, *67*(2), 148–156. https://doi.org/10.1037/h0043022

Shiller, B. R. (2014). *First Degree Price Discrimination Using Big Data*. https://www.researchgate.net/profile/benjamin-shiller/publication/358940523_first-degree_price_discrimination_using_big_data_for_very_useful_feedback_suggestions_and_advice_i_would_also_like_to_thank/links/621e7052395296023159b47d/first-degree-price-discrimination-using-big-data-for-very-useful-feedback-suggestions-and-advice-i-would-also-like-to-thank.pdf

Shiller, B. R. (2020). Approximating Purchase Propensities and Reservation Prices from Broad Consumer Tracking. *International Economic Review*, *61*(2), 847–870. https://doi.org/10.1111/iere.12442

Shugan, S. M. (1980). The Cost of Thinking. *Journal of Consumer Research*, *7*(2), 99. https://doi.org/10.1086/208799

Singh, J. (1988). Consumer Complaint Intentions and Behavior: Definitional and Taxonomical Issues. *Journal of Marketing*, *52*(1), 93–107. https://doi.org/10.1177/002224298805200108

Smith, A. N., Seiler, S., & Aggarwal, I. (2023). Optimal Price Targeting. *Marketing Science*, *42*(3), 476–499. https://doi.org/10.1287/mksc.2022.1387

Stigler, G. J. (1950). The Development of Utility Theory. I. *Journal of Political Economy*, *58*(4), 307–327. https://doi.org/10.1086/256962

Stiving, M., & Winer, R. S. (1997). An Empirical Analysis of Price Endings with Scanner Data. *Journal of Consumer Research*, *24*(1), 57–67. https://doi.org/10.1086/209493

Stole, L. A. (2007). Chapter 34 Price Discrimination and Competition. *Handbook of Industrial Organization*, *3*, 2221–2299. https://doi.org/10.1016/S1573-448X(06)03034-2

Su, X. (2007). Intertemporal Pricing with Strategic Customer Behavior. *Management Science*, *53*(5), 726–741. https://doi.org/10.1287/mnsc.1060.0667

Suri, R., Monroe, K. B., & Koc, U. (2013). Math Anxiety and its Effects on Consumers' Preference for Price Promotion Formats. *Journal of the Academy of Marketing Science*, *41*(3), 271–282. https://doi.org/10.1007/s11747-012-0313-6

Surprenant, C. F., & Solomon, M. R. (1987). Predictability and Personalization in the Service Encounter. *Journal of Marketing*, *51*(2), 86–96. https://doi.org/10.1177/002224298705100207

SurveyCircle. (2022). *Forschungswebseite SurveyCircle. Published 2016.* https://www.surveycircle.com

Taylor, S., & Todd, P. A. (1995). Understanding Information Technology Usage: A Test of Competing Models. *Information Systems Research*, *6*(2), 144–176. http://www.jstor.org/stable/23011007

Treasure Data. (2021). *Muji Grows In-Store Revenue with Digital Savvy*. https://get.treasuredata.com/rs/714-XIJ-402/images/muji-case-study.pdf

Tripathi, A., & Pandey, N. (2017). Price Framing Literature: Past, Present, and Future. *The Marketing Review*, *17*(3), 307–336. https://doi.org/10.1362/146934717X14909733966245

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cognitive Effort and Memory. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 607–617. https://doi.org/10.1037/0278-7393.5.6.607

van Westendorp, P. H. (1976). *NSS-Price Sensitivity Meter (PSM)- A New Approach to Study Consumer Perception of Price.* https://archive.researchworld.com/a-new-approach-to-study-consumer-perception-of-price/

Varian, H. R. (1985). Price Discrimination and Social Welfare. *The American Economic Review*(75), 870-875. https://www.jstor.org/stable/1821366

Varian, H. R. (1989). Chapter 10 Price discrimination. In *Handbook of Industrial Organization. Handbook of Industrial Organization Volume 1* (Vol. 1, pp. 597–654). Elsevier. https://doi.org/10.1016/S1573-448X(89)01013-7

Varian, H. R. (2014). *Intermediate Microeconomics with Calculus: International Student Edition* (1. ed., internat. student ed.). Norton.

Venkatesan, R., & Farris, P. W. (2012). Measuring and Managing Returns from Retailer-Customized Coupon Campaigns. *Journal of Marketing*, *76*(1), 76–94. https://doi.org/10.1509/jm.10.0162

Venkatesan, R., & Leliveld, I. (2009). Catalina Marketing Corporation: Developing a Retailer Value Proposition. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.1420593

Wang, X., Li, X., & Kopalle, P. K. (2023). When Does it Pay to Invest in Pricing Algorithms? *Production and Operations Management*, Article poms.13924. Advance online publication. https://doi.org/10.1111/poms.13924

Weisstein, F. L., Monroe, K. B., & Kukar-Kinney, M. (2013). Effects of Price Framing on Consumers' Perceptions of Online Dynamic Pricing Practices. *Journal of the Academy of Marketing Science*, *41*(5), 501–514. https://doi.org/10.1007/s11747-013-0330-0

Weritz, L. (2023). *Mitigating the Negative Effects of Personalized Pricing* [Unpublished Dissertation]. Ludwig Maximilians Universität, München.

Xia, F., Chatterjee, R., & May, J. H. (2019). Using Conditional Restricted Boltzmann Machines to Model Complex Consumer Shopping Patterns. *Marketing Science*, *38*(4), 711–727. https://doi.org/10.1287/mksc.2019.1162

Xia, L., Monroe, K. B., & Cox, J. L. (2004). The Price is Unfair! A Conceptual Framework of Price Fairness Perceptions. *Journal of Marketing*, *68*(4), 1–15. https://doi.org/10.1509/jmkg.68.4.1.42733

Yoo, B., & Donthu, N. (2002). Testing Cross-Cultural Invariance of the Brand Equity Creation Process. *Journal of Product & Brand Management*, *11*(6), 380–398. https://doi.org/10.1108/10610420210445505

Yoo, C. W., Goo, J., Huang, C. D., Nam, K., & Woo, M. (2017). Improving Travel Decision Support Satisfaction with Smart Tourism Technologies: A Framework of Tourist Elaboration Likelihood and Self-Efficacy. *Technological Forecasting and Social Change*, *123*, 330–341. https://doi.org/10.1016/j.techfore.2016.10.071

Yun, W., & Hanson, N. (2020). Weathering Consumer Pricing Sensitivity: The Importance of Customer Contact and Personalized Services in the Financial Services Industry. *Journal of Retailing and Consumer Services*, *55*, 102085. https://doi.org/10.1016/j.jretconser.2020.102085

Zaichkowsky, J. L. (1985). Measuring the Involvement Construct. *Journal of Consumer Research*, *12*(3), 341. https://doi.org/10.1086/208520

Zhang, H., Zhao, L., & Gupta, S. (2018). The Role of Online Product Recommendations on Customer Decision Making and Loyalty in Social Shopping Communities. *International Journal of Information Management*, *38*(1), 150–166. https://doi.org/10.1016/j.ijinfomgt.2017.07.006

Zhang, J [J.], & Krishnamurthi, L. (2004). Customizing Promotions in Online Stores. *Marketing Science*, *23*(4), 561–578. https://doi.org/10.1287/mksc.1040.0055

Zhang, J [J.], & Wedel, M. (2009). The Effectiveness of Customized Promotions in Online and Offline Stores. *Journal of Marketing Research*, *46*(2), 190–206. https://doi.org/10.1509/jmkr.46.2.190

Zhang, J [Juanjuan] (2011). The Perils of Behavior-Based Personalization. *Marketing Science*, *30*(1), 170–186. https://doi.org/10.1287/mksc.1100.0607

Sandra Baringhorst | Maximilian Niederberger-Kern | Manfred Schwaiger

IV Single- Versus Multi-Item Scales: A Comparison of Test-Retest Reliability

Abstract

Since the 1970s, the use of multi-item scales has been prevalent in marketing research, while single-item scales have been accepted primarily for self-reported facts or demographic information. Nearly 40 years later, the debate over which scale to use has been reopened, and the status quo of utilizing multi-item scales has been challenged. Initial research on reliability and validity was inconclusive. However, single-item and multi-item scales have yet to be compared regarding test-retest reliability. Therefore, this paper examines the intraclass correlation coefficients of several established scales in the marketing literature across two different time frames (two weeks and four months between measurements). The results suggest that single-item scales perform at least as well as multi-item scales in terms of test-retest reliability. In many comparisons, single-item scales even outperform their multi-item counterparts, particularly in the shorter time frame. The paper extends the discussion of the use of single-item versus multi-item scales in marketing research and proposes that measures used in longitudinal research should be carefully selected.

1 Motivation

Until the 1970s, the utilization of single-item (SI) measures was commonplace in marketing research. After Jacoby (1978) challenged marketers' reliance on SI scales, arguing that most constructs are too complex to be effectively measured with an SI scale, seminal works by Churchill (1979) and Peter (1979) led to the tradition of psychometrics (cf. Nunnally & Bernstein, 1994) taking hold in marketing research. Nowadays, SI measures are generally deemed acceptable for self-reported facts and general demographic-type information (e.g., age, education), while the use of multi-item (MI) scales is standard practice in academic marketing research to capture psychological-type constructs such as abilities and traits. The use of MI scales to measure unobservable dispositional phenomena is reflected in scale development guides (e.g., Netemeyer et al., 2003) and handbooks compiling overviews of marketing measures (e.g., Bearden et al., 2011; Bruner, 2021).

Following years of acceptance of established practices, Bergkvist and Rossiter (2007) reignited the debate in marketing about the use of SI and MI measures. Based on Rossiter's (2002) C-OAR-SE procedure, which proposes that *doubly concrete* constructs (i.e., constructs with a singular and concrete object as well as a concrete attribute) do not require MI measures, the authors challenged the status quo of utilizing MI scales on theoretical and empirical grounds. Their findings – SI scales have equally high predictive validity as MI scales for doubly concrete constructs – have been challenged by several researchers (e.g., Diamantopoulos et al., 2012; Sarstedt & Wilczynski, 2009). The question of the use of SI versus MI measures has not been conclusively resolved, and the literature has repeatedly produced conflicting results. While some researchers found comparably good results for SI compared to MI measures in terms of predictive validity (e.g., Bergkvist & Rossiter, 2009), others emphasized the superiority of MI measures (e.g., Diamantopoulos et al., 2012; Sarstedt & Wilczynski, 2009).

An important concept of measurement accuracy that has yet to receive much attention in research on the comparability of SI and MI measures is test-retest reliability. Test-retest reliability is an indicator of consistency and ensures that measurements are representative and stable over time. It is a crucial measurement feature in longitudinal research and concerns the agreement between scores obtained from the same individual on two or more separate occasions, despite intraindividual response variability (i.e., measurement error) (Hays et al., 1993). Since test-retest reliability is a suitable instrument for comparing different types of scales over time but has received little to no attention in the academic literature, the following research question is proposed:

RQ: *Do SI and MI scales differ based on test-retest reliability?*

This study employs a within-subjects design to survey participants about their attitudes toward different brands and products as well as their purchase intentions after exposure to various advertisements. It extends the discussion on SI scales in academic research by providing insight into their test-retest reliability. We assess test-retest reliability by using the intraclass correlation coefficient (ICC), which measures the reliability of a measure between two or more points in time (i.e., periods of two weeks and four months). We compare ICC scores between SI and MI scales using PF and ZPF test statistics. The results indicate that SI scales perform at least as well as MI scales over time. In most comparisons, the SI measures even outperform their MI counterparts. While we do not mean to suggest that researchers should interpret these findings in favor of using SI scales over MI scales, we do urge cautious use of both scales, especially in longitudinal research.

The remainder of the paper is structured as follows. First, we revisit the literature debating SI versus MI scales, focusing on validity and (test-retest) reliability. Thereafter, we introduce and describe our methodology. This is followed by a presentation of the analysis and results. The study concludes with a discussion, implications for marketers and academics, and recommendations for further research.

2 Literature Review

Since the late 1970s, research in management and marketing followed Churchill's (1979, p. 66) paradigm that “marketers are much better served with multi-item than single-item measures of their constructs”. This has led to the conventional wisdom that SI should not be used, resulting in the rejection of research projects using SI scales in review processes. Over the past 30 years, however, the view that SI measures are fatally flawed has been repeatedly challenged, leading to calls for an evaluation of “their appropriateness for a particular piece of research” (Wanous et al., 1997, p. 251). The debate about the psychometric properties of SI and MI measures has centered on their respective advantages and disadvantages, particularly focusing on reliability and validity, which will be discussed below.

