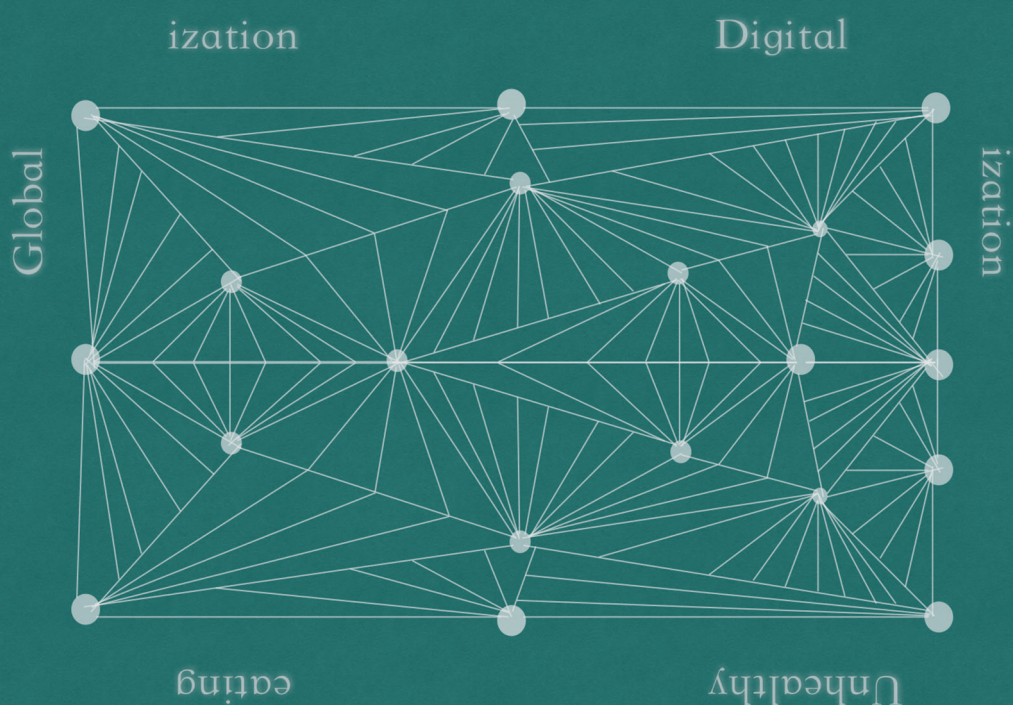PHYLISS JIA GAI

Contextualized Consumers

**Theories and Evidence on Consumer Ethics,
Product Recommendations, and Self-Control**

Contextualized Consumers

Contextualized Consumers:
Theories and Evidence on Consumer Ethics, Product Recommendations, and Self-Control

Gecontextualiseerde Consumenten:
Theorieën en Bewijs van Consumentenethiek,
Productaanbevelingen en Zelfbeheersing

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof. dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on
June 25th 2020 at 13:30 hrs

by

Phyliss Jia Gai
born in Liaoning, China

Erasmus University Rotterdam

Doctoral Committee

Doctoral dissertation supervisor(s):
Prof. dr. S. Puntoni
Prof. dr. S. T. L. Sweldens

Other members:
Prof. dr. C. Goukens
Dr. R. Bhargave
Dr. B. Van den Bergh

Acknowledgements

This dissertation cannot be completed without help from Stefano and Steven. They guided me through this formidable journey. I am indebted to these two in numerous ways, but I especially thank Stefano for cheering me up and Steven for his trust in me.

Chapter 3 would never launch without Anne. I am grateful for her support for my ideas, empathy with my frustrations, and ambition for our project. Her warmth makes me confident and diligence keeps me forward. She is a role model for me to look forward to.

I appreciated the hours Mirjam generously spent with me, sitting together and brainstorming for our projects. Many of them went fruitless but we made one into Chapter 4. I also thank her for her witty suggestions on my other projects.

Many other people have contributed to this work, with or without their awareness. I thank Dan, Jason, and Quentin for their statistical advice at various stages of my research. For helpful comments on the early versions of Chapter 3, I thank Gerrit, Ajay Kohli, Darren Dahl, Christophe Fuchs, Anne her Braak, and Michael Haenlei. I also thank Yuyu, the owner of Jijitang, who let me run field studies with her company, and Hang Yee, who put my study design into neat coding. For lab assistance in Chapter 4, I thank Nuree, Ernesto, and Camilla. Annette kindly ordered the snacks I used for the studies, too. Arnaud and Zining volunteered to help with my online data collection in Chapter 2. Thanks to Rik Pieters, I received training in SEM and was able to apply a tiny bit in Chapter 4. Stephen Raudenbush enlightened me with multilevel models at Chicago, which I used in Chapter 2 and 3.

The idea of news sharing discussed in the last chapter was developed while I was at Columbia. Thanks to Gita, I found a new research question and a bunch of new friends. I appreciated the support and comfort from Shilpa, Yu Ding, Byung, Jennifer, and Alisa.

Much of my perseverance owes to my family and friends. The risks I afforded to take were due to my parents, who gave me freedom and love. The hardship I was able to endure was because of my fond memories with Shen, Kongfan, Taoran, Liuchang, Charlene,

Chapter 1

Introduction

The goal of my dissertation is to examine the issues, arising from globalization, digitalization, and health threats, that are consequential to the collective welfare of consumers. These issues are embedded in the trends that contextualize consumers but have gathered little attention in the literature. In this chapter, I provide an overview of each of the trends and elaborate on the issues that are tied to consumer welfare and await investigation. Then I introduce the three essays and discuss the theoretical and practical relevance of each of them.

The Trends That Contextualize Consumers

The topics covered in this book reflect the age that contextualizes the author and many other consumers. They live in a globalized world, where many people can speak multiple languages and travel across language barriers in no time. A bi- or multi-lingual consumer is constantly switching between different language channels for various purposes. She might fill out a consumer survey in her second language and then shop on a website in her first language. She might watch TV shows in her second language as well as movies in her first language. Whether a company presents the content in English or the local language of the consumer, she can navigate the online world and find what she wants. The seamless shifts from one language context to the other and vice versa could make one wonder whether a consumer always behaves the same way in different languages. We know from the past research that the language a consumer speaks determines the way she makes sense of brand and product names (Luna and Peracchio 2001), serves as a cue for her memories established in that language context (Puntoni, de Langhe, and van Osselaer 2009), and represents a certain set of cultural values that regulates her public self (Lee, Oyserman and Bond 2010). Apparently, language contexts do pose influences on individual consumers, shaping the way they think and behave, either for or against their own benefit. For instance, while the brand name associations can be an irrational factor in consumers' product choices (Wu et al. 2019), the reduced emotional attachment to products and resources can make consumers more rational in transactions (Keysar, Hayakawa, and An 2012).

Consumers nowadays also live in a digitalized world, where the cost of finding and trying new products becomes increasingly marginal (Datta, Knox, and Bronnenberg 2019).

Thanks to the development of algorithms, personalized recommendations are not reserved for luxury consumers anymore but available to anyone who has left their traces of interests and behaviors in the online world. A consumer today does not have to visit brick-and-mortar music or video shops to discover their favorites; instead, a long list of songs or videos is automatically generated and regularly updated by the algorithm, awaiting for her exploration at just one click away. The consumer does not have to tell the salesperson what she likes or dislikes in the hope to get more accurate recommendations; she can communicate her tastes with algorithms by merely clicking on the like or skip button. The consumer does not have to spend hours deliberating over which product to choose but can rely on algorithms to provide a suggestion for her to follow. With more companies adopting recommender systems, numerous studies have looked at how to make recommendations perform better, in terms of consumers evaluation of the recommended products (e.g., Ying, Feinberg, and Wedel 2006), their trust in (Cramer et al. 2008) or interaction with the recommender system (Zhao 2017). These studies primarily focus on the design of recommendation algorithms, attributing consumer reactions to the products being recommended and the underlying system.

Aside from globalization and digitalization, consumers are more inflicted by self-regulation problems than ever. Affluence and abundance of food, once a blessing to consumers living in the age of scarcity, now engender problems such as obesity that threaten the health of consumers. According to the World Health Organization (2016), more than 1.9 billion people around the world are overweight. The epidemic arises not only in the developed world but also in developing countries because of the rising GDPs (Lappo et al. 2015). In response to this acute health crisis, a significant amount of research has been devoted to the development of interventions to curb unhealthy eating. As a result, food marketing and packaging of junk food are increasingly regulated worldwide (Kovic 2019; WCRF 2017); taxes are imposed on unhealthy food and drinks and subsidies on the healthy ones (Powell and Chaloupka 2009). Besides strict policies, benevolent nudges have also been implemented to restaurant menus and cafeteria layout, with the purpose to facilitate consumers to make wiser choices. For instance, calories are explicitly labeled, and traffic lights attached to food options indicating the health value (VanEpps, Downs, and Loewenstein 2016). Healthy options are placed closer to consumers whereas the

unhealthy are moved further away (Rozin et al. 2011). In marketing, particular attention has been paid to the health-taste tradeoff between healthy and unhealthy food. Researchers frequently contrast the healthy yet less tasty food with the tempting unhealthy food as "virtues" versus "vices" and endeavor to understand what drives consumers to select the virtue over the vice food.

The Potential to Improve the Collective Welfare of Consumers

The phenomena discussed above, the multilingual contexts, the availability of automated recommendations, and the relapse of control in food consumption, have been popular areas of inquiries in the last two decades. Although none of these has been researched exhaustively and the current understanding is far from comprehensive, I see one void of research particularly worth notice. That is, the consequences of these binding phenomena on consumers as a collective. I suggest that past research in these contexts has focused much on behaviors that benefit individual consumers but overlooked questions that pertain to consumers as a collective. This dissertation is intended to address some of these questions.

I identify three cases in which the collective welfare of consumers is undermined, giving rise to the questions that are relevant to the trends discussed above and ready to be addressed by consumer research. First, the decision of an individual consumer may be beneficial to herself at the cost of fellow consumers. Such unethical decisions are frequently manifest in the actions of consumers, but even more in their words. One example is opportunistic lying. Dishonest consumers, by taking advantage of the trust of companies, reap the benefits from their fellow consumers and potentially elevate prices of services and products to all consumers. As a linguistic product, the lies of consumers, however, have rarely been examined against their language contexts. Research on bilingual consumers has primarily looked at behaviors that do not harm others yet paid very limited attention to those that concern collective welfare. How does language influence behaviors such as opportunistic lying that is rational yet unfair to other consumers? Is there any regularity and why does language affect consumer dishonesty? Chapter 2 is intended to provide answers to these questions.

Second, utility maximization for each individual consumer may paradoxically undermine the collective benefit due to invisible forces. This is precisely what happens to algorithmic recommendations. Although algorithms strive to recommend new products to each individual, because they typically favor products that have been viewed or consumed more frequently, the same set of popular products might be recommended to many consumers and the overall diversity of consumption be constrained (Fleder and Hosanagar 2009). This means that even when every consumer receives something new, there remains a huge stock of products that consumers might find interesting but never get the chance to see. To solve this problem, past research has been concentrated on the backstage design of recommender systems and conceptualizes the problem as an algorithmic trade-off between accuracy and diversity (Adomavicius and Kwon 2012, 2014). The role of marketing, or the management of the front-end communication with consumers, appears marginal (if not absent) in this realm. Nevertheless, it is possible for marketing to make a unique and important contribution by offering a way to attract consumers towards recommendations of unfamiliar or unpopular products, such that the overall diversity of consumption can be enhanced. Chapter 3 introduces the framing of recommendations as a marketing tool to achieve this goal and suggests that it is not just how the algorithm is designed but also how it is explained to consumers that matters.

Finally, consumer welfare cannot be maximized when research disproportionately focuses on a specific choice setting or population. The research on choices between healthy and unhealthy food categories leaves out a substantial group of consumers who end up choosing within a specific category, especially those who decide to enjoy vice food. These consumers not only deserve the right to derive pleasure from their favorite food but merit more attention from researchers to understand their self-control processes within the vice category. Importantly, even within the vice category, consumers are frequently offered 'lesser vices' that are less harmful to their health, such as Diet Coke. What is the dynamic of choosing between the vice and the lesser vice? Are findings from vice–lesser choices applicable to vice–lesser vice settings? What does the success of self-control mean in the vice–lesser vice setting? Chapter 4 presents a first attempt to conceptualize self-control dynamics when consumers are confronted with options within the vice category and provides experimental evidence for the uniqueness of the vice–lesser vice setting.

To recap, this dissertation is aimed at addressing the questions related to consumer welfare in the current age. As summarized in Figure 1, it presents the investigations of 1) the influence of language on opportunistic lying (Chapter 2), 2) the role of marketing communication in consumer interests in product recommendations (Chapter 3), and 3) the choice and consumption of lesser vices (Chapter 4). It is worth noting that, although the questions raised in this book are oriented towards the collective welfare of consumers, they do not conflict, but dovetail, with the interest of companies. The findings presented in this book provide practical insights regarding how to manage the dishonesty of bilingual consumers to prevent companies' losses, how to market 'long-tail' products without overspending on algorithm development, and how to effectively market lesser vices. The ultimate goal of this dissertation is to invite more research on consumer welfare that benefits companies at the same time. In the remainder of this chapter, I provide a detailed overview of chapters 2 to 4. Chapter 5 will zoom out of the specific questions addressed in each chapter and discuss the overall implications across the chapters. Instead of reiterating the precise implications for the theories and practice, Chapter 5 will focus on the routes to be taken to broaden the current understanding of contextual influences on consumer welfare.

Figure 1: The questions addressed in this book.

Chapter 2: Language and Consumer Dishonesty

Consumers lie and cheat, in domains as diverse as insurance claims, product returns, dating profiles, and tax declarations. When signing up for online services, for example, they often lie about their birthdays to be eligible for discounts (Johnson 2015). Consumer dishonesty produces substantial costs: Automobile insurance fraud costs U.S. companies more than $5 billion annually and prompts significantly higher premiums for all consumers (Insurance Research Council 2012). Although most consumers lie in their native language (L1), they also lie in their second language (L2): Travelers misreport information when they fill out lost-luggage forms; immigrants fail to disclose medical conditions to health insurers; and overseas consumers claim discounts by faking their online identities. Most people in the world speak more than one language, and the number of multilingual people is increasing (Grosjean 2010). Therefore, it is important for consumer researchers to understand how language shapes lying behavior.

In Chapter 2, I examine the effect of language (L2 vs. L1) on lying. Nine studies, spanning several languages and consumption domains, challenge the findings of recent psychological research that indicate second language use leads to lower rates of lying. Second language use does not lead uniformly to more honesty but rather attenuates people's intuitive preferences for lying or telling the truth. Furthermore, stronger intuitive tendencies for (dis)honesty magnify the language effect, and increased feelings of uncertainty in second language contexts contribute to the effect of language on consumer dishonesty. These nuanced findings highlight the fact that language contexts do not automatically activate or inhibit dishonest behavior. Instead, it is the interaction between the moral intuitions that are at least partially established in the native language context and the uncertainty embedded in using a foreign language that jointly shape the language effects.

Chapter 2 makes three main contributions. First, it contributes to consumer behavior and marketing research by linking research on bilingual consumers (e.g., Luna and Peracchio 2005)—which has focused mostly on advertising and branding—to research on consumer dishonesty (e.g., Cowley and Anthony 2018). These have been popular areas of

inquiry in recent years; by considering them together, I produce novel insights about a topic that has become increasingly important as globalization expands.

Second, I contribute to the psychological literature on how language influences decision making (Hayakawa et al. 2016). Although recent research shows that L2 leads to lower rates of lying than L1 (Alempaki, Doğan, and Yang 2017; Bereby-Meyer et al. 2018), I argue that it is premature to conclude that "honesty speaks a second language" (Bereby-Meyer et al. 2018, 1). I propose that whether L2 use decreases or increases lying depends on whether a person's intuitive L1 response is to be honest or to lie. I demonstrate that L2 *decreases* lying when lying is the intuitive response in L1 contexts, but it *increases* lying when honesty is the intuitive response in L1 contexts. I further propose that an important mechanism explaining these effects is that L2 introduces a sense of uncertainty that blunts intuitive responses. I thus qualify existing results and show that the effect of language on lying can often be the opposite of what previously documented.

Third, I provide new insights into phenomena that are of great relevance to managers and policy makers. Lying occurs in countless occasions, for myriad reasons; I focus on opportunistic lying that brings material benefits to consumers. Although lies that bring non-material benefits (e.g., impression management) also are common, lying that is motivated by material gain is especially relevant to business and society, considering its costs to organizations, governments, and honest consumers. I also focus on *private* lying; that is, situations that do not involve face-to-face communication. These conditions characterize much of the deceitful consumer behavior that occurs in computer-mediated environments. Studying the effect of language under these conditions contributes to the understanding of consumer honesty in the digital age.

Chapter 3: Framing Recommendations to Attract More Click-throughs

Many companies provide consumers with product recommendations that have been generated by algorithmic recommender systems: Spotify and Netflix recommend songs or movies for their subscribers, and TripAdvisor and Yelp provide recommendations for hotels or restaurants. Amazon suggests which products consumers might want to buy, and *The New York Times* recommends different news articles. These personalized recommendations help consumers find offerings they likely are interested in and increase

their loyalty (Gupta et al. 2006; Kamakura et al. 2005). According to a survey by Spotify, 65% of consumers find a new favorite song in the personalized playlists they receive (Johnson 2015), and Netflix asserts that its recommender system effectively reduces consumer churn and saves the company more than $1 billion annually (Gomez-Uribe and Hunt 2015).

To improve the accuracy of these algorithmic recommendations, recommender systems frequently adopt a hybrid approach that accounts for both common preferences across consumers and common attributes across products (Amatriain and Basilico 2016). Each recommendation thus is based on both user and product input; it is not straightforward to explain the basis of the recommendation descriptively. Anecdotally, in my interviews with members of a major European e-commerce company, the data scientists expressed different opinions about whether user or product input best described the basis for their recommender system, which actually uses various inputs. In turn, this company, and others alike, could choose which component to emphasize when explaining how it derives recommendations for consumers. Some companies already highlight that their recommendations are *user-based* by focusing on overlaps in consumer preferences, such as "Consumers who viewed this item also viewed…" by Amazon and "Consumers also watched…" by Netflix. In contrast, other companies emphasize that recommendations are *item-based*, such as "Similar to [what you have listened to]" by Spotify and "More in Health" by *The New York Times*.

The question that then arises is which framing, user-based or item-based, is more effective in triggering clicks on a recommendation. Such clicks can increase conversion rates by stimulating consumers to explore other product offerings (Xu, Duan, and Whinston 2014). More importantly, if selecting the proper framing can trigger more clicks on recommendations, framing can be effectively used as a tool to attract consumers to products that are unpopular or unfamiliar, that is, the 'long-tail' products (Anderson 2008). This potentially frees the algorithmic constraint on the aggregate diversity of consumed products (Adomavicius and Kwon 2014), by mitigating the risk of recommending products that have not yet gathered much data and by maximizing the traffic to a variety of products and thus making the chances of them being recommended more even.

Prior research on consumers' responses to recommendations have focused primarily on the underlying recommendation algorithms (Ariely, Lynch, and Aparicio 2004; Hennig-Thurau, Marchand, and Marx 2012; Ying, Feinberg, and Wedel 2006) or characteristics of recommended products (Cooke et al. 2002; Pathak et al. 2010), with limited attention to the framing provided to describe the recommendations. This gap is surprising for two main reasons. First, many recommendations rely on input from both users and items, so companies can choose to highlight different elements. Second, altering recommendation framing is a nearly zero-cost effort. To address this gap, I manipulate recommendation framing (user-based versus item-based) but keep the underlying algorithms and recommended products constant.

My central proposition is that, compared with item-based framing, user-based framing informs consumers that the recommendation is generated through *product matching* (i.e., the recommended product is similar to the focal product) but also indicates *taste matching* between users (i.e., the focal product liked by oneself is also liked by other users). Consumers extract information from similar others' tastes to predict their own liking of unfamiliar products (Yaniv, Choshen-Hillel, and Milyavsky 2011), so this information should provide an additional guarantee to consumers that the product will match their tastes. Consequently, I predict that recommendations framed as user-based (versus item-based) attract more click-throughs, conditional on the assumption that consumers perceive that a recommendation accurately matches their taste.

I test my proposition with six studies, which span a variety of data sources (field, behavioral, and scenario) and consumption domains (articles, paintings, and books). Across these various methods and product domains, I consistently find that user-based framing attracts more click-throughs on recommendations than item-based framing when consumers perceive that others' preferences match their own. I further propose three boundary conditions that potentially cause the recommendation recipient to perceive the taste matching as unsuccessful, so the advantage of user-based framing over item-based framing decreases. These boundary conditions in turn offer substantive guidance for companies on how to adapt the framing of their recommendations to maximize recommendation click-throughs. Importantly, in all these studies, the recommendations

involve products with which consumers are unfamiliar, a design element that establishes insights into how to enhance the aggregate diversity of consumption and how to market novel products.

Chapter 4: When virtues are lesser vices

To overcome overweight and obesity, a significant amount of research has investigated drivers of consumers' choices between healthy and unhealthy food (Hoch and Loewenstein 1991; Metcalfe and Mischel 1999), often also referred to as choices between virtues and vices (Mishra and Mishra 2011; Wertenbroch 1998). Most self-control research on food choices has relied on paradigms where people choose between a healthy and an unhealthy option. For instance, consumers may be asked to choose between the fruit salad, a supposedly 'virtue' and the chocolate cake, a 'vice' (Shiv and Fedorikhin 1999). In recent years, however, researchers have started to focus more on vice foods, the leading cause of obesity (Cecchini et al. 2010). Examples include consumer choice between two vices that differ in visual saliency (Gao, Li, and Wyer 2016) and consumer decision on the portion size of vice food (Haws and Winterich 2013).

Chapter 4 adds to this literature by examining an understudied setting where consumers choose between an unhealthy food (e.g., regular chips) and a slightly less unhealthy version of that food (e.g., light chips). Instead of choosing between virtues and vices, in such situations, consumers are essentially choosing between vices and 'lesser vices'. Practically, studying this context is of great importance as lesser vices abound in the marketplace, such as reduced-fat Cheez-it crackers and sugar-free Hershey's chocolates. These products normally contain fewer calories, while aiming to retain maximum similarity on the taste dimension with the regular vices.

The most important contribution of Chapter 4 is to show that choice conflicts between vices and lesser vices are characterized by their own dynamics. First, vice–lesser vice choices are much less characterized by the gap in short-term gratification. Second, differences in caloric density are uniquely salient in the vice–lesser vice choice setting. These proposed differences are grounded in goal systems theory and result in two predictions. One prediction pertains to the futility of the interventions intended to make the vice look less tempting. I argue that inserting a delay between choice and consumption, a

11

popular nudge towards healthy eating, would no longer be able to shift consumers' perspective from short-term gratification to long-term well-being in the vice–lesser vice setting. Hence, making food choices well in advance would not increase the choice share of lesser vices. In contrast, I predict that consumers who chronically care more about calorie intake are more likely to choose lesser vices. The third characteristic of the vice–lesser vice setting is that consumed quantities matter to the overall success of self-control. To this end, I delineate the competing predictions regarding how choices would influence consumed quantities.

To test these propositions, I conducted two experiments, in which the timing of choice was manipulated, individual differences in restrained eating (indicating the degree of calorie monitoring) measured, and real choices and consumed quantities recorded. The studies corroborate the predictions by showing that decoupling the immediate gratification from food choices by advance ordering does not encourage consumers to substitute the vice with the lesser vice; instead, individual differences in concerns about calorie intake encourage the substitution decision. Critically, the studies also demonstrate that consumers exert limited control over consumed quantities after they make choices. These findings are discussed in relation to previous research and future extensions.

Declaration of contribution

Chapters 1 and 5: I drafted these chapters independently and revised it based on my advisors' feedback.

Chapter 2: I proposed the research project, reviewed the literature, designed and conducted the studies, analyzed the data, and drafted the final paper. My co-author (Stefano Puntoni) provides critical feedback at each stage.

Chapter 3: I formulated the research question, reviewed the literature, designed and conducted the studies, analyzed the data, and drafted the final paper. My co-author (Anne-Kathrin Klesse) provides critical feedback at each stage.

Chapter 4: I developed the research idea, theorizing, and Study design in consultation with my co-authors (Mirjam Tuk and Steven Sweldens). I conducted the studies in the

behavioral lab and analyzed the data. I drafted the paper and revised it based on my co-authors' feedback.

Chapter 2

Lies of Bilingual Consumers

This chapter is based on the paper submitted to Journal of Consumer Research on March 20, 2019:

Chapter 2 is organized as follows. I review research on lying and bilingualism to develop a theory of the role of language in dishonesty, and then present the results of eight experiments and a meta-regression. The studies document the effects of language in both consequential and imagined situations. The investigation spans different languages (Chinese, English, French, and Korean) and consumption domains (insurance, flight delay, advertising, and lotteries). I conclude with theoretical and practical implications as well as methodological limitations.

Theoretical Background

Consumer research on lying and bilingualism

Prior research has examined both consumer lying and bilingualism extensively but separately. Researchers have investigated the content of lies (e.g., emotions) (Andrade and Ho 2009; Sengupta, Dahl, and Gorn 2002), motivation to lie (e.g., social comparison) (Argo, White, and Dahl 2006; Goldsmith, Roux, and Ma 2018; Mazar and Zhong 2010), consequences of lying (Anthony and Cowley 2012; Cowley and Anthony 2018), and moral judgment of lies (Argo and Shiv 2012). They also have investigated specific dishonest consumer behaviors, such as purchases of counterfeit products (Wang, Stoner, and John 2018; Wilcox, Kim, and Sen 2009), lying in consumer surveys (De Jong, Fox, and Steenkamp 2015), the influence of social bonds on lying (Nikolova, Lamberton, and Coleman 2018), and lying to harmful brands (Rotman, Khamitov, and Connors 2018).

Similarly, a great deal of research has been devoted to understanding the role of language in consumer behavior (Carnevale, Luna, and Lerman 2017). Language (L2 vs. L1) influences advertising effectiveness (Krishna and Ahluwalia 2008; Luna and Peracchio 2001, 2005; Puntoni, de Langhe, and van Osselaer 2009) and brand evaluations (Leclerc, Schmitt, and Dube 1994; Shrum et al. 2012; Zhang and Schmitt 2004); L2 use can polarize scale ratings (de Langhe et al. 2011) and reduce impulsive decision making (Klesse, Levav, and Goukens 2015). To my knowledge, no previous consumer research has examined how language relates to lying behaviors, though researchers in other fields have.

Effect of language on lying

Recent research in psychology and economics suggests that L2 decreases lying relative to L1 (Alempaki, Doğan, and Yang 2017; Bereby-Meyer et al. 2018). Bereby-

Meyer et al. (2018) conducted a study in which participants rolled a die and reported the outcome. Participants were incentivized to lie, because they could earn higher rewards if they reported higher outcomes (e.g., rolling a 1 earned participants $1, and rolling a 4 earned $4). Lying was private and anonymous. If participants told the truth, the distribution of reported outcomes should have been roughly even across various payoffs. If participants lied, the distribution of reported outcomes should have been skewed toward higher payoffs. In several experiments, lying occurred to a lesser extent in L2 contexts than L1 contexts. The authors concluded that though people are tempted to lie for self-benefit, L2 weakens this temptation. A separate group of researchers similarly found that L2 decreased lying compared with L1 (Alempaki et al. 2017). These studies provide evidence that lying varies by language. However, does L2 *always* decrease lying, and if not, when does L2 increase rather than decrease lying?

I suggest that L2 does not necessarily suppress lying, and whether it increases or decreases lying depends on people's intuitive L1 responses (i.e., whether lying or telling the truth is their most intuitive behavior). Intuitions are the thoughts that automatically spring to people's minds; they are powerful drivers of decision making (Epley and Gilovich 2006; Risen 2016), particularly moral decision making (Haidt and Joseph 2004). They serve as the default that people cling to unless they are prompted to deviate (Epstein 1994; Thompson 2012). Literature is divided on whether lying or telling the truth tends to be people's intuitive response.

On the one hand, people intuitively may avoid lying, because it is a stressful experience (Caldwell-Harris and Ayçiçeği-Dinn 2009; Dienstbier and Munter 1971). Neural evidence points in this direction: Greene and Paxton (2009) asked participants to privately predict the results of randomly flipped coins (head or tail) and found that control-related areas in the prefrontal cortex are more active when people lie for self-benefit, suggesting that being honest is the intuitive response. On the other hand, lying can be tempting, and people may intuitively pursue it. External stimulation of the brain region involved in behavioral control enhanced honesty in the aforementioned die-rolling task (Maréchal et al. 2017). Time pressure, which encourage fast intuitive responses, increase lying behavior (Shalvi, Eldar, and Bereby-Meyer 2012). People who are depleted of self-

control resources are lie more for monetary gains (Kochaki and Smith 2013; Mead, Baumeister, and Gino 2009), and resisting the temptation to lie depletes self-control resources (Gino et al. 2011). These mixed findings suggest that people can be both intuitively honest and intuitively dishonest, depending on the situation.

If language moderates the extent to which people's behavior follows intuitive responses, I expect L1 and L2 to lead to different behaviors depending on which response—lying or telling the truth—is intuitive. One of the most prominent findings in literature on the role of language in decision making is that L2 decreases intuitive biases in decision making, such as loss aversion (Costa et al. 2014; Keysar, Hayakawa, and An 2012) and the hot hand fallacy (Gao et al. 2015). Accordingly, I predict that L2 results in decreased dishonesty when a person's intuitive L1 response is to lie but increased dishonesty if the intuitive L1 response is to tell the truth. Formally, I refer to this prediction as the reduced-intuition effect:

H1a: L2 increases lying when telling the truth is the intuitive response.

H1b: L2 decreases lying when lying is the intuitive response.

Understanding the reduced-intuition effect

If the effect of language on lying is connected to intuitive responses, the reduced-intuition effect should depend on the strength of people's intuitions. When people lack intuitive preferences for lying or telling the truth, language should have little impact on the occurrence of lying. However, as their intuition becomes stronger, there should be more space for language to exert influence, and the effect of language should become larger. Thus, an important boundary condition for the effects in H1a and H1b should be the strength of intuitive responses. Formally, I predict that the reduced-intuition effect is subject to a strength-of-intuition constraint:

H2: The weaker the intuition, the smaller the reduced-intuition effect.

I further propose that an important causal mechanism that connects language to lying behavior is the greater sense of uncertainty that tends to accompany L2. Most bilinguals are not balanced bilinguals; in their daily lives, they are exposed predominantly to L1 contexts and only sporadically to L2 contexts (e.g., at school). Bilinguals do not have as

18

much experience with L2 as with L1, so L2 likely makes them feel more uncertain about their environments and the appropriateness of their behaviors. For example, L2 contexts lower people's confidence in their moral judgments (Geipel, Hadjichristidis, and Surian 2015).

The claim that L2 tends to generate feelings of uncertainty is unlikely to surprise any introspective bilingual. This notion is exemplified by Yiyun Li, a native Chinese novelist writing in English, her second language: "If you are a native speaker, things are automatic. For me, every time I say or write something (in English), I have to go back and ask, 'Is this what I want to say?'" (Grimes 2018). Many writers have elected to compose in their second languages to challenge feelings of certainty. Samuel Beckett's decision to write in French (in which he wrote *Waiting for Godot*) is often described as an attempt to tame intuitive tendencies: "What seems to attract him about French is the very fact that it is less second nature to him than is English, that his relationship to it is different and makes him more able to manipulate it consciously" (Pattie 2000, 132). American author Jhumpa Lahiri describes her decision to write in Italian as a deliberate quest for uncertainty: "I trade certainty for uncertainty" (Lahiri 2015).

I argue that this sense of uncertainty in L2 contexts also is common outside the literary world; it penetrates the process of ethical decision making. By drawing on research on the impact of meta-cognitive feelings on choices and behavior (Haddock et al. 1999; Novemsky et al. 2007; Schwarz 2012), I suggest that feelings of uncertainty prompt people to rely less on their intuitions. People are less confident about their intuitive choices when they experience greater uncertainty, even when the uncertainty is incidental to the decision (Simmons and Nelson 2006). If L2 triggers feelings of uncertainty, it should move people away from their intuitive preferences for lying or telling the truth:

H3: L2 triggers more feelings of uncertainty than L1.

