UNIVERSITA’ COMMERCIALE “LUIGI BOCCONI”

PhD SCHOOL

PhD program in Business Administration and Management

Cycle: 35

Disciplinary Field (code): SECS-P/07

Less Explored Aspects of Disfluency: Consumer Judgment and Decision Making

Advisor: Kurt Munz

PhD Thesis by

Shahryar MOHSENIN

ID number: 3105048

Year 2024

Extensive research has explored how disfluency affects product judgments and evaluations across various consumption domains such as brand evaluation, price perception, and choice deferral. Generally, experiencing disfluency leads to less favorable judgments and evaluations, often resulting in a delay in decision-making. These effects occur because individuals tend to attribute the negative feeling of difficulty processing information to their evaluation of the product itself. While the role of disfluency in consumer judgments and product evaluation has been well-documented, certain aspects of disfluency have received less attention in the context of consumer choice.

In Chapter 1, I examined the social aspect of subjective disfluency, specifically the gender of a voice. The aim was to determine whether individuals' perception of a digital voice assistant's gender would affect their judgment of a product described by gender-ambiguous voices compared to gender-obvious voices. The findings revealed that people responded unfavorably when assessing products described by narrators with voices that were not clearly identifiable as male or female. This negative reaction stemmed from a sense of discomfort associated with the challenge of determining the narrator's gender, referred to as social disfluency. The negative effect was primarily attributed to a lack of familiarity with gender-ambiguous voices, as they are less commonly encountered in daily life, rather than social categorization or prejudice against LGBTQ individuals. However, the results indicated that increased exposure to gender-ambiguous voices could help overcome initial negative reactions.

In Chapter 2, I aimed to examine the role of subjective disfluency in consumer information processing styles. In contrast to previous findings on the association between disfluency and dual processing modes, this research demonstrated that disfluency leads consumers to simplify complex decisions by relying on mental shortcuts, known as heuristics, rather than engaging in systematic processing of all available information. When information is disfluent and difficult to process, consumers tend to rely more heavily on heuristic product attributes such as brand, country of origin, or recommendations. This tendency to rely on heuristics in the face of disfluency is particularly prominent among fast and effortless decision-makers; however, the effect vanishes when they are informed about the source of disfluency.

Finally, in Chapter 3, I investigated the role of framing information as promotion- or prevention-focused in consumer purchase decisions when subjective difficulty in processing information is present. This research aimed to demonstrate the benefits of framing information as promotion benefits (rather than prevention benefits) under conditions of subjective difficulty, drawing upon regulatory focus theory. It was predicted that when information is hard to process, individuals are more likely to rely on prevention-based information about products to justify their decision to reject. As a result of regulatory mismatch under disfluency, featuring promotion-focused claims reduces choice deferral compared to prevention-based information, which increases regulatory fit.

Overall, these studies shed light on different aspects of disfluency in judgment and decision-making, examining the impact of gender-ambiguous voices, heuristic processing under the feeling of disfluency, and the role of message framing effects on consumer evaluations and choices.

CHAPTER 1

Gender-Ambiguous Voices and Social Disfluency

Abstract

Recently, gender-ambiguous (non-binary) voices have been added to voice assistants to combat gender stereotypes and foster inclusion. However, if people react negatively to such voices, these laudable efforts may be counterproductive. In five preregistered studies (N = 3,684) we find that people do react negatively, rating products described by narrators with gender-ambiguous voices less favorably than when they are described by clearly male or female narrators. This is due to the voices creating a feeling of unease related to difficulty understanding the gender of the narrator, what we call social disfluency, that spills over to affect evaluations of the products being described. These effects are best explained by low familiarity with voices that sound ambiguous. Thus, initial negative reactions can be overcome with more exposure.

Keywords: fluency, voice, gender, familiarity, social categorization, prejudice

In 2022, Apple debuted a new voice for its Siri voice assistant that did not sound obviously male or female. This new gender-ambiguous voice was introduced in the wake of criticism levied by the United Nations and others that subservient voice assistants, being overwhelmingly female, reinforced harmful gender stereotypes (West et al., 2019). The new voice could avoid this issue and help introduce a wide audience to a non-binary voice. While these goals are laudable, open questions remain as to how people may react to hearing voices like this one. Will people react *negatively* to hearing a gender-ambiguous voice? If so, will introducing such voices to digital assistants like Siri be a positive development?