2.1 Status Quo of Research

Advantages and Disadvantages of Single- Versus Multi-Item Scales

The renewed discussion about the use of SI and MI measures ties in with the general scientific debate about the benefits and drawbacks of their use. In a nutshell, the discussion centers on the practical advantages of SI measures versus theoretical concerns regarding their reliability and validity. Due to increasing expenses of data collection and coding (Moore et al., 2002), questionnaires with MI instruments are more costly than those with SI measures. However, scholars should consider more than just monetary costs. In psychology, several researchers emphasized the need to make measurement more efficient, either by substantially reducing the length of measurement instruments or by using SI measures as opposed to MI measures (Nagy, 2002; Russell et al., 2004). By overloading respondents (Wanous et al., 1997), lengthy scales can also lead to lower response and higher dropout rates (Dillman et al., 1993). In addition, they may facilitate sampling bias, as less engaged respondents are also more likely to drop out (Moore et al., 2002). Since SI measures are short, flexible, and easy to administer (Pomeroy et al., 2001), they are less time-consuming and monotonous to complete (Gardner et al., 1998). Research has found that reducing the response time and length of questionnaires can significantly reduce respondent boredom and fatigue (Adigüzel & Wedel, 2008) and lead to higher response rates (Dillman et al., 1993).

While the advantages of SI measures can result in a reduction of potential response biases (Drolet & Morrison, 2001), the use of MI scales has also been linked to a decline in carryover effects. These biases are based on respondents’ state dependency and occur when the response pattern carries over from one item to the next. Carryover effects could be particularly pronounced in the case of SI measures when no other item measures a specific construct (Diamantopoulos et al., 2012). In addition, MI scales can benefit from increased reliability and construct validity on the basis of measurement theory, as they represent a random selection from the set of possible indicators of a construct (Nunnally & Bernstein, 1994). Table 1 presents an overview of the aforementioned advantages of both scale types.

Table 1. Advantages of Using Single- Versus Multi-Item Scales

Advantages of Single-Item Scales		**Advantages of Multi-Item Scales**	
Financial reasons (lower costs, higher response rates, lower dropout rates)	Dillman et al. (1993); Gardner et al. (1998); Moore et al. (2002); Wanous et al. (1997)	Combination of numerous items averages out random error	Churchill and Peter (1984)
Elimination of semantically similar items (no consistency motif bias)	Churchill and Peter (1984)	Reduction of carryover effects	Diamantopoulos et al. (2012)
Practical reasons (short, flexible, easy to administer, mental fatigue)	Pomeroy et al. (2001)	Construct validity: A larger set of adequate indicators covers a larger number of distinct construct facets	Nunnally and Bernstein (1994)

Reliability

Reliability is defined "as the degree of consistency between two measures of the same thing" (Mehrens & Lehmann, 1991, p. 249). In other words, reliability is the extent to which measurements can be replicated. Mathematically, it is represented by the ratio of true variance to true variance plus error variance. Traditionally, SI measurement instruments were considered unreliable because researchers believed that their measurement errors would be inflated. Unlike MI scales, SI measures cannot average out the errors and specificities inherent in individual indicators by summing them up. The misconception that SI scales lack a measure of internal consistency reliability, because standard coefficients for MI scales (e.g., Cronbach's Alpha) cannot be applied, further reinforced this valid critique. Wanous and Reichers (1996) sought to address this misconception by suggesting that SI reliability could be assessed using the correction for attenuation formula. In addition, Wanous and Hudy (2001) used communalities derived from common factor analysis as another way to obtain conservative estimates of SI reliability. Since reliability reflects both the common and unique variance of a variable, they argued that the reliability of an SI measure should at least match its communality. Studies examining reliability estimates for SI measures have consistently found them to perform acceptably (e.g., Sarstedt & Wilczynski, 2009; Wanous & Hudy, 2001).

This finding supports the criticism of many researchers' practical approach to scale development and their overemphasis on reliability. The incremental information provided by additional scale items may be quite small because they inflate across-item error term correlation, which undermines respondent reliability and reduces their informal value (Drolet & Morrison, 2001). In addition, several biases can affect how respondents approach MI scales. When items are semantically similar, respondents tend to make inferences from one item to the next on the same scale without reading them carefully. Specifically, Drolet and Morrison (2001) found that respondents discriminate less between items as they encounter more of them, leading to a stronger influence of earlier items on later ones. They concluded that as the number of items increases, respondents are more likely to engage in mindless response behavior. In addition, the consistency motif bias is more prevalent in MI scales because "subjects tend to try to maintain consistency in their responses of similar questions" (Podsakoff et al., 2003, p. 881).

In summary, (internal consistency) reliability does not appear to be a barrier to the use of SI measures. However, reliability is a necessary but not sufficient condition for validity. The validity of SI measures has also been the subject of intense debate over the past two decades.

Validity

Construct validity refers to the degree to which a measure captures the construct it is intended to capture (Peter, 1981). According to psychometric theory, MI measures are necessary to ensure construct validity because each individual item usually correlates poorly with the construct in question and has a degree of specificity, meaning that the items individually insufficiently capture the conceptual domain of a construct (Nunnally & Bernstein, 1994). Compared to MI scales, SI measures do not illuminate a construct from different perspectives (Baumgartner & Homburg, 1996; Wirtz & Lee, 2003). On the flipside, semantic redundancy, created by the practice of using synonyms as additional items, can lead to reduced content validity (Rossiter, 2002). In a review of scale development in management, Hinkin (1995) found that lengthy measurement instruments had solid reliability but often picked up substance from more than one conceptual domain. Moreover, when MI measures consist of items that are highly similar in focus, and respondents interpret this as redundancy, the result may be a reduced willingness to provide accurate responses, thereby increasing the relative face validity of SI measures (Wanous et al., 1997). These findings led Drolet and Morrison (2001, p. 199, original emphasis) to suggest that "one or two *good* items that elicit *appropriate* respondent behavior [would outperform] multiple, poorly presented items that increase the error term correlations and/or stimulate inappropriate response styles".

Furthermore, studies assessing convergent validity and discriminant validity of SI scales – typically evaluating the correlation between the SI measure and an MI measure or counterpart – generally reported encouraging findings for SI measures (e.g., Boer et al., 2004; Dolbier et al., 2005; Nagy, 2002). The majority of recent marketing publications on the validity of SI measures examined predictive validity, a central criterion for decision-making (Kumar et al., 2018). These publications followed a revival of the debate on the appropriate use of SI and MI measures after the introduction of Rossiter's (2002) C-OAR-SE procedure.[1] According to this procedure, which is aimed at developing scales for marketing constructs, SI measures are sufficient when the object of the construct (e.g., an advertisement, a brand, or an organization) is concrete and singular, and the attribute of the construct (e.g., an attitude or perception) is concrete. Concrete singular objects designate constructs for which the object is described similarly by all raters, and only one object is to be rated. Concrete attributes are characterized by near-unanimous agreement regarding their meaning and singularity (i.e., only one attribute to be rated).

Putting the propositions of C-OAR-SE to the test, Bergkvist and Rossiter (2007, 2009) sparked an intense discussion about the use of SI measures for doubly concrete constructs. Based on their empirical findings, the authors concluded that SI measures have equally high predictive validity as MI scales for doubly concrete constructs and that MI scales should not be used considering their practical disadvantages. Not surprisingly, the implications of the studies were far-reaching, as the results implied that marketing scholars could save substantial resources by using SI scales as predictors. However, the authors' conclusions were based on incorrect tests of significance because they relied on Fisher's z-transformation test to contrast correlation coefficients and R^2-values between SI and MI measures. Since the samples being compared were not independent, a paired samples test should have been used (Diamantopoulos et al., 2012; Sarstedt & Wilczynski, 2009). Moreover, researchers who replicated Bergkvist and Rossiter's (2007, 2009) studies using appropriate test procedures did not find support for their conclusions. While the predictive validity of SI scales for some constructs, product categories, and stimulus objects was equal to (or even better than) MI measures, it generally varied considerably and tended to underperform (Diamantopoulos et al., 2012; Sarstedt & Wilczynski, 2009). Against this backdrop, a simulation study on the influence of different levels of data and measurement characteristics (including inter-item correlations between the items of the predictor and criterion

[1] C-OAR-SE is a procedure for the development of scales to measure marketing constructs, originally proposed by Rossiter (2002). It is the acronym for *Construct definition, Object classification, Attribute classification, Rater identification, Scale formation, and Enumeration and reporting*.

constructs, number of items in the predictor and criterion constructs, as well as sample size) on the performance of SI versus MI measures in predicting an MI criterion showed that MI predictors should be used in most cases. Only under certain circumstances – highly homogeneous items (i.e., Cronbach's alpha > 0.90), small sample sizes (i.e., N < 50), small expected effect sizes, and semantically redundant items – did SI measures match or outperform the predictive validity of MI counterparts and could be legitimately used (Diamantopoulos et al., 2012).

Taking a different analytical approach, Ang and Eisend (2018) conducted a meta-analysis to examine whether effect sizes from 189 advertising studies depended on the measurement scale of the dependent variable. Specifically, all studies considered in the meta-analysis used attitudes as the dependent variable and differed in whether they were measured using SI or MI scales (with different numbers of items). The authors found only a single significant difference in effect size between measurement instruments (which could be attributed to a Type I error), suggesting that the number of items used to measure the dependent variables is inconsequential and that SI measures are on par with MI measures.

In summary, there has been a lively discussion in the literature over the past decades about the different types of validity of SI and MI measures. However, it has not provided definitive evidence or guidelines for the use of SI or MI scales. Given that reliability is a necessary condition for validity, it seems all the more surprising that no studies to date have examined the difference in consistency of different scales over time, i.e., the so-called test-retest reliability.

2.2 Test-Retest Reliability as an Unexplored Criterion

To date, few studies have examined SI measures in terms of test-retest reliability (e.g., Boer et al., 2004; Fisher et al., 2016; Shamir & Kark, 2004). Test-retest reliability focuses on the "consistency of scores across two separate measurements over time, and is sometimes referred to as stability or reproducibility" (Polit, 2014, p. 1713). It requires researchers to collect data on the same sample at two or more points in time and calculate correlations (Mehrens & Lehmann, 1991) or differences between them (Polit, 2014). While this method appears time-consuming and costly (Sarstedt & Mooi, 2014), it is of particular interest in the context of this study because it is the only stand-alone measure of reliability that can be applied to SI measures (Sarstedt & Wilczynski, 2009).

Background

Several systematic factors undermine reliability assessments for both short and long retest intervals. These include the desire for consistency or memory effects, which refer to the risk that respondents will remember questions and their responses, leading to carryover effects between measurements. The challenges inherent in retest assessments have led many traditional psychometricians to avoid test-retest reliability. For example, Cronbach (1947) and Nunnally and Bernstein (1994) have pointed out difficulties of meeting all the requirements or cautioned against the inappropriate use of retest methods. Therefore, researchers should design surveys in such a way as to minimize the likelihood that participants will be able to recall their answers from a previous leg of examination. Subjects may also become familiar with the instrument and therefore learn how to respond in an expected way (Hendrickson et al., 1993). Such changes in responses due to better comprehension, more efficient retrieval of data, or reflection on an item, when seen a second time, are referred to as rehearsal. As a result, retest intervals that are too short can have adverse effects on reliability assessments.

However, if too much time elapses between measurement intervals, differences in the reliability could be attributed to the subjects themselves. Such response shifts are changes in respondents' evaluation of the construct (rather than a change in the construct itself) as a result of, for example, altered priorities or a reconceptualization of the target construct. A further challenge is the possibility of genuine changes in attitudes towards the measured items in the interval between measurement legs (Polit, 2014). Therefore, the time intervals between measurements in retest studies must be chosen appropriately. Additionally, sampling should be closely monitored. Varying sample sizes between measurements and homogeneity of the retest sample reduce the potential for variation and thus reliability estimates.