H4: The greater uncertainty associated with L2 mediates the reduced-intuition effect.

A different explanation for the reduced-intuition effect is that L2 weakens the affective appeal of intuitive responses; people have weaker emotional experiences when reading L2 expressions than L1 expressions (Puntoni et al. 2009). Yet it is unclear whether

L2 attenuates people's feelings about their own behaviors, in that they are not reading descriptions of their behaviors (e.g., "lying," "cheating"). Therefore, I examine the post-lying feelings associated with L1 and L2. I expand the discussion of alternative theories in the General Discussion.

The studies reported here aim to answer three questions: (1) Do L2 contexts always decrease lying, as prior research suggests? (2) If not, when does language choice increase or decrease lying? (3) Why does the effect of language occur? I conducted nine studies to answer these questions. Studies 1a–1c answer the first question, in car insurance and flight delay contexts; studies 2a–2c (advertising contexts) and study 3 (flight delay) examine the second question and test the reduced-intuition effect (hypotheses 1a and 1b). Study 4 addresses the final question in a lottery context and tests the proposed mechanism (hypotheses 3 and 4), as well as excluding some alternative explanations. Finally, study 5 involves a meta-regression across studies to test the strength-of-intuition constraint (H2).

In all procedures, participants chose between true and false statements. Following prior research (Simmons and Nelson 2006; Simmons et al. 2011), I identified the intuitive response (to lie or not lie) by comparing the proportion of liars (studies 1a, 1b, 3, and 4) and the likelihood of lying (studies 2a, 2b, and 2c) against chance, assuming that the level of chance represents indifference to lying or telling the truth. Moreover, to avoid self-selection biases, I did not mention language as a study component when recruiting participants. To ensure that the data reflect lying rather than random noise, I excluded data points according to two considerations. First, in studies 1a, 1b, and 3, in which participants choose among multiple options, I exclude those who select non-truthful options that do not offer any benefits. Second, in studies 2a, 2b, 2c, and 4, in which participants choose between two options—lying or telling the truth—I could not tell whether participants understood the study based on their choices, so I exclude those who report that their L2 proficiency is below a minimum requirement (see studies 2a–2c). (Note that self-reported L2 proficiency has no relationship to lying across studies, ps > .10.) Participants should have no reason to lie on the English proficiency question, which always appears before the language manipulation and is embedded among other L1 filler questions. In most of the studies, Chinese is L1 and English is L2, reflecting the world's most common L1 and L2

respectively. For generalizability, I also extend the results to other languages (French as L1 and Korean as L2).

<div align="center">**Does L2 Always Decrease Lying?**</div>

Studies 1a–1c tested the language effect on lying behavior in two consumer-relevant scenarios: purchase of car insurance and flight delay compensation. These scenarios involved monetary incentives for people to lie with little risk of being caught and punished. All three studies used English as L2 and Chinese as L1. According to Bereby-Meyer et al. (2018), foreign language decreases lying; however, I consistently find the opposite result.

Study 1a: Car insurance

Study 1a was a pre-registered test of the effect of language on lying (registered at https://goo.gl/ADbXm4). Inspired by Shu et al. (2012) (experiment 3), I designed a car insurance scenario. I asked participants to imagine that they were purchasing car insurance, filling out the policy form, and reporting the current odometer reading. Because higher odometer readings mean more driving, higher risk of accidents, and more expensive insurance, there is an incentive for consumers to underreport odometer readings.

Participants and study design. To reach native Chinese speakers in an anonymous environment, I distributed the study online to users of one of the largest Internet companies in China, Baidu.com. I predetermined the recruitment of 400 participants from the website (171 males, Mage = 30, SD = 7.7) in the hope of having at least 100 participants per cell after data exclusions. Most participants were employed in various industries; unemployed (including students) people constituted 12% of the sample.

Procedure. In the first part, participants answered demographic questions and then read a brief introduction to the car insurance scenario. This part was always written in Chinese (L1). In the second part, I randomly assigned participants to enter the English (L2) or Chinese (L1) scenarios. In their designated scenarios, they viewed the policy form (see Appendix); I told them that the actual odometer reading on their car was 4,501 kilometers. I also told participants that the insurance company could not check their actual odometer

readings. Participants then chose an odometer reading to report from a pricing list. I arranged the pricing such that the insurance premium would increase by ¥3,000 (Chinese yuan) when the odometer readings exceeded 3,200 kilometers but not if they were lower than this level. There were two options that did not exceed 3,200 kilometers (see Table A1 in Appendix). If people are indifferent to lying or telling the truth, there should be an equal chance for them to select one of the three options (truth and the two lying options), such that the chance level of lying is approximately 67%. I measured whether participants chose to tell the truth or lie to evade the additional payment and compared the proportion of liars against the chance level.

Results. I excluded participants who failed to understand the scenario (n = 147; see preregistration form). These participants chose the untrue options that did not bring monetary benefits (e.g., choosing 3,400–3,600 kilometers, which was not the truth but did not reduce the payment). Counter to Bereby-Meyer et al.'s (2018) finding, the proportion of liars was higher in the L2 condition (49 of 130, 38%) than in the L1 condition (25 of 123, 20%) (χ_2 (1, N = 253) = 8.39, p = .003). As shown in figure 1, lying in both the L1 and L2 conditions fell below the chance level (ps < .001).

Study 1b: Replication of study 1a

Study 1b was a replication of study 1a with slightly different stimuli. The goal was to test the robustness of the findings in study 1a.

Participants and study design. This Study had 333 participants (146 males; age not recorded; 11% unemployed, including students). It was a 2 (L2 vs. L1) × 2 (number of lying options: 6 vs. 1) between-participants design. I included the second factor to vary the chance level of lying. If there are six lying options, in addition to the truthful statement, the chance level of lying is approximately 86%. In contrast, when there is only one lying option, the chance level drops to 50%. By varying the chance level of lying, I also varied the strictness of data exclusion (i.e., more data exclusions and lower chance level with fewer lying options), allowing us to assess the robustness of the language effect to the strictness of data exclusion.

Procedure. I arranged the pricing such that the insurance premium increased by ¥3,000 when the odometer reading exceeded 3,000 kilometers. I created two lists with the

same pricing but differing numbers of lying options. One list had six options that did not exceed 3,000 kilometers (i.e., chance level lying = 86%), whereas the other had only one option that did not exceed 3,000 kilometers (chance level lying = 50%). I randomly assigned participants to one of the two lists and directed them to make their choices. For details, see Table A1 in Appendix.

Results. Following study 1a, I excluded participants who failed to understand the scenario (n = 29 when there were six lying options, n = 73 when there was one lying option). As figure 1 shows, the proportion of liars always fell below the chance level ($p <$ .001). Logistic regressions showed that the number of lying options did not moderate the language effect ($\beta = 0.30$, $SE = 0.71$, $z = 0.43$, $p = .671$). Therefore, I collapsed the observations across pricing lists to focus on the role of language. Consistent with study 1a, the proportion of liars was higher in the L2 condition (41%) than in the L1 condition (17%) (χ_2 (1, N = 201) = 12.19, $p < .001$).

Study 1c: Flight delay

Study 1c generalizes the findings of studies 1a and 1b to a flight delay context. Consumers frequently receive compensation for delays, which increases with longer delays. This situation creates an incentive for consumers to over-report delays.

Participants and study design. This study had 372 participants (168 males; age not recorded; 13% unemployed, including students). It was a 2 (L2 vs. L1) × 2 (number of lying options: 8 vs. 3) between-participants design. As in study 1b, the second factor varied the chance level of lying.

Procedure. The procedure followed studies 1a and 1b. The only difference was the scenario; in this study, participants viewed a complaint form similar to that in figure 1. They were told (in L1 or L2) that their flight had a 30-minute delay. Participants then chose the delay to report from a list of compensation options. The compensation was ¥95 if they reported a delay longer than 180 minutes and ¥50 otherwise. Similar to study 1a, I created two lists with the same structure of compensation options but different numbers of lying options (see Appendix). One list had eight options longer than 180 minutes (chance level lying = 89%), whereas the other had three options (chance level lying = 75%). Each

participant randomly viewed one of the two lists and made a choice. I measured whether participants chose to tell the truth or lie for additional compensation.

Results. As in studies 1a and 1b, I excluded participants who failed to understand the scenario (no exclusions for eight lying options, n = 27 when there were three lying options). Consistent with study 1b, the number of lying options did not moderate the language effect (p = .691). Crucially, lying was again more likely in the L2 condition (47%) than in the L1 condition (16%) (χ_2 (1, N = 345) = 35.22, p < .001).

Discussion

Studies 1a–1c test the effect of language on lying behavior in two consumer-relevant contexts. Counter to Bereby-Meyer et al.'s (2018) finding that L2 contexts decrease lying, they show that L2 makes lying more likely. This contradiction motivated us to investigate when L2 decreases lying behavior and when it does the opposite.

When Does L2 Increase (vs. Decrease) Lying?

According to the reduced-intuition effect, L2 increases lying when people intuitively prefer not to lie (H1a) and decreases it when the intuition is to lie (H1b). I also predicted that these language effects would depend on the strength of intuition (H2). To test these predictions, I had to identify a factor that systematically affects the intuitive preference between lying and telling the truth. In the L1 condition, the intuitive response could vary with the gap between truth and lies. Although most people lie when a small deviation from the truth brings self-benefits (e.g., claiming to have successfully solved three problems in a task, after actually solving two), most people avoid lies that are far from the truth (e.g., claiming to have successfully solved three problems after solving none). In a die-rolling task, lying decreases as the truth–lie gap increases (Hilbig and Hessler 2013). The rationale for this effect is that small gaps are easy to justify (e.g., "I almost made it") and rationalize as non-violations of the norm to be honest, but larger gaps warn people of the risk of being blatant liars and threatening their self-images (Mazar, Amir, and Ariely 2008). I exploited this factor to manipulate the intuitive response in an otherwise identical task.

Figure 1: Results of studies 1a-1c.
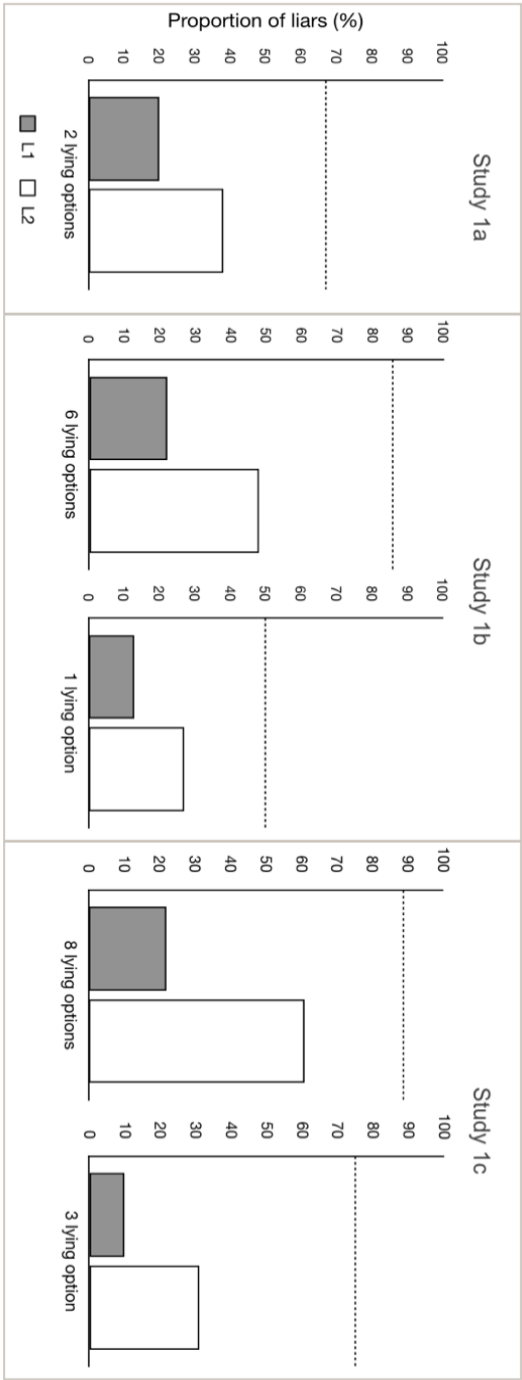
Notes: Y-axes indicate the proportion of liars (0–100%). Grey bars represent L1, and white bars represent L2. Horizontal dashed lines indicate chance levels of lying.

In studies 2a and 2c, I employed a spot-the-difference task to test the theory. The task adopted a 3 (minor vs. moderate vs. major truth-lie gap) × 2 (L2 vs. L1) mixed design, with the truth–lie gap manipulated within participant and language manipulated between participants. The task required participants to find three differences between two advertising images, ostensibly as a test of visual perception of advertisements. Participants observed a pair of images for 5 seconds and then indicated whether they found three differences by selecting "yes" or "no" (no time limit to answer). Participants repeated this task in 12 trials with different image pairs; they were promised a fixed monetary reward each time they managed to find three differences. Unbeknownst to participants, only 3 of the 12 trials had three differences; the rest had 0–2 differences. Table 1 summarizes the types of trials and their interpretation. Participants must have been lying if they claimed to have found three differences in the zero-to-two difference trials. Moreover, the truth–lie gap increased from minor to major as the number of actual differences decreased from two to zero.

Table 1: Trials in the spot-the-difference task and the interpretation of responses.

"Did you find all three differences?"	3-difference trials (n = 3)	2-difference trials (n = 3)	1-difference trials (n = 3)	0-difference trials (n = 3)
"Yes" (with reward)	Honest or lying [a]	Lying (minor)	Lying (moderate)	Lying (major)
"No" (no reward)	Honest			

[a] Ambiguous interpretation if the error rate is non-negligible.

Studies 2a-2c: pretest

I pretested the validity of my paradigm. If the manipulation of truth–lie gaps is valid, I should observe that (1) participants could accurately spot the real number of differences when there is no incentive to lie and (2) their performance is independent of language.

Participants. I posted the link to the pretest on Witmart, a Chinese crowdsourcing platform similar to Amazon Mechanical Turk. Before entering the pretest, participants had to answer demographic questions in Chinese (L1). They also rated their abilities to read in

English (L2) by selecting an option from the following: "Elementary: I can read a few words and phrases," "Intermediate: I can read simple paragraphs like emails," "Upper-intermediate: I can read long essays," or "Advanced: I can read original works." To ensure that participants understood L2 instructions, I screened out those who selected the elementary level; 67 Witmart workers (34 males; M_{age} = 26 years, SD = 5.6) passed the screening and participated in the pretest.

Procedure. I randomly assigned participants to complete the pretest in either their L1 or L2. I asked them to indicate how many differences they found between images, as accurately as possible. After reading the instructions, they went through two practice trials with three differences. They then completed the 12-trial task. In each trial, participants observed two images side by side for 5 seconds and then selected the number of differences they found (0–3). There was no incentive for them to lie. I measured participants' accuracy by calculating the percentage of trials in which they found the right number of differences.

Results. Table 2 summarizes participants' performance. As expected, accuracy was significantly above the chance level (25% if they randomly select a number from 0–3 differences) across types of trials (ps < .001). That is, participants could accurately tell the number of differences when there was no incentive to lie. In addition, performance did not differ by language condition (p = .577). These results testified to the validity of my paradigm. Notably, the error rates in the three-difference trials were non-negligible, making it difficult to tell whether participants would lie in these trials when incentivized (see footnote to Table 1). For this reason, I did not analyze three-difference trials in the main studies.

Table 2: Accuracy of spotting the correct number of differences in the pretest

	3-difference trials	2-difference (minor) trials	1-difference (moderate) trials	0-difference (major) trials
L1	69%	83%	83%	78%
L2	68%	83%	83%	74%

Study 2a: English versus Chinese

Participants and procedure. An additional 153 Witmart workers (who passed the screening of English proficiency, 86 males, M_{age} = 26 years, SD = 4.5) participated in study 2a. About 25% were unemployed (including students). The majority (54%) indicated an "intermediate" level of English reading proficiency. The procedure was the same as in the pretest except that participants were incentivized to lie. Every time participants claimed to have found three differences, I rewarded them with ¥0.65 (USD $0.10). At the end of the study, I debriefed participants about the study purpose using a funneled procedure.

Results. Given the nested structure of the data, I conducted a multilevel analysis with the following model, estimated with the lme4 package in R:

Level 1: $Response_{ij} = \alpha_{ij} + \epsilon_{ij}$.

Level 2: $\alpha_{ij} = \alpha_i + \beta_i \times gap + \epsilon_i$.

Level 3: $\alpha_i = \alpha + \beta \times condition + \epsilon$

Subscript *i* refers to the individual, and *j* refers to the type of trial within individuals; *E*s represent random errors. The outcome variable "response" is binary with a logit link (successful find/lying = 1, unsuccessful find/honest = 0). The truth–lie gap is the trial-level predictor (continuous), and language condition is the individual-level predictor (binary). Coefficients were estimated log odds: Lying was more likely than honesty (i.e., exceeded the 50% chance level) when coefficients were above 0, and vice versa.

I find a significant interaction between the truth–lie gap and language (z = -5.02, p < .001). The negative sign indicates that the tendencies in the L1 condition were attenuated in the L2 condition. In the L1 condition, minor lying was significantly higher than the chance level (β = 2.20, SE = 0.33, z = 6.71, p < .001). In contrast, major lying was significantly lower than the chance level (β = -1.02, SE = 0.30, z = -3.38, p = .001). The L2 condition significantly reduced these tendencies, decreasing minor lying (z = -4.30, p < .001) and increasing major lying (z = 3.59, p < .001). Moderate lying fell between minor and major lying and was not influenced by language (z = -0.07, p = .948). Table 3 summarizes the estimated likelihoods. Figure 2 plots the results.

Table 3: Estimated likelihoods of lying across sizes of lies and languages.

| Study	Language condition	Minor lying	Moderate lying	Major lying
Study 2a	L1	82% a	57%	35% a
	L2	59% a	57%	56% a
Study 2b	L1	47%	25% b	12% a
	L2	39%	34% b	30% a
Study 2c	L1	59% b	49%	41% b
	L2	48% b	53%	55% b

a Significant language effect at $p < .001$. b Language effect at $p < .1$.

Study 2b: English versus French

Participants and procedure. To generalize the findings to a different L1, I conducted study 2b with native French speakers. I recruited 70 participants (40 males, $M_{age} = 31$ years, SD = 7.7; all passed English-proficiency screening) via social networks. They were located in France and working in various industries (13% were students). The majority of participants (50%) reported an "advanced" level of English proficiency. The procedure followed study 2a. Participants received €0.20 (USD $0.23) each time they claimed to have found three differences.

Results. I conducted the same analysis as in study 2a and obtained a significant interaction in the same direction ($z = -3.37$, $p < .001$). In the L1 condition, major lying was significantly below chance ($\beta = -2.94$, $SE = 0.48$, $z = -6.14$, $p < .001$); the L2 condition reduced this tendency, increasing major lying ($z = 3.47$, $p < .001$). In contrast with study 2a, I did not find that the rate of minor lying significantly differed from the chance level ($\beta = -0.21$, $SE = 0.47$, $z = -0.44$, $p = .657$). Unsurprisingly, given the theorizing, the L2 condition had little influence on minor lying ($z = -1.26$, $p = .209$). Moderate lying fell between minor and major lying and was less likely in the L2 condition than in the L1 condition ($z = 1.93$, $p = .054$).

Figure 2: Results of studies 2a-2c.
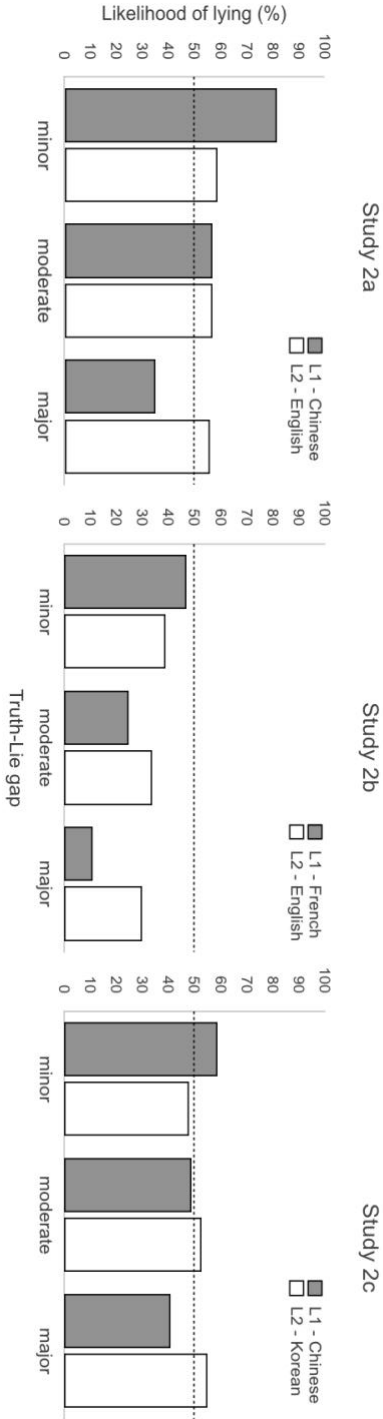
Notes: Y-axes are estimated likelihoods of lying (0–100%), and x-axes the truth–lie gap (minor, moderate, major). Horizontal dashed lines represent the chance level of lying (50%).

Study 2c: Korean versus Chinese

Participants and procedure. To generalize the findings to a different L2, I conducted study 2c with adult Chinese learners of Korean (using a panel from the survey company wjx.cn). I used the same screening question for Korean proficiency as in studies 2a and 2b. The majority (81%) indicated an "intermediate" level of proficiency. There were 117 participants (30 males; M_{age} = 30 years, SD = 4.3; 7% were students). The procedure followed studies 2a and 2b, with a reward of ¥1.5 (USD $0.22) for each claim of successful find.

Results. Using the same model, I again obtained a significant interaction between the truth–lie gap and language (z = -3.95, p < .001). In the L1 condition, minor lying was significantly above chance (β = 0.71, SE = 0.30, z = 2.34, p = .019). Major lying fell below the chance level, though the difference was marginal (β = -0.66, SE = 0.41, z = -1.61, p = .108); the L2 condition decreased minor lying (z = -1.82, p = .069) and increased major lying (z = 1.76, p = .079). Moderate lying fell between minor and major lying and was not influenced by language (z = 0.27, p = .785).

Discussion of studies 2a–2c

Studies 2a–2c demonstrate the reduced-intuition effect: When lying is intuitive (i.e., minor deviation from the truth is required to obtain the monetary reward), L2 decreases lying (H1b); when honesty prevails (i.e., major deviation from the truth is required to obtain the monetary reward), L2 increases lying (H1a). Language does not (or only weakly, as in study 2c) affect lying behavior when lying does not differ (or marginally differs) from the chance level (H2). This pattern is robust across language pairs. These crossover effects also suggest that L2 does not alter the perceived value of material benefits. If L2 increased lying in studies 1a–1c by making rewards look more appealing, I should have observed that L2 always increased lying relative to L1. To test the generalizability of the findings, in the next study (study 3) I add a manipulation of the truth–lie gap to the flight-delay scenario used in study 1c and implement the manipulation between, rather than within, participants.

Study 3: Revised flight-delay scenario

Participants and design. I used the flight-delay scenario from study 1c and randomly assigned 613 participants from Baidu (294 males, M_{age} = 30 years, SD = 7.8; 12% were students or unemployed) to read the scenario in L1 (Chinese) or L2 (English). The study was a 2 (L2 vs. L1) × 2 (minor vs. major truth–lie gap) between-participants design. In the scenario, participants could receive ¥50 as compensation when reporting the flight delay truthfully ("≤ 30 minutes") but ¥95 when overreporting its length. In the minor truth–lie gap condition, to receive ¥95, participants needed to select "31–60 minutes," which was the category closest to the truth. In the major condition, participants instead needed to select the "> 300 minutes" category, which was furthest from the truth.

Results and discussion. Similar to studies 1a to 1c, I excluded participants who failed to understand the scenario, so the final data set contained 363 participants. A logistic regression (1 = lying, 0 = honest) showed a significant interaction between language (L2 vs. L1) and the truth–lie gap (minor vs. major) on the likelihood of lying (z = -2.46, p = .014). In the major truth–lie-gap condition, participants avoided lying in the L1 condition (proportion of liars = 12%, significantly below the 50% chance level, p < .001). However, this tendency toward honesty decreased in the L2 condition (proportion of liars = 25%, z = 1.98, p = .048). The lying rate in the minor truth–lie- gap condition did not significantly differ from the chance level in the L1 condition (49%, p = .860) and thus, unsurprisingly, was not affected by the L2 condition (39%, z = -1.49, p = .136). These results are consistent with the findings of studies 2a–2c. Figure 3 plots the results.

Why Does the Reduced-Intuition Effect Occur?

I also tested the proposed mechanism, by which L2 increases feelings of uncertainty (H3), which mediates the reduced intuition effect (H4). I also tested alternative explanations. For example, L2 may reduce emotional reactions to lying, such that lying feels positive (negative) in L1 contexts, but L2 contexts reduce the intensity of such feelings and result in less intuitive responses. In a final meta-analysis of all the experiments, I formally test the strength-of-intuition constraint (H2).

Figure 3: Results of study 3.
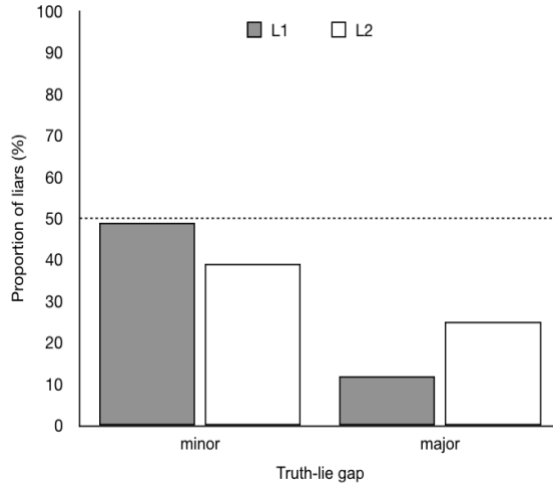
Notes: The y-axis is the proportion of liars (0–100%), and the x-axis the truth–lie-gap. The horizontal dashed line denotes the chance level of lying (50%).

Study 4: The role of uncertainty

Participants and study design. I recruited 500 participants over 25 years of age (i.e., targeted those with a minimum age of 26 years) from Baidu for a study about online shopping. In the online shopping survey (in L1), I embedded the English proficiency-screening question from studies 2a–2c. Those who passed the screening continued to the question that allowed us to detect lying. In the end, 302 participants (164 men, $M_{age} = 31$ years, SD = 5.8) completed the focal measures. Most participants (68%) reported an "intermediate" level of proficiency. The study was a 2 (L2 vs. L1) × 2 (lottery value: low vs. high) between-participants design. I included the second factor for exploratory purposes, because prior research shows that larger incentives increase lying (Kajackaite and Gneezy 2017).

Procedure. Participants completed a series of questions in either L1 or L2. The first two questions referred to gender and monthly income. The critical question was the third one, which incentivized participants to lie. Specifically, I told participants that in addition to their participation fee, they could enter a lottery for a gift card from Amazon.com if they were undergraduate students. Participants randomly viewed the low-value (¥5, about USD $0.72) or high-value (¥100, about USD $14.40) gift card. I asked them to indicate whether

33

they were currently undergraduate students by selecting "yes" or "no." The latest government census (NBS 2012) shows that less than 1% of Chinese people older than 25 years are still undergraduates. Therefore, if the participants, who were all over 25 years, claimed to be undergraduates, they likely were lying (motivated by the prospect of receiving a gift card). (Note that I recruited participants according to birthday information provided by Baidu; I did not ask participants their age.)

Following the question about undergraduate status, I asked participants to indicate whether the gift card was appealing to them, whether they felt bad, and whether they felt good right now (1 = definitely not, 7 = definitely yes). Finally, I measured participants' perceived uncertainty. Based on Geipel et al.'s (2015) measurement of confidence in moral judgments, I created two items to measure generic feelings of uncertainty: "Are you confident about your responses above?" and "Are you sure about your responses above?" (1 = definitely not, 7 = definitely yes, α = .81); I computed an uncertainty score by averaging participants' reverse-coded responses.

Results: Proportion of liars. I submitted language condition (L2 vs. L1) and reward size (high vs. low) to a logistic regression to predict the likelihood of lying. Reward size did not moderate the effect of language ($\beta = 0.40$, $SE = 0.57$, $z = 0.48$, $p = .488$). Chi-square analyses showed that the proportion of liars was slightly higher when the value of the gift card was high (36%) versus low (26%) (χ_2 (1, N = 302) = 3.79, $p = .052$). Participants avoided lying in the L1 condition (16% were liars, significantly below the 50% chance level, $p < .001$), but the percentage of liars increased in the L2 condition (43% were liars), (χ_2 (1, N = 302) = 24.81, $p < .001$), confirming the reduced-intuition effect.

Results: Attractiveness of the gift card. I submitted language and reward size to a two-way analysis of variance (ANOVA) on the perceived attractiveness of the gift card. Unsurprisingly, the high-value gift card was rated more attractive ($M = 4.78$) than the low-value gift card ($M = 4.42$; $F(1, 298) = 4.07$, $p = .045$, partial $\boldsymbol{\eta^2} = .01$). Moreover, the L2 condition decreased the perceived attractiveness of the gift card ($M = 4.87$) relative to L1 ($M = 4.33$; $F(1, 298) = 9.38$, $p = .002$, partial $\boldsymbol{\eta^2} = .03$). According to a marginal interaction between language and reward size, the attractiveness gap between high- and low-value gift cards was smaller in the L2 condition ($M_{high} = 4.34$, $M_{low} = 4.33$) than in the

L1 condition (M_{high} = 5.21, M_{low} = 4.52; $F(1, 298)$ = 3.79, p = .053, partial η^2 = .01).

Attractiveness of the gift card did not predict the likelihood of lying (β = -0.04, SE = 0.08, z = 0.26, p = .611). This finding negates the possibility that language affects lying by altering the perceived value of gains.

Results: Liars' feelings. I next examined how people felt after lying. Because language and lying were related, I conducted subgroup analyses. In the L1 condition, liars did not feel more positive (M = 4.86) than non-liars (M = 4.68; $t(134)$ = 0.52, p = .600). However, liars felt more negative (M = 3.64) than non-liars (M = 2.93; $t(134)$ = 1.95, p = .053). The same pattern emerged in the L2 condition: Liars did not feel more positive (M = 4.82) than non-liars (M = 5.05; $t(164)$ = 1.04, p = .302) but felt more negative (M = 3.93) than non-liars (M = 3.26; $t(164)$ = 2.51, p = .013). Overall, language had no effect on liars' or non-liars' feelings (p > .15), except that non-liars in the L2 condition felt slightly more positive than those in the L1 condition ($t(207)$ = 1.76, p = .079).

Results: Perceived uncertainty. Finally, I assessed the role of uncertainty. The L2 condition *(M = 2.79)* increased feelings of uncertainty relative to the L1 condition (M = 2.24; $t(300)$ = 3.82, p < .001). Moreover, uncertainty positively predicted the likelihood of lying (β= 0.53, SE = 0.11, z = 25.77, p < .001). I next tested a mediation model with 5,000 bootstrapped samples, using Hayes's (2013) PROCESS Macro in SPSS (model 4). Figure 4 illustrates the results. In support of H4, the indirect path through uncertainty is significant (β = 0.26, SE = 0.09, 95% CI = [0.12, 0.47]). Accounting for the role of uncertainty, the effect of language on lying shrinks in magnitude (from β = 1.35, SE = 0.28, z = 4.82 to z = 4.00) but remains significant (path c in figure 5, p < .001).

Discussion. Study 4 provides evidence for the increased-uncertainty account (H3 and H4). It also shows that emotional responses toward lying and perceived attractiveness of material benefits cannot explain the effect of language.