On the one hand, reactions to gender-ambiguous voices may be similar to those with obvious genders. Voice-only communication offers richer social information about a narrator than printed text. For example, in hiring situations, evaluators rated job candidates as more competent when they heard their presentations rather than read them (Schroeder & Epley, 2015). This is because speakers use paralinguistic cues such as intonation, pausing, and inflection to convey emotions (Kruger et al., 2005) which can foster greater connections between people, even when the speaker is not actually human (Schroeder & Epley, 2016). Since gender-ambiguous digital voices can communicate with similar paralinguistic cues as gender-obvious voices, reactions to them may be similar too.

On the other hand, reactions may differ. Vocal cues automatically signal various aspects of one's identity, including gender (Cartei et al., 2020). People can use this information to apply gender stereotypes (Ko, Judd, & Blair, 2006), even to the synthetic voices that sound non-human (Nass et al., 1997; Eyssel & Hegel, 2012). For example, laboratory participants preferred female computerized voices for stereotypically feminine tasks such as designing and male computerized voices for stereotypically masculine tasks such as computation and engineering (Nass et al., 1997).

But how will people react when they cannot easily evaluate the gender of the speaker from their voice? A substantial body of research in the area of processing fluency suggests that ease or difficulty experienced when processing information affects people's judgments (Alter & Oppenheimer, 2009; Reber, Winkielman, & Schwarz, 1998). Specifically, people generally judge hard-to-process (disfluent) information less favorably than information presented in easier-to-process formats. For example, when a speaker is difficult to understand due to a foreign accent (Lev-Ari & Keysar, 2010) or poor audio quality (Newman & Schwarz, 2018), people tend to downgrade their assessment about the quality of the information presented. Could a similar process also apply to difficulty processing *social* information? That is, when the information itself is easy to understand, could difficulty processing the social information, such as the gender of the speaker's voice, lead to a similar feeling of disfluency? We predict that it could, and we call difficulty processing such social information *social disfluency*.

There are at least two underlying mechanisms that could lead to social disfluency. First, since non-binary voices are less commonly encountered in day-to-day life, people may react negatively toward them merely because they seem unfamiliar. It is well-documented that familiarity leads to liking (Zajonc, 1968). In contrast, a low level of familiarity has been identified as a source of disfluency (Alter & Oppenheimer, 2008; Song & Schwarz, 2009; Reber, Schwarz, & Winkielman, 2004). Thus, if a non-binary voice sounds unfamiliar, people may experience more disfluency compared to when they encounter a more familiar gender-obvious voice.

Second, it may be difficult to categorize a non-binary voice, and this feeling of difficulty could spill-over into evaluations of the speaker or evaluations about what the speaker is talking about (Nass & Brave, 2005). Although some researchers have shown that gender perception for voices is not categorical (Mullennix et al., 1995), people generally do not like it when they cannot easily categorize something (Peracchio & Tybout, 1996). Furthermore, if a stimulus cannot be categorized as an example of an already defined category, then associations based on a category cannot be quickly retrieved and applied (Sujan, 1985). Therefore, difficulty categorizing the gender of the speaker may create a negative affective state that could color judgments about the speaker's words.

Indeed, this idea about categorization has been tested for visual perceptions. For example, Winkielman, Olszanowski, and Gola (2015) asked participants to make attractiveness and trust judgments for faces that varied in gender-typicality. Contrary to our hypothesis, they found that gender-ambiguous faces were rated *more* favorably than gender-obvious ones overall. However, when participants were first tasked with categorizing the gender of the face, ratings for gender-ambiguous faces declined. They argued that the categorization task was difficult for gender-ambiguous faces, leading to disfluency that depressed evaluations.

While both low familiarity and difficulty categorizing could reduce social fluency, as we have proposed, an alternative explanation may be that people harbor prejudice against members of the LGBTQ community (Lick & Johnson, 2013). Since non-binary voices are often associated with this community (Schwarz et al., 2018), negative reactions to a non-binary voice may reflect this prejudice, rather than be caused by social disfluency, as we have suggested.

In this research, we disentangle the above-mentioned possibilities. We show that narration in a gender-ambiguous voice can lead to less favorable judgments about consumer products, consistent with our theorizing about social disfluency. Furthermore, we show that these adverse effects occur above and beyond prejudice and are best explained by low levels of familiarity with non-binary voices.