Preliminary Evidence

Regarding test-retest reliability, there is preliminary evidence suggesting that SI scales can be reliable (e.g., Boer et al., 2004; Shamir & Kark, 2004). Shamir and Kark (2004) examined SI scales for the identification with organizations and organizational units, such as departments or work teams. Three samples were employed to investigate the distribution of responses, convergent validity, and concurrent construct validity. Two additional student samples were surveyed to calculate test-retest reliability. These respondents were administered the newly developed SI scale on two separate occasions, two weeks apart. With high correlations between measurements (i.e., over .73 and .80), Shamir and Kark (2004) found evidence for the reliability of the

SI scale. Both samples ware also rather small (i.e., N=53 and N=68). Moreover, as Shamir and Kark (2004) created a new SI scale for organizational identification, a comparison with the results of a test-retest of an MI measure was not feasible. Similarly, Fisher et al. (2016) used correlations to assess the test-retest reliability of 18 newly developed items and 19 SI measures selected from MI scales in organizational and occupational health psychology research using correlations. While one-month retest intervals yielded high test-retest reliability, re-measuring the SI constructs after a three-month interval still produced stable results.

Although test-retest reliability is commonly judged based on Pearson's product-moment correlation between test and retest scores, the procedure is not recommended for calculating test-retest reliability (Liu et al., 2016; Yen & Lo, 2002). Theoretically, correlation measures the strength of a relationship between variables, not the agreement between them. Agreement and correlation both indicate the strength of association between variables of interest, but they require different statistics due to their conceptual differences. Relatedly, correlation is intended to determine the relationship between two different variables, not between test and retest scores for the same variable. Agreement emphasizes the degree of concordance in scores between two or more assessments of a construct of interest (or, alternatively, in the opinions of different individuals). In addition, correlation cannot detect the presence of a systematic error. For example, two different studies may show perfect correlations between test and retest scores. If the scores are very close together in one study (small difference between test and retest scores), while they are far apart in the other study (large difference between test and retest scores), then the correlation coefficient would not detect this systematic bias. In contrast, the ICC would be high in the first case and lower in the second case.

The ICC is a popular measure of reliability, reflecting both the degree of correlation and agreement between measurements. Compared to Pearson's product-moment correlation, the ICC can be applied to three or more separate scores and accounts for intra-individual response variability or rater bias, the element that distinguishes agreement from correlation. That is, the ICC considers the degree to which respondents are distinguishable despite the presence of measurement error. Therefore, a high ICC requires not only a high correlation but also a low rater bias (Liu et al., 2016). The ICC has the advantage that it allows for multiple scores to be considered at the same time. For example, if a construct is measured three times, three scores are obtained. A simple correlation coefficient cannot be calculated for three scores – instead, three bivariate correlations would have to be calculated. In contrast, the ICC can reflect the test-retest reliability between all three measurement points.

Using the ICC, Boer et al. (2004) assessed the test-retest reliability of SI quality of life scales within a clinical trial. They evaluated the scores of disease-free patients, with measurements being taken at two separate times, yielding a high overall ICC score (i.e., .87). However, SI and MI measures were not compared and the study was not without limitations. In particular, the two measurements were based on two different elicitation methods (postal questionnaires and interviews), which could have introduced error.

In summary, studies that have investigated the test-retest reliability of SI scales have reported encouraging results, but they have also suffered from various shortcomings. They relied on simple correlations (Fisher et al., 2016; Shamir & Kark, 2004), used small sample sizes (e.g., Shamir & Kark, 2004), or were inconsistent in their measurement approaches (Boer et al., 2004). Furthermore, none of the studies compared the test-retest reliability of MI and SI measures at different time points. The experimental design described below aims to address this research gap.

3 Methodology

The purpose of this study is to investigate how SI scales perform compared to MI scales in terms of their test-retest reliability. In line with research examining the predictive validity of SI and MI measures (Bergkvist & Rossiter, 2007, 2009; Diamantopoulos et al., 2012), the study uses data from consumer responses to advertisements across different constructs and stimulus objects. Subjects were asked to indicate their attitude toward the ad (A_{Ad}), attitude toward the brand (A_{Brand}), and brand purchase intention (PI_{Brand}). According to Rossiter (2002) and Bergkvist (2015), A_{Ad}, A_{Brand}, and PI_{Brand} are doubly concrete constructs and can therefore be measured using SI scales.

Subjects and Design

Participants were recruited via MTurk, using the crowdsourcing platform's option to collect panel data. Since test-retest reliability decreases with an increasing time between measurements (Geere et al., 2013), short retest intervals (e.g., one week) have been suggested to reduce the risk that external factors may compromise accurate measurement of reliability and that that attitudes toward the measured attribute may have changed. However, the shorter the interval, the greater the risk that participants will recall their previous responses (Polit, 2014). Since this study was intended to be exploratory and to provide a first impression of the performance of SI scales versus MI scales regarding their test-retest reliability, we decided to take a middle course and gathered data at three points in time, resulting in a short and a long retest interval. Two weeks after respondents were presented with the stimuli for the first time (T0), we conducted the second elicitation (T1). The third elicitation took place after four months (T2).

A total of 659 responses were collected in the first round. However, 265 participants were removed because they did not pass an attention check.[2] The sample was then randomly divided in half. One half was invited to participate in the second leg of the study (i.e., first subsample, two weeks later), and the other half was invited to the third leg of the study (i.e., second subsample, four months later). We collected 394 valid responses in the initial wave of the survey, 113 (out of 200 invited) valid responses in the second elicitation (57%), and 55 (out of 194 invited) valid responses after four months (28%). In the first wave (T0 – T1), the plurality of respondents was between 26 – 35 years old (44%), male (52%), working full-time (84%) with an undergraduate degree (59%), and with a net annual household income below €50,000 (60%). In the second wave (T0 – T2), the demographic structure was similar. However, the plurality

[2] Participants had to state the product categories that were presented in the advertisements.

of respondents was between 36 and 45 years old (36%) and had a net annual income of between €50,001 and €90,000 (42%). A complete overview of all assessed demographics can be found in Table 2.

Table 2. Sample Demographics

		T0-T1		T0-T2	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Age					
	18 - 25	17	15 %	10	18 %
	26 - 35	50	44 %	19	35 %
	36 - 45	32	28 %	20	36 %
	46 - 60	11	10 %	5	9 %
	above 60	3	3%	1	2 %
Gender					
	Male	59	52 %	39	71 %
	Female	54	48 %	16	29 %
Education					
	High school	16	14 %	3	5 %
	Undergraduate	67	59 %	36	66 %
	Postgraduate	28	25 %	13	24 %
	Diploma	2	2 %	3	5 %
	Others				
Employment					
	Full-time	95	84 %	47	85 %
	Part-time	7	6 %	6	11 %
	Student	2	2 %	0	0 %
	Stay-at-home-parent	3	3 %	2	4 %
	Unemployed	3	3 %	0	0 %
	Retired	2	2 %	0	0 %
	Others	1	0 %	0	0 %
Income (voluntary)					
	< 10,000 €	11	10 %	9	16 %
	10,000 € – 50,000 €	56	50 %	16	29 %
	50,001 € – 90,000 €	34	30 %	23	42 %
	90,001 € - 150,000 €	8	7 %	5	9 %
	> 150,001 €	4	3 %	2	4 %
		N = 113		**N = 55**	

Stimuli and Procedure

The study featured a within-subjects design. In line with studies by Bergkvist and Rossiter (2007, 2009) and Sarstedt, Diamantopoulos, Salzberger, and Baumgartner (2016), we chose advertisements from different product categories (i.e., insurance, jeans, pain relievers, coffee) as stimuli. To ensure that participants had no prior knowledge of the brands or advertisements and were not being influenced between the different measurement points, we selected advertisements from diverse countries (e.g., New Zealand, Australia, South Africa) to ensure that most advertisements were unfamiliar to participants (Sarstedt, Diamantopoulos, Salzberger, & Baumgartner, 2016).

The advertisements can be found in Appendix 1. Each advertisement was presented to the participants at the beginning of a question block. Subjects were asked to look at each advertisement carefully before answering several questions regarding their impression of the ad. In order to prevent participants from viewing the ad superficially, a 15-second timer was set that did not allow participants to move on until the timer expired. Furthermore, a smaller version of the advertisement was displayed with the measures to assist in answering the scales. To prevent attitude toward the brand from influencing evaluation of the ads, the brand measures were presented in reverse hierarchy-of-effects sequence (i.e., measures to assess the advertisements were asked first) (Bergkvist & Rossiter, 2007).

Measures

As described above, following the lead of previous authors (e.g., Bergkvist & Rossiter, 2007; Sarstedt, Diamantopoulos, Salzberger, & Baumgartner, 2016) who compared the predictive validity of MI and SI measures, we used the doubly concrete constructs A_{Ad}, A_{Brand}, and PI_{Brand} to assess test-retest reliability. After each advertisement, participants were first asked to indicate their familiarity with the advertisement or the advertised brand on seven-point Likert-type scales, using the items “I am familiar with this ad” and “I am familiar with this brand” (1 = strongly disagree, 7 = strongly agree). Participants then responded to the A_{Ad}, A_{Brand}, and PI_{Brand} scales (SI and MI). All measures were presented immediately after exposure to an ad before subjects moved on to the next ad. In order to minimize potential bias between measures, participants were asked to respond to the SI scale first before being presented with the MI scale (Bergkvist & Rossiter, 2009).

Attitude Toward the Ad. For the SI scale of A_{Ad}, participants indicated their agreement with the statement “Thinking about the ad for *//brand name//*, which of the following statements best describes your feeling about the ad?” on a seven-point Likert-type scale (1 = I disliked it extremely, 7 = I liked it extremely) adapted from Bergkvist and Rossiter (2009). For the MI scale, subjects indicated how well the four adjectives “dislike/like”, “good/bad”, “pleasant/unpleasant”, and “uninformative/informative” described their perception of the displayed ad on a seven-point semantic differential scale (Bergkvist & Rossiter, 2009).

Purchase Intention. We used the SI statement “If you were going to buy *//product category//*, how likely would you be to try *//brand name//*?” to measure PI_{Brand} on a seven-point Likert scale (1 = No chance or almost no chance, 7 = Certain or practically certain) (Bergkvist & Rossiter, 2009). On the MI scale, participants indicated how well the adjectives “unlikely/likely”, “probable/improbable”, “uncertain/certain”, and “impossible/possible” described the likelihood that they would buy the brand (Bergkvist & Rossiter, 2009).

Brand Attitude. To assess A_{Brand} on the SI scale, participants reported their feelings about the brand on a seven-point Likert-type scale (1 = I think it is extremely bad, 7 = I think it is extremely good) using the item “Thinking about *//product category//*, which of the following statements best describes your feeling about the *//brand name//* brand?” (Bergkvist & Rossiter, 2009). The MI measure consisted of a four-item seven-point semantic differential scale using the four pairs of adjectives “bad/good”, “like/dislike”, “pleasant/unpleasant”, and “useful/useless” based on Bergkvist and Rossiter (2009).

Attention Check. As an attention check, participants selected the product categories of the advertisements presented to them, choosing from a list of eight different options: fashion; automobile, insurance, medication/drugs, coffee, airline, consumer electronics, travel agency.

Demographics. Participants indicated their gender, age, education, and employment status. Income information was voluntary.

4 Analysis and Results

The following chapter is divided into two parts. First, the ICCs for T0 – T1 and T0 – T2 are calculated. Subsequently, the ICC values of the SI and MI scales are compared descriptively. In the second part, the difference between the test-retest reliability of the SI and MI scales is examined for statistical significance using PF and ZPF test statistics.

4.1 Test-Retest Reliability Based on the Intraclass-Correlation Coefficient

Intraclass-Correlation Coefficient

As explained earlier, the ICC is a recommended approach for assessing test-retest reliability (Shrout & Fleiss, 1979). While the ICC was introduced as a modification of the Pearson correlation coefficient, modern ICC is usually obtained by mean squares through analysis of variance (ANOVA), i.e., estimates of the population variances based on the variability among a given set of measures (Koo & Li, 2016). Therefore, the assumptions and data requirements of an ANOVA also apply to the ICC. Our data violated the assumption of normality. However, ANOVA has been shown to be robust against this violation (Schmider et al., 2010). Since the sample is sufficiently large, it can be regarded as approximately normally distributed based on the central limit theorem. In addition, variances of the different groups should be approximately equal to show homoscedasticity. Data were tested for homoscedasticity using Levene's Test and Bland Altman analysis and found to be generally homogeneous.