Figure 4: Mediation model in study 4.
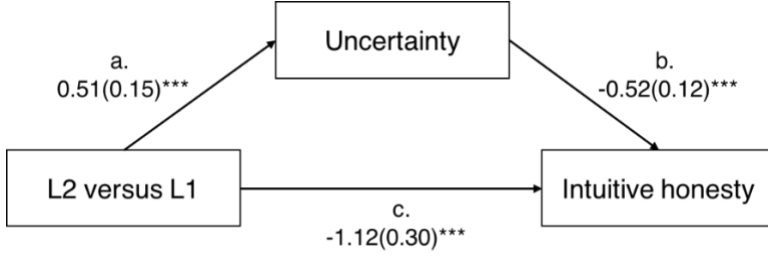
Notes: Standardized coefficients and standard errors are in parentheses. Coefficients for paths b and c are log odds ratios. ***$p < .001$.

Study 5: The intuitive-strength constraint

The reduced-intuition effect implies that the language effect should become stronger as the strength of intuitive responses increases (H2, strength-of-intuition constraint). Although studies 2a–3 offered suggestive evidence, I also seek to assess H2 using meta-regression. I tested the relationship between the strength of intuition in the L1 condition and the magnitude of the reduced intuition effect across eight studies. I operationalized strength of intuition in terms of the deviation from the chance level of lying, as the absolute value of the standardized distance between the L1 response and the chance level of lying:

$$\text{Intuitive strength} = \left|\frac{\text{Prob}_{L1} - \text{Chance}}{\text{Chance}}\right|$$

where $Prob_{L1}$ is the proportion of liars (studies 1a, 1b, 3, and 4) or the estimated mean likelihood of lying (studies 2a–2c) in the L1 condition, and *Chance* refers to the chance level of lying, which varies across conditions and studies. According to this formula, intuitive strength ranges from 0 to 1, with 0 equal to the chance level.

I operationalized the magnitude of the language effect as the absolute value of Cohen's *d*, calculated with the formula by Borenstein et al. (2009):

$$\text{Magnitude of language effect} = \left|\text{LogOddsRatio} \times \frac{\sqrt{3}}{\pi}\right|.$$

Table 4: Data points used for meta-regression.

Study	Condition	N	Chance-level lying (%)	L1 lying (%)	L2 lying (%)	Log odds ratio
1a	-	253	67	20	38	0.86
1b	6 lying options	117	86	22	48	1.23
1b	1 lying options	84	50	13	27	0.93
1c	8 lying options	183	89	22	61	1.69
1c	3 lying options	162	75	10	31	1.47
2a	Minor truth–lie gap	153	50	82	59	-1.67
2a	Moderate truth–lie gap	153	50	57	57	-0.14
2a	Major truth–lie gap	153	50	35	56	1.38
2b	Minor truth–lie gap	70	50	47	39	-0.38
2b	Moderate truth–lie gap	70	50	25	34	0.44
2b	Major truth–lie gap	70	50	11	30	1.27
2c	Minor truth–lie gap	119	50	59	48	-0.77
2c	Moderate truth–lie gap	119	50	49	53	0.12
2c	Major truth–lie gap	119	50	41	55	1.02
3	Minor truth–lie gap	232	50	49	39	-0.40
3	Major truth–lie gap	131	50	12	25	0.93
4	Low-value gift card	152	50	16	39	1.54
4	High-value gift card	150	150	21	46	1.15

The calculations use the data points in Table 4. With the calculated results, I fitted a random-effects model, regressing the magnitude of the language effect on the strength of intuition measure, with adjustment to standard errors (Suurmond, van Rhee, and Hak 2017). In support of H2, the strength of intuition was a strong predictor of the magnitude of the language effect ($\beta = 0.71$, SE $= 0.17$, $z = 4.11$, $p < .001$). The intercept was not significant ($\beta = 0.16$, SE $= 0.10$, $z = 1.63$, $p = .104$), suggesting that language has little effect when people are indifferent to lying or telling the truth.

Figure 5: Relationship of the intuitive strength of response in L1 (x-axis, 0–1) and the magnitude of the language effect (y-axis).
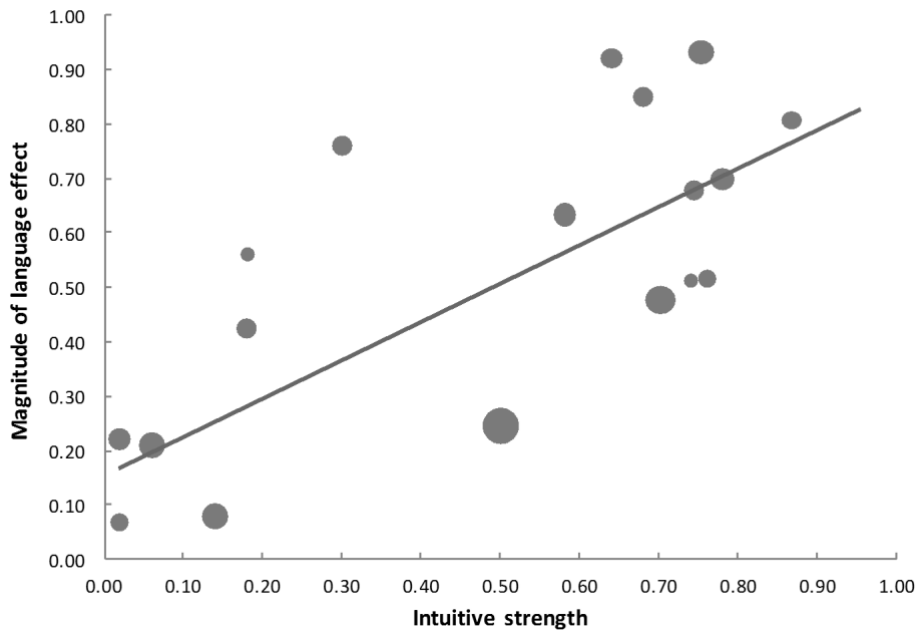
Notes: Dots are data points from studies 1a–4, and dot size represents relative weight. Line is the fitted meta-regression line.

General Discussion

This chapter explores three questions about the effect of language on lying. First, does language always decrease lying, as shown by Bereby-Meyer et al. (2018) and by Alempaki et al. (2017)? I answer in the negative. In car insurance and flight delay contexts, I find that L2 increases the proportion of liars relative to L1, in which most people choose to be honest (H1a). Second, when does L2 increase versus decrease the occurrence of lying? The answer depends on intuitive responses in the L1 context. Using a novel spot-the-difference task in the context of print advertising, I find that L2 makes lying more likely when people intuitively avoid lying (H1a) and less likely when they intuitively pursue lying (H1b). This finding is confirmed by a replication with a flight delay scenario. Third, what explains the reduced-intuition effect? The answer is that L2 engenders uncertainty (H3), which in turn reduces their intuitive attraction or aversion to lying (H4). I do not find evidence that L2 alters people's valuations of the material benefits of lying or attenuates their feelings about

lying or telling the truth. In a final meta-regression, I test the role of intuitive responses by combining all the data; I find that strength of intuition significantly accounts for the variation in the language effect.

This research complements prior research on lying and bilingualism in three ways. First, whereas prior research focuses on whether or not people lie and on how much they lie (Mazar et al. 2008), I investigate people's preferences for lying or telling the truth. By differentiating whether lying or telling the truth is intuitive, I am able to resolve inconsistencies between the current findings and previous research. Second, the main finding (the reduced-intuition effect) adds to prior research showing that L2 reduces intuitive biases in decision making (Keysar et al. 2012), hinders the intuitive choice of unhealthy food (Klesse et al. 2015), and decreases intuitive aversion to drinking recycled water (Geipel, Constantinos, and Klesse 2018). Moreover, the current study offers preliminary evidence on what drives the effect of language on lying behavior. To my knowledge, this research the first to examine the meta-cognitive feeling elicited by L2. Third, the consumer-related studies address the paucity of research on consumer frauds and have useful implications for consumption situations, discussed next. I encourage continued studies to generalize the results from simplified paradigms (e.g., die-roll, coin-flip) to real-life situations.

Alternative theories

In this research, I propose that L2 reduces intuitive responses by introducing a generic sense of uncertainty. In light of these findings, I discuss some alternative theories.

Cultural priming. Language can activate norms and associations originating from the cultural background of the language (Chen and Bond 2010). For example, Lee, Oyserman, and Bond (2010) find that English triggers self-enhancement motives among native Chinese speakers whose L1 is associated with a more collective culture that de-emphasizes individuality. In a context of lying, the effect of language depends on perceived cultural norms of honesty: L2 decreases (increases) lying if it is associated with more (less) honest cultures compared with L1 cultures. Although this tendency may be true when a population has a strong, uniform belief about cultural values, it cannot explain why, in the

spot-the-difference task, I observed for the same language pair both an increase and a decrease of lying within the same participants (studies 2a and 2c).

Attenuated emotionality. The "intuitive responses" that I refer to also may describe people's positive emotions. Because L2 tends to decrease emotional intensity to stimuli, it may decrease the dominant responses for which people have positive (or non-negative) feelings. Corey et al. (2017) find that in hypothetical moral dilemmas, L2 makes people less aversive to "killing" someone to save others. In contrast, I do not find that lying triggers weaker emotional response in L2 contexts; rather, people who tell the truth in L2 contexts feel more positive than those in L1 contexts (study 4). Emotional experience per se thus is unlikely to explain the effects I observe. Instead, the increased-uncertainty account suggests that language changes the weighting of intuition, which could be emotional information in moral decisions.

Enhanced analytical thinking. In contrast to prior findings that foreign language reduces decision biases (Gao et al. 2015), I do not find that L2 increases "rational" responses. If so, I should have observed that L2 increases lying across the board, given the anonymity of lying and its sure gains. This research echoes the recent finding that L2 does not enhance utilitarianism when people respond to moral dilemmas (Hayakawa et al. 2017) and does not promote analytical reasoning (e.g., reducing the Moses illusion) (Costa et al. 2014; Geipel et al. 2015). In general, rationality seems to be a byproduct of reduced intuition in L2 contexts, manifested only when the alternative to the intuitive choice is normatively superior (e.g., binary choice between lotteries) and absent when deliberative calculation is required (Vives, Aparici, and Costa 2018).

To summarize, these findings provide little support for alternative theories about the effect of language on lying. A failure to account for the current findings does not mean these alternative mechanisms cannot underlie the effect of language in other circumstances though; cultural priming does not appear influential in the populations of the current sutdy, but it may have a significant role among people who strongly, homogeneously perceive a cultural gap with regard to honesty (e.g., undergraduates in elite schools). Similarly, I do not find that L2 weakens emotional responses, but emotionality may determine when the

outcome of lying is more salient (e.g., lying face-to-face versus on a computer). Further research can examine these possibilities.

Practical implications

Companies are concerned about consumers' dishonesty; they invest heavily in identifying and deterring potential fraud (FRISS 2019; Ma 2018). This research reveals that language influences dishonesty, so the algorithms and indicators used to identify liars in L1 context may not work as well when people are native speakers of another language. Companies thus should consider language contexts in their efforts to prevent fraudulent behavior.

In some cases, language might be used to nudge people toward more desirable behavior without coercion (e.g., drinking more recycled water) (Geipel et al. 2018). At least for ethical behavior, changing the language context from L1 to L2 does not necessarily help and even can backfire. These findings suggest caution with using language as a tool to curb lying behavior; instead, the recommendation is to understand the intuitive behavior of the targeted population and determine how robust the behavior is, before gauging whether language serves the goal of reducing lying behavior.

Relatedly, this research suggests that interventions that have proved effective in reducing lying in L1 contexts may not work as well in L2 contexts. Reminders of moral identity may transform people's intuitions from acquiring resources to being honest, but this shift might be weaker or even absent in L2 contexts. This possibility is mirrored in the finding that people's intuitions about dishonesty are malleable, but the malleability is much lower in L2 contexts because of the weaker intuitive response in these contexts (studies 2a–3). Policy makers and executives should consider the language to be adopted for the planned intervention and determine whether it matches the native language of the targeted population.

Limitations

The use of online panels allowed us to reach diverse samples of participants and maximize the privacy of lying. However, it also posed challenges, such as ensuring that participants understood the study requirements (especially difficult in bilingual studies). I applied a posteriori data exclusions to half the studies (studies 1a, 1b, 1c, and 3) and a

priori data exclusions to the other half (studies 2a, 2b, 2c, and 4). Both criteria have limitations. A posteriori data exclusions could result in high attrition rates and potentially introduce biases. The variation in chance levels of lying across studies, and across conditions in some of the studies, attenuates this concern, because the pattern of results persists across levels of leniency of data exclusion (e.g., no data exclusion when eight of nine were lying options versus very strict exclusion when one of nine was; studies 1a–1c). A priori data exclusions should eliminate the bias, because the randomization is implemented after data exclusion. Nevertheless, they are less strict than a posteriori data exclusions and introduce more noise, which is less worrisome, given the pretest results showing that response quality does not vary across language conditions. Despite these efforts and results, I encourage studies that replicate the current findings with alternative methods.

Another challenge is that participants might have dropped out of the study, which could have introduced bias to the findings, especially in the L2 conditions, in which people might have experienced greater difficulties. To minimize dropouts, I used L1 as much as possible, especially at the beginning of the studies, and made the response formats as simple as possible (i.e., choice rather than typing sentences). I could not identify the dropout rates from Baidu, because Baidu does not record partial responses. However, in studies 2a–2c, Qualtrics recorded all data and revealed no partial responses (i.e., no dropout after participants started responding). Therefore, it seems reasonable to anticipate no impact from dropouts on the findings. Researchers could try to identify who is likely to drop out of bilingual studies in general and how they differ from other participants.

Finally, though I confirm the reduced-intuition effect, I have not delved into the question of how intuitions are established in L1 contexts in the first place. Although previous research offers some hints (e.g., justifiability), I are not aware of any systematic research on this. This question lies outside the scope of the current research, but it would be a fruitful path of investigation for continued research on consumer lying behavior.

To conclude, bilingual consumers can lie in more than one language; this research shows that the language they use has a systematic effect on their propensity to lie. I predicted and established a novel effect that I refer to as the reduced-intuition effect: L2

reduces people's intuitive preferences for lying or telling the truth, such that they either increase or decrease lying, depending on their L1 intuitive responses.

Figure A1: English (L2) policy forms (upper) used in studies 1a–1b and complaint (lower) used in study 1c.

Table A1: Pricing lists for studies 1a–1c.

Study 1	Study 1b		Study 1c	
2 lying options	6 lying options	1 lying option	8 lying options	3 lying options
>10,000km (pay ¥3,000 more)	>10,000km (pay ¥3,000 more)	>10,000km (pay ¥3,000 more)	> 300 min (receive ¥95)	> 300 min (receive ¥95)
5,000–10,000km (pay ¥3,000 more)	5,000–10,000km (pay ¥,3000 more)	5,000–10,000km (pay ¥3,000 more)	240–300 minutes (receive ¥95)	240–300 minutes (receive ¥95)
4,000–5,000km (pay ¥3,000 more)	3,000–5,000km (pay ¥3,000 more)	4,000–5,000km (pay ¥3,000 more)	230–240 minutes (receive ¥95)	180–240 minutes (receive ¥95)
3,800–4,000km (pay ¥3,000 more)	2,500–3,000km (pay ¥0 more)	3,800–4,000km (pay ¥3,000 more)	220–230 minutes (receive ¥95)	150–180 minutes (receive ¥50)
3,600–3,800km (pay ¥3,000 more)	2,000–2,500km (pay ¥0 more)	3,600–3,800km (pay ¥3,000 more)	210–220 minutes (receive ¥95)	120–150 minutes (receive ¥50)
3,400–3600km (pay ¥3,000 more	1,500–2,000km (pay ¥0 more)	3,400–3,600km (pay ¥3,000 more	200–210 minutes (receive ¥95)	90–120 minutes (receive ¥50)
3,200–3,400km (pay ¥3,000 more)	1,000–1,500km (pay ¥0 more)	3,200–3,400km (pay ¥3,000 more)	190–200 minutes (receive ¥95)	60–90 minutes (receive ¥50)
3,000–3,200km (pay ¥0 more)	500–1000km (pay ¥0 more)	3,000–3,200km (pay ¥3,000 more)	180–190 minutes (receive ¥95)	30–60 minutes (receive ¥50)
≤ 3,000km (pay ¥0 more)	≤ 500km (pay ¥0 more)	≤ 3,000km (pay ¥0 more)	≤ 180 minutes (receive ¥50)	≤ 30 minutes (receive ¥50)

Notes: Participants were excluded if they selected options in gray (off-truth and no-benefit).

Figure A2: A one-difference trial in the L1 (upper) and the L2 condition (lower) in study 2a.

Native (Mandarin) condition

Foreign (English) condition
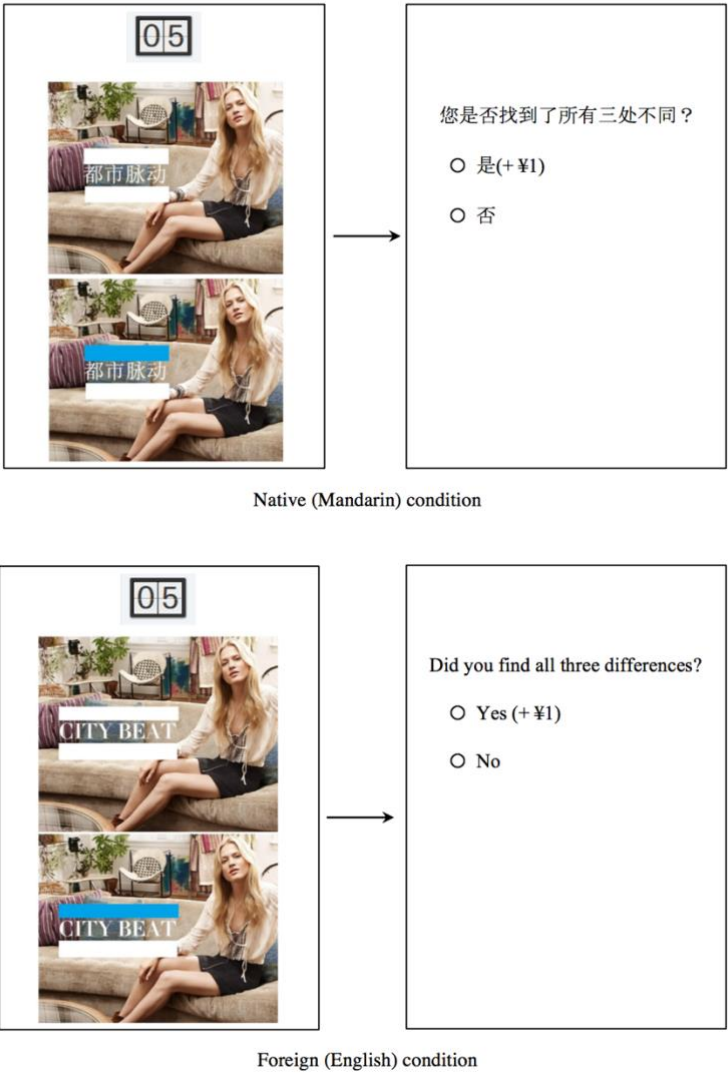
Notes: A pair of images was shown for 5 seconds with a countdown timer, followed by the target question. The incremental re-ward was displayed next to the success claim "Yes".

Chapter 3

Making Recommendations More Effective through Framings

This chapter is based on the original paper published on Journal of Marketing:

Gai, Phyliss Jia and Anne-Kathrin Klesse (2019), "Making Recommendations More
Effective through Framings: Impacts of User- versus Item-based Framings on
Recommendation Click-throughs," *Journal of Marketing*, 83 (6), 61-75.

Chapter 3 is organized as follows. I first review prior research on recommender systems and explanations, define user-based and item-based framings, and develop my conceptual framework with concrete predictions. I report six studies, including two online field experiments, that test my predictions. This chapter ends with a discussion of the theoretical and managerial implications and promising avenues for future research.

Theoretical Background

Recommendation systems and explanations to consumers

Recommendation systems are an automated, data-driven tool that companies frequently adopt to fulfill their consumers' personalization needs (Hinz and Eckert 2010; Ricci et al. 2011). Depending on what consumers have viewed, liked, or purchased, these systems predict what other products they could be interested in and deliver instant suggestions. Research in marketing and information systems highlights such recommender systems as important determinants of sales (Bodapati 2008; Fleder and Hosanagar 2009; Pathak et al. 2010). Two typical methods inform these recommendations. First, collaborative filtering identifies consumers who are similar in their product rating history and recommends items that one consumer likes to similar other consumers. The product ratings might be explicitly provided by consumers or inferred from their online behavior. Second, content-based filtering identifies the product attributes that a consumer likes and recommends products with similar attributes (Ansari, Essegaier, and Kohli 2000). Because each method has shortcomings, companies often combine them to improve the performance of their hybrid recommender systems. Examples include Amazon's "item-to-item collaborative filtering" (Linden, Smith, and York 2017), and the New York Times' collaborative topic modeling (Spangher 2015). Extensive research suggests ways to improve the prediction accuracy of recommendation algorithms using hybrid frameworks (Zhang et al. 2018).

The computationally complex algorithms pose challenges for explaining recommendations to consumers. A clear, concise, accurate explanation is crucial, because it promotes consumers' trust in the recommender systems (Wang and Benbasat 2007) and acceptance of recommendations (Cramer et al. 2008; Kramer 2007). To the best of my knowledge, no research in marketing suggests the optimal methods for explaining

recommendations. In information systems literature, Tintarev and Masthoff (2015) identify five recommendation explanation types. Two explanations are particularly relevant to the current research: collaborative-based and content-based. As their names imply, collaborative-based explanations such as "Consumers who bought this item also bought…" rely on recommender systems that adopt collaborative filtering, whereas content-based explanations, such as "Recommended because you said you owned…," involve recommender systems that use content-based filtering. The other explanation types either overlap with the content-based explanation (e.g., cased-based that specifies the items compared by the underlying algorithm) or assume unique inputs (e.g., demographic-based; Tintarev and Masthoff 2015).

Rather than addressing explanation styles yoked to distinct, specific recommendation algorithms (Tintarev and Masthoff 2015), I define recommendation framing according to the various explanations that might be provided, even with the same recommender system. Because most recommender systems take a hybrid approach that combines the input from users (i.e., inter-user similarity in preferences) and the input from items (i.e., inter-item similarity in attributes), I compare framings that highlight one input over the other, user-based framing versus item-based framing. Accordingly, the goal of the current research is to establish the causal impact of alternating between the user-based and item-based framings on click-throughs of recommendations, rather than to provide an exhaustive categorization of explanation styles (Tintarev and Masthoff 2015).

Comparisons of item-based and user-based framing

As detailed in Figure 1, with user-based framing, the provided explanations draw attention to the shared tastes of consumers of a focal item. This framing describes how the target user (u) is similar to other users (u'), due to their shared interest in the focal item (i), and it indicates that the focal item (i) and recommended item (i') are related because they attract the same users (u'). Item-based framing instead highlights the match between the focal and the recommended products (i and i'), either with or without specifying their shared properties. For example, "More in Health" suggests that recommended articles will be similar to the focal item, because they fall in the same news category; "Similar to this

item" also emphasizes the relationship between the items but does not cite specific product attributes.

Figure 1: Illustration of the definitions of user-based and item-based framing.
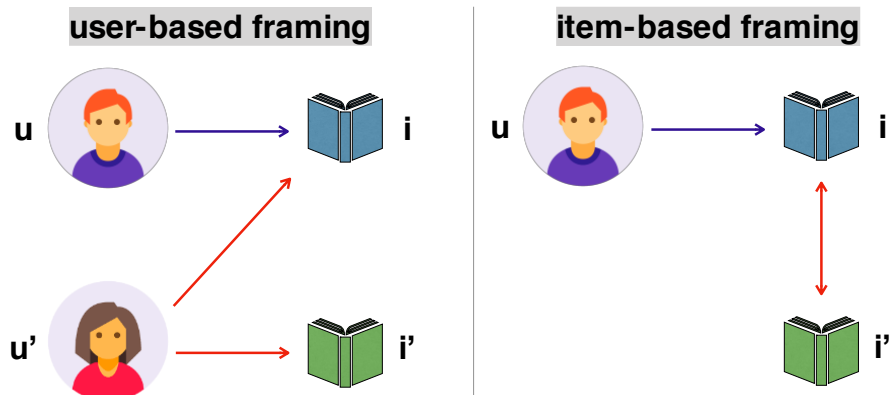
Notes: In user-based framing, consumers u and u' match in their liking of product i, and products i and i' match in their consumer u'. Item-based framing suggests products i and i' are related.

As these definitions make clear, both user-based and item-based framings suggest *product matching* between items i and i' as the basis for the recommendation. User-based framing matches products by consumers; item-based framing suggests that products are matched on their attributes. Notably, user-based framing also suggests *taste matching* (users' shared taste in the focal product) as the basis for recommendation, such that it offers informational value beyond that provided by item-based framing. According to advice-taking research, consumers extract information from others' tastes to predict their own satisfaction with unfamiliar products (Yaniv, Choshen-Hillel, and Milyavsky 2011; Morvinski, Amir, and Muller 2017) and tend to adopt others' preferences if they believe those others' tastes match their own (Hilmert, Kulik, and Christenfeld 2006; Naylor, Lamberton, and Norton 2010). Therefore, I reason that user-based framing offers additional information (i.e., about others' tastes) that can reduce consumers' uncertainty about whether they will like or dislike the recommended item. By offering additional information about taste matching beyond product matching, user-based framing can serve

as a sort of double-guarantee that consumers will enjoy the recommended item and thus should be more effective in triggering click-throughs. Formally,

H1: User-based framing increases recommendation click-throughs relative to item-based framing.

This predicted advantage of user-based framing is premised on consumers' perception that the taste matching is successful. Taste matching provides valid information for consumers to infer their liking of the recommended item only if they believe others' preferences reflect their personal tastes. With automated recommendations, many factors could influence the extent to which consumers perceive taste matching as successful and potentially reduce or even reverse the framing effect, such that user-based framing actually becomes disadvantageous compared with item-based framing. I consider three such factors that might provide important boundary conditions to the framing effect. I purposefully select a range of factors related to the consumer segment (i.e., more or less consumption experience), the products (i.e., more or less attractive focal products), and other users (i.e., more or less similar to the recommendation recipient).

Consumption experience, focal attractiveness, and dissimilarity cues

User-based framing differs from item-based framing in the implication that the recommender system attempts to match users on the basis of their tastes in the focal product. Consumers who have accumulated more experience in a consumption domain may be less likely to perceive this taste matching as successful, for two reasons. First, consumers develop more refined and sophisticated tastes as they acquire more experience within a consumption category (Bettman, Luce, and Payne 1998). With more experience, consumers are better able to differentiate products and develop a more complex understanding of the category (Alba and Hutchinson 1987). Second, more experienced consumers have accrued more observations of individual differences in tastes and therefore likely regard their own taste as idiosyncratic (Packard and Berger 2017). Accordingly, they might deem a shared interest in a single or a limited set of products (i.e., focal products) as insufficient for taste matching, leaving them reluctant to converge with or rely on other users' preferences. In contrast, inexperienced consumers whose tastes are still coarse (Hoeffler and Ariely 1999) may be less skeptical of a match between their own and others'

tastes (Becker 1991), leading to the advantage of user-based over item-based framing. I predict:

H2: The advantage of user-based framing relative to item-based framing decreases for consumers with more consumption experience in the focal domain.

Consumers' perceptions of taste-matching success also likely depend on the products themselves. I propose that taste matching may appear less accurate if the focal product is less attractive, because consumers constantly learn about their own preferences through their reactions to different products (Ariely and Hoeffler 1999; West, Brown, and Hoch 2002). More attractive focal products would serve as salient and diagnostic signals of personal preferences (Zunick, Teeny, and Fazio 2017), which in turn should promote perceived success in taste matching with other users who presumably also like the attractive focal product. In contrast, people tend to view less attractive products as less indicative of their taste or even a negative signal of preferences, lowering the perceived accuracy of taste matching and resulting in a smaller advantage or even a disadvantage of user-based framing relative to item-based framing. Specifically:

H3: The advantage of user-based framing over item-based framing diminishes for unattractive focal products.

Finally, in ambiguous situations, in which the identities of other consumers are not revealed, people tend to assume self–other similarity (Naylor, Lamberton, and West 2012). However, some companies provide information about the users who are the basis for the recommendation, explicitly (e.g., location of other users on booking.com) or implicitly (e.g., books of "teen's choice" on Amazon). When this information points to a dissimilarity between users, it may undermine the value of taste matching. As existing research shows, dissimilarity on certain dimensions (e.g., gender) activates thoughts of self–other dissimilarity in other domains (e.g., product attitudes; Tuk et al. 2019). Consumers thus might categorize a recommendation as reflecting "non-self" tastes if it is associated with dissimilar others and deem taste-matching efforts unsuccessful. In this case, I no longer expect an advantage of user-based framing over item-based framing but rather predict that it becomes disadvantageous, because consumers tend to avoid dissimilar others' tastes (Berger and Heath 2008). Formally:

H4: User-based framing decreases recommendation click-throughs relative to item-based framing in the presence of cues suggesting self–other dissimilarity.

To summarize, I posit that, compared with item-based framing, user-based framing provides additional information about the preferences of other users that consumers can use to reduce their uncertainty about the recommendation; as a result, it encourages them to click on it. The informational value of user-based framing and whether it benefits or harms recommendation click-throughs depends on the perceived success of taste matching. I conducted six studies to test these predictions and the conceptual framework (see Appendix for the full results of all the studies). Studies 1a and 1b test H1 (main effect) in field experiments with article recommendations. The results affirm the advantages of user-based framing over item-based framing in a managerially relevant setting. Study 2 tests H2 that consumption experience functions as a moderator, in a setting that provides painting recommendations. For studies 3 and 4, I created book-shopping scenarios to test H3. I find consistent support for the hypotheses, whether the attractiveness of the focal product is rated by a separate batch of consumers (study 3; analogous to data gathered by companies from prior consumers) or by the same consumers (study 4). In study 4, I also leverage information about the ages of other consumers to establish a dissimilarity cue that leaves user-based framing disadvantageous relative to item-based framing, as predicted in H4. Study 5 strengthens the support for H4 by using gender as a different cue of dissimilarity. These findings thus add unique theoretical insights and suggest managerial strategies for companies.

Studies 1a and 1b

I conducted studies 1a and 1b in collaboration with a media company that regularly pushes articles to its subscribers on WeChat, the top mobile app in China (Novet 2017). These two field studies differ primarily in the item-based framing, which I varied to ensure that the user-based framing is responsible for the increased click-throughs. This company also offers an ideal context to test the predicted main effect (H_1) for two reasons. First, it primarily publishes articles about social science research, and its subscribers represent a highly homogeneous community. In this context, readers likely view taste matching as

successful in general. Second, this company had not used recommendations before I ran study 1a, so I could observe the unique effects of framing, unaffected by prior practice.

Study 1b, conducted 14 months after study 1a, then offers a conceptual replication with completely new stimuli. During the 14-month interval, the company did not adopt any other article recommendation and witnessed a 52% increase in the number of subscribers (from 70,488 to 107,338) on WeChat. These changes should minimize carry-over effects from study 1a to study 1b. Because I had no access to individual users' data, I conducted both experiments at the article level (for a similar design, see Gong et al. 2017).

Study 1a

Article selection. Before the experiment started, I carefully selected 71 original articles that had been previously pushed to all subscribers, according to four criteria. First, the number of times people had read each article could not exceed 400, which is low compared with the overall average 3,071 (as of August 2017, immediately before study 1a). Second, the article had been pushed to subscribers at least three months ago, to ensure that it was likely to be unfamiliar to most readers. Third, it reported on research on human beings, which is the main content the company disseminates, to avoid the risk that the article topic would seem odd to readers. Fourth, the article could not contain time- or event-specific content (e.g., "Top research of 2016"), because timeliness might interfere with the framing effects.