As there are various forms of the ICC that can produce different results, it is important to select the appropriate type based on the study. ICC types differ in terms of the underlying model (one-way random effects, two-way random effects, or two-way fixed effects), the type of raters or measurements (single versus mean of multiple), and the definition of the relationship being analyzed (consistency versus absolute agreement). Since we cannot consider our repeated measurements (i.e., after two weeks or after four months) as randomized, we chose a two-way mixed effects model. Moreover, we compared several measurements of the same constructs and thus opted for an ICC based on the mean of multiple measurements. Furthermore, we were interested in the absolute agreement between measurements, as measurements of the constructs in this study would be meaningless if there was no agreement between repeated measurements (Koo & Li, 2016).

Results and Interpretation

ICC estimates and their 95%-confidence interval (CI) were calculated based on a mean-rating (k = 3), absolute agreement, and two-way mixed-effects model (Koo & Li, 2016). We interpreted the ICC average measures as we analyzed the average of multiple measurements.[3] There is no standard procedure for interpreting ICC results. As a rule of thumb, a higher ICC indicates better consistency (Mehta et al., 2018). In clinical research, Portney (2020) and Koo and Li (2016) suggest that ICC values below 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values above 0.9 indicate exceptional reliability.

Table 3. Overview of ICC Scores

			T0 - T1			T0 - T2	
	Product Category	**SI ICC**	**MI ICC**	**Δ (SI - MI)**	**SI ICC**	**MI ICC**	**Δ (SI - MI)**
Attitude toward the Ad	Insurance	.828	.684	.144	.690	.818	-.128
	Jeans	.849	.743	.106	.767	.826	-.059
	Pain relievers	.750	.625	.125	.548	.638	-.090
	Coffee	.667	.667	.000	.752	.702	.050
Purchase Intention	Insurance	.817	.749	.068	.735	.754	-.019
	Jeans	.838	.826	.012	.711	.839	-.128
	Pain	.774	.685	.089	.609	.584	.025
	Coffee	.718	.746	-.028	.837	.688	.149
Attitude toward the Brand	Insurance	.784	.520	.264	.677	.751	-.074
	Jeans	.802	.697	.105	.774	.727	.047
	Pain relievers	.717	.584	.133	.474	.477	-.003
	Coffee	.657	.585	.072	.708	.672	.036

Notes. ICC calculation based on two-way mixed effects model, column effects fixed and row effects random.
Type A intraclass correlation coefficients using an absolute agreement definition. Two-way mixed effect model, interaction absent (i.e., average measures).

As outlined in Table 3, the ICC scores indicate moderate to good test-retest reliability for the SI and MI scales across both retest intervals. Results show that over the two-week period (T0 – T1) all but one SI measure (i.e., $PI_{\text{Brand, coffee}}$) perform better than the MI measures in terms of test-retest reliability. The difference in ICC values between SI and MI measures is greatest for attitude toward the brand and lowest for purchase intention. There is no clear pattern across product categories, however, the differences between SI and MI measures tend to be greatest for the insurance and pain relievers ads. For the four-month period (T0 – T2), the results paint a slightly different picture regarding test-retest reliability of the scale types. While some SI measures outperform their MI counterparts, such as the measures for the coffee brand, MI have

[3] A comparison of ICCs for single measurements did not change our findings. The respective values can be found in the Appendix.

higher ICC values in about half of the comparisons. Particularly, this is the case for attitude toward the ad and for the insurance ad. The results suggest that SI scales tend to perform better than MI scales at shorter retest intervals, whereas MI scales show a relative improvement in performance at longer retest intervals. Although Pearson's correlation is not the appropriate measure for comparing the SI and MI scales, it is still worthwhile as a robustness check to examine the correlation coefficients. Appendices 10 and 11 show that the observed pattern in the ICCs is also reflected in a similar form in the correlation coefficients.

An overview of all ICC scores for period T0 – T1 and T0 – T2 can be found in Appendices 2-7. Whether these differences are significant will be investigated in the following chapter.

4.2 Test for Significance

In order to compare the ICC scores between SI and MI measures it is important to note that they are nonindependent parameters, because they have been computed using the same sample. In addition, they are *nonoverlapping* coefficients, meaning that they do not have a variable in common. In contrast, when comparing *overlapping* correlations, one of the two variables being correlated is also involved in the other correlation. For example, the variables x_1, x_2, and x_3 yield three overlapping correlations (i.e., r_{12}, r_{13}, r_{23}), any two of which can be compared. The correlations are overlapping because in each comparison, they have one variable in common (e.g., x_1 when comparing r_{12} and r_{13}). This is also the case, when one wants to compare a correlation between Test A and a certain dependent variable with the correlation between Test B and the same dependent variable (Raghunathan et al., 1996). Extending the example to four variables, one can illustrate the difference between overlapping and nonoverlapping correlations. Four variables (i.e., x_1, x_2, x_3, x_4) yield six different correlations with 15 possible comparisons, three of which are nonoverlapping (e.g., r_{12} vs. r_{34}). Since we are interested in comparing the ICC scores of SI measures with the ICCs of MI measures over either the period T0 – T1 or T0 – T2, there is no single variable that is involved in the measurement of SI and MI measures performance. Therefore, we have to resort to a procedure to compare nonoverlapping correlations. We employ a test procedure dating back to Pearson and Filon (1898). The PF test statistic allows to compare any two nonoverlapping correlations, as represented in Table 4.

Table 4. Intercorrelations Among Variables

Variable	1	2	3	4
Time T0				
1 = measurement 1	1	r_{12}	r_{13}	r_{14}
2 = measurement 2		1	r_{23}	r_{24}
Time T1				
3 = measurement 1			1	r_{34}
4 = measurement 2				1

Note. Based on Raghunathan et al. (1996)

The PF statistic can be calculated by using the following equation, where r_{12} and r_{34} represent any two nonoverlapping correlations that are to be compared:

$$PF = \frac{\sqrt{N}(r_{12}-r_{34})}{\sqrt{(1-r_{12}^2)^2+(1-r_{34}^2)^2-k}} \quad (1)$$

k, denoting twice the large sample covariance between r_{12} and r_{34} (Raghunathan et al., 1996), is given by the following formula:

$$\begin{aligned} k = \; & (r_{13} - r_{23}*r_{12})*(r_{24} - r_{23}*r_{34}) \\ & + (r_{14} - r_{13}*r_{34})*(r_{23} - r_{13}*r_{12}) \\ & + (r_{13} - r_{14}*r_{34})*(r_{24} - r_{14}*r_{12}) \\ & + (r_{14} - r_{12}*r_{24})*(r_{23} - r_{24}*r_{34}). \end{aligned} \quad (2)$$

However, several authors (e.g., Weaver & Wuensch, 2013) recommend the ZPF procedure, based on Fisher z-transformation of Pearson's r (Weaver & Wuensch, 2013), which is more accurate and a theoretically better test statistic than the PF test statistic (Raghunathan et al., 1996). Thus, the ZPF statistic relies on r' values obtained via Fisher's r-to-z transformation[4], while k is still given by formula (2):

$$ZPF = \sqrt{\frac{N-3}{2}} \times \frac{r'_{12}-r'_{34}}{\sqrt{1-\frac{k}{2(1-r_{12}^2)(1-r_{34}^2)}}} \quad (3)$$

The expression in the denominator $\sqrt{1-\frac{k}{2(1-r_{12}^2)(1-r_{34}^2)}}$ represents the adjustment factor for nonindependence of any nonoverlapping correlations r_{12} and r_{34} underlying r'_{12} and r'_{34} (Raghunathan et al., 1996). This adjustment factor is the difference between the ZPF test (for

[4] Note that r-to-z transformed correlations should not be confused with z-scores or test values, which is why r' is a common representation.

comparing nonindependent correlations) and the independent samples Z-test (for comparing independent correlations). Appendix 9 provides an exemplary calculation of the PF and ZPF test statistics.

Table 5 summarizes the results of the PF and ZPF tests for all constructs across all categories. The test statistics can be treated as standard normal deviates. The results indicate that the SI and MI measure ICCs differ significantly in all but four comparisons in the T0 – T1 period. Consequently, most of the SI scales outperformed the MI counterparts. In the comparisons where no significant difference was found (i.e., $A_{Ad, coffee}$, $PI_{Brand, jeans}$, $PI_{Brand, coffee}$, $A_{Brand, coffee}$), SI scales performed on par with the MI scales. In the period T0 – T2, one SI measure outperformed its MI counterpart (i.e., $PI_{Brand, coffee}$), while two MI measures had a significantly higher ICC than the SI measures (i.e., $A_{Ad, insurance}$, $PI_{Brand, jeans}$). The remaining SI scales performed as well as the MI scales. Appendix 8 provides a detailed overview of the PF and ZPF test procedures.

Table 5. Results of Significance Test

		T0-T1 (N = 113)				T0-T2 (N = 55)			
	Product	**SI ICC**	**MI ICC**	**PF**	**ZPF**	**SI ICC**	**MI ICC**	**PF**	**ZPF**
Attitude Toward the Ad	Insurance	.828	.684	3.41***	3.76***	.690	.818	-2.38**	-2.61***
	Jeans	.849	.743	3.59***	4.14***	.767	.826	-1.22	-1.22
	Painkillers	.750	.625	3.11***	3.33***	.548	.638	-1.40	-1.39
	Coffee	.667	.667	.00	.00	.752	.702	.83	.81
Purchase Intention	Insurance	.817	.749	2.01**	2.07**	.735	.754	-.29	-.28
	Jeans	.838	.826	.45	.44	.711	.839	-2.66***	-3.06***
	Painkillers	.774	.685	2.21**	2.27**	.609	.584	.41	.40
	Coffee	.718	.746	-.66	-.66	.837	.688	2.59***	2.91***
Brand Attitude	Insurance	.784	.520	4.52***	5.15***	.677	.751	-1.01	-1.00
	Jeans	.802	.697	3.16***	3.44***	.774	.727	1.03	1.02
	Painkillers	.717	.584	2.82***	2.94***	.474	.477	-.04	-.04
	Coffee	.657	.585	1.40	1.40	.708	.672	.62	.60

Notes. ICC calculation based on two-way mixed effects model, column effects fixed and row effects random.
Type A intraclass correlation coefficients using an absolute agreement definition. Two-way mixed effect model, interaction absent (i.e., average measures).
*** p < .01, ** p < .05, * p < .10

5 Discussion, Limitations, and Future Research

Discussion

SI measures have both many benefits and drawbacks compared to MI measures. Consequently, the academic debate about the appropriate use of operationalizations has been ongoing for a long time. In marketing, the debate was reignited after the introduction of Rossiter's (2002) C-OAR-SE procedure and the publication of studies claiming that SI scales are as valid as MI scales (Bergkvist & Rossiter, 2007, 2009). While the scientific community is still debating whether this conclusion is accurate, there is a clear classification of both operationalizations: practical advantages of SI are pitted against the theoretical and statistical advantages of MI. On average, SI measures are cheaper, more efficient, more flexible, and less prone to measurement or sampling bias. In contrast, MI instruments typically benefit from higher reliability and validity. As marketing research emphasized psychometric properties in recent decades, psychometric scale development has become the dominant measurement paradigm in marketing. A tenet of psychometrics is that MI measures are always superior to their SI counterparts. Therefore, the use of SI measures has been considered a fatal flaw.

This study contributed to the debate by assessing test-retest reliability, a criterion previously overlooked by scholars. As a measure of consistency, high test-retest reliability scores ensure that measurements are stable and representative. Using a within-subject design, we measured consumer responses to advertisements (T0) and repeated the measurement after two weeks (T1) and after four months (T2), yielding 113 (T0 – T1) and 55 (T0 – T2) observations for the comparison periods. Specifically, participants indicated their attitudes toward the advertised product and the advertising brand as well as their purchase intention across ads from four different product categories. Comparing ICCs between SI and MI measures for the three constructs across the four categories, we found that the test-retest reliability of the SI measures was significantly higher or equal for all comparisons in the shorter interval (two weeks) and significantly higher or equal for all but two comparisons in the longer interval (four months). Specifically, the SI scales performed significantly better than the MI scales in the T0 – T1 period, with the exception of four comparisons. However, even in these comparisons, the ICC values of the SI measures were lower than those of the MI counterparts in only one case. In the T0 – T2 period, two MI measures were found to outperform its SI counterpart (i.e., $A_{Ad,\ insurance}$ and $PI_{Brand,\ jeans}$) and one SI measure outperformed its MI counterpart (i.e., $PI_{Brand,\ coffee}$). In the other comparisons, the two scale types performed equally well. While SI measures tend to have higher test-retest reliability than MI measures at the short retest interval, the results are more mixed at the long retest interval. Time effects could account for the difference between the short and long retest interval such that differences in reliability are attributed to the subjects rather than the measures. While SI measures can demonstrate their strengths in test-retest reliability at short measurement intervals, they suffer from response shifts in respondents' evaluation at long intervals. When respondents change their evaluation (e.g., due to altered priorities) on an SI scale, even a small adjustment carries a certain implication. In contrast, an adjustment of one or two items on an MI scale has a smaller impact on the overall evaluations. However, the results of our study show that SI and MI measures still tend to perform equally well at long retest intervals.

We do not mean to interpret the results of this study to conclude that SI scales should be used over MI scales. There are many studies that rightfully caution against the use of SI scales (e.g., Sarstedt, Diamantopoulos, & Salzberger, 2016). However, this study extends the discussion to test-retest reliability, and the findings show that SI measures indeed have equal or higher test-retest reliability scores than MI measures. Previous research on internal consistency reliability, predictive validity, and other criteria have each contributed to the discussion of using SI instead of MI scales. We arrive at similar results with respect to convergent and predictive validity. Convergent validity between the SI and MI scales at all three time points is consistently adequate to high (i.e., all above >.60). Predictive validity is lower for SI than for MI measures in our study, however the results show that the correlations between SI predictors and MI criterion measures are substantial (ranging between .40 and .90). Details can be found in Appendix 12. Overall, however, there is no definitive conclusion because the decision about which scale type to use depends on many different factors related to study design, research purposes, and many more. Our investigation of test-retest reliability of SI versus MI scales is one further piece of the puzzle in the discussion, providing impetus for researchers to be receptive to the use of SI scales when conducting longitudinal studies with repeated measures. Nevertheless, especially in short retest periods, SI scales should be preferred over MI scales if high test-retest reliability is sought.

Limitations and Future Research

Our study ties into the ongoing discussion about the use of SI and MI scales in marketing and therefore builds on the conceptual considerations underlying C-OAR-SE. As described earlier, the introduction of Rossiter's (2002) C-OAR-SE procedure has reinvigorated the debate and generated a steady stream of studies. It is important to keep this context in mind when interpreting and classifying the results of this study. In line with earlier studies (e.g., Bergkvist &

Rossiter, 2007; Diamantopoulos et al., 2012), we used doubly concrete constructs in our investigation. Therefore, the findings only apply to such constructs. Specifically, the findings should not be transferred to SI measures of constructs that are not doubly concrete. Our results do not suggest, and we do not believe, that such constructs perform on par with MI counterparts.

In addition, scholars have disagreed about whether and how appropriate SI measures can be identified for doubly concrete constructs. It remains unclear how to determine the *best fitting* SI from a set of MI scales (Diamantopoulos, 2005). Neither Rossiter (2002) nor Bergkvist and Rossiter (2007) made any recommendations on how to make this choice. Sarstedt, Diamantopoulos, Salzberger, and Baumgartner (2016) tested different identification approaches (rater assessments and statistical criteria, such as item-to-total correlations or principal components analysis loadings) and found that even in the best-case scenario, there was only a 50% chance of identifying the best item. In addition, the majority of expert raters concluded that SI measures were inadequate for measuring doubly concrete constructs, such as attitude toward the ad and attitude toward the brand. The results also showed that expert and non-expert raters differed significantly in their judgments of item suitability. For many constructs, however, there are no MI scales from which SI can be extracted. This raises the question of how the best tailored SI measures can be developed in the absence of an MI scale, as there are no established procedures for creating SI scales (Diamantopoulos et al., 2012). However, such tailor-made global SI measures did not outperform other approaches when tested (Sarstedt, Diamantopoulos, Salzberger, & Baumgartner, 2016). In response to these findings, Bergkvist (2016) argued that the solution to selecting the best SI measure is to rely on correct "expert judgment of the content validity of items given how the construct has been defined" (p. 3429). That is, expert raters should indicate whether the object and the attribute of a construct are concrete, given the researcher's definition of the construct.[5] If the construct is doubly concrete, the researcher should subsequently decide which SI measure to use to capture it. However, some argued that the disagreement about what makes a construct doubly concrete stems from different evaluations of objects and attributes – for example, the difference between considering an object's denotative meaning versus its connotative meaning (Diamantopoulos, 2005; Sarstedt, Diamantopoulos, Salzberger, & Baumgartner, 2016). Given this disagreement, future research is needed to clarify how to identify SI measures for doubly concrete constructs.

[5] Regarding the object, Rossiter (2002, p. 321) suggests asking experts whether "there is only one object to be rated and nearly everyone would describe this object identically", and regarding the attribute, whether "nearly everyone would describe this attribute identically". He considers a construct as doubly concrete, if most experts answer both questions with "yes".

Furthermore, an important caveat to our findings arises from one of the central assumptions of the study. That is, we assumed that the constructs we measured in response to different advertisements would remain consistent across measurement intervals. There is no evidence to suggest that this assumption was violated, but we cannot conclusively test it. However, there is no indication that one type of scale is a better reflection of changes in constructs over time. Since we employed a within-subjects design, any potential change in the constructs should be evenly reflected in both SI and MI measures. Additionally, future research should investigate whether the results hold when conducting a study with more stable or inherent constructs (i.e., personality types).

Moreover, the present study focuses on examining the test-retest reliability of SI vs. MI scales by using different doubly concrete advertisements as stimuli. In contrast, Fisher et al. (2016) examined the test-retest reliability of (presumably non-doubly concrete) constructs in organizational research (e.g., work centrality, supervisor support) without manipulations. However, they did not compare SI and MI measures. As an extension of our findings and the study of Fisher et al. (2016), an investigation of test-retest-reliability of doubly concrete constructs without the use of manipulations, such as Internet shopping motivation (e.g., Ganesh et al., 2010), might reveal that the comparison between SI and MI scales yields different results in the absence of visual stimuli.

Conclusion

In summary, our study extends the discussion on the usage of SI versus MI scales. The results regarding the test-retest reliability of SI measures are promising. The doubly concrete SI constructs used in our study performed better than MI counterparts in the short retest interval of two weeks and equally well in the long retest interval of four months. We do not mean to recommend the use of SI over MI measures in general. Researchers and practitioners should carefully consider whether the individual constructs they wish to measure are doubly concrete and then weigh the advantages and disadvantages of SI and MI scales. However, we conclude that SI scales perform at least as well as MI scales regarding their test-retest reliability.

Appendix

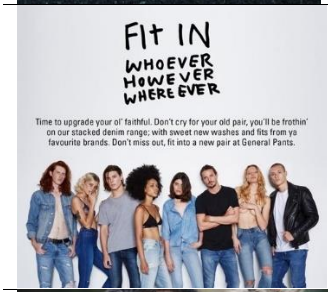
Appendix 1. Overview of Measures and Stimuli

Construct/ Variable	SI vs. MI	Item/Proxy	Precedents/ Sources
Attitude Toward the Ad	SI	7-point Likert scale (1 = I disliked it extremely, 7 = I liked it extremely) (1) Thinking about the ad for *//brand name//*, which of the following statements best describes your feeling about **the ad?**	Bergkvist & Rossiter, 2009
Attitude Toward the Ad	MI	7-point semantic differential scale Below you will find four pairs of adjectives. Indicate how well one or the other adjective in each pair describes how you perceived **the ad** for *//brand name//*. (1) Dislike / Like (2) Good / Bad (3) Pleasant / Unpleasant (4) Uninformative / Informative	Bergkvist & Rossiter, 2009
Purchase Intention	SI	7-point Likert scale (1 = No chance or almost no chance, 7 = Certain or practically certain) (1) If you were going to buy //product category//, how likely would you be to try *//brand name//*?	Bergkvist & Rossiter, 2009
Purchase Intention	MI	7-point semantic differential scale Below you will find four pairs of adjectives. Indicate how well one or the other adjective in each pair describes the likelihood that you would try //brand name// if you were to buy *//product category//*. (1) Unlikely / Likely (2) Probable / Improbable (3) Uncertain / Certain (4) Impossible / Possible	Bergkvist & Rossiter, 2009
Brand Attitude	SI	7-point Likert scale (1 = I think it is extremely bad, 7 = I think it is extremely good) (1) Thinking about *//product category//*, which of the following statements best describes your feeling about the *//brand name//* **brand**.	Bergkvist & Rossiter, 2009
Brand Attitude	MI	7-point semantic differential scale Below you will find four pairs of adjectives. Indicate how well one or the other adjective in each pair describes your overall feeling of the *//brand name//* **brand**. (1) Bad / Good (2) Like / Dislike (3) Pleasant / Unpleasant (4) Useful / Useless	Bergkvist & Rossiter, 2009
		Control Variables	
Familiarity with the Ad	SI	7-point Likert scale (1 = strongly disagree, 7 = strongly agree) (1) I am familiar with this ad.	Self-devel-oped
Familiarity with the Brand	SI	7-point Likert scale (1 = strongly disagree, 7 = strongly agree) (1) I am familiar with this brand.	Self-devel-oped
		Attention Check	
Attention Check		A dropdown with eight different categories was provided for each advertisement: fashion; automobile, insurance, mediation/drugs, coffee, airline, consumer electronics, travel agency Please select the type of product or service that was advertised in the ads you saw. (1) Ad#1 (2) Ad#2 (3) Ad#3 (4) Ad#4	Self-devel-oped

Appendix 1. Overview of Measures and Stimuli (continued)

Construct/ Variable	SI vs. MI	Item/Proxy	Precedents/ Sources
		Demographics	
Gender		What gender do you identify as? • Male • Female • Non-binary/third gender • Prefer not to say	Self-developed
Age		How old are you? • Below 18 • 18 – 25 • 26 – 35 • 36 – 45 • 46- 60 • Above 60	Self-developed
Education		What is your highest degree? • High school • Undergraduate • Postgraduate • Diploma • Other	Lo et al., 2019
Employment		What is your current occupation? • Full-time • Part-time • Student • Stay-at-home parent • Unemployed • Retired	Lo et al., 2019
Income		What is your average yearly net household income? (voluntary information) • < $10,000 • $10,000 – $50,000 • $50,001 – $90,000 • $90,001 - $150,000 • > $150,001	Lo et al., 2019

Appendix 1. Overview of Measures and Stimuli (continued)

<table>
<tr><th>Advertisement</th><th>Company/Brand – Country</th><th>Source</th></tr>
<tr><td>
</td><td>Pacific Blue Cross – Canada</td><td>https://funeasypopular.com/obamacare-survived-now-what-5-things-insurers-should-know/</td></tr>
<tr><td>
</td><td>General Pants – Australia</td><td>https://www.gen-eralpants.com/au/blog/fit-in</td></tr>
<tr><td>
</td><td>BHI Migraine Relief - USA</td><td>n/a (as of November, 23, 2022)</td></tr>
<tr><td>
</td><td>7-Eleven – Australia</td><td>https://www.7eleven.com.au/</td></tr>
</table>

Appendix 2. Results of ICC Calculation for Attitude Toward the Ad (T0 – T1)

Attitude Toward the Ad (T0 – T1)			ICC[b]	95% Confidence Interval		F Test with True Value 0			
				Lower Bound	Upper Bound	Value	df1	df2	Sig
Insurance	SI	Single Measures	.707[a]	.591	.792	6.186***	112	112	.000
		Average Measures	.828[c]	.743	.884	6.186***	112	112	.000
	MI	Single Measures	.520[a]	.372	.642	3.161***	112	112	.000
		Average Measures	.684[c]	.542	.782	3.161***	112	112	.000
Jeans	SI	Single Measures	.738[a]	.641	.811	6.598***	112	112	.000
		Average Measures	.849[c]	.781	.896	6.598***	112	112	.000
	MI	Single Measures	.591[a]	.457	.700	3.873***	112	112	.000
		Average Measures	.743[c]	.627	.823	3.873***	112	112	.000
Pain	SI	Single Measures	.600[a]	.467	.706	4.104***	112	112	.000
		Average Measures	.750[c]	.636	.828	4.104***	112	112	.000
	MI	Single Measures	.455[a]	.296	.589	2.668***	112	112	.000
		Average Measures	.625[c]	.457	.742	2.668***	112	112	.000
Coffee	SI	Single Measures	.501[a]	.350	.626	3.035***	112	112	.000
		Average Measures	.667[c]	.518	.770	3.035***	112	112	.000
	MI	Single Measures	.500[a]	.348	.627	2.993***	112	112	.000
		Average Measures	.667[c]	.517	.771	2.993***	112	112	.000