Study design. I assigned these preselected articles to three conditions: no recommendation (N = 9), user-based framing (N = 31), or item-based framing (N = 31). The assignment used stratified randomization; each condition includes approximately the same percentage of articles published in different years (12% published in 2014, 20% in 2015, 48% in 2016, and 20% in 2017). This approach helped exclude bias due to publication timing. With the control condition (no recommendation), I test whether a recommendation per se is effective, regardless of its specific framing. I limit the sample size for this control condition, because it is not the focal interest and to maximize the statistical power of the contrast between user-based and item-based framings. The recommended articles, with user-based or item-based framing, attracted more reads than

non-recommended articles ($p < .001$; see Appendix). I do not discuss the non-recommended articles further.

   Procedure. I randomly paired one article in the user-based framing with another in the item-based framing. The 31 pairs of recommendations then were distributed randomly across 31 days. Every weekday, the company pushed one set of articles to all subscribers. Each set had a headline article that was most salient to readers, which served as the focal article. Each pair of recommended articles was inserted toward the end of the focal article. Therefore, the readers would only see the recommendations if they were really interested in the focal article and finished reading it. The recommendation consisted of the recommendation framing and the title of the recommended article (a hyperlink to click on and read). The user-based framing read, "People who like this article also like…," and the item-based framing specified the category that both the focal and the recommended articles fell in, "More analyses of scientific research" (all focal articles were in this category). The order in which the two framings appeared (one preceded the other) was counterbalanced across days.

   To measure click-throughs on the recommendations, I calculated the click-through rate (CTR) for each recommended article:

$$CTR = \frac{CurrentRead - InitialRead}{FocalRead} \times 100$$

InitialRead is the number of reads of the recommended article before the experiment started. It does not differ by the framing condition ($p = .543$). CurrentRead is the number of reads after the experiment started, recorded at four time points of 24 hours, 48 hours, 72 hours, and 2 weeks after the recommendation, which enables us to determine whether the framing effect varies over time. The number of reads of the focal article (FocalRead) also was recorded at these four time points. Table 1 presents the descriptive statistics.

Table 1: Means (standard deviations) of number of article reads in study 1a.

	Focal Articles	User-Based Framing	Item-Based Framing	Non-Recommended
Before experiment	0 (0)	316 (93)	294 (93)	274 (103)
24 hours	2595 (2280)	343 (125)	306 (96)	-
48 hours	2610 (2404)	344 (126)	307 (97)	-
72 hours	2793 (2564)	345 (126)	308 (97)	-
2 weeks	3071 (2693)	349 (132)	310 (98)	275 (103)

Results. There were 17 missing cases because I could not observe the reads of some articles at some time points. Furthermore, I excluded one outlier article in the user-based condition from the analysis, because its CTR ($M = 19.25\%$ across the time points) was disproportionately higher than the average of all the other articles (0.61%). The final data set contains 228 observations: 112 in the user-based condition and 116 in the item-based condition. Due to the nested structure of the data (articles nested within days), I constructed a multilevel model with CTR as the outcome variable and random intercepts at the day level. The recommendation faming served as the predictor (0 = item-based, 1 = user-based). Because time did not moderate the framing effect ($p = .919$), I focus on the overall effect. Consistent with H1, CTR was significantly higher in the user-based condition than in the item-based condition ($M = 0.72\%$ versus 0.51%; $b = 0.22$, $SE = 0.06$, $t(196) = 3.79$, $p < .001$). Including the outlier article added to the error of estimation but also magnified the framing effect ($M = 1.26\%$ versus 0.56%; $b = 0.70$, $SE = 0.22$, $t(199) = 3.11$, $p = .002$).

Follow-up survey. These results provide initial evidence that user-based framing outperforms item-based framing. Recall that I propose this effect arises because, unlike item-based framing, user-based framing offers additional informational value by suggesting taste matching as part of the recommendation strategy. To determine whether readers interpret the two framings in this way, I distributed a follow-up survey to the subscribers ($N = 780$, 67% females, $M_{age} = 24.4$ years, $SD_{age} = 5.7$). Note that I do not know whether the survey participants also participated in the experiment, because the

experiment was conducted on the article level. The survey participants were randomly assigned to read the user-based framing (N = 409) or item-based framing (N = 371) that I used in the field experiment, then rated the extent to which they agreed with eight statements (1 = "strongly disagree," 6 = "strongly agree"). Half of the statements referred to product matching as the basis for the recommendation (e.g., "The recommendation is based on articles that are similar to what I have read," "The recommendation is based on the categorization of articles"; Cronbach's α = .68), whereas the other half referred to taste matching (e.g., "The recommendation is based on readers who have similar preferences with me," "The recommendation is based on the categorization of readers"; Cronbach's α = .70). See appendix for all the items.

To test whether both user-based and item-based framings imply product matching to consumers but only user-based framing suggests taste matching as a recommendation strategy, I submitted the perceived product-matching and perceived taste-matching scores to a 2 (two dependent measurements) × 2 (recommendations framing: user-based vs. item-based) mixed analysis of variance. A main effect of the measurement arose; participants more readily recognized product matching than taste matching as the basis for recommendations ($F(1,778) = 226.04, p < .001$), suggesting that product matching is the default perceived recommendation strategy. In addition, I found a significant interaction between measurement and framing ($F(1, 778) = 9.10, p = .003$). In support of my reasoning, participants recognized product matching as the basis for the recommendation equally in both user-based and item-based conditions (M_{user} = 4.83, M_{item} = 4.82, $t(778)$ = -0.08, p = .941). However, participants in the user-based framing condition agreed that taste matching was a basis for the recommendation to a greater extent than participants in the item-based framing condition (M_{user} = 4.38, M_{item} = 4.18, $t(778)$ = 3.44, p = .001). That is, user-based framing (versus item-based framing) offers information about taste matching, in addition to product-matching information.

Discussion. Consistent with H1, study 1a demonstrates that framing recommendations as user-based rather than item-based attracts more click-throughs in a field setting. It also provides support for the notion that perceived taste matching differentiates user- from item-based framing. It remains unclear, however, whether the framing effect really is due

to the additional informational value of user-based framing or if readers instead avoid reading more articles in the same category, a response potentially evoked by the item-based framing that read "More analyses of scientific research." In study 1b, I thus use a different item-based framing operationalization but keep the user-based framing constant. If the framing effect in study 1a is due to the informational value of user-based framing, it should emerge regardless of whether the item-based framing specifies the article category.

Study 1b

Article selection and study procedure. study 1b contains a new set of articles and a more generic item-based framing ("Similar to this article"). I selected the recommended articles using criteria similar to those I applied in study 1a, except I also required that they had not been recommended in study 1a. I increased the constraint on the number of reads before recommendation, from 400 to 480 reads, to ensure a decent sample size and account for the substantial increase in the number of subscribers to the company. With these criteria, I identified 66 articles, half randomly assigned to the user-based and the other half to the item-based framing condition. The procedure is the same as in study 1a, and the experiment lasted for 33 days.

Results. Similar to study 1a, I excluded an outlier article in the item-based condition that had a peculiarly high CTR (M = 24.35%) relative to the average of all the other articles (0.95%). Thus I retain 258 observations in the final dataset. Unlike study 1a, I did not balance the year of publication across conditions; more articles published in 2018 were assigned to the user-based condition than to the item-based condition (49% vs. 37%, p = .073). Therefore, I controlled for publication bias (0 = published before 2018, 1 = published in 2018) in the analysis. Using the same multilevel modeling approach as in study 1a (time did not moderate the framing effect, p = .945), I found a lower CTR for articles published in 2018 than for those published before 2018 (M = 0.58% vs. 1.06%; b = -0.49, SE = 0.13, $t(223)$ = -3.80, $p < .001$). More importantly, controlling for the publication year, CTR was higher in the user-based condition than in the item-based condition (M = 1.25% vs. 1.06%; b = 0.19, SE = .08, $t(223)$ = 2.44, p = .015). The advantage of user-based framing persisted but shrunk in magnitude without the covariate (M = 1.01% vs. 0.88%; b = 0.14, SE = 0.08, $t(224)$ = 1.79, p = .075). Including the outlier

article made the framing effect insignificant (b = -0.32, SE = 0.28, $t(227)$ = -1.12, p = .262).

Discussion. Study 1b strengthens the support for H1 by showing that the advantage of user-based framing over item-based framing persists when the item-based framing does not specify the category of the articles. Taken together, the framing effects observed in studies 1a and 1b cannot be accounted for by avoidance of same-category items. I next seek to provide evidence for the conceptualization by testing boundary conditions on the framing effect.

Study 2

With study 2 I examine the prediction that user-based framing outperforms item-based framing in terms of recommendation CTR for inexperienced consumers but not for experienced consumers within a consumption domain (H_2). I displayed the article recommendations in studies 1a and 1b at the end of the focal article, which guaranteed that readers liked the focal article, because they would only see the recommendation if they read it to the end. In line with this element, in study 2 I provide product recommendations to only to participants who indicate that they like the focal product.

Participants and design

I recruited 403 participants located in the United States from Amazon Mechanical Turk (MTurk; 186 females, M_{age} = 37.3 years, SD_{age} = 11.6). Data from MTurk offer reliability comparable to those gathered from offline laboratories (Horton, Rand, and Zeckhauser 2011; Paolacci, Chandler, and Ipeirotis 2010). After giving their informed consent, participants entered an "Online Museum" study and viewed 50 paintings created between the 17th and 20th centuries. Each painting was paired with a hidden recommended painting with the same theme (e.g., seascape). The recommendations feature either a user-based (N = 195) or item-based (N = 208) framing. All paintings were obtained from Google Art Project. I operationalized the proposed moderator, consumption experience, as the frequency of visiting art museums, measured on a continuous scale.

Procedure

Participants viewed 50 focal paintings in a random sequence. Next to each focal painting, there was a "like" button in the shape of a heart. I told participants to mark their favorite paintings by pressing the button. To make sure they provided their honest opinions, I told them that they would enter into a lottery for postcards of their favorite paintings. After participants clicked on a "like" button, another button appeared, indicating either "People who like this painting also like…" (user-based framing) or "Similar painting to this" (item-based framing), depending on the randomly assigned condition. They could click on this button to view the recommended painting in a pop-up window, as well as exit the pop-up window any time to continue viewing the focal paintings. The CTR for each recommendation was calculated as

$$\text{CTR} = \frac{N \text{ who viewed the recommended}}{N \text{ who liked the focal}} \times 100$$

The denominator is the number of participants who liked the focal painting, and the numerator indicates how many of them chose to view the recommended painting. The CTR varied between 0% and 100%. After participants finished viewing all the focal paintings, they saw both the user-based and item-based framing and indicated which one they encountered in the "Online Museum" (94% answered correctly). To measure participants' consumption experience, I asked them to indicate how often they visited art museums in their life (1 = never, 2 = seldom, 3 = sometimes, 4 = often, 5 = very often). I deliberately chose this single-item measurement to maximize the number of observations per level of consumption experience and thus to obtain reliable CTRs to estimate the effects of framing. I ended the study with a few demographic questions.

Results

Similar to study 1, I conducted the analyses at the level of each recommended painting. I calculated the CTR for each recommendation per framing and per level of consumption experience, resulting in a dataset with 499 observations. One observation was missing because one focal painting in the item-based condition was not liked by any participants who never went to art museums. I regressed CTR on framing (0 = item-based, 1 = user-based), consumption experience (continuous from 1 to 5), and their interaction. The results, as plotted in Figure 2, indicated a significant interaction between framing and

consumption experience (b = -2.23, SE = 0.97, $t(495)$ = -2.30, p = .022). In support of H2, the CTR was higher in the user-based condition than in the item-based condition among people who never (b = 7.83, SE = 2.38, $t(495)$ = 3.29, p = .001), seldom (b = 5.60, SE = 1.68, $t(495)$ = 3.33, p < .001), or sometimes (b = 3.36, SE = 1.37, $t(495)$ = 2.45, p = .015) visited art museums. I found no significant difference across framings for those who often (b = 1.13, SE = 1.68, $t(495)$ = 0.67, p = .501) or very often visited art museums (b = -1.10, SE = 2.37, $t(495)$ = -0.47, p = .642). On average, the CTR was slightly but significantly higher in the user-based condition than in the item-based condition (M = 46.63% vs. 43.23%, $F(1,495)$ = 6.15, p = .014), probably because most participants had rather limited experience with arts (median = 2 of 5). Moreover, the CTR decreased with more consumption experience (b = -6.70, SE = 0.68, $t(495)$ = -9.78, p < .001) in the user-based condition, but this trend was attenuated in the item-based condition (b = -4.46, SE = 0.69, $t(495)$ = -6.49, p < .001).

Figure 2: Regression results of study 2.
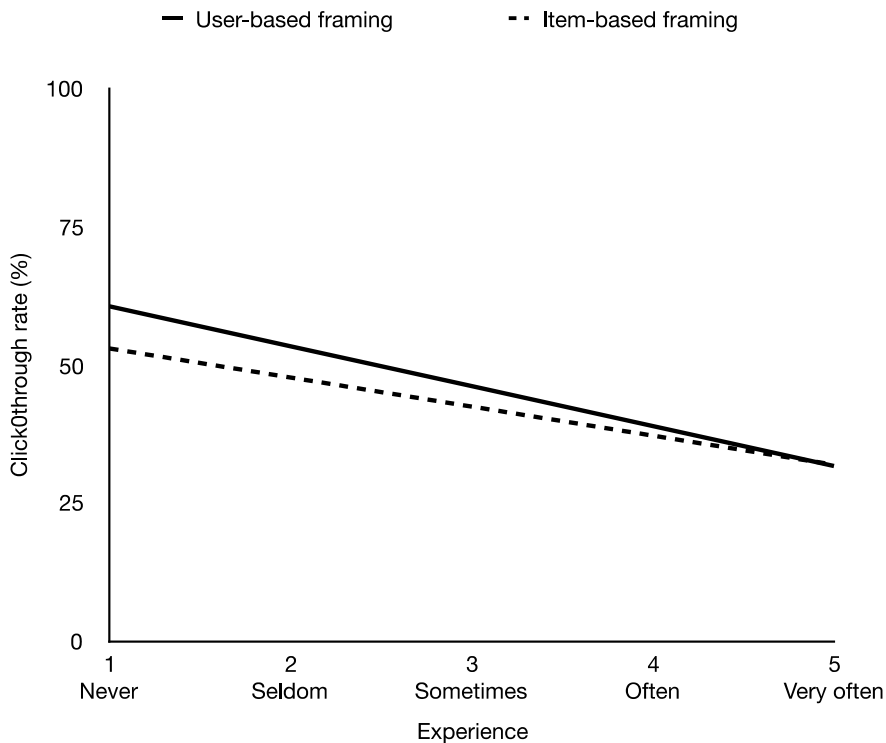
Discussion

Study 2 provides support for H2; the advantage of user-based framing over item-based framing diminishes for people with more consumption experience. Also consistent with the theorizing that more experienced people are less likely to perceive taste matching as accurate, I find that greater consumption experience induces a greater decrease in the recommendation CTR when it is framed as user-based as opposed to item-based.

The paradigms I use in studies 1a, 1b, and 2 guarantee that participants like the focal product; they only see the recommendation if they finish reading the article or like the focal painting. In study 3, I relax this criterion so that all participants receive a product recommendation regardless of whether they expressed interest in the focal product; this allows us to test for a moderating role of the attractiveness of the focal product (H3).

Study 3

According to the theorizing, liking the focal product is a necessary prerequisite for taste matching to be perceived as successful and thus for the advantage of user-based framing over item-based framing to arise. To test this assumption explicitly, I vary the attractiveness of the focal products and inquire into people's intentions to click on the recommendation. I expect the advantage of user-based framing to diminish or even reverse for less attractive focal products (H3). Moreover, if focal attractiveness affects perceived success in taste matching, it should relate more positively to CTR when the recommendations are framed as user-based as opposed to item-based.

Participants and study design

Fifty participants located in the United States were recruited from MTurk to participate in a study about shopping for novels on Amazon (18 females, M_{age} = 37.57 years, SD_{age} = 12.32). The majority (56%) had never purchased a novel on Amazon. I manipulated the framing within participants. Unlike the previous studies, the user-based framing emphasized users' actions ("Consumers who viewed this book also viewed…") rather than likes. This variation is purposeful; the decision to view a book's webpage might be driven merely by the appearance of the cover, but liking a book requires understanding it. The item-based framing followed study 2 ("Similar to this item"). This

setup might be less engaging than previous studies. However, it taps into an important situation in which consumers are merely browsing products without concrete goals.

Book selection

I took a convenience sample of 50 novels from the "Literature and Fiction" category on Amazon that had garnered fewer than 200 reviews before the experiment started, which were presumably unfamiliar to most of the participants. A pretest with a separate batch of 50 participants from MTurk (23 females, M_{age} = 33.6 years, SD_{age} = 8.4) confirms that these books are unfamiliar to MTurk workers (maximum mean familiarity is 2.14 of 10, where higher values indicate more familiarity). From the 50 books, I randomly selected 25 candidates as focal books and the other 25 candidates as recommended books. Then I randomly paired a candidate from the focal set with another from the recommended set. I aimed for an equal number of focal–recommended pairs per framing condition for the Study.

The pretest demonstrates that the distribution of mean attractiveness scores across the 25 focal books centered around the scale midpoint (M = 5.21 of 10, SD = 0.64). I selected 6 focal books (3 per framing condition) that represent this distribution for extrapolation (M = 5.04, SD = 0.64). The attractiveness scores of the selected books do not differ by framing condition (p = .941).

Procedure

In the main study, participants viewed the preselected focal books in random sequences, each accompanied by a preassigned recommended book. For each recommendation, participants indicated whether they would click on the recommended book, on a 10-point scale (1 = "Definitely not," 10 = "Definitely yes"). After they finished viewing all the recommendations, they selected the reasons that they had seen for recommendation: (1) user-based framing, (2) item-based framing, (3) both, or (4) neither. The study ended with demographic questions.

Results

I excluded eight participants who recalled neither the user-based nor the item-based framing. This exclusion is important, because it rules out the possibility that the framing effect shrinks due to participants' lack of attention to the recommendations associated with

less attractive focal books. The final data set includes 252 observations (6 books nested within 42 participants). I regressed participants' intention to click on the recommended book on three predictors: recommendation framing, the score of focal attractiveness as obtained from the pretest, and their interaction. The regression model allowed for a random intercept for each participant.

I found a significant interaction effect between framing and focal attractiveness (b = 0.65, SE = 0.31, $t(207)$ = 2.09, p = .039). Consistent with H3, the advantage of user-based framing decreased for less attractive focal books. To illustrate, when focal attractiveness was one standard deviation (SD) above the mean, user-based framing increased people's intention to click on the recommendation relative to item-based framing (b = .87, SE = .42, $t(207)$ = 2.07, p = .039). No framing effect emerged at the mean level of focal book attractiveness (b = 0.22, SE = 0.29, $t(207)$ = 0.77, p = .441) or at one SD below the mean (b = -0.43, SE = 0.43, $t(207)$ = -1.00, p = .317). For very unattractive books (1 out of 10), the model even predicted that user-based framing lowered click-through intentions compared with item-based framing (b = -4.22, SE = 2.16, $t(207)$ = -1.96, p = .052). Furthermore, in support of the theorizing, focal book attractiveness predicted intentions to click for the user-based framing (b = 1.09, SE = 0.44, $t(207)$ = 2.48, p = .014), but this trend was absent for item-based framing (b = -0.01, SE = 0.29, $t(207)$ = -0.02, p = .981).

When all cases were included, the moderation by focal attractiveness was in the same direction and marginally significant ($t(247)$ = 1.76, p = .080). See Figure A2 in the Appendix for the similar patterns with and without data exclusion.

Discussion

In support of H3, study 3 establishes focal product attractiveness as a boundary condition for the advantage of user-based framing over item-based framing. It renders insights into the framing effect in a setting where consumers are merely browsing products without explicit signals of their interest in the focal product. In study 4, I aim to replicate this finding using a different procedure; I also test whether presenting a salient cue of self–other dissimilarity makes user-based framing disadvantageous relative to item-based framing (H4).

Study 4

Cue of self–other dissimilarity

The majority of MTurk workers are at least 25 years of age (Ipeirotis 2010), so I use the age group "18–24 years" as a dissimilarity cue. That is, for this Study, a bar graph indicates other consumers' ages, under the title "Age of interested consumers," with three bars: "18–24," "25–55," and "above 55." I highlighted the "18–24" bar and informed participants that it represented the age of consumers who also viewed the recommended book. A pretest (N = 101; 62 females, M_{age} = 36.7 years, SD_{age} = 12.4) confirmed that most MTurk workers (89%) are older than 24 years, who also perceive themselves as more similar to other consumers in their age group than to people in the 18–24 age group (p < .001).

Participants and study design

I recruited 360 participants from MTurk, who are at least 25 years old (169 women; M_{age} = 37.6 years, SD_{age} = 1.19), and randomly assigned them to three conditions: user-based framing ("Consumers who viewed this item also viewed…"), item-based framing ("Similar to this item"), and user-based framing with the age group dissimilarity cue.

Procedure

The procedure is similar to that in study 3, with two differences. First, instead of presenting participants with preselected books, I allowed them to self-select three focal books to view from nine books, thereby simulating browsing behavior in online stores. Second, the attractiveness of focal books was rated by the participants, not based on the score from the pre-test, which captures the heterogeneity of ratings across individuals. At the end of the study, participants evaluated how attractive they found each focal book using a 10-point scale (1 = "Not at all," 10 = "Very attractive"; M = 6.94, SD = 1.96). The attractiveness was not influenced by the assigned conditions (p = .353; overall M = 6.94, SD = 1.96).

Results

As in study 3, I excluded participants (N = 133) who could not recall the framing they saw, leaving a data set with 680 observations (3 books nested within 227 participants). I took the same analysis approach as in study 3, with two dummy predictors: user-based

condition and dissimilarity cue condition, each of which could interact with the rating of the focal book's attractiveness. Because the dissimilarity (age group) cue did not interact with focal attractiveness (p = .845), and including this interaction term did not increase model fit (p = .364), I dropped it from the analysis to focus on the main effect of dissimilarity. Figure 3 plots the results.

Figure 3: Regression Results of study 4.
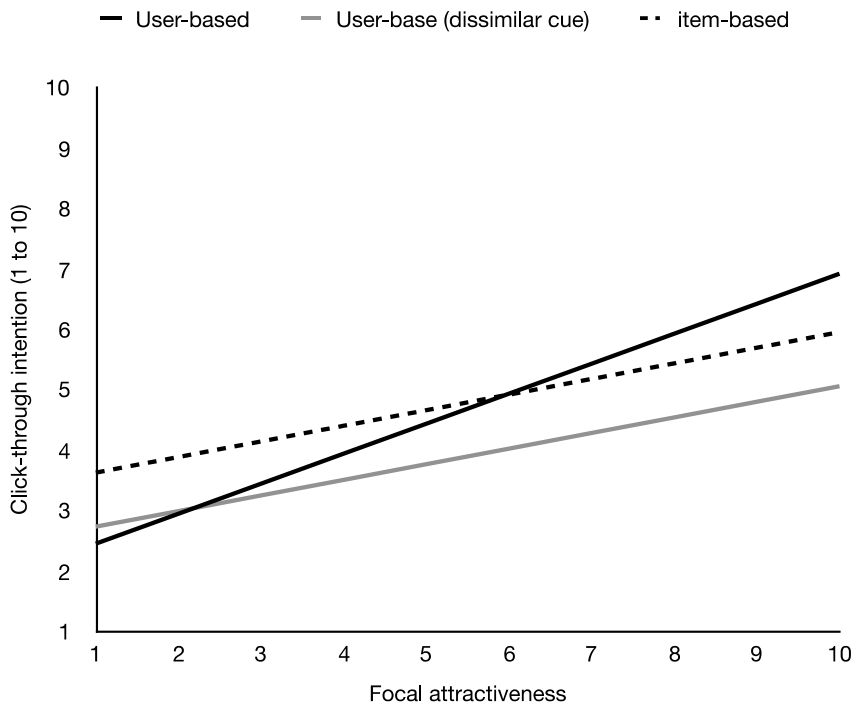
In line with the Study 3 results, I found a significant interaction of focal book attractiveness and recommendation framing when the dissimilarity cue was absent (b = 0.24, SE = 0.10, $t(451)$ = 2.32, p = .021). Specifically, user-based framing (versus item-based framing) increased participants' intention to click on the recommended book when focal attractiveness scored one SD above the mean (b = 0.70, SE = 0.36, $t(224)$ = 1.93, p = .055) but not when it scored at the mean (b = 0.24, SE = 0.30, $t(224)$ = 0.78, p = .434) or one SD below the mean (b = -0.23, SE = 0.36, $t(224)$ = -0.64, p = .526). For very unattractive books (1 out of 10), user-based framing even lowered click-through intentions

relative to item-based framing ($b = -1.18$, $SE = 0.68$, $t(224) = -1.74$, $p = .084$). In addition, focal attractiveness related more positively to click-through intentions in the user-based condition ($b = 0.50$, $SE = 0.12$, $t(451) = 6.51$, $p < .001$) than in the item-based condition ($b = 0.26$, $SE = 0.06$, $t(451) = 4.01$, $p < .001$). Critically, when the dissimilar cue was present, user-based framing (versus item-based framing) decreased intentions to click on recommended books ($b = -0.89$, $SE = 0.32$, $t(224) = -2.84$, $p = .005$).

When all cases were included, I replicated the reversal of the framing effect ($t(357) = -3.11$, $p = .002$). The moderation by focal attractiveness was in the same direction but not significant ($t(717) = 1.58$, $p = .113$). See Figure A3 in the Appendix for the similar patterns with and without data exclusion.

Discussion

Study 4 replicates the findings of Study 3 with a different procedure, strengthening the support for H2. Furthermore, consistent with H3, I find that the presence of a cue suggesting dissimilarity with other users makes user-based framing disadvantageous compared with item-based framing, regardless of the attractiveness of the focal books. To provide additional support for H4 and in line with prior research (Naylor, Lamberton, and West 2012), in Study 5 I used gender composition as a different cue of self-other dissimilarity. Moreover, I include both dissimilar (most other users are a different gender) and similar (most other users are the same gender) cue. In line with the theorizing and prior research (Naylor, Lamberton, and West 2010), I anticipate that cueing consumers with their similarity to other users will have an effect similar to user-based framing that lacks information about the identity of other users.

Study 5

Study design

Study 5 follows the design of study 2 (painting) but in the domain of books. I selected 57 books of various genres (e.g., comics, thrillers, philosophy) that were not available on the market when the study was conducted (i.e., "coming soon" category), so participants were unlikely to be familiar with them. I selected another 57 coming-soon books as recommendations and paired them with the focal books. Participants viewed the book covers, titles, author names, and genres, then marked books they would like to read by

clicking on a heart button. Then the recommendation button popped up, indicating either "Consumers who like this also like…" in the user-based condition or "Similar book to this" in the item-based condition. In both conditions, I told participants that the recommendation came from readers on Amazon. Moreover, participants had the chance to win a book that they marked as "would like to read."

In the user-based condition, next to the recommendation button, participants also saw the gender composition of people who liked the focal book. Of the 57 focal books, 21 were predominantly liked by males (95%–100%), and 21 were mainly liked by females (95%–100%), so 42 books offered a cue of self–other similarity, and 42 provided a cue of self–other dissimilarity (see Table 2). In addition, 15 neutral books were liked about equally by both genders (45%–55% males). These neutral books serve two purposes. First, their presence creates a more realistic book-shopping scenario, in which consumers encounter books that attract either gender and those that appeal to both genders. Second, the neutral books, combined with similar-cue and dissimilar-cue books, increase the power of the contrasts relative to the item-based condition (i.e., same 57 books compared across conditions). For the similar-cue books, I expect to replicate the moderating role of consumption experience from study 2. For the dissimilar-cue books, in line with study 4, I anticipate that user-based framing will decrease CTR.

Table 2: Design of study 5 (user-based framing)

	Similar cue	Dissimilar cue
Male participants	21 books liked by 95% to 100% men	21 books liked by 95% to 100% women
Female participants	21 books liked by 95% to 100% women	21 books liked by 95% to 100% men
Total	57 books (42 plus 15 neutral books)	57 books (42 plus 15 neutral books)

Participants and procedure

Three-hundred sixteen MTurk workers participated in the study (159 males, M_{age} = 35.51 years, SD_{age} = 1.61). After viewing the focal books, participants indicated their

experience with book shopping on the item, "How often do you visit books stores (online or offline) in general?" with the same scale from Study 2. Compared with participants in Study 2 (paintings), participants in Study 5 had more experience with books (median = 3 versus 2; significantly higher mean, $p < .001$). The study ended with a few demographic questions.

Figure 4: Regression results of Study 5.
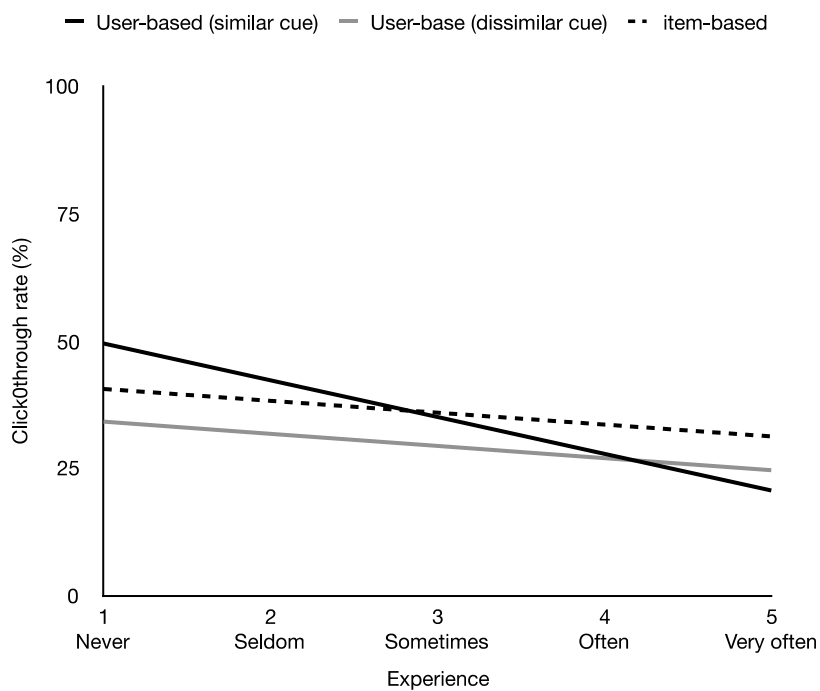
Results

I calculated separate CTRs for similar-cue books, dissimilar-cue books, and the item-based frame books. I then regressed the CTRs on two dummy predictors: similar-cue books and dissimilar-cue books, each of which could interact with consumption experience. Similar to study 4, the interaction between the dissimilar cue and consumption experience was insignificant ($p = .975$), and including the interaction term did not increase model fit ($p = .975$). I thus dropped the interaction (see Table A3 in the Appendix for the

full regression results). The results, as plotted in Figure 4, showed that for similar-cue books, as in study 2, there was a significant interaction between framing and experience (b = -4.88, SE = 1.63, $t(623)$ = -3.00, p = .003). Specifically, user-based framing was more advantageous for participants who never visited bookstores (b = 8.91, SE = 4.35, $t(623)$ = 2.05, p = .041), but this advantage decreased and even reversed as they gained more experience (seldom b = 4.03, SE = 2.98, $t(623)$ = 1.35, p = .176; sometimes b = -0.85, SE = 2.03, $t(623)$ = -0.42, p = .676; often b = -5.73, SE = 2.15, $t(623)$ = -2.66, p = .008; very often b = -1.61, SE = 3.24, $t(623)$ = -3.28, p = .001). In support of H4, user-based framing became disadvantageous, relative to item-based framing, for dissimilar-cue books (b = -6.55, SE = 1.94, $t(623)$ = -3.38, p < .001).

Discussion

Using the paradigm from study 2, study 5 conceptually strengthens support for H4. Cueing consumers to recognize self–other dissimilarity leads to a disadvantage of user-based framing relative to item-based framing. This study also generalizes the role of consumption experience to the domain of books.