Notes. Two-way mixed effects model, column effects fixed and row effects random.
[a] Two-way mixed effect model, with interaction (i.e., the estimator is the same, whether the interaction effect is present or not).
[b] Type A intraclass correlation coefficients using an absolute agreement definition.
[c] Two-way mixed effect model, interaction absent (i.e., estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise).
*** p < .01, ** p < .05, * p < .10

Appendix 3. Results of ICC Calculation for Purchase Intention (T0 – T1)

Purchase Intention (T0 – T1)			ICC[b]	95% Confidence Interval		F Test with True Value 0			
				Lower Bound	Upper Bound	Value	df1	df2	Sig
Insurance	SI	Single Measures	.691[a]	.546	.789	6.089***	112	112	.000
		Average Measures	.817[c]	.706	.882	6.089***	112	112	.000
	MI	Single Measures	.599[a]	.462	.707	4.155***	112	112	.000
		Average Measures	.749[c]	.632	.828	4.155***	112	112	.000
Jeans	SI	Single Measures	.721[a]	.619	.799	6.264***	112	112	.000
		Average Measures	.838[c]	.765	.888	6.264***	112	112	.000
	MI	Single Measures	.704[a]	.597	.786	5.729***	112	112	.000
		Average Measures	.826[c]	.748	.880	5.729***	112	112	.000
Pain	SI	Single Measures	.631[a]	.501	.732	4.596***	112	112	.000
		Average Measures	.774[c]	.668	.845	4.596***	112	112	.000
	MI	Single Measures	.521[a]	.374	.643	3.189***	112	112	.000
		Average Measures	.685[c]	.544	.783	3.189***	112	112	.000
Coffee	SI	Single Measures	.560[a]	.412	.677	3.725***	112	112	.000
		Average Measures	.718[c]	.584	.808	3.725***	112	112	.000
	MI	Single Measures	.594[a]	.460	.702	3.910***	112	112	.000
		Average Measures	.746[c]	.631	.825	3.910***	112	112	.000

Notes. Two-way mixed effects model, column effects fixed and row effects random.
[a] Two-way mixed effect model, with interaction (i.e., the estimator is the same, whether the interaction effect is present or not).
[b] Type A intraclass correlation coefficients using an absolute agreement definition.
[c] Two-way mixed effect model, interaction absent (i.e., estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise).
*** $p < .01$, ** $p < .05$, * $p < .10$

Appendix 4. Results of ICC Calculation for Attitude Toward the Brand (T0 – T1)

Brand Attitude (T0 – T1)			ICC[b]	95% Confidence Interval		F Test with True Value 0			
				Lower Bound	Upper Bound	Value	df1	df2	Sig
Insurance	SI	Single Measures	.645[a]	.520	.743	4.812***	112	112	.000
		Average Measures	.784[c]	.684	.852	4.812***	112	112	.000
	MI	Single Measures	.352[a]	.181	.502	2.102***	112	112	.000
		Average Measures	.520[c]	.307	.669	2.102***	112	112	.000
Jeans	SI	Single Measures	.669[a]	.554	.759	5.059***	112	112	.000
		Average Measures	.802[c]	.713	.863	5.059***	112	112	.000
	MI	Single Measures	.535[a]	.389	.655	3.281***	112	112	.000
		Average Measures	.697[c]	.560	.791	3.281***	112	112	.000
Pain	SI	Single Measures	.559[a]	.417	.674	3.629***	112	112	.000
		Average Measures	.717[c]	.588	.805	3.629***	112	112	.000
	MI	Single Measures	.412[a]	.247	.554	2.399***	112	112	.000
		Average Measures	.584[c]	.396	.713	2.399***	112	112	.000
Coffee	SI	Single Measures	.489[a]	.335	.618	2.909***	112	112	.000
		Average Measures	.657[c]	.502	.764	2.909***	112	112	.000
	MI	Single Measures	.414[a]	.248	.556	2.403***	112	112	.000
		Average Measures	.585[c]	.398	.715	2.403***	112	112	.000

Notes. Two-way mixed effects model, column effects fixed and row effects random.
[a] Two-way mixed effect model, with interaction (i.e., the estimator is the same, whether the interaction effect is present or not).
[b] Type A intraclass correlation coefficients using an absolute agreement definition.
[c] Two-way mixed effect model, interaction absent (i.e., estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise).
*** p < .01, ** p < .05, * p < .10

Appendix 5. Results of ICC Calculation for Attitude Toward the Ad (T0 – T2)

Attitude Toward the Ad (T0 – T2)			ICC[b]	95% Confidence Interval		F Test with True Value 0			
				Lower Bound	Upper Bound	Value	df1	df2	Sig
Insurance	SI	Single Measures	.527[a]	.305	.694	3.199***	54	54	.000
		Average Measures	.690[c]	.468	.820	3.199***	54	54	.000
	MI	Single Measures	.692[a]	.524	.808	5.433***	54	54	.000
		Average Measures	.818[c]	.688	.894	5.433***	54	54	.000
Jeans	SI	Single Measures	.622[a]	.431	.760	4.307***	54	54	.000
		Average Measures	.767[c]	.602	.864	4.307***	54	54	.000
	MI	Single Measures	.704[a]	.539	.816	6.004***	54	54	.000
		Average Measures	.826[c]	.701	.899	6.004***	54	54	.000
Pain	SI	Single Measures	.377[a]	.129	.582	2.221***	54	54	.002
		Average Measures	.548[c]	.229	.735	2.221***	54	54	.002
	MI	Single Measures	.469[a]	.238	.651	2.787***	54	54	.000
		Average Measures	.638[c]	.384	.788	2.787***	54	54	.000
Coffee	SI	Single Measures	.603[a]	.403	.748	3.984***	54	54	.000
		Average Measures	.752[c]	.574	.856	3.984***	54	54	.000
	MI	Single Measures	.541[a]	.322	.704	3.319***	54	54	.000
		Average Measures	.702[c]	.488	.826	3.319***	54	54	.000

Notes. Two-way mixed effects model, column effects fixed and row effects random.
[a] Two-way mixed effect model, with interaction (i.e., the estimator is the same, whether the interaction effect is present or not).
[b] Type A intraclass correlation coefficients using an absolute agreement definition.
[c] Two-way mixed effect model, interaction absent (i.e., estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise).
*** $p < .01$, ** $p < .05$, * $p < .10$

Appendix 6. Results of ICC Calculation for Purchase Intention (T0 – T2)

Purchase Intention (T0 – T2)			ICC[b]	95% Confidence Interval		F Test with True Value 0			
				Lower Bound	Upper Bound	Value	df1	df2	Sig
Insurance	SI	Single Measures	.582[a]	.376	.733	3.748***	54	54	.000
		Average Measures	.735[c]	.546	.846	3.748***	54	54	.000
	MI	Single Measures	.605[a]	.405	.749	4.007***	54	54	.000
		Average Measures	.754[c]	.576	.856	4.007***	54	54	.000
Jeans	SI	Single Measures	.551[a]	.337	.711	3.428***	54	54	.000
		Average Measures	.711[c]	.504	.831	3.428***	54	54	.000
	MI	Single Measures	.722[a]	.561	.830	6.582***	54	54	.000
		Average Measures	.839[c]	.719	.907	6.582***	54	54	.000
Pain	SI	Single Measures	.438[a]	.195	.629	2.529***	54	54	.000
		Average Measures	.609[c]	.326	.772	2.529***	54	54	.000
	MI	Single Measures	.413[a]	.172	.608	2.429***	54	54	.001
		Average Measures	.584[c]	.293	.756	2.429***	54	54	.001
Coffee	SI	Single Measures	.719[a]	.563	.826	6.056***	54	54	.000
		Average Measures	.837[c]	.720	.905	6.056***	54	54	.000
	MI	Single Measures	.525[a]	.307	.691	3.252***	54	54	.000
		Average Measures	.688[c]	.469	.818	3.252***	54	54	.000

Notes. Two-way mixed effects model, column effects fixed and row effects random.
[a] Two-way mixed effect model, with interaction (i.e., the estimator is the same, whether the interaction effect is present or not).
[b] Type A intraclass correlation coefficients using an absolute agreement definition.
[c] Two-way mixed effect model, interaction absent (i.e., estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise).
*** p < .01, ** p < .05, * p < .10

Appendix 7. Results of ICC Calculation for Attitude Toward the Brand (T0 – T2)

Brand Attitude (T0 – T2)			ICC[b]	95% Confidence Interval		F Test with True Value 0			
				Lower Bound	Upper Bound	Value	df1	df2	Sig
Insurance	SI	Single Measures	.512[a]	.285	.684	3.060***	54	54	.000
		Average Measures	.677[c]	.444	.812	3.060***	54	54	.000
	MI	Single Measures	.601[a]	.400	.746	3.964***	54	54	.000
		Average Measures	.751[c]	.572	.855	3.964***	54	54	.000
Jeans	SI	Single Measures	.632[a]	.444	.767	4.547***	54	54	.000
		Average Measures	.774[c]	.615	.868	4.547***	54	54	.000
	MI	Single Measures	.571[a]	.364	.724	3.691***	54	54	.000
		Average Measures	.727[c]	.534	.840	3.691***	54	54	.000
Pain	SI	Single Measures	.310[a]	.048	.532	1.884**	54	54	.011
		Average Measures	.474[c]	.092	.694	1.884**	54	54	.011
	MI	Single Measures	.313[a]	.052	.533	1.898**	54	54	.010
		Average Measures	.477[c]	.098	.696	1.898**	54	54	.010
Coffee	SI	Single Measures	.548[a]	.334	.709	3.416***	54	54	.000
		Average Measures	.708[c]	.501	.830	3.416***	54	54	.000
	MI	Single Measures	.506[a]	.279	.680	3.020***	54	54	.000
		Average Measures	.672[c]	.436	.809	3.020***	54	54	.000

Notes. Two-way mixed effects model, column effects fixed and row effects random.
[a] Two-way mixed effect model, with interaction (i.e., the estimator is the same, whether the interaction effect is present or not).
[b] Type A intraclass correlation coefficients using an absolute agreement definition.
[c] Two-way mixed effect model, interaction absent (i.e., estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise).
*** $p < .01$, ** $p < .05$, * $p < .10$

Appendix 8. Overview of Results of PF and ZPF Test Statistics

Variable	Product	SI ICC[a]	MI ICC[a]	k	Z SI ICC	Z MI ICC	PF	ZPF
T0 – T1 (N = 113)								
Attitude Toward the Ad	Insurance	.828	.684	.180	1.182	.837	3.41***	3.76***
	Jeans	.849	.743	.180	1.253	.957	3.59***	4.14***
	Painkillers	.750	.625	.381	.973	.733	3.11***	3.33***
	Coffee	.667	.667	.354	.805	.805	.00	.00
Purchase Intention	Insurance	.817	.749	.174	1.148	.971	2.01**	2.07**
	Jeans	.838	.826	.108	1.214	1.175	.45	.44
	Painkillers	.774	.685	.259	1.030	.838	2.21**	2.27**
	Coffee	.718	.746	.230	.904	.964	-.66	-.66
Attitude Toward the Brand	Insurance	.784	.520	.295	1.056	.576	4.52***	5.15***
	Jeans	.802	.697	.267	1.104	.861	3.16***	3.44***
	Painkillers	.717	.584	.420	.901	.669	2.82***	2.94***
	Coffee	.657	.585	.458	.788	.670	1.40	1.40
T0 – T2 (N = 55)								
Attitude Toward the Ad	Insurance	.690	.818	.225	.848	1.151	-2.38**	-2.61***
	Jeans	.767	.826	.141	1.013	1.175	-1.22	-1.22
	Painkillers	.548	.638	.614	.616	.755	-1.40	-1.39
	Coffee	.752	.702	.245	.978	.871	.83	.81
Purchase Intention	Insurance	.735	.754	.156	.940	.982	-.29	-.28
	Jeans	.711	.839	.205	.889	1.218	-2.66***	-3.06***
	Painkillers	.609	.584	.626	.707	.669	.41	.40
	Coffee	.837	.688	.185	1.211	.844	2.59***	2.91***
Attitude Toward the Brand	Insurance	.677	.751	.189	.824	.975	-1.01	-1.00
	Jeans	.774	.727	.268	1.030	.922	1.03	1.02
	Painkillers	.474	.477	.878	.515	.519	-.04	-.04
	Coffee	.708	.672	.362	.883	.814	.62	.60

Notes. ICC calculation based on two-way mixed effects model, column effects fixed and row effects random. PF and ZPF test statistics based on nonoverlapping correlations.

[a] Type A intraclass correlation coefficients using an absolute agreement definition. Two-way mixed effect model, interaction absent (i.e., average measures).

*** p < .01, ** p < .05, * p < .10

Appendix 9. Exemplary Calculation of PF and ZPF Test Statistics

We reorganized Table 4 to be able to compare the SI (r_{12}) and MI (r_{34}) measures in the periods T0 – T1 and T0 – T2. Table 3 exemplary presents the values of attitude toward the ad for the product category insurance in period T0 – T1.

Table 3. Intercorrelations Among SI and MI Measures During T0 and T1

Attitude toward the Ad (Insurance)	1	2	3	4
1 = SI, T0	1	.828	.861	.724
2 = SI, T1		1	.694	.814
3 = MI, T0			1	.684
4 = MI, T1				1

Notes. Exemplary calculation for attitude toward the ad (insurance).
N = 113.

After reorganizing the table, we calculated k using formula (2):

$$\begin{aligned} k_{\text{(AAd. Insurance)}} = \; & (.861 - .694 * .828) * (.814 - .694 * .684) \\ & + (.724 - .861 * .684) * (.694 - .861 * .828) \\ & + (.861 - .724 * .684) * (.814 - .724 * .828) \\ & + (.724 - .828 * .814) * (.694 - .814 * .684) \\ & = .180 \end{aligned}$$

Further, to calculate the PF test statistics, equation (1) is used:

$$\mathrm{PF} = \frac{\sqrt{113}(.828 - .684)}{\sqrt{(1 - .828^2)^2 + (1 - .684^2)^2 - .180}} = 3.406$$

The z-based PF (ZPF) test statistics can be calculated using equation (3):

$$\mathrm{ZPF} = \sqrt{\frac{113-3}{2}} \times \frac{1.182 - .837}{\sqrt{1 - \frac{.180}{2(1-.828^2)(1-.684^2)}}} = 3.765.$$

Appendix 10. Correlations (T0 – T1)

↓ T1		T0→	1	2	3	4	5	6	7	8	9	10	11	12
Single-Item Measures														
1	Insurance	A_{Ad}	.726***	.599***	.645***	.480***	.523***	.442***	.452***	.538***	.400***	.451***	.433***	.445***
2		PI_{Brand}	.532***	.721***	.630***	.475***	.577***	.465***	.452***	.476***	.384***	.436***	.436***	.394***
3		A_{Brand}	.589***	.615***	.658***	.506***	.504***	.466***	.510***	.514***	.498***	.468***	.383***	.416***
4	Jeans	A_{Ad}	.521***	.479***	.494***	.737***	.711***	.697***	.580***	.629***	.584***	.438***	.380***	.496***
5		PI_{Brand}	.521***	.552***	.530***	.676***	.725***	.634***	.530***	.630***	.506***	.409***	.463***	.471***
6		A_{Brand}	.556***	.517***	.549***	.706***	.677***	.671***	.584***	.628***	.570***	.458***	.416***	.491***
7	Pain	A_{Ad}	.456***	.426***	.399***	.423***	.486***	.382***	.608***	.608***	.569***	.366***	.308***	.357***
8		PI_{Brand}	.562***	.599***	.506***	.541***	.649***	.550***	.510***	.644***	.531***	.390***	.461***	.407***
9		A_{Brand}	.607***	.573***	.541***	.456***	.525***	.460***	.531***	.610***	.569***	.381***	.368***	.394***
10	Coffee	A_{Ad}	.320***	.330***	.293***	.369***	.424***	.316***	.364***	.437***	.374***	.506***	.506***	.430***
11		PI_{Brand}	.347***	.386***	.396***	.525***	.613***	.523***	.426***	.524***	.421***	.515***	.581***	.557***
12		A_{Brand}	.366***	.353***	.310***	.413***	.465***	.396***	.485***	.582***	.481***	.513***	.453***	.492***
Multi-Item Measures														
1	Insurance	A_{Ad}	.519***	.459***	.412***	.301***	.273***	.330***	.318***	.324***	.279***	.335***	.337***	.305***
2		PI_{Brand}	.377***	.616***	.495***	.360***	.375***	.391***	.327***	.322***	.309***	.299***	.425***	.314***
3		A_{Brand}	.273***	.274***	.356***	.298***	.213**	.339***	.275***	.285***	.294***	.336***	.333***	.352***
4	Jeans	A_{Ad}	.422***	.311***	.295***	.592***	.600***	.578***	.538***	.526***	.511***	.404***	.397***	.409***
5		PI_{Brand}	.461***	.437***	.372***	.668***	.703***	.626***	.510***	.528***	.494***	.449***	.503***	.414***
6		A_{Brand}	.378***	.274***	.276***	.508***	.521***	.534***	.461***	.389***	.443***	.352***	.331***	.364***
7	Pain	A_{Ad}	.290***	.300***	.292***	.313***	.318***	.338***	.455***	.451***	.449***	.258***	.269***	.308***
8		PI_{Brand}	.360***	.442***	.373***	.441***	.520***	.422***	.462***	.523***	.463***	.327***	.406***	.326***
9		A_{Brand}	.320***	.252***	.293***	.241**	.242***	.308***	.362***	.387***	.412***	.276***	.249***	.354***
10	Coffee	A_{Ad}	.249***	.186**	.242***	.286***	.276***	.317***	.365***	.231**	.341***	.499***	.361***	.460***
11		PI_{Brand}	.204**	.297***	.225**	.316***	.351***	.300***	.306***	.314***	.275***	.463***	.595***	.429***
12		A_{Brand}	.124	.064	.162*	.227*	.182*	.269***	.335***	.200*	.359***	.431***	.310***	.412***

Note. *** $p < .01$, ** $p < .05$, * $p < .10$

Appendix 11. Correlations (T0 – T2)

↓ T2		T0→	1	2	3	4	5	6	7	8	9	10	11	12
Single-Item Measures														
1	Insurance	A_{Ad}	.525***	.324**	.285**	.257*	.347***	.282**	.310**	.333**	.464***	.436***	.209	.343**
2		PI_{Brand}	.488***	.580***	.525***	.413***	.472***	.527***	.356***	.476***	.525***	.476***	.483***	.343**
3		A_{Brand}	.411***	.455***	.508***	.274**	.337**	.351***	.394***	.528***	.612***	.470***	.346***	.364***
4	Jeans	A_{Ad}	.199	.238*	.256*	.638***	.584***	.758***	.485***	.405***	.497***	.583***	.389***	.437***
5		PI_{Brand}	.205	.234*	.172	.540***	.553***	.669***	.501***	.508***	.566***	.572***	.461***	.464***
6		A_{Brand}	.203	.187	.249*	.534***	.490***	.659***	.483***	.460***	.552***	.579***	.416***	.447***
7	Pain	A_{Ad}	.077	.394***	.167	.371***	.524***	.484***	.410***	.393***	.393***	.401***	.530***	.436***
8		PI_{Brand}	.263*	.555***	.235*	.379***	.451***	.409***	.384***	.438***	.383***	.398***	.418***	.384***
9		A_{Brand}	.266**	.380***	.140	.292**	.415***	.364***	.329**	.281**	.307**	.268**	.513***	.451***
10	Coffee	A_{Ad}	.265*	.465***	.318**	.272**	.402***	.392***	.356***	.416***	.538***	.615***	.512***	.509***
11		PI_{Brand}	.302**	.512***	.314**	.357***	.499***	.425***	.271**	.422***	.360***	.462***	.718***	.517***
12		A_{Brand}	.449***	.375***	.193	.332**	.452***	.388***	.162	.218	.291**	.340**	.586***	.559***
Multi-Item Measures														
1	Insurance	A_{Ad}	.605***	.335**	.428***	.387***	.378***	.407***	.324**	.360***	.555***	.408***	.304**	.435***
2		PI_{Brand}	.556***	.368***	.372***	.261	.317**	.324**	.139	.321**	.257	.129	.279**	.184
3		A_{Brand}	.578***	.396***	.498***	.244	.321**	.263	.164	.321**	.458***	.266**	.382***	.282**
4	Jeans	A_{Ad}	.301**	.207	.292**	.638***	.574***	.678***	.434***	.355***	.476***	.587***	.440***	.479***
5		PI_{Brand}	.232	.212	.136	.574***	.525***	.665***	.466***	.493***	.506***	.525***	.454***	.428***
6		A_{Brand}	.282**	.159	.264	.576***	.466***	.607***	.480***	.360***	.494***	.547***	.435***	.489***
7	Pain	A_{Ad}	.199	.324**	.209	.464***	.484***	.484***	.339**	.345***	.369***	.397***	.543***	.535***
8		PI_{Brand}	.357***	.554***	.296**	.495***	.564***	.518***	.329**	.386***	.314**	.298**	.538***	.403***
9		A_{Brand}	.307**	.277**	.147	.293**	.348***	.330**	.227	.241	.304**	.282**	.487***	.421***
10	Coffee	A_{Ad}	.326**	.410***	.252	.225	.248	.287**	.185	.328**	.533***	.488***	.436***	.383***
11		PI_{Brand}	.311**	.421***	.267**	.398***	.532***	.456***	.250	.389***	.383***	.402***	.677***	.479***
12		A_{Brand}	.433***	.276**	.165	.241	.336**	.275**	.090	.138	.296**	.370***	.527***	.500***

Note. *** p < .01, ** p < .05, * p < .10

Appendix 12. Psychometric Attributes

Variable	Product	Internal Reliability		Convergent Validity			Predictive Validity					
		C.A.	C.f.A.	SI MI[a]			MI → MI[b]			SI → MI[c]		
		T0		T0	T1	T2	T0	T1	T2	T0	T1	T2
Insurance	A_{Ad}	.796	.555	.665	.708	.810	.808	.680	.774	.546	.408	.584
	PI_{Brand}	.780	.582	.674	.802	.669						
	A_{Brand}	.893	.426	.617	.609	.686	.712	.665	.677	.653	.725	.578
Jeans	A_{Ad}	.851	.665	.752	.851	.887	.845	.901	.871	.646	.799	.784
	PI_{Brand}	.862	.682	.767	.842	.917						
	A_{Brand}	.913	.546	.706	.814	.850	.760	.833	.822	.799	.835	.857
Pain-killers	A_{Ad}	.831	.586	.698	.740	.825	.825	.842	.829	.582	.576	.779
	PI_{Brand}	.808	.650	.725	.780	.855						
	A_{Brand}	.907	.483	.662	.711	.837	.785	.793	.808	.740	.754	.795
Coffee	A_{Ad}	.804	.489	.627	.696	.759	.831	.839	.721	.475	.634	.729
	PI_{Brand}	.781	.608	.689	.760	.891						
	A_{Brand}	.887	.409	.602	.623	.872	.688	.720	.744	.637	.597	.785

Notes. C.A. = Cronbach's α. C.f.A. = Internal reliability estimate based on correction for attenuation formula.
[a] Correlation between SI and MI measures.
[b] Correlation between MI measures of A_{Ad} (predictor) and A_{Brand} (criterion) and between A_{Brand} (predictor) and PI_{Brand} (criterion).
[c] Correlation between the SI measure of A_{Ad} (predictor) and MI measure of A_{Brand} (criterion) as well as between SI measure of A_{Brand} (predictor) and MI measure of PI_{Brand} (criterion).

References

Adigüzel, F., & Wedel, M. (2008). Split Questionnaire Design for Massive Surveys. *Journal of Marketing Research*, *45*(5), 608–617. https://doi.org/10.1509/jmkr.45.5.608

Ang, L., & Eisend, M. (2018). Single Versus Multiple Measurement of Attitudes. *Journal of Advertising Research*, *58*(2), 218–227. https://doi.org/10.2501/JAR-2017-001

Baumgartner, H., & Homburg, C. (1996). Applications of Structural Equation Modeling in Marketing and Consumer Research: A Review. *International Journal of Research in Marketing*, *13*(2), 139–161. https://doi.org/10.1016/0167-8116(95)00038-0

Bearden, W. O., Netemeyer, R. G., & Haws, K. L. (Eds.). (2011). *Handbook of Marketing Scales: Multi-Item Measures for Marketing and Consumer Behavior Research* (3rd ed.). SAGE Publications Inc.