General Discussion

Consumers frequently receive product recommendations from recommender systems, and companies often frame them as user-based (e.g., "People who like this also like…") or item-based (e.g., "Similar to this item"). I compare these two framings while keeping the actual recommendation constant (or randomized, as in the field studies) and thereby demonstrate the advantages of user-based framing over item-based framing in terms of recommendation CTR. In two field experiments with the mobile app WeChat (study 1a and 1b), I establish that recommending articles with user-based (versus item-based) framing increases recommendation CTR. Study 2 identifies consumption experience as an important boundary condition for the framing effect; studies 3 and 4 show that the effect shrinks and even reverses for unattractive focal products. Finally, studies 4 and 5 reveal that cueing consumers to their dissimilarity with other users makes user-based framing less effective than item-based framing. Table 3 summarizes the studies and the hypotheses they support. I took care to test the predictions using various product categories (articles, books, and paintings) and different paradigms mimicking real recommendation

70

practices to establish the generalizability and robustness of the effects. The results in turn offer several contributions to literature, practical suggestions for companies that use product recommendations in their marketing strategy, and directions for further research.

Table 3: Overview of studies.

Study	Data Source	Outcome Variable	Supported Hypotheses
Study 1a and 1b	Field experiment	Click-through rate of recommended books (0% to 100%)	H1
Study 2	Behavioral experiment	Click-through rate of recommended paintings (0% to 100%)	H2
Study 3	Scenario experiment	Intention to click on recommended books (1 to 10)	H3
Study 4	Scenario experiment	Intention to click on recommended books (1 to 10)	H3 and H4
Study 5	Behavioral experiment	Click-through rate of recommended books (0% to 100%)	H2 and H4

Theoretical implications

Prior investigations of recommender systems primarily focus on technical designs (e.g., Ansari, Essegaier, and Kohli 2000; Ariely, Lynch, and Aparicio 2004; Hennig-Thurau, Marchand, and Marx, 2012) or the consequences of their use (e.g., Bodapati 2008; Fleder and Hosanagar 2009; Pathak et al. 2010). Little research has explored the ideal ways for companies to communicate the basis of recommendations to their consumers. This research represents an initial attempt to fill this gap by comparing the effects of user-based and item-based framings on recommendation CTR. Simply changing the framing of recommendations can have an impact on this metric. I thus emphasize the importance of studying the effect of framing, in addition to the technical aspects of the underlying algorithms.

The current findings also advance understanding of consumers' interpretations of recommendations. As the follow-up survey in study 1a shows, consumers recognize product matching more readily than taste matching, regardless of the recommendation

framing. In two pilot studies (see Appendix), I also find that product matching is perceived as a more dominant recommendation strategy than taste matching. This primacy of product matching might result from the visual salience of products, relative to the latency of consumers: On a typical product webpage, consumers see products, not other consumers, and can directly compare the products but not themselves with others. The results of the survey show that the difference between the two framings is due to taste matching. By signaling that taste matching is part of the recommendation strategy, beyond product matching, user-based framing offers additional informational value for consumers that, presumably, mitigates their uncertainty about their satisfaction with the recommendation.

More broadly, this work contributes to advice-taking research (Iyengar, Van den Bulte, and Lee 2015; Müller-Trede et al. 2018; Sinan and Walker 2012). Prior studies focus on how consumers take advice from other users; I investigate consumers' tendency to follow recommendations generated by algorithms. Consistent with findings that indicate that consumers adopt others' choices (Morvinski, Amir, and Muller 2017) and opinions (e.g., online reviews; Chen and Xie 2008; Zhu and Zhang 2010), I demonstrate that mentioning others' preferences can encourage consumers to click on recommended products. However, a fundamental difference between following recommendations and adopting others' preferences is that the former depends on consumers' understanding of the "black box" of recommender systems, whereas the latter pertains to how consumers navigate the social world. Recommendations framed as user-based (versus item-based) might exert more influence on consumers by adding a social component to the recommender system.

Managerial implications

Companies heavily invest in recommender systems; global spending is estimated at $5.9 billion in 2019 (International Data Corporation 2019). This research suggests that it is not only the technical aspects of recommender systems that matter; the framing of recommendations exerts a notable influence as well. Companies might fail to maximize recommendation click-throughs if they rely only on item-based framing. Managers must develop effective recommender systems but also devote attention to how to frame the recommendations for consumers. Adapting the framing, while keeping the underlying

algorithm and the recommended product constant, comes with nearly zero cost, unlike developing and improving technical aspects of recommender systems.

The field studies suggest a general advantage of user-based framing over item-based framing in a setting where consumers' tastes are homogenous and they show deep interest in the focal item (e.g., they read the entire article). Studies 2 to 5 document situations in which this advantage can diminish or even reverse. These boundary conditions are particularly important for companies to consider when deciding on the framing that they want to utilize. First, consumers with less consumption experience are particularly susceptible to the impact of recommendation framing. Managers can identify these consumers by analyzing their past behavior and infer the degree to which they possess consumption experience in a specific domain. Consumers who seldom listen to classical music probably know little about this genre, for example, so they likely follow the lead of other classical music fans and exhibit high responsiveness to user-based framing.

Second, in situations in which consumers are merely browsing on a website and do not necessarily express interest in focal products (as it was the case in the paradigms of studies 3 and 4), utilizing a user-based framing is unlikely to be advantageous compared to an item-based framing. Conversely, a user-based framing is more advantageous than item-based framing for attractive products; it can trigger consumers to click the recommendation when they already have expressed some interest in the focal product, such as by reading an article or watching a video to the end. Managers can infer the attractiveness of focal products by tracking consumers' real-time behavior and thereby decide whether to prioritize user-based framing. Moreover, considering that user-based framing appears particularly beneficial for products that receive high ratings from prior consumers, if managers cannot easily infer a particular target consumer's attitude toward the focal product, they still can decide whether to prioritize user-based framing, depending on prior consumers' reactions to it.

Third, user-based framing is less effective than item-based framing when it is coupled with a cue suggesting that others (on whom the recommendation is based) are dissimilar to the recommendation recipient. This insight is critical for companies that present prior consumers' information to target consumers (e.g., "teens' choices"). If these selected

others differ from the target consumer in salient ways, the target consumer might avoid a recommendation framed as user-based. To maximize the value of user-based framing, managers either should not display any cues suggestive of differences or else should selectively emphasize other consumers who are similar to the target in some important aspect. If these displays of information cannot be adjusted, managers might compare the backgrounds of the target and others, then choose a user-based framing only if a match exists and item-based framing if not.

Fourth, consumers more readily recognize product matching than taste matching (as shown in the follow-up survey for study 1a). However, the advantages of user-based framing stem from consumers' awareness of the taste-matching effort and their recognition of successful taste matching. Therefore, it is important for companies to make user-based framings salient, such as by increasing the font size or underscoring the framing, if they intend to leverage its value to the fullest.

Caveats and calls for further research

I purposefully compare generic user-based and item-based framings, which are common in the marketplace, to generate externally valid and practically relevant insights. However, both framings can vary in their specificity. For example, user-based framing can refer to a specific group of users, such as friends (e.g., Spotify's "what friends are listening to"), which may alter how likely consumers are to perceive taste matching as successful. A generic user-based framing is unlikely to prompt consumers to question their similarity with ambiguous other users, but referring to specific friends could more easily trigger perceptions of dissimilarity. Typically, consumers know their friends' tastes and therefore recognize fine-grained differences in them. In that sense, referring to friends' preferences might backfire for user-based framing, making it less effective than item-based framing. I encourage continued research into this practically relevant issue.

Similarly, companies might specify standards for item categorization. Instead of merely mentioning that the recommended item is similar to a focal item or that the two fall in a rather broad category (e.g., romantic novels), companies might emphasize books by the same author or movies by the same director. Noting the primacy of product matching as the perceived recommendation strategy, I speculate that the width of the category exerts

little influence on the difference between user-based and item-based framing. However, it is possible that item categorization variations could affect certain consumers; for example, those with greater consumption experience within a product category might find item-based framing more attractive if the item categorization is narrower, because they are motivated to deepen their knowledge of specific categories (Clarkson, Janiszewski, and Cinelli 2013).

Alternatively, recommendation framing might be analyzed along dimensions other than an emphasis on different inputs (i.e., users or items), such as whether it refers to the target consumer's own past behavior as a basis for recommendation. Spotify uses "Because you have listened to X" in parallel with a more generic "Similar to X" to explain its recommendations. Does explicitly referring to consumers' own tastes make a difference? On the one hand, personalized explanations ("you" and "your" behavior) might cause consumers to perceive greater effort by the recommender system and the recommendation as more self-relevant. On the other hand, personalization could raise consumers' awareness that their private information has been collected and prompt reactance to the recommendations. Additional research could compare different recommendation framings along multiple dimensions to achieve a fuller understanding of their roles.

Although I explore three theoretically derived, practically relevant moderators, a variety of factors could shift the perceived success of taste matching and thus moderate the framing effect. According to social influence literature, for example, consumers tend to perceive more self–other dissimilarity as their distance grows (Meyners et al. 2017). Their perceptions of taste-matching success thus might depend on their geographical distance. Another pertinent factor is consumers' perception of the size of the group of other users (Argo, Dahl, and Manchanda 2005), as defined by the type of product. Consumers interested in a niche product may infer a small group of interested other users; those considering a mainstream product likely presume a large group. Larger groups can be more influential but also appear more heterogeneous in their tastes (Latane 1981). Studies of such influences could deepen understanding of framing effects across communities.

I suggest user-based framing is advantageous compared to item-based framing because it signals taste matching and provide support for this theorizing in product domains in which taste is an important decision criterion (articles, paintings, and books). I speculate that for products primarily differentiated by quality (for instance, utilitarian products such as laptops), consumers' reaction to recommendations could be less sensitive to their perception of taste matching; in such instances, the informational value of taste matching is likely to diminish. Future research could examine if the advantage of user-based framing relative to item-based framing depends on whether taste or quality is the more salient decision criterion for a particular product.

Importantly, the more consumers are familiar with the digital world, the more experienced they are with recommender systems and might develop their own understanding of how these systems work. For instance, ethnographic work on recommendations shows that experienced consumers tend to game with the recommender system to generate desired recommendations (Devendorf and Goodman 2014). This suggests that experienced consumers might interact with recommender systems more rationally and might deliberately choose to click or not to click on recommendations with the purpose to improve the quality of future recommendations. For instance, consumers might resist a recommendation related to the opposite gender's taste not only because they perceive a mismatch with their own taste, but also to avoid misidentification by the recommender system and to prevent any future recommendations associated with the other gender. The implication is that consumers who are more experienced with recommender systems could be more likely to scrutinize taste matching efforts. I see this as a fruitful avenue for future research.

As a concluding remark, in a blog post, Netflix has acknowledged that it provides explanations for why it has recommended a movie or show in order to gain consumers' trust (Amatriain and Basilico 2012). The current research reiterates this notion, by revealing that when companies explain a recommendation to their consumers, the decision of which framing to use, user-based or item-based, is crucial in terms of its impact on recommendation click-throughs.

76

Appendix

Study 1a: recommended versus non-recommended articles

To ensure the recommendation is effective, I compared the percentage increase in reads of recommended and non-recommended articles (increase = current read – initial read). The initial reads of all articles were recorded before the experiment started, and then the current reads were recorded two weeks after the recommendation. The current reads of non-recommended articles were recorded after the experiment ended.

In Figure A1, the distribution of the percentage increase is extremely positively skewed in each condition. Therefore, I conducted a Kruskall-Wallis test to compare the distribution of the increase across conditions; it differed significantly across conditions (H(2) = 23.59, $p < .001$). Pairwise comparisons revealed that non-recommended articles had lower mean rank (5.67) than articles recommended with the user-based framing (43.42; $Z = -4.83$, $p < .001$) and those with the item-based framing (37.39; $Z = -4.06$, $p < .001$). A comparison of the medians indicates the same pattern of results ($p = .001$).

Figure A1: Distribution of Percentage Increases.
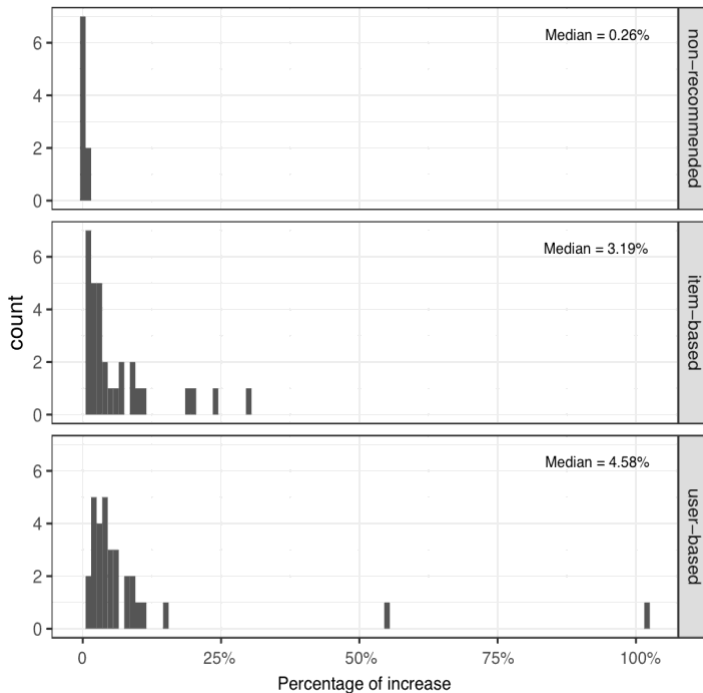
Measurements of taste matching [product matching]

1 = strongly disagree, 6 = strongly agree

The recommendation is based on readers who have similar preferences with me [articles
that are similar to what I have read].

The recommendation is based on the categorization of people [articles].

The recommendation is based on readers [content].

The recommendation system takes into account my preferences [the type of articles].

Full regression results

Table A1: Regression results of studies 1a and 1b.

	Study 1a	Study 1b
Intercept (with random intercepts within days)	0.51*** (0.07)	1.06*** (0.12)
Framing (1 = user-based, 0 = item-based)	0.22*** (0.66)	0.19* (0.08)
Publication bias (1 = published in 2018, 0 = published before 2018)	-	-0.49*** (0.13)

Notes: Coefficients (0 to 100%) with standard errors in parentheses. ***$p < .001$, **$p <$
.01, *$p < .05$

Table A2: Regression results of studies 2 and 3.

Study 2: Click-through rate (0 to 100%)				
	b	*SE*	*t*	*p*
Intercept	46.63	0.97	25.25	<.001
Framing (1 = user-based, 0 = item-based)	3.35	1.37	2.45	.015
Experience (mean-centered)	-4.46	0.69	-6.49	<.001
Framing by experience	-2.23	0.97	-2.30	.022
Study 3: Click-through intention (1 to 10)				
	b	*SE*	*t*	*p*
Intercept (with random intercepts within individuals)	5.44	0.27	20.43	<.001
Framing (1 = user-based, 0 = item-based)	0.22	0.29	0.77	.441
Focal attractiveness (mean-centered)	-0.01	0.29	-0.02	.981
Framing by focal attractiveness	1.10	.53	2.08	.039

Table A3: Results of studies 4 and 5.

Study 4: Click-through intention (1 to 10)		
	Model 1 (in the main text)	Model 2
Intercept (with random intercepts within individuals)	5.16***(.21)	5.16***(.22)
User-based (no cue)	0.24(.30)	0.24(.30)
User-based (dissimilar cue)	-0.89**(.32)	-0.90**(.32)
Focal Attractiveness (mean-centered)	0.26***(.06)	0.25**(.08)
User-based (no cue) by Focal Attractiveness	0.24*(.10)	0.25*(.11)
User-based (dissimilar cue) by focal attractiveness	-	0.03(.13)
Study 5: Click-through rate (0 to 100%)		
	Model 1 (in the main text)	Model 2
Intercept	35.03***(1.39)	35.03***(1.39)
User-based (similar cue)	-2.75(1.92)	-2.75(1.92)
User-based (dissimilar cue)	-6.55**(1.94)	-6.55**(1.94)
Experience (mean-centered)	-2.35*(.98)	-2.33*(1.14)
User-based (similar cue) by Experience	-4.88**(1.63)	-4.88**(1.63)
User-based (dissimilar cue) by Experience	-	-0.05(1.69)

Notes: Coefficients (1 to 7 in Study 4 and 0 to 100% in Study 5) with standard errors in the parentheses. Model 1 is also reported in the main text. ***$p < .001$, **$p < .01$, *$p < .05$.

Studies 3 and 4: plotted multilevel regression results with and without data exclusion

Figure A2: Results of study 3 without (left) and with (right) data exclusions.
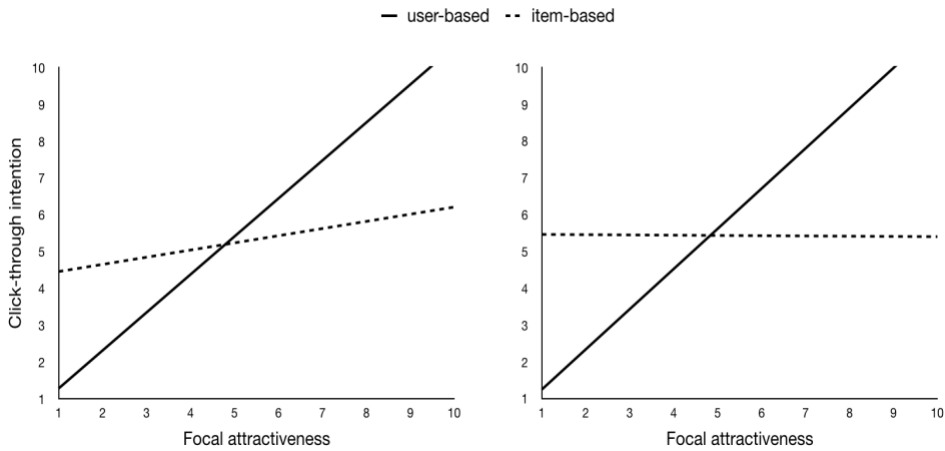
Figure A3: Results of Study 4 without (left) and with (right) data exclusions.
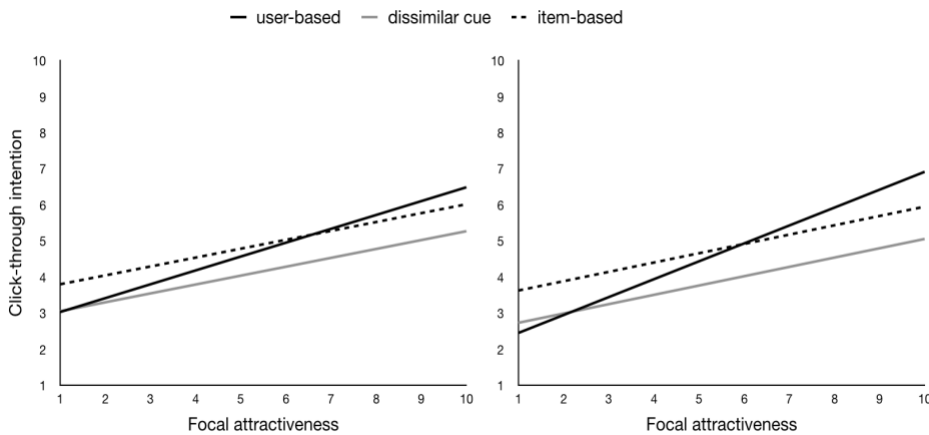
Pilot studies

Participants were randomly assigned to read either the user-based framing or the item-based framing after watching a focal video or listening to a focal song. After reading the framing, they responded to a question, "The product was recommended to me based on…," by selecting from three options: taste matching, product matching, or random. The

cells in Table A4 below display the numbers of participants selecting each option. Product matching remains dominant for user-based framing; most participants selected it, even when they could only choose one option. Yet user-based framing increases awareness of taste matching (video: $\chi^2(N = 117) = 11.20$, $p = .004$; music: $\chi^2(N = 178) = 7.36$, $p = .025$). I also asked participants about their perceptions of the overlap between the focal and recommended products. In the bottom row of Table A4, perceived overlap does not differ by the framing condition. Additional analyses indicate no reliable relationship between perceived overlap and desire to watch the recommended video (Pearson's $r(117) = .28$, $p = .033$) or listen to the recommended song ($r(179) = .06$, $p = .444$).

Table A4: Summary of the results of the pilot studies.

	Video recommendation (MTurk participants)	
	User-based framing "Based on what people who like … also like, we recommend the following video to you"	Item-based framing "Based on the content of this video/genre of…, we recommend the following video to you"
Based on product matching	36	53
Based on taste matching	13	2
Random	7	6
Perceived product overlap (1–7) by framing	M_{user} = 4.85, M_{item} = 5.16; $t(115)$ = -0.83, p = .406	
	Song recommendation (Prolific participants)	
	"People who like … also like"	"Music of the same style as …"
Based on product matching	52	49
Based on taste matching	27	13
Random	14	24
Perceived product overlap (0–100) by framing	M_{user} = 27.88; M_{item} = 32.85; $t(176)$ = 1.29, p = .200	

Chapter 4

When Virtues are Lesser Vices

Chapter 4 is organized as follows. I start by analyzing the conceptual difference between the vice–virtue and vice–lesser vice choice settings and identify the factors that would or would not influence the choice between the vice and the lesser vice. I also emphasize the importance of studying consumed quantities in the vice–lesser vice choice setting. The proposed model was tested in two well-controlled experiments.

Theoretical Background

Virtues, vices, and lesser vices

Past research on self-control of food consumption has primarily focused on the choice between the vice and the virtue. Vices are unhealthy and hedonic while virtues are healthy and utilitarian (for a review, see Vosgerau, Scopelleti, and Huh 2019). Lesser vices that remain in the unhealthy category but are less unhealthy than the typical vice, such as sugar-free donuts versus regular donuts, are much less studied and conceptually indifferentiable from virtues (Werterbroch 1998). Per definition, a vice scores higher on short-term gratification but lower on long-term health benefits relative to a virtue (Ainslie 1975; Myrseth and Fishbach 2009). Whether the virtue falls within the healthy or the unhealthy category should not make any difference. In this research, however, I put forward that lesser vices are uniquely different from healthy virtues. Throughout the paper, 'vice' exclusively refers to the unhealthy, hedonic food and 'lesser vice' to the food that is similarly indulgent but incurs lower costs on health.

Drivers of choice in vice–virtue versus vice–lesser vice contexts

Classic self-control conflicts of the vice–virtue type are characterized by a palpable conflict between short-term and long-term consequences and self-control success lies in maximizing one's long-term outcomes (Wertenbroch 1998). One crucial reason why people struggle doing so is that they are greatly present-focused. As people hyperbolically discount future gains, they often favor the option that provides the greatest immediate pay-off (Loewenstein and Elster 1992). Yet, when substituting a vice with a lesser vice (e.g., a sugar-free donut), the perceived conflict between short-term gratification and long-term health outcomes should be much weaker or even absent. Because companies try to make lesser-vice products taste as similar to the original vice products as possible to keep consumers' interest (Nielson 2014), consumers still derive pleasure from consuming the

lesser vice. That is, regardless of whether consumers choose the vice or the lesser vice, they can have their short-term craving satisfied[1]. Goal systems theory suggests that the fulfillment of the short-term goal should trigger consumers to pursue the competing long-term goal (Förster, Liberman, and Higgins 2005; Shah, Kruglanski, and Friedman 2003). Following this reasoning, inasmuch as consumers see both the vice and the lesser vice as indulgent, they would focus on the health outcomes of their choices and the choice of lesser vices would be determined by the strength of the health goal.

This distinction is important, as previous work consistently showed that factors that draw attention to the immediate, short-term outcomes tend to weaken self-control, whereas those that direct attention to the future, long-term outcomes tend to enhance self-control (Fujita et al. 2007; Loewenstein 1996). For example, encouraging people to make their food choices well in advance (i.e., advance ordering) separates the moment of choice from the moment of consumption, thereby reducing the potential impact of immediate gratification while increasing the saliency of long-term goals. As a result, it is a highly effective intervention to increase the choice share of healthy over unhealthy food options (Miller et al. 2017; Read and van Leeuwen 1998; VanEpps et al. 2017). As outlined above, contrary to the vice–virtue trade-off, the vice–lesser vice setting is much less characterized by the consideration of immediate gratification. Therefore, I expect that an intervention such as advance ordering, of which the effectiveness is predicated on changing the relative saliency of the short-term versus long-term outcomes, is unlikely to influence the choice between the vice and the lesser vice.

If consumers are choosing between the vice and the lesser vice according to their health consequences, what would differentiate the two in terms of healthiness? I posit that the crucial characteristic that distinguishes lesser vices from vices is their lower calorie density. For healthy and unhealthy food, their primary difference lies in the nutritional value such as the supply of fibers, protein, and vitamins. Calorie density is not a good

---

[1] Although lesser vices are likely viewed as less tasty than regular vices, the argument here is that choosing either would bring immediate pleasure. Therefore, even when selecting lesser vices does not maximize the immediate pleasure, I suggest that consumers would primarily focus on whether a food option brings immediate pleasure rather than how much the pleasure is. Further support for this argument (from an ancillary study) will be discussed in General Discussion.

proxy of healthiness here, as healthy products could even contain more calories than unhealthy ones (e.g., 100 grams of almonds contain more calories than 100 grams of crisps). Hence, in vice–virtue choice contexts, the calories differ qualitatively, rendering their total amount of little importance. Conversely, in the vice–lesser vice choice context, calories differ mostly quantitatively, and calorie density becomes of unique importance. Prior research shows that individuals differ in their sensitivity to calorie density, which is captured in the tendency of restrained eating (van Strien et al. 1986; Elfhag, Tynelius, and Rasmussen 2007). Thus, I predict that consumers who are more restrained in their eating behavior would be more inclined to substitute the vice with a lesser vice.

To summarize, I predict that eliminating the present bias by decoupling choice and consumption (advance versus immediate ordering) would not bolster the preference for the lesser vice but that the sensitivity to calorie intake would. Next, I turn to another equally, if not more, important behavioral consequence in the vice–lesser vice setting: consumed quantities.

Post-choice consumption: effects of choices, timing of choice, and restrained eating

Typically, self-control success (versus failure) is operationalized as consumers' propensity to choose the healthy (versus unhealthy) food. The qualitative difference in nutritional value between healthy and unhealthy food renders the quantity of calorie intake of little relevance relative to the exact choice. Therefore, in the vice–virtue choice setting, choice is often viewed as a better proxy for self-control success compared to consumed quantities. In contrast, when an unhealthy food is substitutable by a lesser vice, the total amount of consumed calories is a crucial proxy for the overall success in self-control; if the lesser-vice choosers end up consuming more calories, self-control still fails.

What is the impact of the choice between the vice and the lesser vice on the overall calorie intake? To answer this question, we must consider the effect of choice on consumption volume (e.g., grams). One possibility is that choosing the lesser vice would decrease consumed grams and thus reduce consumed calories. This is because consumers might behave in consistency with their choices, monitoring their eating process more closely if they select the lesser vice. Past research hints at this possibility by showing that choice in general triggers thoughts about the consumption experience (Redden, Haw, and

Chen 2017). Yet, there is no explicit evidence for the choice-consumption consistency in self-regulation. Another possibility is that choosing the lesser vice would increase consumed grams and even consumed calories. There is evidence that serving consumers lighter food can lower their anticipated guilt and subsequently increase their consumed quantities (e.g., low-fat M&M's, Wansink and Chandon 2006; Koenigstorfer and Baumgartner 2015). Whether this finding is generalizable to consumers' proactive choices, however, remains an open question. The last possibility is that choosing the lesser vice would have no impact on the number of grams consumed but would still lower the calorie intake by grace of the lesser vice's lower caloric density. The rationale is that while some deliberation might go into the choice between the vice (e.g., regular chips) and the lesser vice (e.g., light chips), the eating behavior itself is a much more mindless, automatic process (Galak and Loewenstein 2013). This prediction is also consistent with recent research showing that consumers under-weigh the influence of consumed quantities on healthiness, suggesting that they may not perceive the health threat from larger consumed quantities as much as that from the food of choice (Liu et al. 2019). These competing predictions are tested in this chapter.

Apart from the effect of choices on consumption, would there be any impact from the timing of choice and restrained eating on consumed quantities? Given my prediction of a null effect of advance ordering on choices, I do not expect it to influence consumption. As for restrained eating, I predict that it would lower the calorie intake by encouraging the choice of lesser vices and reducing cravings for unhealthy ingredients (Burton, Smit, and Lightowler 2007). Note that it is also possible to examine the *direct* effects of the timing of choice and restrained eating on consumption, which I do not draw any prediction of given the lack of literature. However, it would be reasonable to expect the absence of direct effects if consumers indeed pay little attention to their eating processes after making a choice. The conceptual model is presented in Figure 1, including the paths with and without specific predictions.

General Method

I conducted two studies to examine the predictions and questions posed in Figure 1. The studies involve snacking while watching an entertaining video, a common experience
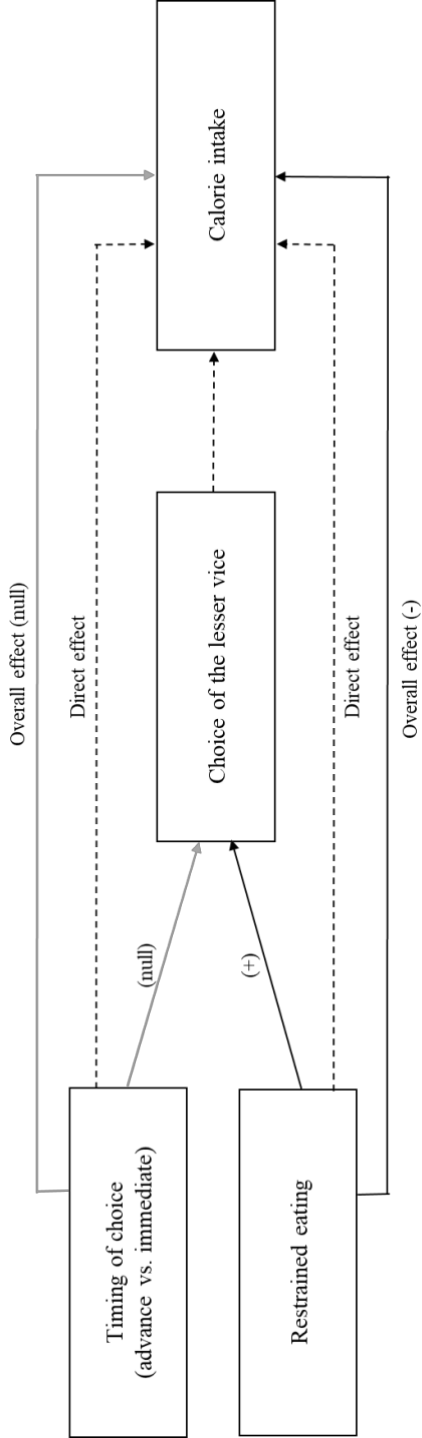
among consumers (Coelho do Vale, Pieters, and Zeelenberg 2008). In both studies, I manipulated the timing of choice (advance versus immediate before consumption) between participants and measured the individual difference in restrained eating after the main study. The studies were conducted with standardized procedures in a laboratory, eliminating potential confounds such as social influence and minimizing noises in the measurements (see study procedure for details). To limit the variance in choices and consumed quantities and to make sure participants care about their health, I conducted the studies with females only (Argo and White 2012). I aimed to reach at least 100 participants for each timing condition.

Two methodological challenges have to be addressed for the current research. One is how to test one path while controlling for the others in the model (Figure 1). I tackle this with a structural equation modeling approach and estimating all the paths simultaneously (Muthén 2011). Second, significance testing alone would not suffice to support null effects. As a complement, I calculated the Bayes factor (BF_{01}) for all insignificant effects based on the model fit ($BF_{01} = e^{(BIC_{H_1} - BIC_{H_0})/2}$, Jarosz and Wiley 2014). BF_{01} is the ratio of the likelihood of a null hypothesis given the data to the likelihood of the alternative hypothesis. There is evidence for the null hypothesis when BF_{01} is above one and the evidence is stronger as the value of BF_{01} increases. Based on Jeffreys' (1961) criteria, BF_{01} above 10 indicates strong evidence for the null hypothesis and BF_{01} above 100 is decisive.

Study 1

In study 1, I operationalized the vice as regular potato chips and the lesser vice as light chips. The choice set includes two popular flavors of chips (classical/paprika), each with a regular version (5.31/5.41kcal per gram; 181g per pack) and a light version (4.8/4.9kcal per gram, about 9% lower in calorie density than the regular versions; 176g per pack), all of the same brand. A pretest with a separate group of participants from the same population confirms that light chips are viewed as the 'lesser vice': they are rated as unhealthy in general (M = 2.90 on a 7-point scale with 1 = not healthy at all and 7 = extremely healthy; $p < .001$ compared to the midpoint of the scale) but less unhealthy compared to regular chips (M = 1.95; difference from light chips: $t(42) = 5.39$, $p < .001$).

Figure 1: The conceptual model.

Overall effect (null)

Direct effect

Calorie intake

Overall effect (-)

Direct effect

Choice of the lesser vice

Timing of choice
(advance vs. immediate)

(null)

Restrained eating

(+)

Notes: Gray lines represent predicted null effects, black lines represent predicted positive and negative effects, and the dashed lines are the open questions.

Participants and procedure

    Two-hundred and twenty female undergraduates from Rotterdam School of Management (Mean of age = 20.16, SD = 1.33) participated in the study in exchange of credits. I informed participants that they would snack on potato chips while watching a video in the lab. The purpose of the study was described as to provide feedback for a snack company.

    The study consisted of two parts, a morning survey and a late afternoon lab session, always administered on the same day (see Figure 2). Participants received the morning survey at around 9:00AM, which remained open until 11:30AM. In the survey, I asked participants in the advance (versus immediate) condition to choose the chips (versus the video) that they would like to eat (versus watch) in the afternoon lab session. Participants then reported when and what they ate last time. None of them had lunch before filling out the morning survey.

    The afternoon lab sessions were scheduled between 3:00PM and 5:30PM, when the afternoon crunch is likely to kick in. Each participant was picked up individually and seated in a separate cubicle in front of a computer. All participants first re-read the cover story about the snack company. Afterwards, I asked participants in the advance (versus immediate) condition to choose a video to watch (versus the chips to eat). Thereby, participants in both conditions made a choice immediately before consumption, which excludes any effect from the mere action of choosing on subsequent self-control (Baumeister et al. 1999). Participants then received a full pack of the chips of their choice. They could eat as much as they like while watching the video. While the video was presented as four different titles at the choice stage, in reality all participants watched the same video that lasted 650 seconds. This ensures that consumed quantities are not affected by the content of the video or the duration of eating. When the video ended, the experimenter took away the remaining chips and let participants complete a questionnaire. The experimenter weighed the remaining chips with a kitchen scale at one-gram intervals.

Figure 2: Study procedure.
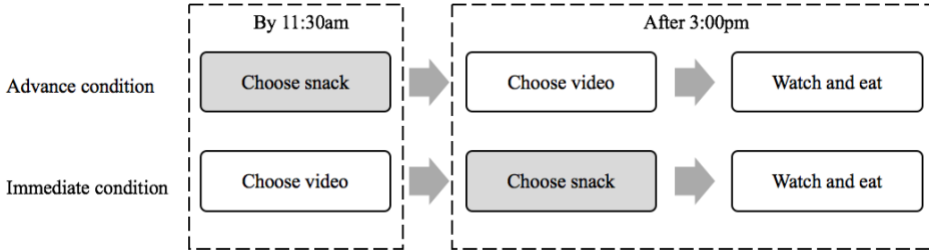
The post-consumption questionnaire started with filler items on how participants thought about the chips they ate and the video they watched, followed by questions on when participants ate last time and whether they had chips before coming to the lab. None of the measures differed between the timing condition ($ps > .5$). More importantly, participants completed the measurement of restrained eating on a five-point scale (e.g., "Do you watch exactly what you eat?", Cronbach's $\alpha = .90$; a 10-item subscale of Dutch Eating Behavior Questionnaire by van Strien et al. 1986). In the end, participants were asked about their age and debriefed.

Results

I specified a model in Mplus as shown in Figure 1. Calorie intake was the product of the consumed grams and the calorie content of the selected pack of chips (M = 199.28 kcal, SD = 98.40). The model was estimated with the maximum-likelihood approach and a probit link for the choice of chips (0 = regular, 1 = light). For purposes of efficiency, I summarized the path estimates and standard errors in Figure 3 and focus on the statistical significance below. For null effects, I reported both p values and Bayes factors (BF_{01}). Table 1 presents the descriptives.

Advance ordering. As expected, the timing of choice did not predict the choice of chips ($p = .441$, $BF_{01} = 11.02$). In both timing conditions, the share of light choosers was about the same as the share of regular choosers (57 out of 111, or 51% in the advance condition and 50 out of 109, or 46% in the immediate condition). Also consistent with the prediction, the timing of choice did not have an overall effect on the calorie intake ($p = .916$, $BF_{01} = 128.44$; no direct effect either: $p = .487$, $BF_{01} = 11.66$).

Restrained eating. In support of the prediction, more restrained eaters were more likely to choose light chips ($t(213) = 2.54, p = .011$). In turn, choosing light chips led to lower calorie intake ($t(213) = -4.23, p < .001$). Overall, restrained eaters consumed less calories ($t(213) = -2.19, p = .029$) and 67% of the total variance was accounted for by the indirect effect via the choice of light chips. There was no direct effect of restrained eating on calorie intake ($p = .295,$ $BF_{01} = 9.92$).

Consumed grams. To examine the consumption process more closely, I replaced calorie intake with consumed grams and re-estimated the model. I see that choice of light chips reduced consumed grams ($t(213) = -2.72, p = .006$). The timing of choice and restrained eating had no direct effect on consumed grams ($p = .437$, $BF_{01} = 10.97$ and $p = .382$, $BF_{01} = 10.10$ respectively).

Discussion

Study 1 confirms the predictions that restrained eating encourages the choice of the lesser vice and, through choice, lowers overall calorie intake, while the timing of choice has no effect. According to Jeffreys (1961), there is strong evidence for the null effect of advance ordering on the choice of the lesser vice (BF_{01} between 10 and 20) and decisive evidence for its null effect on calorie intake ($BF_{01} > 100$). I also find that neither advance ordering nor restrained eating has any direct impact on consumed calories.

In addition, study 1 shows that the choice of the lesser vice reduces both consumed calories and consumed grams, suggesting behavioral consistency among the lesser-vice choosers. Nevertheless, caution is warranted here. Participants in this study received the chips in their original packaging. As a consequence, the 'light' label remained salient in the consumption stage and might serve as a reminder for healthy eating. Indeed, past studies frequently utilize such visual reminders to prompt consumers to mind their eating processes (e.g., posters highlighting calorie intake by Haws and Winterich 2013; flyers featuring a healthy recipe by Papies et al. 2014). In study 2, I remove the difference in packaging between the vice and the lesser vice.

Another explanation for the lower consumption among lesser-vice choosers is based on the means-to-goal priming, which suggests that choosing the lesser vice, as a mean to meet the health goal, could temporarily make the health goal more salient (Shah and

Kruglanski, 2003; van Osselaer & Janiszewski, 2012). To test this, I manipulated the accessibility of the health goal immediately before consumption in study 2. If the effect of choice on consumption that I observed in study 1 is due to the means-to-goal priming, priming consumers with the health goal should 1) lower the consumed quantities and 2) eliminate the effect of choice on consumption. However, if the effect is due to the salience of packaging, the goal-priming procedure should exert no effect.

Study 2

Participants and procedure

Three-hundred and forty-nine female undergraduates from the same participant pool of study 1 (Mean of age = 20.09, SD = 1.13) participated in the study in exchange of credits. Instead of potato chips, participants were asked to choose between two bowls of M&M's of the same flavor (chocolate), one presented as the regular version (4.80 kcal per gram) and the other as the light version (4.58 kcal per gram, about 5% lower in calorie density than the regular) based on their incidental caloric difference. The bowls looked exactly the same and each bowl had 213 grams of M&M's. The procedure followed study 1. The pretest shows that both light and regular M&M's were viewed as unhealthy (M = 1.75 and 1.27 out 7 with the lower value indicating more unhealthiness; $ps < .001$ compared to the midpoint) but the light version appeared less unhealthy ($t(39) = 4.69$, $p < .001$).

Before participants in the main study received the M&M's of their choice, they went through a scrambled-sentence task. Half of the participants were randomly assigned to scramble 10 sentences with health-related words (e.g., "She has a healthy baby", "She is on a diet"), while the others assigned to sentences with neutral words (e.g., "She has a beautiful baby", "She is on a flight"). The priming procedure did not affect or moderate the effect of other factors on consumption ($ps > .7$), suggesting that the post-choice saliency of the health goal does not explain the effect of choice on consumption that was observed in study 1 (also see McCarthy et al. 2018 for the difficulty to detect priming effects with the paradigm). I collapsed the data across the priming conditions and followed the same analyses as in study 1.

Results

I estimated the same model as in study 1 and summarized the results in Figure 3. Table 1 presents the descriptives.

Advance ordering. Consistent with study 1, the timing of choice did not influence the choice of M&M's ($p = .577$, $BF_{01} = 15.93$; 66 out of 173, or 38% versus 77 out 176, or 44% chose the light over regular in the advance versus immediate condition). In addition, it did not have an overall effect ($p = .805$, $BF_{01} = 296.78$) or a direct effect on calorie intake ($p = .907$, $BF_{01} = 18.65$).

Restrained eating. Restrained eating increased the preference for light M&M's ($t(342) = 8.45$, $p < .001$). In turn, choosing light M&M's resulted in slightly lower calorie intake ($t(342) = -1.52$, $p = .141$; although the effect was not significant, the p value was close to 10% and much lower than the null effects; $BF_{01} = 0.33$ rendering substantial evidence for the alternative hypothesis). Overall, restrained eaters consumed fewer calories ($t(342) = -2.04$, $p = .041$) and 77% of the total variance was accounted for by the indirect effect via the choice of light M&M's. There was no direct effect of restrained eating on calorie intake ($p = .540$, $BF_{01} = 15.71$).

Consumed grams. Different from study 1, choosing the lesser vice (light M&M's) did not lower consumed grams ($p = .360$, $BF_{01} = 12.29$). In line with study 1, the timing of choice and restrained eating had no direct effect on consumed grams ($p = .914$, $BF_{01} = 18.57$ and $p = .542$, $BF_{01} = 15.42$ respectively).

Table 1: Descriptives of choice and consumption in studies 1 and 2.

Study 1		
	Light choosers	Regular choosers
Advance condition (N = 111)	n = 57 Mean (SD) of calorie = 186.34 (89.65) Mean (SD) of grams = 38.39 (18.46)	n = 54 Mean (SD) of calorie = 218.57 (100.35) Mean (SD) of grams = 40.72 (18.74)
Immediate condition (N = 109)	n = 50 Mean (SD) of calorie = 152.58 (74.31) Mean (SD) of grams = 31.46 (15.25)	n = 59 Mean (SD) of calorie = 233.69 (100.48) Mean (SD) of grams = 43.61 (19.88)
Retrained eating (M = 2.87 out of 5; SD = 0.89) did not differ by the timing of order (p = .741).		
Study 2		
	Light choosers	Regular choosers
Advance condition (N = 173)	n = 66 M (SD) of consumed calorie = 189.65 (120.53) M (SD) of consumed grams = 41.41 (26.32)	n = 107 M (SD) of consumed calorie = 212.10 (131.53) M (SD) of consumed grams = 44.19 (27.40)
Immediate condition (N = 176)	n = 77 M (SD) of consumed calorie = 181.06 (137.24) M (SD) of consumed grams = 39.53 (29.97)	n = 99 M (SD) of consumed calorie = 215.18 (129.69) M (SD) of consumed grams = 44.83 (27.02)
Restrained eating (M = 2.75 out of 5, SD = 0.93) did not differ by the timing of order (p = .259).		

Figure 3: Path estimates with standard errors in parentheses in study 1 (upper panel) and study 2 (lower panel).

Study 1 | N = 220

Timing of choice
(advance vs. immediate)

Overall: -3.57(4.72)

Direct: 4.40(6.34)

0.07(0.09)

Restrained eating

0.25(0.10)

Choice of the lesser vice

Direct: -7.61(7.26)

Overall: -19.76(9.03)

-54.36(12.86)

Calorie intake

Study 2 | N = 349

Timing of choice
(advance vs. immediate)

Overall: 1.78(7.18)

Direct: -0.73(6.95)

-0.04(0.07)

Restrained eating

0.76(0.09)

Choice of the lesser vice

Direct: -4.72(8.51)

Overall: -23.05(11.29)

-23.53(16.00)

Calorie intake

Notes: Gray lines are null effects.

96

Discussion

Study 2 replicates the main findings of study 1 in that restrained eating strengthens the preference for the lesser vice and lowers the overall calorie intake. In contrast, the timing of choice has no influence. Also consistent with study 1, I do not find any direct effect of the independent factors on consumed quantities. Importantly, keeping the packaging constant, I find a null effect of choice on consumed grams. This suggests that the effect observed in study 1 is likely driven by the saliency of the packaging of the lesser vice during consumption.

Taken together, studies 1 and 2 provide strong to decisive evidence (Jeffreys 1961) for the proposed null effect of advance ordering on the preference for lesser vices. To exclude the possibility that the null effects are due to methodological flaws, I evaluate the studies against the criteria established by Frick (1995). As shown in Table 2, the study features and results satisfy the relevant criteria, lending further credence to the null effects. The only unlisted criterion is to have many trials within each participant, which I deem not feasible for the manipulation of timing in the lab setting.

Table 2: Frick's (1995) criteria for validating the null effect and the support.

Criteria	Support
Large number of participants	109 to 176 participants per timing condition
Major sources of variance are controlled for	Well-controlled laboratory setting; fixed duration of eating; same-gender participants
Large and effective manipulation	Unambiguous manipulation of the timing of choice (for similar manipulation, see vanEpps et al. 2016)
Sensitive measurements	Choice and calorie intake are sensitive to the influence of restrained eating.
No floor or ceiling effect	The choice share of lesser-vice products falls between 38% and 51% across timing conditions; average calorie intake falls between 196.49 and 203.53 kcal, out of the entire range from 4.58 to 676.8kcal.

General Discussion

This research strives to understand both the drivers and the consequences of substituting vice food with a lesser vice. It moves away from the unhealthy-healthy dichotomy and captures a more nuanced picture of food consumption. Two studies consistently show that, contrary to what has been found in the vice–virtue domain, inserting a delay between choice and consumption, which is supposed to reduce the present-bias, does not prompt consumers to substitute the vice with a lesser vice. Instead, restrained eating is consistently associated the choice of the lesser vice. Moreover, restrained eaters end up consuming less calories, an advantage largely driven by their stronger preference for the lesser vice rather than their direct control over the consumption process.

The current research makes notable contributions to the literature on consumer self-control. First, it draws a distinction between the lesser vice and healthy virtues. The classic definition of vice and virtue is continuous and relative, equating the lesser vice with the virtue and disregarding the healthy or unhealthy categorization of food (Wertenbroch 1998). Accordingly, the factors that have been shown to enhance the preference for healthy food are assumed applicable to the lesser-vice substitutes. This research challenges this assumption. I suggest that the trade-off between long-term health goals and short-term indulgence goals is much less salient in the vice–lesser vice choice setting, and indeed I find that tuning down the focus on immediate gratification does not increase the preference for lesser-vice products. On the contrary, I argue that the calorie difference is much more salient in the vice–lesser vice trade-off and find that consumers who tend to be more concerned with calorie intake make wiser decisions.

This argument, grounded in Goal System Theory, has much broader implications than what the current study focuses on. It implies that, similar to the choice between the vice and the lesser vice, the choice between the virtue and the greater virtue should also differ from the traditional vice–virtue choice. Since both the virtue and the greater virtue (e.g., soymilk and no-sugar soymilk) can satisfy the health goal, consumers should look for the difference in tastes between the options. Indeed, my ancillary study shows that, consumers primarily consider the difference in tastiness between food options when choosing between

the virtue and the greater virtue but consider the difference in healthiness when choosing between the vice and the lesser vice (see Appendix for details of the study). Put differently, whether consumers mentally categorize the food as healthy or unhealthy fundamentally changes the way they make decisions. This highlights the need for a refined theorizing of self-control problems in the food domain. It also suggests that the interventions and campaigns aimed at improving consumers' interests in the lesser vices or the greater virtues should selectively focus on the health benefit or the superior taste. I plan to continue my research along this line.

This research also highlights the value of studying choice and consumption as chain effects. As alluded to earlier, there is little research that examines food choice and consumption in conjunction. For instance, while it has been shown that misleading labels on unhealthy food can liberate consumers to consume more (Irmak et al. 2011), it is unclear whether the pattern would persist if the labeled food is actively chosen rather than passively accepted by consumers. The current research suggests that consumers may not think beyond choices and the exact influence of choice on consumption is subject to product features such as the packaging. By investigating choice and consumption jointly, future research would not only be able to generate novel insights into consumer psychology but provide more relevant recommendations to companies and policy makers.

Finally, the reported findings illustrate the difficulty of rationing food intake at the consumption stage. Although restrained eaters tend to substitute the unhealthy food with a lesser vice, they do not exert much direct control over their eating process. This dovetails with the notion that consumers are insensitive to their physiological changes during food consumption and the subjective feeling of satiation is largely constructed by the ongoing consumption experience (Redden 2015). Apparently, people find it easier to substitute a vice for a lesser vice than to limit the amount they consume. This only reinforces the value of lesser vices in the marketplace: restrained eaters do end up consuming fewer calories, not by limiting their consumption amount, but by grace of the product they choose.

Appendix

Figure A1: snack choices in study 1 (upper) and study 2 (lower).

Ancillary study

The study was to test whether there is any asymmetry in the predictive power of perceived healthiness and tastiness on the choice of food within healthy and unhealthy categories.

I first conducted a pretest to determine the choice sets for the main study. I collected 100 common grocery items and presented their names and images in random sequences to 50 workers from Amazon Mechanical Turk (19 females; mean of age = 36.18; SD of age = 11.69). For each item, I asked participants to report whether they think it falls within the healthy or the unhealthy category. They are allowed to select "I don't know" if they cannot decide. I selected 53 food items where the categorization achieves at least 80% consensus among the participants. This selection resulted in 12 unique pairs of unhealthy food options, 12 unique pairs of healthy food options, and 12 unique healthy-unhealthy mixed pairs.

The main study was conducted in two phases. In phase one, I recruited 114 participants (36 females; Mean of age = 35.72, SD of age = 10.51; none had taken part in the pretest). They viewed all the pre-selected food items, one item per page, and evaluated the healthiness as well as the tastiness of each item (1 = Not healthy/tasty at all, 9 = Very healthy/tasty). Two days later, the same participants were recruited into phase two. They viewed all the food pairs, one pair per page, in three blocks (unhealthy, healthy, and mixed pairs). Because the main purpose was to compare within-category choices, I randomized the sequence of unhealthy and healthy blocks and kept the mixed block in the end.

Participants were asked to imagine they were doing regular grocery shopping for themselves and select the food they prefer within each pair. Importantly, I told them that the options within each pair are comparable in price and portion size. To encourage participants to indicate their real preferences, I also promised a lottery for a $50 Walmart e-giftcard on top of the regular payment.

For analyses, I excluded 11 participants who reported allergies to or avoidance of one or more food options in the pre-selected set. Including them into analyses did not alter the pattern of results. I utilized the differences in perceived healthiness and tastiness between the food options within each pair, which was observed in phase 1 of the study, to predict

food choices that were observed in phase 2. For instance, within a pair of food options A and B, choice would be the dependent variable (choosing A = 1, choosing B = 0) and (Healthiness of A - Healthiness of of B) would be the difference score for healthiness and (Tastiness of A - Tastiness of B) be the difference score for tastiness. All scores were standardized before analyses. Given that the predictors (healthiness vs tastiness) and food choices are nested within the type of choice-set (unhealthy, healthy, or mixed) within each individual, I conducted a multilevel regression with choice being the binary dependent variable.

There was an overall significant three-way interaction among the difference score, the measurement (tastiness or healthiness), and the type of choice ($p < .001$). Within each type of food pairs, there was significant two-way interaction between the difference score and the measurement ($ps < .004$). Specifically, within healthy pairs, the taste-difference score predicted the food choice ($\beta = 0.45$, $SE = 0.05$, $z = 9.56$, $p < .001$) but the health-difference score did not ($\beta = 0.01$, $SE = 0.05$, $z = 0.13$, $p = .897$). In contrast, within unhealthy pairs, the health-difference score significantly predicted the food choice ($\beta = -0.11$, $SE = 0.04$, $z = 2.81$, $p = .005$) but the taste-score did not ($\beta = -0.04$, $SE = 0.03$, $z = -1.39$, $p = .165$). Critically, the variance of the health score does not differ from the variance of the taste score in the healthy (Levene's test $F(1, 2470) = 2.32$, $p = .128$) and unhealthy pairs ($F(1, 2470) = 2.12$, $p = .146$), suggesting that the observed asymmetry cannot be explained by the difference in the variance of the scores.

The results of mixed pairs resemble the pattern of healthy pairs. However, the variance of the health score was significantly larger than the variance of the taste score within mixed pairs (Levene's $F(1, 2470) = 37.77$, $p < .001$). This suggests that the results of mixed pairs may be a statistical artifact.

Chapter 5

General Discussion

The world consumers are living in is increasing globalized, digitalized, and saturated with temptations. As a result, consumers are surrounded by multiple languages, automated recommendations, and health reminders. Substantial research has been conducted to understand how consumers navigate in this world, by looking at how they evaluate products and make decisions in different language contexts (e.g., Karataş 2019; Vidal, Costa, and Foucart 2019), how to generate the best recommendations to individual users (Kim et al. 2017), and how to drive consumers towards a healthy diet (Cadario and Chandon 2019).

Despite the wealth of research in these fields, I put forward that more research needs to be done in order to promote the welfare of consumers as a collective. First, although there is extensive research on (un)ethical behavior and bilingualism as separate fields, few have linked the two. Second, automated recommendations contribute tremendously to consumers' new discoveries, but it also engenders the concern that consumers may never be exposed to a broader set of new products. This problem has been considered by computer scientists looking at the algorithmic 'black box', yet rarely viewed as a problem addressable by marketing communications. Third, research on healthy eating has been dominated by how to improve the share of healthy choosers, leaving out a large proportion of cases in which consumers are deciding between two hedonic options that do not qualitatively yet quantitatively differ in healthiness. To fill these gaps, this dissertation uncovers the intricate influence of language on consumer dishonesty (Chapter 2), develops framing of recommendation as a tool to maximize consumer traffic to recommendations of novel products (Chapter 3), and pinpoints the self-control dynamic when consumers are choosing between the vice and the lesser vice (Chapter 4). By doing so, these essays take a step further in realizing the potential of consumer research to advance the collective welfare of consumers. Meanwhile, the advancement is accompanied by greater benefits to companies, in terms of the management of losses incurred by dishonest consumers, consumer engagement with recommendations, and the marketing of lesser-vice products. In the remainder of the chapter, I would not reiterate the implications of each chapter (as well summarized in the general discussion sections of chapters 2 to 4) but provide a vision about what these studies mean to future research on consumer welfare and the research of contextual influence in general.

Implications for Research in the Backdrop of Globalization, Digitalization, and Self-Control Problems

The viewpoint of consumer welfare

A focus on the collective welfare of consumers would broaden the horizon of the current consumer research and bring about exciting questions. One phenomenon that concerns consumer welfare throughout history and particularly in this digital age is the spread of misinformation. Consumers are motivated to stay informed and to receive high-quality information. However, this is frequently sabotaged by social factors. One factor that has been overlooked yet of high relevance to the rise of digitalization is the device consumers use to share information. Nowadays consumers have news at their fingertips, either shown on their computers or mobile devices, and they are bombarded by information from various sources. Assessing the quality of the information could be a daunting task, especially when sharing is just one click or tap away. How would device influence consumers' sharing of information of different quality? One of my ongoing projects (not collected in this book) is to answer this question. My analysis of a large-scale Twitter dataset shows that, counter to the lay intuition that using mobile devices lowers the quality of shared content, the use of mobile devices increases the proportion of content from high-quality sources being shared as compared to non-mobile devices. Follow-up experiments reveal that the change of device does not alter consumers' ability to discern the quality of content, but rather magnifies their intention to share high-quality content with others. Paradoxically, this magnified intention can result in less sharing of objectively high-quality content that is subjectively low-quality. This work shows that a contextual shift (mobile versus non-mobile device) does not always lead to desirable behavior (sharing high-quality content) but strengthens consumers' goodwill against the undesirable which can lead to either desirable or undesirable behavioral outcomes.

Another phenomenon that has been touched upon in Chapter 4 but deserves more expanded research is the marketing of greater virtues and lesser vices that abound in the marketplace and help move consumers inches toward a healthier life. The exact meaning of a healthy diet is heterogeneous across individuals and situations (Vosgerau, Scopelliti, and Huh 2019). While some may find themselves struggling between the healthy and unhealthy categories, others may choose within the virtue or the vice category. Chapter 4

implies that there is an asymmetry in consumers' concern over healthiness and tastiness of food options when they are choosing within the vice or the virtue categories. Drawing on the goal systems theory, I predict that consumers would weigh tastiness more when choosing between the virtue and the greater virtue but healthiness more when choosing between the vice and the lesser vice. Chapter 4 suggests one way to test this asymmetry (the ancillary study). Here I discuss two other ways to test it. The first is to examine individual differences in their focus on the immediate hedonic value and the long-term health value of food consumption. Prior research suggests that the present focus and the future focus are related yet dissociable: They differ in the neurological roots (McClure et al. 2004) and can be captured by different constructs (e.g., external eating versus restrained eating, Van Strien 1986). Following my prediction, individuals who are more mindful of healthiness should be more likely to choose the lesser vice over the vice. In contrast, individual differences in their focus on the hedonic value of consumption should matter less to the choice of lesser vices. When the choice is between the virtue and the greater virtue, the pattern should reverse, such that individuals who are more mindful of tastiness should be more likely to choose the greater virtue over the virtue, yet individual differences in their focus on the healthiness of consumption should matter less. Another way to test the asymmetry is by comparing the effectiveness of marketing slogans that emphasize taste or health benefits. Specifically, health-oriented slogans should be more effectively moving consumers from virtues to greater virtues as compared to from vices to lesser vices, whereas taste-oriented slogans should do the opposite. Empirical work along these lines would better inform managers and policy makers of how to market the increasingly available healthier alternatives.

In addition, I believe consumer research can contribute a lot more to the understanding of the restricted variety of consumption in the presence of automated recommendations. While this is frequently a result of algorithm bias, it is possible that consumers react to algorithmic recommendations in a self-handicapping way that undermines the diversity of recommendations they receive. Nowadays algorithms are flexible enough to incorporate consumer behavioral signals to personalize recommendations. This means that consumers can tell recommender systems if they want more diversity, such as by searching for products that are not typically recommended or

avoiding recommended products of the same category. In two preliminary studies, I find that the extent to which consumers proactively signal their preference for diversity depends on their perception of the recommender being human- or algorithm-driven. In specific, consumers who tend to believe their recommenders are algorithm-driven are less likely to seek variety for more diverse recommendations in return. Importantly, this effect is specific to variety-seeking; it does not emerge for other types of interactions with the recommender system, such as consumers' searching for what they like and deleting what they dislike, for the purpose of receiving more accurate recommendations. This suggests that consumers may miss out opportunities to receive diverse recommendations from algorithms. Consumers' suboptimal interaction with algorithms can result from two different perceptions. On the one hand, consumers may perceive their interaction with algorithms as more private, which would lower their motivation to manage an interesting and balanced self-image and make them cling to their defaults (Ratner and Khan 2002; Young, Vosgerau, and Morewedge 2014). In this case, the solution to the self-imposed constraint on diversity would lie in enhancing the perceived publicity of interaction with recommender systems, such as by humanizing the algorithmic recommender. On the other, consumers may see algorithmic recommenders as worse at learning the diversity of preferences as compared to human recommenders. If this is true, explicit cues of algorithmic learning capacity may be required to encourage consumers to pursue variety proactively. Understanding how to encourage consumers to signal their preference for diversity would benefit consumers in the long term, allowing them to discover the beauty in a wide range of products, to acquire knowledge from different sources, and to see the opinions of various parties.

The intersections between globalization, digitalization, and overconsumption

The movement towards a more globalized, digitalized, and health-minded world paves the way for novel research that bridges these fields. The crossovers give rise to burning questions that, to my knowledge, not being answered but would concern consumers even more in the future. One area of my interest is consumer self-regulation of food choices in the presence of recommendations. Recent research shows that consumers tend to follow recommendations even when the recommended is an inferior option (Banker and Khetani 2019). Would the ubiquity of recommendations make consumers pick

whatever is recommended without much deliberation? If so, consumers may feel less accountable for their choices and attribute their self-control failures or successes to the recommender system. This is likely to disrupt self-regulation and to make the consumption of vices guilt-free (Hagen, Krishna, and Mcferran 2017). Meanwhile, it is possible that the mere act of taking recommendations creates an illusion of choice (Hadi and Block 2014), which can make consumers over-claim their self-control credits (if the recommendation is filled up with virtues) and loosen their subsequent control (Khan and Dhar 2006). Future research can look at self-control behavior in the presence of recommendations and examine the long-term effects of relying on recommendations on eating habits.

Another interesting phenomenon is how consumers interact with artificial intelligence (AI) in different languages. An increasing number of consumers own voice assistants, such as Alexa, Siri, and Google assistant, which they can talk to and give commands to. How would the language consumers speak shape their interaction with AI? Would they outsource different types of tasks to their voice assistants when using their first or second languages? For instance, would using a second language liberate consumers to outsource tasks that they feel uncomfortable outsourcing to machines in the first language, such as the tasks related to their core identity (Garcia-Rada et al 2018; Leung, Puntoni, and Paolocci 2018)? In addition, would the language of choice change consumers' non-verbal communication with AI, such as by changing their pitch, volume, and tempo of their speech? Would consumers feel less repulsive by machines attempting to imitate humans speaking when the interaction is in their second language as opposed to the first language? Are non-native speakers less able to tell the difference between AI and human voices? These are all interesting questions for future research.

Lastly, little research has examined how language affects ongoing food consumption, although eating frequently involves social interaction in different languages. While previous research shows that ordering in a foreign language prompts consumers to choose the virtue over the vice (Klesse et al. 2015), it is unknown whether choosing vices in a second language dissociates negative self-conscious feelings, such as guilt and shame, from choosing the vice. Reduced guilt has been shown to increase the intake of unhealthy food (Duke and Amir 2018), suggesting that foreign language may lead to more unhealthy

consumption. However, it is also possible that speaking a foreign language facilitates healthy eating: because speaking a foreign language is cognitively taxing, it can activate inhibition in general (Tuk, Trampe, and Warlop 2011) and make consumers more mindful of their eating processes. Future research can disentangle these competing predictions.

Implications for the Approach to Studying Contextual Influences on Consumer Behavior

From the average towards the idiosyncratic

The findings presented in this book speak to the view that the influence of contextual factors (such as the language context) on consumer behavior is subject to consumers' idiosyncrasies in cognition: The effect of language on consumers' opportunistic lying is as malleable as the moral intuition of individuals (Chapter 2). In addition, consumers do not unanimously flock to recommendations with a more informative explanation, even when the explanation implies a social norm. Instead, they proactively evaluate the quality of recommendations by reflecting upon who they are, who the other users are, and product features (Chapter 3). Finally, temporal distancing from consumption does not bring a more rational self out of the same individual when the food options are equally tempting (Chapter 4).