Bergkvist, L. (2015). Appropriate Use of Single-Item Measures Is Here to Stay. *Marketing Letters*, *26*(3), 245–255. https://doi.org/10.1007/s11002-014-9325-y

Bergkvist, L. (2016). The Nature of Doubly Concrete Constructs and How to Identify Them. *Journal of Business Research*, *69*(9), 3427–3429. https://doi.org/10.1016/j.jbusres.2016.02.001

Bergkvist, L., & Rossiter, J. R. (2007). The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs. *Journal of Marketing Research*, *44*(2), 175–184. https://doi.org/10.1509/jmkr.44.2.175

Bergkvist, L., & Rossiter, J. R. (2009). Tailor-Made Single-Item Measures of Doubly Concrete Constructs. *International Journal of Advertising*, *28*(4), 607–621. https://doi.org/10.2501/S0265048709200783

Boer, A. de, van Lanschot, J., Stalmeier, P., van Sandick, J. W., Hulscher, J., Haes, J. de, & Sprangers, M. (2004). Is a Single-Item Visual Analogue Scale as Valid, Reliable and Responsive as Multi-Item Scales in Measuring Quality of Life? *Quality of Life Research*, *13*(2), 311–320. https://doi.org/10.1023/B:QURE.0000018499.64574.1f

Bruner, G. C. (Ed.). (2021). *Marketing Scales Handbook: Multi-Item Measures for Consumer Insight Research* (11th ed.). GCBII Productions.

Churchill, G. A. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research*, *16*(1), 64–73. https://doi.org/10.1177/002224377901600110

Churchill, G. A., & Peter, J. P. (1984). Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis. *Journal of Marketing Research*, *21*(4), 360–375. https://doi.org/10.1177/002224378402100402

Cronbach, L. J. (1947). Test “Reliability”: Its Meaning and Determination. *Psychometrika*, *12*(1), 1–16. https://doi.org/10.1007/BF02289289

Diamantopoulos, A. (2005). The C-OAR-SE Procedure for Scale Development in Marketing: A Comment. *International Journal of Research in Marketing*, *22*(1), 1–9. https://doi.org/10.1016/j.ijresmar.2003.08.002

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for Choosing Between Multi-Item and Single-Item Scales for Construct Measurement: A Predictive Validity Perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449. https://doi.org/10.1007/s11747-011-0300-3

Dillman, D. A., Sinclair, M. D., & Clark, J. R. (1993). Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys. *Public Opinion Quarterly*, *57*(3), 289–304. https://doi.org/10.1086/269376

Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and Validity of a Single-Item Measure of Job Satisfaction. *American Journal of Health Promotion*, *19*(3), 194–198. https://doi.org/10.4278/0890-1171-19.3.194

Drolet, A. L., & Morrison, D. G. (2001). Do We Really Need Multiple-Item Measures in Service Research? *Journal of Service Research*, *3*(3), 196–204. https://doi.org/10.1177/109467050133001

Fisher, G. G., Matthews, R. A., & Gibbons, A. M. (2016). Developing and Investigating the Use of Single-Item Measures in Organizational Research. *Journal of Occupational Health Psychology*, *21*(3), 3–23. https://doi.org/10.1037/a0039139

Ganesh, J., Reynolds, K. E., Luckett, M., & Pomirleanu, N. (2010). Online Shopper Motivations, and e-Store Attributes: An Examination of Online Patronage Behavior and Shopper Typologies. *Journal of Retailing*, *86*(1), 106–115. https://doi.org/10.1016/j.jretai.2010.01.003

Gardner, D. G., Cummings, L. L., Dunham, R. B., & Pierce, J. L. (1998). Single-Item Versus Multiple-Item Measurement Scales: An Empirical Comparison. *Educational and Psychological Measurement*, *58*(6), 898–915. https://doi.org/10.1177/0013164498058006003

Geere, J. H., Geere, J.-A. L., & Hunter, P. R. (2013). Meta-Analysis Identifies Back Pain Questionnaire Reliability Influenced More by Instrument Than Study Design or Population. *Journal of Clinical Epidemiology*, *66*(3), 261–267. https://doi.org/10.1016/j.jclinepi.2012.06.024

Hays, R. D., Anderson, R., & Revicki, D. (1993). Psychometric Considerations in Evaluating Health-Related Quality of Life Measures. *Quality of Life Research*, *2*(6), 441–449. https://doi.org/10.1007/BF00422218

Hendrickson, A. R., Massey, P. D., & Cronan, T. P. (1993). On the Test-Retest Reliability of Perceived Usefulness and Perceived Ease of Use Scales. *MIS Quarterly*, *17*(2), 227. https://doi.org/10.2307/249803

Hinkin, T. R. (1995). A Review of Scale Development Practices in the Study of Organizations. *Journal of Management*, *21*(5), 967–988. https://doi.org/10.1177/014920639502100509

Jacoby, J. (1978). Consumer Research: How Valid and Useful Are All Our Consumer Behavior Research Findings? A State of the Art Review. *Journal of Marketing*, *42*(2), 87–96. https://doi.org/10.1177/002224297804200213

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kumar, V., Leone, R. P., Aaker, D. A., & Day, G. S. (2018). *Marketing Research* (13th ed.). John Wiley & Sons.

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and Agreement: Overview and Clarification of Competing Concepts and Measures. *Shanghai Archives of Psychiatry*, *28*(2), 115–120. https://doi.org/10.11919/j.issn.1002-0829.216045

Lo, C. J., Tsarenko, Y., & Tojib, D. (2019). To Tell or Not to Tell? The Roles of Perceived Norms and Self-Consciousness in Understanding Consumers' Willingness to Recommend Online Secondhand Apparel Shopping. *Psychology & Marketing*, *36*(4), 287–304. https://doi.org/10.1002/mar.21179

Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and Evaluation in Education and Psychology* (4th ed.). Harcourt Brace College Publishers.

Mehta, S., Bastero-Caballero, R. F., Sun, Y., Zhu, R., Murphy, D. K., Hardas, B., & Koch, G. (2018). Performance of Intraclass Correlation Coefficient (ICC) as a Reliability Index Under Various Distributions in Scale Reliability Studies. *Statistics in Medicine*, *37*(18), 2734–2752. https://doi.org/10.1002/sim.7679

Moore, K. A., Halle, T. G., Vandivere, S., & Mariner, C. L. (2002). Scaling Back Survey Scales: How Short is Too Short? *Sociological Methods & Research*, *30*(4), 530–567. https://doi.org/10.1177/0049124102030004003

Nagy, M. S. (2002). Using a Single-Item Approach to Measure Facet Job Satisfaction. *Journal of Occupational and Organizational Psychology*, *75*(1), 77–86. https://doi.org/10.1348/096317902167658

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. SAGE Publications Inc.

Nunnally, J. C., & Bernstein, I. R. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.

Pearson, K., & Filon, L. N. G. (1898). Vii. Mathematical Contributions to the Theory of Evolution.— Iv. On the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *191*, 229–311. https://doi.org/10.1098/rsta.1898.0007

Peter, J. P. (1979). Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *Journal of Marketing Research*, *16*(1), 6–17. https://doi.org/10.1177/002224377901600102

Peter, J. P. (1981). Construct Validity: A Review of Basic Issues and Marketing Practices. *Journal of Marketing Research*, *18*(2), 133–145. https://doi.org/10.1177/002224378101800201

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, *88*, 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Polit, D. F. (2014). Getting Serious About Test-Retest Reliability: A Critique of Retest Research and Some Recommendations. *Quality of Life Research*, *23*(6), 1713–1720. https://doi.org/10.1007/s11136-014-0632-9

Pomeroy, I. M., Clark, C. R., & Philp, I. (2001). The Effectiveness of Very Short Scales for Depression Screening in Elderly Medical Patients. *International Journal of Geriatric Psychiatry*, *16*(3), 321–326. https://doi.org/10.1002/gps.344

Portney, L. G. (2020). *Foundations of Clinical Research: Applications to Evidence-Based Practice* (4th ed.). F.A. Davis. http://babymariam.gm/sites/default/files/webform/pdf-foundations-of-clinical-research-applications-to-practice-3rd-e-leslie-g-portney-mary-p-watkins-pdf-download-free-book-d3094a3.pdf

Raghunathan, T. E., Rosenthal, R., & Rubin, D. B. (1996). Comparing Correlated but Nonoverlapping Correlations. *Psychological Methods*, *1*(2), 178–183. https://doi.org/10.1037/1082-989X.1.2.178

Rossiter, J. R. (2002). The C-OAR-SE Procedure for Scale Development in Marketing. *International Journal of Research in Marketing*, *19*(4), 305–335. https://doi.org/10.1016/S0167-8116(02)00097-6

Russell, S. S., Spitzmüller, C., Lin, L. F., Stanton, J. M., Smith, P. C., & Ironson, G. H. (2004). Shorter Can Also be Better: The Abridged Job in General Scale. *Educational and Psychological Measurement*, *64*(5), 878–893. https://doi.org/10.1177/0013164404264841

Sarstedt, M., Diamantopoulos, A., & Salzberger, T. (2016). Should We Use Single Items? Better Not. *Journal of Business Research*, *69*(8), 3199–3203. https://doi.org/10.1016/j.jbusres.2016.02.040

Sarstedt, M., Diamantopoulos, A., Salzberger, T., & Baumgartner, P. (2016). Selecting Single Items to Measure Doubly Concrete Constructs: A Cautionary Tale. *Journal of Business Research*, *69*(8), 3159–3167. https://doi.org/10.1016/j.jbusres.2015.12.004

Sarstedt, M., & Mooi, E. (2014). *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics* (2nd ed. 2014). *Springer Texts in Business and Economics*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-53965-7

Sarstedt, M., & Wilczynski, P. (2009). More for Less? A Comparison of Single-Item and Multi-Item Measures. *Die Betriebswirtschaft*, *69*(2), 211–227.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is It Really Robust? *Methodology*, *6*(4), 147–151. https://doi.org/10.1027/1614-2241/a000016

Shamir, B., & Kark, R. (2004). A Single-Item Graphic Scale for the Measurement of Organizational Identification. *Journal of Occupational and Organizational Psychology*, *77*(1), 115–123. https://doi.org/10.1348/096317904322915946

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Wanous, J. P., & Hudy, M. J. (2001). Single-Item Reliability: A Replication and Extension. *Organizational Research Methods*, *4*(4), 361–375. https://doi.org/10.1177/109442810144003

Wanous, J. P., & Reichers, A. E. (1996). Estimating the Reliability of a Single-Item Measure. *Psychological Reports*, *78*(2), 631–634. https://doi.org/10.2466/pr0.1996.78.2.631

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall Job Satisfaction: How Good Are Single-Item Measures? *Journal of Applied Psychology*, *82*(2), 247–252. https://doi.org/10.1037/0021-9010.82.2.247

Weaver, B., & Wuensch, K. L. (2013). Spss and SAS Programs for Comparing Pearson Correlations and OLS Regression Coefficients. *Behavior Research Methods*, *45*(3), 880–895. https://doi.org/10.3758/s13428-012-0289-7

Wirtz, J., & Lee, M. C. (2003). An Examination of the Quality and Context-Specific Applicability of Commonly Used Customer Satisfaction Measures. *Journal of Service Research*, *5*(4), 345–355. https://doi.org/10.1177/1094670503005004006

Yen, M., & Lo, L.-H. (2002). Examining Test-Retest Reliability: An Intra-Class Correlation Approach. *Nursing Research*, *51*(1). https://journals.lww.com/nursingresearchonline/Fulltext/2002/01000/Examining_Test_Retest_Reliability__An_Intra_Class.9.aspx

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Ferner erkläre ich, dass mir die Gelegenheit zum Promotionsverfahren nicht kommerziell vermittelt wurde - d.h. ich habe keine Organisation eingeschaltet, die Betreuer für die Anfertigung von Dissertationen sucht oder die mir obliegende Prüfungsleistungen ganz oder teilweise erledigt hat. Mir ist bekannt, dass die Inanspruchnahme eines sog. „Promotionsberaters" den Ausschluss der Annahme als Doktorand sowie der Zulassung zum Promotionsverfahren, den Abbruch des Promotionsverfahrens und die Rücknahme des erlangten Grades wegen Täuschung zur Folge haben kann.

Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

München, 20.12.2023

Sandra Barbara Baringhorst