The social-cognitive approach advanced in this book contrasts with the other popular approach that portrays contextual factors as mechanical determinants of consumer behavior. This competing view is reflected in how subtle shifts in the context can induce behavioral changes, and even more famously, in the power of Nudges (Thaler and Sunstein 2008). This mechanical approach may have a particular appeal when applied to marketing. Part of the appeal is economic; if consumer behavior can be transformed by a uniform contextual change, other costly strategies such as promotions or personalization would seem unnecessary. It has an even greater appeal at the psychological level, as it promises the potential to change consumers in a way that is immune to their conscious control.

The appeal of the competing perspective, however, obscures the nature of consumers, who are sophisticated, thoughtful, and capricious. Indeed, recent research has started to look at why Nudges fail in some cases on some individuals (Sunstein 2017) and how consumers interpret choice architectures (Job, Tannenbaum, and Fox 2017). In the same

vein, this dissertation advocates the agency of consumers and suggests that averaging contextual effects across individuals disguises their intricacies. The most obvious example is Chapter 4, which shows that advance ordering is not as effective in the vice–lesser vice choice setting as in the vice–virtue setting. Another example is the role of language. The published research on bilingualism often suggests a unidirectional effect of using a foreign language. Contradictory results are discarded as "no effect" instead of a fruitful opportunity for refined theorizing (e.g., Alempaki et al. 2019). Chapter 2, however, takes into account the variation of moral intuitions and enriches the understanding of the language effect.

The benefit of considering consumer idiosyncrasies extends beyond the theoretical level. Companies gather and hoard data about consumers and the contexts they are in. Understanding the interaction between consumers and their contexts would help companies maximize the usage of their data and generate deeper insights. For instance, Chapter 3 suggests that it is not wise to frame all recommendations as user-based; instead, similar to personalized recommendations, the framing of them should also be tailored to who the users are and the attractiveness of focal products.

From the static towards the dynamic

The view that consumer-context interaction determines behavior also implies that the contextual influence should evolve across time or situations instead of being static. I identify two kinds of dynamics that deserve more investigation. First, it is not well understood how consumers, once situated in a specific context, behave differently when they have different choices to make. This is partially captured in Chapter 2, where consumers assigned to different language conditions had to decide whether to cheat at various magnitudes. This kind of dynamic can be purely behavioral in the sense that it does not necessarily affect how contexts exert psychological influences (e.g., meta-cognitive uncertainty) but rather change the observed effects on behaviors.

The other type of dynamic, in contrast, is at the cognitive level and much less studied in consumer research. With repeated exposures to the same context, consumers could develop their own understanding of the context, especially when they have little experience with the context or any established knowledge about that. Take algorithmic

110

recommendations as an example. One common scenario that is not captured in any of my studies in Chapter 3 is when consumers keep clicking on the recommendations, such as binge-watching videos recommended by a website. In this situation, consumers could slip into a narrow category of products, be fed up with the recommendations, and even develop an aversion to the underlying algorithm. This implies that the framing effect of recommendations would evolve with consumers' knowledge and experience with recommendations. Another example is the use of mobile versus non-mobile devices. Recent research shows that consumers trust reviews sent from mobile devices more than those from computers (Grewal and Stephen 2019). However, longitudinal evidence shows the opposite in the long term (Ransbotham, Lurie, and Liu 2019). This kind of dynamic, resulting from consumer learning processes, is prevalent in reality yet much less studied in the lab. I wish to extend the time span of my future investigations and examine the dynamic of consumer behavior.

Conclusion

In conclusion, this dissertation investigates how to promote consumer welfare in the age of globalization, digitalization, and overconsumption. The findings and thoughts presented in this book hopefully shed some light, bright or dim, on the complex causes of consumer behavior, and provide some useful insights for managers and policy makers alike. They represent the start, instead of the closure, of a larger, more ambitious program of research in the coming years.

References

Adomavicius, Gediminas and Youngok Kwon (2012), "Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques," *IEEE Transactions on Knowledge and Data Engineering*, 24 (5), 1–15.

Adomavicius, Gediminas, Youngok Kwon, and Gediminas Adomavicius (2014), "Optimization-Based Approaches for Maximizing Aggregate Recommendation Diversity," *INFORMS Journal on Computing*, 26 (2), 351–69.

Ahn, Jungsun and Carrie La Ferle (2008), "Enhancing Recall and Recognition for Brand Names and Body Copy: A Mixed-Language Approach," *Journal of Advertising*, 37 (3), 107–17.

Ainslie, George (1975), "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control," *Psychological Bulletin*, 82 (4), 463–96.

Alba, Joseph W. and J. Wesley Hutchinson (2002), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (4), 411-54.

Alempaki, Despoina, Gönül Doğan, and Yang Yang (2017), "When in Rome: Lying in a Foreign Language," working paper, University of Nottingham.

Alempaki, Despoina, Gönül Doğan, and Yang Yang (2019), "Lying in a Foreign Language?," working paper, University of Warwick.

Amatriain, Xavier and Justin Basilico (2012), "Netflix Recommendations: Beyond the 5 Stars (Part 1)," *Netflix Technology Blog*, [available at https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429].

——— and ——— (2016), "Past, Present, and Future of Recommender Systems: An Industry Perspective," *Proceedings of the 10th ACM Conference on Recommender Systems*, 211–14.

Anderson, Chris (2008), *The Long Tail: Why the Future of Business is Selling Less of More*, New York: Hyperion Books.

Andrade, Eduardo B., and Teck-Hua Ho (2009), "Gaming Emotions in Social Interactions," *Journal of Consumer Research,* 36 (4), 539–52.

Ansari, Asim, Skander Essegaier, and Rajeev Kohli (2000), "Internet Recommendation Systems," *Journal of Marketing Research*, 37, 363–75.

Anthony, Christina I. and Elizabeth Cowley (2012), "The Labor of Lies: How Lying for Material Rewards Polarizes Consumers' Outcome Satisfaction," *Journal of Consumer Research*, 39 (3), 478–92.

Argo, Jennifer J., Darren W. Dahl, and Rajesh V. Manchanda (2005), "The Influence of a Mere Social Presence in a Retail Context," *Journal of Consumer Research*, 32 (2), 207–12.

Argo, Jennifer J., Katherine White, and Darren W. Dahl (2006), "Social Comparison Theory and Deception in the Interpersonal Exchange of Consumption Information," *Journal of Consumer Research*, 33 (1), 99–108.

Argo, Jennifer J. and Kathrine White (2012), "When Do Consumers Eat More? The Role of Appearance Self-Esteem and Food Packaging Cues", *Journal of Marketing,* 76 (2), 67–80.

Ariely, Dan and Steve Hoeffler (1999), "Constructing Stable Preferences: A Look Into Dimensions of Experience and Their Impact on Preference Stability*," Journal of Consumer Psychology*, 8 (2), 113–39.

Ariely, Dan, John G. Lynch, and Manuel Aparicio IV (2004), "Learning by Collaborative and Individual-Based Recommendation Agents," *Journal of Consumer Psychology*, 14 (1&2), 81–95.

Bandura, Albert (2001), "Social Cognitive Theory: An Agentic Perspective*," Annual Review of Psychology*, 52 (1), 1–26.

Banker, Sachin and Salil Khetani (2019), "Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being," *Journal of Public Policy and Marketing*, 38 (4), 500–15.

Bargh, John A. and Tanya L. Chartland (2000), "The Mind in the Middle: A Practical Guide to Priming and Automaticity Research," in *Handbook of Research Methods in Social and Personality Psychology,* Harry T. Reis and Charles M. Judd, eds. New York: Cambridge University Press, 253–85.

Becker, Gary S. (1991), "A Note on Restaurant Pricing and Other Examples of Social Influences on Price," *Journal of Political Economy*, 99 (5), 1109–16.

Bereby-Meyer, Yoella, Sayuri Hayakawa, Shaul Shalvi, Joanna D. Corey, Albert Costa, and Boaz Keysar (2018), "Honesty Speaks a Second Language," *Topics in Cognitive Science*, 1–12.

Berger, Jonah and Chip Heath (2008), "Who Drives Divergence? Identity Signaling, Outgroup Dissimilarity, and the Abandonment of Cultural Tastes," *Journal of Personality and Social Psychology*, 95 (3), 593–607.

Blasi, Augusto (2005), "Moral Character: A Psychological Approach," in *Character psychology and character education,* Daniel. K. Lapsley and F. Clark Power, eds. Notre Dame, IN: University of Notre Dame Press, 67–100.

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187–217.

Bodapati, Anand. V (2008), "Recommendation Systems with Purchase Data," *Journal of Marketing Research*, 45 (1), 77–93.

Borenstein, Michael, Larry V. Rothstein, Julian P. T. Hedges, and Hannah R. Higgins (2009), "Converting Among Effect Sizes," in *Introduction to Meta-Analysis*. Chichester: John Wiley & Sons, 45–9.

Burton, Pat, Hendrik J. Smit, and Helen J. Lightowler (2007), "The Influence of Restrained and External Eating Patterns on Overeating," *Appetite*, 49 (1), 191–97.

Caldwell-Harris, Catherine L. and Ayşe Ayçiçeği-Dinn (2009), "Emotion and Lying in a Non-Native Language," *International Journal of Psychophysiology*, 71 (3), 193–204.

Carnevale, Marina, David Luna, and Dawn Lerman (2017), "Brand Linguistics: A Theory-Driven Framework for the Study of Language in Branding," *International Journal of Research in Marketing*, 34 (2), 572–91.

Cecchini, Michele, Franco Sassi, Jeremy A. Lauer, Yong Y. Lee, Veronica Guajardo-Barron, and Daniel Chisholm (2010), "Tackling of Unhealthy Diets, Physical Inactivity, and Obesity: Health Effects and Cost-Effectiveness," *The Lancet*, 376, 1775–84.

Chen, Sylvia Xiaohua and Michael Harris Bond (2010), "Two Languages, Two Personalities? Examining Language Effects on the Expression of Personality in a Bilingual Context," *Personality and Social Psychology Bulletin*, 36 (11), 1514–28.

Chen, Yubo and Jinhong Xie (2008), "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science*, 54 (3), 477–91.

Clarkson, Joshua, J. Janiszewski, and Melissa D. Cinelli (2013), "The Desire for Consumption Knowledge," *Journal of Consumer Research*, 39 (6), 1313–29.

Coelho do Vale, Rita, Rik Pieters, and Marcel Zeelenberg (2008), "Flying under the Radar: Perverse Package Size Effects on Consumption Self-Regulation," *Journal of Consumer Research*, 35 (3), 380–90.

Cooke, Alan D. J., Harish Sujan, Mita Sujan, and Barton A. Weitz (2002), "Marketing the Unfamiliar: The Role of Context and Item-specific Information in Electronic Agent Recommendations," *Journal of Marketing Research*, 39 (4), 488–97.

Corey, Joanna D., Sayuri Hayakawa, Alice Foucart, Melina Aparici, Juan Botella, Albert Costa, and Boaz Keysar (2017), "Our Moral Choices Are Foreign to Us," *Journal of Experimental Psychology: Learning Memory and Cognition*, 43 (7), 1109–28.

Costa, Albert, Alice Foucart, Inbal Arnon, Melina Aparici, and Jose Apesteguia (2014), "'Piensa' Twice: On the Foreign Language Effect in Decision Making," *Cognition*, 130 (2), 236–54.

Cowley, Elizabeth and Christina I Anthony (2018), "Deception Memory: When Will Consumers Remember Their Lies?," *Journal of Consumer Research,* 46 (1), 180-99.

Cramer, Henriette, Vanessa Evers, Satyan Ramlal, Maarten Van Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob Wielinga (2008), "The Effects of Transparency

on Trust in and Acceptance of a Content-based Art Recommender," *User Modeling and User-Adapted Interaction*, 18 (5), 455-96.

Dayan, Eran and Maya Bar-Hillel (2011), "Nudge to Nobesity II: Menu Positions Influence Food Orders," *Judgment and Decision Making*, 6 (4), 333–42.

de Langhe, Bart, Stefano Puntoni, Daniel Fernandes, and Stijn M.J van Osselaer (2011), "The Anchor Contraction Effect in International Marketing Research," *Journal of Marketing Research*, 48 (2), 366–80.

De Jong, Martijn G., Jean-Paul Fox, and Jan-Benedict E.M. Steenkamp (2015), "Quantifying Under- and Overreporting in Surveys through a Dual-Questioning-Technique Design," *Journal of Marketing Research*, 52 (6), 737–53.

DePaulo, Bella M., Deborah A. Kashy, Susan E. Kirkendol, Melissa M. Wyer, and Jennifer A. Epstein (1996), "Lying in Everyday Life," *Journal of Personality and Social Psychology*, 70 (5), 979–95.

Devendorf, Laura and Elizabeth Goodman (2014), "*The Algorithm Multiple, the Algorithm Material*," Contours of Algorithmic Life, UC Davis, May 15.

Dienstbier, Richard A. and Pamela Osbourne Munter (1971), "Cheating as a Function of the Labeling of Natural Arousal," *Journal of Personality and Social Psychology,* 17 (2), 208–13.

Dovey, Terence M, Tina Torab, Dorothy Yen, E. J. Boyland, Jason C. G. Halford (2017), "Responsiveness To Healthy Advertisements in Adults: An Experiment Assessing Beyond Brand Snack Selection and the Impact of Restrained Eating," *Appetite*, 112, 102–6.

Elfhag, K., P. Tynelius, and F. Rasmussen (2007), "Sugar-Sweetened and Artificially Sweetened Soft Drinks in Association to Restrained, External and Emotional Eating," *Physiology & behavior*, 91 (2–3), 191–95.

Epley, Nicholas and Thomas Gilovich (2006), "The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient," Psychological Science, 17 (4), 311–18.

Epstein, Seymour (1994), "Integration of the Cognitive and the Psychodynamic Unconscious," American Psychologist, 49 (8), 709–24.

Fleder, Daniel and Kartik Hosanagar (2009), "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity," *Management Science*, 55 (5), 697–712.

Frick, Robert W. (1995), "Accepting the Null Hypothesis," *Memory & Cognition*, 23 (1), 132–38.

FRISS (2018), "Insurance Fraud Report," https://apprissretail.com/wp-content/uploads/sites/4/2017/02/wp_TRE4013-WhitePaper-ReturnFraud101_Feb2013.pdf.

Fujita, Kentaro, Yaccov Trope, Nira Liberman, and Maya Levin-Sagi Levin-Sagi (2006), "Construal Levels and Self-Control," *Journal of Personality and Social Psychology*, 90 (3), 351–67.

Galak, Jeff and George Loewenstein (2013), "Slow Down! Insensitivity to Rate of Consumption Leads to Avoidable Satiation," *Journal of Consumer Research,* 39 (5), 933–1009.

Gai, Phyliss Jia and Anne-Kathrin Klesse (2009). "Making Recommendations More Effective Through Framings: Impacts of User-Versus Item-Based Framings on Recommendation Click-Throughs," *Journal of Marketing,* 83 (6), 61-75.

Gai, Phyliss Jia and Amit Bhattacharjee (2009). "The Moral Advantage of Successful Self-Control (and vice versa)," working paper, Erasmus University.

Gao, Lei, Yanjie Li, and Robert S. Wyer Jr. (2016), "Choosing Between Two Evils: The Determinants Of Preferences for Two Equally Goal-Inconsistent Options," *Journal of Consumer Research*, 42 (6), 1002–12.

Gao, Shan, Onderij Zika, Robert D. Rogers, and Guillaume Thierry (2015), "Second Language Feedback Abolishes the 'Hot Hand' Effect during Even-Probability Gambling," *Journal of Neuroscience, 35* (15), 5983–89.

Garcia-Rada, Ximena, Mary Steffel, Elanor F. Williams, and Michael I. Norton (2018), "A Preference for Effort When Caring for Close Others," in *Advances in Consumer Research (Volume 64),* A. Gershoff, R. Kozinets, and T. White, eds., 346–7.

Geipel, Janet, Constantinos Hadjichristidis, and Anne-Kathrin Klesse (2018), "Barriers to Sustainable Consumption Attenuated by Foreign Language Use," *Nature Sustainability,* 1, 31–33.

Geipel, Janet, Constantinos Hadjichristidis, and Luca Surian (2015), "How Foreign Language Shapes Moral Judgment," *Journal of Experimental Social Psychology*, 59, 8–17.

Gino, Francesca, Maurice E. Schweitzer, Nicole L. Mead, and Dan Ariely (2011), "Unable to Resist Temptation: How Self-Control Depletion Promotes Unethical Behavior," *Organizational Behavior and Human Decision Processes*, 115 (2), 191–203.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel (2018), "Lying Aversion and the Size of the Lie," *American Economic Review*, 108 (2), 419–53.

Goldsmith, Kelly, Caroline Roux, and Jingjing Ma (2018), "When Seeking the Best Brings Out the Worst in Consumers: Understanding the Relationship between a Maximizing Mindset and Immoral Behavior," *Journal of Consumer Psychology*, 28 (2), 293–309.

Gomez-Uribe, Carlos A. and Neil Hunt (2015), "The Netflix Recommender System," *ACM Transactions on Management Information Systems*, 6 (4), 1–19.

Gong, Shiyang, Juanjuan Zhang, Ping Zhao, and Xuping Jiang (2017), "Tweeting as a Marketing Tool–Field Experiment in the TV Industry," *Journal of Marketing Research*, 54(6), 833-50.

Greene, Joshua D. and Joseph M. Paxton (2009), "Patterns of Neural Activity Associated with Honest and Dishonest Moral Decisions," *Proceedings of the National Academy of Sciences*, 106(30), 12506–11.

Grewal, Lauren, and Andrew T. Stephen, (2019) "In Mobile We Trust: The Effects of Mobile versus Nonmobile Reviews on Consumer Purchase Intentions," *Journal of Marketing Research,* 56 (5), 791-808.

Grimes, William (2018), "Using the Foreign to Grasp the Familiar," *New York Times,* https://www.nytimes.com/2014/04/26/books/writing-in-english-novelists-find-inventivesnew-voices.html.

Grosjean, Francois (2010), "Why are People Bilingual," in *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press.

Gupta, Sunil, Dominique Hanssens, Bruce Hardie, Wiliam Kahn, V. Kumar, Nathaniel Lin, Nalini Ravishanker, and S. Sriram (2006), "Modeling Consumer Lifetime Value," *Journal of Service Research*, 9 (2), 139–55.

Haddock, Geoffrey, Alexander J. Rothman, Rolf Reber, and Norbert Schwarz (1999), "Forming Judgments of Attitude Certainty, Intensity, and Importance: The Role of Subjective Experiences," *Personality and Social Psychology Bulletin*, 25 (7), 771–82.

Hadi, Rhonda and Lauren Block (2014), "I Take Therefore I Choose? The Impact of Active vs. Passive Acquisition on Food Consumption," *Appetite*, 80, 168–73.

Hagen, Linda, Aradhna Krishna, and Brent Mcferran (2017), "Rejecting Responsibility: Low Physical Involvement in Obtaining Food Promotes Unhealthy Eating," *Journal of Marketing Research*, 54 (4), 589–604.

Haidt, Jonathan and Craig Joseph (2004), "Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues," *Dædalus* (Fall), 55–66.

Haig, Brian D. (2005), "An Abductive Theory of Scientific Method," *Psychological Methods*, 10 (4), 371-88.

Hayakawa, Sayuri, Albert Costa, Alice Foucart, and Boaz Keysar (2016), "Using a Foreign Language Changes Our Choices," *Trends in Cognitive Sciences*, 20 (11), 791–93.

Hayakawa, Sayuri and Boaz Keysar (2018), "Using a Foreign Language Reduces Mental Imagery," *Cognition*, 173, 8-15.

Haws, Kelly L. and Karen Page Winterich (2013), "When Value Trumps Health in a Supersized World," *Journal of Marketing*, 77 (3), 48–64.

Hayakawa, Sayuri, David Tannenbaum, Albert Costa, Joanna D. Corey, and Boaz Keysar (2017), "Thinking More or Feeling Less? Explaining the Foreign-Language Effect on Moral Judgment," *Psychological Science*, 28 (10), 1387–97.

Hayes, Andrew F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, New York: Guilford Press.

Hennig-Thurau, Thorsten, André Marchand, and Paul Marx (2012), "Can Automated Group Recommender Systems Help Consumers Make Better Choices?," *Journal of Marketing*, 76 (5), 89–109.

Hilbig, Benjamin E. and Corinna M. Hessler (2013), "What Lies beneath: How the Distance between Truth and Lie Drives Dishonesty," *Journal of Experimental Social Psychology,* 49 (2), 263–66.

Hilmert, Clayton J., James a. Kulik, and Nicholas J.S. Christenfeld (2006), "Positive and Negative Opinion Modeling: The Influence of Another's Similarity and Dissimilarity," *Journal of Personality and Social Psychology*, 90 (3), 440–52.

Hinz, Oliver and Jochen Eckert (2010), "The Impact of Search and Recommendation Systems on Sales in Electronic Commerce," *Business & Information Systems Engineering*, 2 (2), 67–77.

Hoch, Stephen J. and George Loewenstein (1991), "Time-Inconsistent Preferences and Consumer Self-Control," *Journal of Consumer Research*, 17 (4), 492-507.

Hoeffler, Steve and Dan Ariely (1999), "Constructing Stable Preferences: A Look into Dimensions of Experience and Their Impact on Preference Stability," *Journal of Consumer Psychology*, 8 (2), 113–39.

Horton, John J., David G. Rand, and Richard J. Zeckhauser (2011), "The Online Laboratory: Conducting Experiments in a Real Labor Market," *Experimental Economics*, 14 (3), 399–425.

Huang, Rui and Muzhe Yang (2013), "Buy What Is Advertised on Television? Evidence from Bans on Child-Directed Food Advertising," *Journal of Public Policy and Marketing*, 32 (2), 207–22.

Insurance Research Council (2012), "Insurance Research Council Finds That Fraud and Buildup Add Up to $7.7 Billion in Excess Payments for Auto Injury Claims," https://www.insurancefraud.org/downloads/InsuranceResearchCouncil02-15.pdf.

International Data Corporation (2019), "Worldwide Spending on Artificial Intelligence Systems Will Grow to Nearly $35.8 Billion in 2019, According to New IDC Spending Guide" (accessed May 23, 2019), [available at https://www.idc.com/getdoc.jsp?containerId=prUS43662418].

Ipeirotis, Panagiotis G. (2010), "Demographics of Mechanical Turk," CeDER-10–01 working paper, New York University.

Irmak, Calgar, Beth Vallen, and Stefanie Rosen Robinson (2011), "The Impact of Product Name On Dieters' and Nondieters' Food Evaluations and Consumption," *Journal of Consumer Research*, 38 (2), 390–405.

Ivaz, Lela, Albert Costa, and Jon Andoni Duñabeitia (2016), "The Emotional Impact of Being Myself: Emotions and Foreign-Language Processing," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42 (3), 489-96.

Iyengar, Raghuram, Christophe Van den Bulte, and Jae Young Lee (2015), "Social Contagion in New Product Trial and Repeat," *Marketing Science*, 34 (3), 408–29.

Jarosz, Andrew F. and Jennifer Wiley (2014), "What Are The Odds? A Practical Guide to Computing and Reporting Bayes Factors," *Journal of Problem Solving,* 7, 2–9.

Jeffreys, Harold (1961), *The Theory of Probability*, Oxford: Oxford University Press.

Krijnen, Job, David Tannenbaum, and Craig R. Fox (2017), "Choice architecture 2.0: Behavioral policy as an implicit social interaction," *Behavioral Science & Policy*, 3 (2), 1–18.

Johnson, Chris (2015), "From Idea to Execution: Spotify's Discover Weekly," (accessed August 5, 2018), [available at https://www.slideshare.net/MrChrisJohnson/from-idea-to-execution-spotifys-discover-weekly/26-Google_Form_1_Results].

Johnson, Lauren (2015), "26% of Millennials Give Fake Birthdays to Score Better Online Deals," *Adweek*, https://www.adweek.com/digital/26-millennials-give-fake-birthdays-score-better-online-deals-165624/

Kajackaite, Agne and Uri Gneezy (2017), "Incentives and Cheating," *Games and Economic Behavior*, 102, 433–44.

Kamakura, Wagner, Carl F. Mela, Anand Bodapati, Pete Fader, Prasad Naik, Scott Neslin, Peter C. Verhoef, Michel Wedel, and Ron Wilcox (2005), "Choice Models and Consumer Relationship Management," *Marketing Letters*, 16, 279–91.

Keysar, Boaz, Sayuri L. Hayakawa, and Sun Gyu An (2012), "The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases," *Psychological Science*, 23 (6), 661–68.

Khan, Uzma and Ravi Dhar (2006), "Licensing Effect in Consumer Choice," *Journal of Marketing Research*, 43 (2), 259–66.

Klesse, Anne-Kathrin, Jonathan Levav, and Caroline Goukens (2015), "The Effect of Preference Expression Modality on Self-Control," *Journal of Consumer Research,* 42, 535-50.

Koenigstorfer, Joerg and Hans Baumgartner (2015), "The Effect of Fitness Branding on Restrained Eaters' Food Consumption and Post-Consumption Physical Activity," *Journal of Marketing Research*, 53, 124–38.

Kouchaki, M. and I. H. Smith (2014), "The Morning Morality Effect: The Influence of Time of Day on Unethical Behavior," *Psychological Science*, 25 (1), 95–102.

Kovic, Yumi, (2019) "The Impact of Junk Food Marketing Regulations on Food Sales: An Ecological Study," *Master's Thesis*, University of Connecticut.

Kramer, Thomas (2007), "The Effect of Measurement Task Transparency on Preference Construction and Evaluations of Personalized Recommendations," *Journal of Marketing Research*, 44 (2), 224–33.

Krishna, Aradhna and Rohini Ahluwalia (2008), "Language Choice in Advertising to Bilinguals: Asymmetric Effects for Multinationals versus Local Firms," *Journal of Consumer Research*, 35 (4), 692–705.

Lahiri, Jhumpa (2015), "Teach Yourself Italian," *New Yorker*, https://www.newyorker.com/magazine/2015/12/07/teach-yourself-italian.

Lappo, Alena, Trond Bjorndal, José Fernandez-Polanco, and Audun Lem (2015), "Consumers' Concerns and External Drivers in Food Markets," *FAO*.

Latane, Bibb (1981), "The Psychology of Social Impact," *American Psychologist*, 36 (4), 343–56.

Leclerc, France, Bernd H. Schmitt, and Laurette Dube (1994), "Foreign Branding and Its Effects on Product Perceptions and Attitudes," *Journal of Marketing Research*, 31 (2), 263.

Lee, Spike W. S., Daphna Oyserman, and Michael Harris Bond (2010), "Am I Doing Better than You? That Depends on Whether You Ask Me in English or Chinese: Self-Enhancement Effects of Language as a Cultural Mindset Prime," Journal of Experimental Social Psychology, 46 (5), 785–91.

Leung, Eugina, Gabriele Paolacci, and Stefano Puntoni (2018), "Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior," *Journal of Marketing Research*, 55 (6), 818–31.

Liberman, Nira, Jens Förster, and E. Tory Higgins (2007), "Completed vs. Interrupted Priming: Reduced Accessibility from Post-fulfillment Inhibition," J*ournal of Experimental Social Psychology*, 43 (March), 258–64.

Linden, Greg, Brent Smith, and Jeremy York (2017), "Two Decades of Recommender Systems at Amazon.com," *IEEE Internet Computing*, 7 (1), 12–18.

Liu, Peggy J., Kelly L. Haws, Karen Scherr, Joseph P. Redden, James R. Bettman, and Gavan J. Fitzsimons (2019), "The Primacy of 'What' over 'How Much': How Type and Quantity Shape Healthiness Perceptions of Food Portions," *Management Science*, 65 (7), 3353–81.

Loewenstein, George and John Elster (Eds.) (1992), *Choice over Time*, New York: Russell Sage Foundation.

Loewenstein, George (1996), "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, 65 (3), 272–92.

Luna, David and Laura A. Peracchio (2001), "Moderators of Language Effects in Advertising to Bilinguals: A Psycholinguistic Approach," Journal of Consumer Research, 28 (2), 284–95.

_____ (2005), "Advertising to Bilingual Consumers: The Impact of Code-Switching on Persuasion," Journal of Consumer Research, 31 (4), 760–65.

Ma, Wayne (2018), "Inside Apple's War on iPhone Fraud in China," *The Information*, https://www.theinformation.com/articles/inside-apples-war-on-iphone-fraud-in-china.

Maréchal, Michel André, Alain Cohn, Giuseppe Ugazio, and Christian C. Ruff (2017), "Increasing Honesty in Humans with Noninvasive Brain Stimulation," *Proceedings of the National Academy of Sciences*, 114 (17), 4360–64.

Mazar, Nina, On Amir, and Dan Ariely (2008), "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance," *Journal of Marketing Research*, 45, 633–44.

Mazar, Nina and Chen-Bo Zhong (2010), "Do Green Products Make Us Better People?," *Psychological Science*, 21 (4), 494–98.

McClure, Samuel M., David I. Laibson, George Loewenstein, and Jonathan D. Cohen (2004), "Separate Neural Systems Value Immediate and Delayed Monetary Rewards," *Science*, 306, 503–8.

Mead, Nicole L, Roy F Baumeister, and Francesca Gino (2009), "Too Tired to Tell the Truth: Self-Control Resource Depletion and Dishonesty," J*ournal of Experimental Social Psychology*, 45 (3), 594–97.

Metcalfe, Janet and Walter Mischel (1999), "A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower," *Psychological Review*, 106 (1), 3–19.

Meyer, Andrew, Shane Frederick, Terence C. Burnham, Juan D. Guevara Pinto, Ty W. Boyer, Linden J. Ball, Gordon Pennycook, Rakefet Ackerman, Valerie A. Thompson, and Jonathon P. Schuldt (2015), "Disfluent Fonts Don't Help People Solve Math Problems," *Journal of Experimental Psychology: General*, 144 (2), 16–30.

Meyners, Jannik, Christian Barrot, Jan U. Becker, and Jacob Goldenberg (2017), "The Role of Mere Closeness: How Geographic Proximity Affects Social Influence," *Journal of Marketing*, 81, 49–66.

Morvinski, Coby, On Amir, and Eitan Muller (2017), "'Ten Million Readers Can't Be Wrong!,' or Can They? On the Role of Information about Adoption Stock in New Product Trial," *Marketing Science*, 36 (2), 290–30.

Milkman, Katherin, Todd Rogers, and Max H. Bazerman (2010), "I'll Have the Ice Cream Soon and the Vegetables Later: A Study of Online Grocery Purchases and Order Lead Time," *Marketing Letters*, *21* (1), 17–35.

Miller, Gabrielle F., Sonam Gupta, Jaclyn D. Kropp, Kelly A. Grogan, and Anne Mathews (2016), "The Effects of Pre-Ordering and Behavioral Nudges on National School Lunch Program Participants' Food Item Selection," *Journal of Economic Psychology*, 55, 4–16.

Mishra, Arul, and Himanshu Mishra (2011), "The Influence of Price Discount versus Bonus Pack on the Preference for Virtue and Vice Foods," *Journal of Marketing Research*, 48 (1), 196–206.

Müller-Trede, Johannes, Shoham Choshen-Hillel, Meir Barneron, and Ilan Yaniv (2018), "The Wisdom of Crowds in Matters of Taste," *Management Science*, 64 (4), 1779–1803.

Muthén, Bengt (2011), *Applications of Causally Defined Direct and Indirect Effects in Mediation Analysis Using SEM in Mplus,* Unpublished manuscript.

Myrseth, Kristian Ove R. and Ayelet Fishbach (2009), "Self-Control: A Function of Knowing When and How To Exercise Restraint," *Current Directions in Psychological Science*, 18 (4), 247–52.

National Bureau of Statistics of China (NBS) (2012), *Tabulation on the 2010 Population Census of the People's Republic of China*, Beijing: China Statistics Press.

Naylor, Rebecca Walker, Cait Poynor Lamberton, and David A. Norton (2010), "Seeing Ourselves in Others: Reviewer Ambiguity, Egocentric Anchoring, and Persuasion," *Journal of Marketing Research*, 48(3), 617–31.

Naylor, Rebecca Walker, Cait Poynor Lamberton, and Patricia M. West (2012), "Beyond the 'Like' Button: The Impact of Mere Virtual Presence on Brand Evaluations and Purchase Intentions in Social Media Settings," *Journal of Marketing*, 76 (6), 105–2.

Nielson (2014), *Snack Attack*, https://www.nielsen.com/content/dam/nielsenglobal/kr/docs/global-report/2014/Nielsen%20Global%20Snacking%20Report%20September%202014.pdf.

Nikolova, Hristina, Cait Lamberton, and Nicole Verrochi Coleman (2018), "Stranger Danger: When and Why Consumer Dyads Behave Less Ethically Than Individuals," *Journal of Consumer Research*, 45 (1), 90–108.

Novemsky, Nathan, Ravi Dhar, Norbert Schwarz, and Itamar Simonson (2007), "Preference Fluency in Choice," *Journal of Marketing Research*, 44 (3), 347–56.

Novet, Jordan (2017), "China's WeChat Captures Almost 30% of the Country's Mobile App Usage: Meeker Report," *CNBC*, (accessed March 1, 2018), [available at https://www.cnbc.com/2017/05/31/wechat-captures-about-30-percent-of-chinas-mobile-app-usage-meeker-report.html].

Oyserman, Daphna and Spike Wing-Sing Lee "(2007). "Priming 'Culture': Culture as Situated Cognition," in *Handbook of Cultural Psychology*, S. Kitayama and D. Cohen, eds. New York: Guilford, 255–79.

Packard, Grant and Jonah Berger (2016), "How Language Shapes Word of Mouth's Impact," *Journal of Marketing Research*, 54 (4), 572–88.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis (2010), "Running Experiments on Amazon Mechanical Turk," *Judgment and Decision Making*, 5 (5), 411–19.

Pathak, Bhavik, Robert Garfinkel, Ram D. Gopal, Rajkumar Venkatesan, and Fang Yin (2010), "Empirical Analysis of the Impact of Recommender Systems on Sales," *Journal of Management Information Systems*, 27 (2), 159–88.

Papies, E. K., I. Potjes, M. Keesman, S. Schwinghammer, and G. M. Van Koningsbruggen (2014), "Using Health Primes to Reduce Unhealthy Snack Purchases among Overweight Consumers in a Grocery Store," *International Journal of Obesity*, 38 (4), 597–602.

Pattie, David (2000), *Samuel Beckett*, London: Routledge.

Pennycook, Gordon, Adam Bear, Evan T. Collins, and David G. Rand (2020), "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines without Warnings." *Management Science*, articles in advance, 1-14.

Powell, Lisa M. and Frank J. Chaloupka (2009), "Food Prices and Obesity: Evidence and Policy Implications for Taxes And Subsidies," *Milbank Quarterly*, 87 (1), 229–57.

Puntoni, Stefano, Bart de Langhe, and Stijn M. J. van Osselaer (2009), "Bilingualism and the Emotional Intensity of Advertising Language," *Journal of Consumer Research*, 35 (6), 1012–25.

Ransbotham, Sam, Nicholas H. Lurie, and Hongju Liu. (2019), "Creation and Consumption of Mobile Word of Mouth: How Are Mobile Reviews Different?," *Marketing Science*, 38 (5), 773–92.

Ratner, Rebecca K. and Barbara E. Kahn (2002), "The Impact of Private versus Public Consumption on Variety-Seeking Behavior," *Journal of Consumer Research*, 29 (2), 246–57.

Read, Daniel and Barbara Van Leeuwen (1998), "Predicting Hunger: The Effects of Appetite and Delay on Choice," *Organizational Behavior and Human Decision Processes*, 76 (2), 189–205.

Redden, Joseph P. (2015), "Desire Over Time: The Multi-Faceted Nature of Satiation," in *The Psychology of Desire,* William Hoffmann and Loran F. Nordgen, eds. New York: The Guilford Press, 82–103.

Redden, Joseph P., Kelly L. Haws, and Jinjie Chen (2017), "The Ability to Choose Can Increase Satiation," *Journal of Personality and Social Psychology*, 112 (2), 186–200.

Ricci, Francesco, Lior Rokach, and Bracha Shapira (Eds.) (2015), *Recommender Systems Handbook*. New York: Springer.

Risen, Jane L. (2016), "Believing What I Do Not Believe: Acquiescence to Superstitious Beliefs and Other Powerful Intuitions," *Psychological Review*, 123 (2), 182–207.

Rotman, Jeff D., Mansur Khamitov, and Scott Connors (2018), "Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically," *Journal of Consumer Psychology*, 28 (2), 353–61.

Schwarz, Norbert (2012), "Feelings-as-Information Theory," in *Handbook of Theories of Social Psychology*, ed. Paul Van Lange, Arie Kruglanski, and Edward Tory Higgins, Sage, 289–308.

Sengupta, Jaideep, Darren W. Dahl, and Gerald J. Gorn (2002), "Misrepresentation in the Consumer Context," *Journal of Consumer Psychology*, 12 (2), 69–79.

Shah, James Y. and Arie W. Kruglanski (2003), "When Opportunity Knocks: Bottom-up Priming of Goals by Means and Its Effects on Self-Regulation," *Journal of Personality and Social Psychology*, 84 (6), 1109–22.

Shah, James Y., Arie W. Kruglanski, and Ron Friedman (2003), "Goal Systems Theory: Integrating the Cognitive and Motivational Aspects of Self-Regulation," in *Motivated Social Perception: The Ontario Symposium*, Vol. 9, Ontario Symposium on Personality and Social Psychology, ed. Steven J. Spencer and Steven Fein, Mahwah, NJ: Erlbaum, 247–75.

Shalvi, Shaul, Ori Eldar, and Yoel Bereby-Meyer (2012), "Honesty Requires Time (and Lack of Justifications)," *Psychological Science*, 23 (10), 1264–70,

Shiv, Baba, and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of consumer Research*, 26 (3), 278-292.

Shrum, L. J., T. M. Lowrey, David Luna, D. B. Lerman, and Min Liu (2012), "Sound Symbolism Effects across Languages: Implications for Global Brand Names," *International Journal of Research in Marketing*, 29 (3), 275–79.

Shu, Lisa L., Nina Mazar, Francesco Gino, Dan Ariely, and Max H. Bazerman (2012), "Signing at the Beginning Makes Ethics Salient and Decreases Dishonest Self-Reports in Comparison to Signing at the End," *Proceedings of the National Academy of Sciences,* 109 (38), 15197–200.

Simmons, Joseph P. and Leif D. Nelson (2006), "Intuitive Confidence: Choosing Between Intuitive and Nonintuitive Alternatives," *Journal of Experimental Psychology: General*, 135 (3), 409–28.

Simmons, Joseph P., Leif D. Nelson, Jeff Galak, and Shane Frederick (2011), "Intuitive Biases in Choice versus Estimation: Implications for the Wisdom of Crowds," *Journal of Consumer Research*, 38 (1), 1–15.

Spangher, Alexander (2015), "Building the Next New York Times Recommendation Engine," *The New York Times*, https://open.blogs.nytimes.com/2015/08/11/building-the-next-new-york-times-recommendation-engine/.

Sunstein, Cass R (2017), "Nudges that fail," *Behavioural Public Policy*, 1 (1), 4–25.

Suurmond, Robert, Henk van Rhee, and Tony Hak (2017), "Introduction, Comparison, and Validation of Meta-Essentials: A Free and Simple Tool for Meta-Analysis," *Research Synthesis Methods*, 8 (4), 537–53.

Thaler, Richard H. and Cass R. Sunstein (2008), *Nudge*, New Haven: Yale University Press.

Thompson, Valerie A. (2012), "Dual-Process Theories: A Metacognitive Perspective," In *Two minds: Dual Processes and Beyond*, Jonanthan Evans and Keith Frankish, eds. New York: Oxford University Press, 171-96.

Tintarev, Nava and Judith Masthoff (2015), "Designing and Evaluating Explanations for Recommender Systems," in *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira, eds. New York: Springer, 353–82.

Tuk, M. A., D. Trampe, and L. Warlop (2011), "Inhibitory Spillover: Increased Urination Urgency Facilitates Impulse Control in Unrelated Domains," *Psychological Science*, 22 (5), 627–33.

Tuk, Mirjam A., Peeter W.J. Verlegh, Ale Smidts, and Daniël H.J. Wigboldus (2019), "You and I Have Nothing in Common: The Role of Dissimilarity in Interpersonal Influence," *Organizational Behavior and Human Decision Processes*, 151, 49–60.

VanEpps, Eric. M., Julie S. Downs, and George Loewenstein (2016), "Advance Ordering For Healthier Eating? Field Experiments on the Relationship between the Meal Order–Consumption Time Delay and Meal Content," *Journal of Marketing Research*, 53 (3), 369–80.

VanEpps, Eric M., Julie S. Downs, and George Loewenstein (2016), "Calorie Label Formats: Using Numeric and Traffic Light Calorie Labels to Reduce Lunch Calories," *Journal of Public Policy and Marketing*, 35 (1), 26–36.

Van Osselaer, Stijn. M. J. and Chris Janiszewski (2012), "A Goal-Based Model of Product Evaluation and Choice," *Journal of Consumer Research*, 39 (2), 260–92.

Van Strien, Tatjana, Jan E. R. Frijters, Gerard Bergers, Peter B. Defares (1986), "The Dutch Eating Behavior Questionnaire (DEBQ) for Assessment of Restrained, Emotional, and External Eating Behavior," *International Journal of Eating Disorders*, 5 (2), 295–315.

Vives, Marc Lluís, Melina Aparici, and Albert Costa (2018), "The Limits of the Foreign Language Effect on Decision-Making: The Case of the Outcome Bias and the Representativeness Heuristic," *PLoS ONE*, 13 (9), 1–14.

Wang, Yajin, Jennifer L. Stoner, and Deborah Roedder John (2018), "Counterfeit Luxury Consumption in a Social Context: The Effects on Females' Moral Disengagement and Behavior," *Journal of Consumer Psychology*, forthcoming.

Wang, Weiquan and Izak Benbasat (2007), "Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs," *Journal of Management Information Systems*, 23 (4), 217–46.

Wansink, Brian and Pierre Chandon (2006), "Can "Low-Fat" Nutrition Labels Lead to Obesity?," *Journal of Marketing Research*, 43 (4), 605–17.

Wertenbroch, Klaus (1998), "Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice," *Marketing Science*, *17* (4), 317–37.

West, Patricia M., Christina L. Brown, and Stephen J. Hoch (2002), "Consumption Vocabulary and Preference Formation," *Journal of Consumer Research*, 23 (2), 120–35.

WHO (2016), *Obesity and Overweight*, http://www.who.int/mediacentre/factsheets/fs311/en/.

Wilcox, Keith, Hyeong Min Kim, and Sankar Sen (2009), "Why Do Consumers Buy Counterfeit Luxury Brands?" *Journal of Marketing Research*, 46 (2), 247–59.

World Cancer Research Fund International (WCRF) (2017), "Restrict food advertising and other forms of commercial promotion," *NOURISHING framework*.

Xu, Lizhen, Jason A Duan, and Andrew Whinston (2014), "Path to Purchase: A Mutually Exciting Point Process Model for Online Advertising and Conversion," *Management Science*, 60 (6), 1392–1412.

Yaniv, Ilan, Shoham Choshen-Hillel, and Maxim Milyavsky (2011), "Receiving Advice on Matters of Taste: Similarity, Majority Influence, and Taste Discrimination," *Organizational Behavior and Human Decision Processes*, 115 (1), 111–2.

Ying, Yuanping, Fred Feinberg, and Michel Wedel (2006), "Leveraging Missing Ratings to Improve Online Recommendation Systems," *Journal of Marketing Research*, 43 (3), 355–65.

Young, Eun Huh, Joachim Vosgerau, and Carey Morewedge (2014), "Social defaults: Observed Choices Become Choice Defaults," *Journal of Consumer Research*, 41, 746–60.

Zhang, Shi and Bernd H. Schmitt (2004), "Activating Sound and Meaning: The Role of Language Proficiency in Bilingual Consumer Environments," *Journal of Consumer Research*, 31 (1), 220–28.

Zhang, Shuai, Lina Yao, Aixin Sun, and Yi Tay (2018), "Deep Learning Based Recommender System: A Survey and New Perspectives," *ACM Computing Surveys*, 1 (1), 1–35.

Zhao, Qian, Gediminas Adomavicius, F. Maxwell Harper, Martijn Willemsen, and Joseph A. Konstan (2017), "Toward Better Interactions in Recommender Systems: Cycling and Serpentining Approaches for Top-N Item Lists," *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 1444–53.

Zhu, Feng and Xiaoquan (Michael) Zhang (2010), "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, 74, 133–48.

Zunick, Peter V., Jacob D. Teeny, and Russell H. Fazio (2017), "Are Some Attitudes More Self-Defining Than Others? Assessing Self-Related Attitude Functions and Their Consequences," *Personality and Social Psychology Bulletin*, 43 (8), 1136–49.

Summary (English)

The world consumers are living in is increasing globalized, digitalized, and saturated with temptations. As a result, consumers are surrounded by multiple languages, automated recommendations, and health reminders. Substantial research has been conducted to understand how consumers navigate in this world, by looking at how they evaluate products and make decisions in different language contexts, how to generate the best recommendations to individual users, and how to drive consumers towards a healthy diet.

Despite the wealth of research in these fields, I put forward that more research needs to be done in order to promote the welfare of consumers as a collective. First, although there is extensive research on (un)ethical behavior and bilingualism as separate fields, few have linked the two. Second, automated recommendations contribute tremendously to consumers' new discoveries, but it also engenders the concern that consumers may never be exposed to a broader set of new products. This problem has been considered by computer scientists looking at the algorithmic 'black box', yet rarely viewed as a problem addressable by marketing communications. Third, research on healthy eating has been dominated by how to improve the share of healthy choosers, leaving out a large proportion of cases in which consumers are deciding between two hedonic options that do not qualitatively yet quantitatively differ in healthiness. To fill these gaps, this dissertation uncovers the intricate influence of language on consumer dishonesty (Chapter 2), develops framing of recommendation as a tool to maximize consumer traffic to recommendations of novel products (Chapter 3), and pinpoints the self-control dynamic when consumers are choosing between the vice and the lesser vice (Chapter 4). By doing so, these essays take a step further in realizing the potential of consumer research to advance the collective welfare of consumers. Meanwhile, the advancement is accompanied by greater benefits to companies, in terms of the management of losses incurred by dishonest consumers, consumer engagement with recommendations, and the marketing of lesser-vice products.

Samenvatting (Nederlands)

De wereld waarin consumenten leven raakt in toenemende mate geglobaliseerd, gedigitaliseerd en verzadigd met verleidingen. Dit zorgt ervoor dat consumenten omringd zijn door meerdere talen, geautomatiseerde aanbevelingen en reminders over hun gezondheid. Er is veel onderzoek gedaan om te begrijpen hoe consumenten hun weg vinden in deze wereld, door te kijken naar de manier waarop zij producten beoordelen en besluiten nemen in verschillende talige contexten, op welke manier de beste aanbevelingen aan individuele gebruikers kunnen worden gegenereerd, en hoe consumenten kunnen worden aangezet tot een gezond voedingspatroon.

Ondanks de overvloed aan onderzoek op deze gebieden stel ik dat er meer onderzoek nodig is om het welzijn van consumenten als collectief te bevorderen. Ten eerste is er weliswaar uitgebreid onderzoek gedaan naar (on)ethisch gedrag en tweetaligheid als afzonderlijke gebieden, maar beide onderwerpen zijn maar zelden aan elkaar verbonden. Ten tweede dragen geautomatiseerde aanbevelingen weliswaar enorm bij aan nieuwe ontdekkingen door consumenten, maar het fenomeen geeft ook aanleiding tot de bezorgdheid dat consumenten nooit worden blootgesteld aan een breder scala van nieuwe producten. Dit probleem is bestudeerd door computerwetenschappers die zich bezig houden met de algoritmische 'zwarte doos', maar wordt zelden gezien als een probleem dat kan worden aangepakt door marketingcommunicatie. Ten derde wordt onderzoek naar gezonde voeding gedomineerd door de vraag hoe het aandeel van mensen dat gezonde keuzes maakt kan worden vergroot. Daarbij wordt een groot deel van de gevallen buiten beschouwing gelaten waarin consumenten tussen twee hedonistische opties kiezen die qua gezondheid niet kwantitatief maar wel kwalitatief van elkaar verschillen. Om deze hiaten op te vullen wordt in dit proefschrift de complexe invloed blootgelegd die taal heeft op de oneerlijkheid van de consument (hoofdstuk 2). Verder wordt een framing van aanbevelingen ontwikkeld als instrument om consumentenverkeer naar aanbevelingen van nieuwe producten te maximaliseren (hoofdstuk 3) en wordt de dynamiek van zelfbeheersing getoond wanneer consumenten kiezen voor de minste van twee kwaden (hoofdstuk 4). Hiermee gaan deze essays een stap verder in de realisatie van de mogelijkheden van consumentenonderzoek om het collectieve welzijn van consumenten te verbeteren. Ondertussen gaat deze verbetering gepaard met grotere voordelen voor

bedrijven met betrekking tot het beheer van verliezen veroorzaakt door oneerlijke consumenten, de betrokkenheid van consumenten bij aanbevelingen en de marketing van minder ongezonde producten.

Hoofdstuk 2: Taal en oneerlijkheid van consumenten

Consumenten liegen en bedriegen in sterk uiteenlopende domeinen, bijvoorbeeld als het gaat om verzekeringsclaims, het retourneren van producten, datingprofielen en belastingaangiftes. Hoewel de meeste consumenten liegen in hun moedertaal (L1), liegen ze ook in hun tweede taal (L2): reizigers geven informatie verkeerd weer wanneer ze formulieren voor verloren bagage invullen; immigranten geven geen informatie over medische aandoeningen aan ziektekostenverzekeraars; en buitenlandse consumenten claimen kortingen door middel van een valse online-identiteit.

In hoofdstuk 2 onderzoek ik het effect van taal (L2 vs. L1) op liegen. Negen studies, betrekking hebbend op verschillende talen en consumptiedomeinen, stellen de bevindingen van recent psychologisch onderzoek ter discussie die erop wijzen dat het gebruik van een tweede taal leidt tot minder liegen. Het gebruik van een tweede taal leidt niet in alle gevallen tot meer eerlijkheid, maar zwakt eerder de intuïtieve voorkeur af die mensen hebben voor liegen of het vertellen van de waarheid. Verder wordt het effect van taal vergroot door sterkere intuïtieve neigingen naar (on)eerlijkheid. Ook draagt een extra groot gevoel van onzekerheid in tweedetaalcontexten bij aan het effect van taal op de oneerlijkheid van consumenten. Deze genuanceerde bevindingen markeren het feit dat taalcontexten oneerlijk gedrag niet automatisch activeren of remmen. In plaats daarvan is het de wisselwerking tussen de morele intuïtie, die ten minste gedeeltelijk is gevormd in de context van de moedertaal, en de onzekerheid die eigen is aan het gebruik van een vreemde taal, die het effect van taal vormgeeft.

Hoofdstuk 2 biedt nieuwe inzichten in fenomenen die van groot belang zijn voor managers en beleidsmakers. Liegen komt in talloze situaties voor en om vele verschillende redenen. Ik richt me op opportunistische leugens die materiële voordelen opleveren voor consumenten. Hoewel leugens die niet-materiële voordelen opleveren (bijv. impressiemanagement) ook vaak voorkomen, zijn met name leugens die zijn gemotiveerd door materieel gewin relevant voor het bedrijfsleven en de maatschappij, gezien de kosten

voor organisaties, overheden en eerlijke consumenten. Ik richt me ook met name op *private lying*; dat wil zeggen, liegen in situaties waarin geen sprake is van face-to-face communicatie. De bovengenoemde condities kenmerken een groot deel van het misleidende consumentengedrag dat voorkomt in computergestuurde omgevingen. Het bestuderen van het effect van taal onder deze omstandigheden draagt bij aan een beter begrip van de eerlijkheid van de consument in het digitale tijdperk.

Hoofdstuk 3: Framing van aanbevelingen voor het aantrekken van meer click-throughs

Veel bedrijven geven consumenten productaanbevelingen die zijn gegenereerd door algoritmische aanbevelingssystemen: Spotify en Netflix raden liedjes of films aan en TripAdvisor en Yelp geven aanbevelingen voor hotels of restaurants. Amazon doet suggesties welke producten consumenten mogelijk zouden willen kopen, en *The New York Times* beveelt verschillende nieuwsartikelen aan. Deze gepersonaliseerde aanbevelingen helpen consumenten bij het vinden van aanbod waarin ze waarschijnlijk geïnteresseerd zijn en zorgen voor meer loyaliteit.

Om de nauwkeurigheid van deze algoritmische aanbevelingen te verbeteren, hanteren aanbevelingssystemen vaak een hybride aanpak die rekening houdt met zowel gemeenschappelijke voorkeuren van consumenten als gemeenschappelijke kenmerken van producten. Elke aanbeveling is dus gebaseerd op input over zowel gebruikers als producten; het is niet eenvoudig in woorden uit te leggen waar de aanbeveling op gebaseerd is. Sommige bedrijven geven reeds aan dat hun aanbevelingen *user-based* zijn, door de nadruk te leggen op de overlap tussen consumentenvoorkeuren, zoals "Consumenten die dit product bekeken, bekeken ook …" van Amazon en "Klanten bekeken ook …" van Netflix. Sommige andere bedrijven benadrukken daarentegen dat aanbevelingen *item-based* zijn, zoals "Vergelijkbaar met [waar je naar hebt geluisterd]" van Spotify en "More in Health" van *The New York Times*.

Dit leidt tot de vraag welke vorm van framing van aanbevelingen, user-based of item-based, effectiever is bij het triggeren van clicks. Mijn centrale stelling is dat user-based framing, in vergelijking met item-based framing, de consument laat weten dat de aanbeveling wordt gegenereerd door middel van *product matching* (het aanbevolen

132

product is vergelijkbaar met het gekozen product) maar ook wijst op *taste matching* tussen gebruikers (het gekozen product dat je leuk vindt, wordt ook leuk gevonden door andere gebruikers). Consumenten halen informatie uit voorkeuren van vergelijkbare anderen om hun eigen waardering van onbekende producten te voorspellen. Deze informatie moet de klant dus een extra garantie bieden dat het product aansluit bij de eigen voorkeuren. Om deze reden voorspel ik dat aanbevelingen die zijn geframed als user-based (in tegenstelling tot item-based) meer click-throughs genereren, als de consument tenminste de indruk heeft dat de aanbevelingen goed aansluiten bij de eigen voorkeuren.

Ik test mijn stelling met zes onderzoeken die verschillende soorten gegevens (veld-, gedrags- en scenario-gebaseerd onderzoek) en consumptiedomeinen (artikelen, schilderijen en boeken) beslaan. Bij deze verschillende methoden en productdomeinen vind ik steeds dat user-based framing voor meer click-throughs op aanbevelingen zorgt dan item-based framing, wanneer consumenten de indruk hebben dat de voorkeuren van anderen met die van henzelf overeenkomen. Ik stel verder drie randvoorwaarden voor die er voor kunnen zorgen dat de ontvanger van de aanbeveling de *taste matching* als niet succesvol ervaart, waardoor het voordeel van de user-based framing ten opzichte van de item-based framing afneemt. Deze randvoorwaarden zorgen op hun beurt voor belangrijke handvatten voor bedrijven over hoe de framing van hun aanbevelingen kan worden aangepast om het aantal click-throughs te maximaliseren. Belangrijk is dat de aanbevelingen in al deze onderzoeken betrekking hebben op producten waarmee de consument niet bekend is. Dit is een ontwerpelement dat inzicht verschaft in de manier waarop de totale consumptieve diversiteit kan worden vergroot en hoe nieuwe producten op de markt gebracht kunnen worden.

Hoofdstuk 4: Wanneer de goede keuze de minst slechte is

Om overgewicht en obesitas te bestrijden, is er een aanzienlijke hoeveelheid onderzoek gedaan naar de drijvende krachten achter de keuzes van consumenten tussen gezonde en ongezonde voeding, vaak ook wel goede en slechte keuzes genoemd. Het meeste onderzoek over zelfbeheersing met betrekking tot voedselkeuzes is gebaseerd op een model waarin mensen een keuze maken tussen een gezonde en een ongezonde keuze.

Hoofdstuk 4 vult deze literatuur aan door te kijken naar een weinig onderzochte setting waarin consumenten kiezen tussen een ongezond voedingsproduct (bijv. gewone chips) en een iets minder ongezonde versie ervan (bijv. light chips). In dergelijke situaties kiezen consumenten dus tussen slecht en minder slecht, in plaats van tussen goed en slecht. Het bestuderen van deze context is van groot praktisch belang, omdat minder slechte keuzes in overvloed aanwezig zijn in de markt, zoals kaascrackers met minder vet en suikervrije chocola. Deze producten bevatten normaal gesproken minder calorieën terwijl ze qua smaak zo veel mogelijk op het gewone product proberen te lijken.

De belangrijkste bijdrage van hoofdstuk 4 is dat het laat zien dat keuzeconflicten tussen slechte en minder slechte keuzes worden gekenmerkt door hun eigen dynamiek. Ten eerste worden keuzes tussen slecht en minder slecht veel minder gekenmerkt door het verschil in directe behoeftebevrediging. Ten tweede is het verschil in calorische dichtheid het enige opvallende verschil in de setting 'slecht versus minder slecht'. Deze voorgestelde verschillen zijn gebaseerd op goal systems theory en leiden tot twee voorspellingen. Een van de voorspellingen heeft betrekking op de zinloosheid van interventies die bedoeld zijn om de slechte keuze minder aantrekkelijk te laten lijken. Ik beargumenteer dat de toevoeging van een vertraging tussen keuze en consumptie, een populaire nudging-tactiek in de richting van gezond eten, in de keuzesetting 'slecht versus minder slecht' niet meer genoeg is om het perspectief van de consument te verleggen van directe behoeftebevrediging naar welzijn op de lange termijn. Het ruim van tevoren maken van voedselkeuzes zou er dan niet voor zorgen dat er minder slechte keuzes worden gemaakt. Daar tegenover staat mijn voorspelling dat consumenten die zich structureel meer bezighouden met calorie-inname meer geneigd zijn om voor de minder slechte keuzes te gaan. Het derde kenmerk van de keuzesetting 'slecht versus minder slecht' is dat de geconsumeerde hoeveelheden van belang zijn voor het algehele succes van zelfbeheersing. Met het oog hierop schets ik de concurrerende voorspellingen over de manier waarop keuzes invloed uitoefenen op geconsumeerde hoeveelheden.

Om deze stellingen te testen heb ik twee experimenten uitgevoerd waarin het moment van keuze werd gemanipuleerd, individuele verschillen in terughoudendheid bij het eten (een indicatie van de mate waarin op calorieën wordt gelet) werden gemeten, en de

daadwerkelijke keuzes en geconsumeerde hoeveelheden werden vastgelegd. De onderzoeken bevestigen de voorspellingen door te laten zien dat het loskoppelen van de directe bevrediging van voedselkeuzes door van tevoren te bestellen er niet toe leidt dat consumenten de slechte keuze vervangen door de minder slechte keuze; in plaats daarvan zorgen individuele verschillen in de mate waarin men zich bezighoudt met de calorie-inname voor de keuze voor het vervangende product. Belangrijk is ook dat de onderzoeken aantonen dat consumenten slechts beperkte controle uitoefenen over de geconsumeerde hoeveelheden nadat ze een keuze hebben gemaakt. Deze bevindingen worden besproken in relatie tot eerder onderzoek en vervolgonderzoek.

About the Author

Phyliss Jia Gai (盖 嘉) was born in Liaoning, China on September 18, 1991. Before moving to the Netherlands, she received her bachelor's degree in psychology with first class honors from the Chinese University of Hong Kong and her master's degree in social sciences from the University of Chicago. Phyliss joined RSM to pursue her PhD in marketing in 2015. She studies consumer behavior and her work has been presented at a number of conferences and research seminars in different countries. The article version of the chapter "Making recommendations more effective with framings" has been published on Journal of Marketing and the chapter "Language and dishonesty" under revision at Journal of Consumer Research. In Spring 2019, Phyliss visited Columbia Business school for research collaboration. Apart from research, she has supervised over 50 Bachelor students' and Master students' theses at RSM. In Fall 2020, she will join the Guanghua School of Management at Peking University in China as an assistant professor in marketing. She will continue her research in digital consumption, self-control, and ethics.

Portfolio

MANUSCRIPTS

Published and Under Revision

Gai, Phyliss Jia and Anne-Kathrin Klesse (2019), "Making Recommendations More
Effective through Framing: Impacts of User- versus Item-based Framings on
Recommendation Click-throughs", *Journal of Marketing,* 83 (6), 61-75.

Gai, Phyliss J. and Stefano Puntoni, "Lies of Bilingual Consumers", revising for
resubmission at *Journal of Consumer Research.*

In Preparation

Gai, Phyliss J., Mirjam Tuk, and Steven Sweldens, "When Virtues are Lesser Vices: The
Impact of Advance Ordering and Restrained Eating on Choice and Consumption"

Gai, Phyliss J. and Amit Bhattacharjee, "Saints or Sinners? How Self-control Affects
Evaluation of Moral Character"

Research in Progress

"The Impact of Mobile Devices on the Sharing of Misinformation", with Gita V. Johar.

"Asymmetrical Influences of Health and Taste on Within-Category Food Choices" with
Mirjam Tuk and Steven Sweldens.

"Consumer Beliefs in Algorithms Learning Capability and Behavioral Consequences",
with Anne-Kathrin Klesse and Eugina Leung.

CONFERENCE PRESENTATIONS

Making Recommendations More Effective through Framing

October 2019, ACR, Atlanta, USA

June 2019, La Londe Conference, France

May 2019, TADC, London Business School, UK

June 2018, European ACR, Ghent, Belgium

Lies of Bilingual Consumers

October 2018, ACR, Dallas, USA

November 2017, SJDM, Vancouver, Canada

SELECT HONORS AND AWARDS

AMA-Sheth Doctoral Consortium Fellow, 2019

ERIM competitive 5th-year PhD funding, Erasmus University, 2019-20

Erasmus Trustfonds Grant for Research Visit, 2019

EMAC Doctoral Colloquium Fellow, 2018 & 19

TEACHING AND ADVISING

Master Thesis in Business Administration, Coach and co-reader, RSM, 2015-20

Bachelor Thesis in Business Administration, Instructor, RSM, 2018

PROFESSIONAL AND UNIVERSITY SERVICES

Trainee reviewer, Journal of Consumer Research, 2018-20

Paper reviewer, La Londe Conference, 2019

Competitive & working paper reviewer, European ACR, 2018

Competitive & working paper reviewer, SCP, 2018

ACR volunteer, Berlin, Germany, 2016

Behavioral lab administrator, Erasmus University, 2016-17

INVITED TALKS

Academic

Peking University

Renmin University

Fudan University

Shanghai University of Finance and Economics

University of Miami

Imperial College Business School

Warwick Business School

Amsterdam Business School

Tilburg University

University Carlos III de Madrid

Industrial

Frontiers in Marketing, RSM, Erasmus University

Coolblue (with Anne Klesse)

PROFESSIONAL AFFILIATIONS

American Marketing Association

Association for Consumer Research

Society for Consumer Psychology

Society for Judgment and Decision Making

SELECT GRADUATE COURSEWORK

Marketing and Behavioral Research

Behavior Economics (with Prof. G. Loewenstein at NHH)

Behavioral Decision Theory (with Prof. P. Wakker)

Current Topics in Marketing Research (with Prof. S. Puntoni et al.)

Specialization in Consumer Behavior (with Dr. M. Tuk et al.)

Topics in Judgment and Decision-Making (with Prof. W. Goldstein at Uchicago)

Statistics and Methods

Structural Equation Modeling (with Prof. R. Pieters at Tilburg University)

Experimental Design (with Dr. M. Wubben)

Data Visualization, Web Scraping, and Text Analysis in R (with Dr. J. Roos)

Applied Hierarchical Model (with Prof. S. Raudenbush at UChicago)