Documenting this phenomenon and investigating the root cause is important for several reasons. First, we contribute to theory related to disfluency. We explore how social disfluency for voices can spill over to affect product judgements, joining related work exploring how facial-gender perception can affect social judgments (e.g., Owen et al., 2016; Winkielman, et al., 2015). Second, the research contributes to our understanding of prejudice. We suggest that social disfluency may be a contributing factor leading to outcomes that appear to be prejudiced, joining research that has explored ambiguity in areas such as race (Eberhardt et al., 2003; Hugenberg & Bodenhausen, 2004; Lick & Johnson, 2015). Understanding the root causes can also offer potential solutions for how to alleviate these negative responses. Our research supports the idea that advocating for more exposure to non-binary voices like the ones we study will be an effective tool to help reduce the negative reactions people may initially experience, aligning with research on clinical therapies (Birtel & Crisp, 2012).

Study 1

Do people judge products less favorably when they have difficulty discerning the gender of the voice describing them? Study 1 investigated this question by comparing the gender-ambiguous voice for Siri against both male and female versions of Siri's voice.

Method

For all studies, data, preregistrations and survey materials are posted on ResearchBox (https://researchbox.org/695). Additional analyses are available in the Supplemental Online Materials (SOM) when specified.

Participants and Design

Because of a lack of comparable prior studies, we could not base our target sample size on an empirical effect size. We therefore opted for a relatively large target sample size (N = 600; 300 per between-participants condition). Six hundred and two Mturk participants (55.6% female; age: M = 40.66 years; SD = 12.33) were paid a small sum for completing this pre-registered study (https://aspredicted.org/Y89_JPJ). No participants were excluded. All participants rated two hypothetical brands of toothpaste in a 2 (*gender-ambiguity*: gender-ambiguous, gender-obvious; within) × 2 (*gender of the gender-obvious voice*: female vs. male; between) mixed design. *Stimulus order* (which toothpaste was presented first) and *voice-gender order* (whether the gender-ambiguous or gender-obvious voice was first) were independently counterbalanced.

Procedure

All participants completed a CAPTCHA bot check at the beginning of the study. After completing the CAPTCHA, participants were randomly assigned to conditions. They were asked to rate their perception of each of two toothpaste brands. Toothpaste was chosen as a stimulus because past research found this product category to be gender-neutral (Iyer & Debevec, 1986). In each condition, participants heard two voices with different genders, the Siri gender-ambiguous voice and one of either a male or a female voice that Siri can be made to sound like. The voices described two similar toothpaste products. We predicted that when the toothpaste was described by the gender-obvious voice (either male or female) it would be preferred to the toothpaste described by the gender-ambiguous voice, regardless of whether the gender-obvious voice was male or female-sounding. The dependent variable was the judgment of liking for each toothpaste.

First, participants heard 12-seconds of audio information about one of the two brands, either in a gender-obvious voice (male or female) available for Siri or the gender-ambiguous voice for Siri depending on counterbalancing condition. They could listen as many times as they wished. Then, on a subsequent screen, they were asked to answer the primary dependent variable, *product judgment*, about the first toothpaste brand by asking on a 7-point scale ranging from 1 (not at all) to 7 (very much), "How much do you like this toothpaste?" Afterward, they heard about the second brand presented with the other voice and were subsequently asked to answer their judgment about the second brand.

Next, they were asked to evaluate the ease or difficulty evaluating the gender of each voice, *social disfluency*, on a 7-point scale, asking "How difficult was it to evaluate the gender of each speaker?" (1 = Not at all difficult; 7 = Extremely difficult). In addition, we administered a self-reported scale to measure *difficulty processing the information* as an alternative explanation of the effect by asking, "How difficult was it to evaluate the information provided by each speaker?" (1 = Not at all difficult; 7 = Extremely difficult). We also asked participants to evaluate the gender of each speaker's voice as a manipulation check on a 7-point scale, "What is your best guess about the gender of the speaker?" (1 = Clearly male; 4 = Ambiguous; 7 = Clearly female). We recoded the response as a measure of *voice-gender ambiguity* based on the distance from the midpoint (1 = Highly clear; 4 = Highly ambiguous). That is, a response at the midpoint of the gender question (4) was coded as 4, one position away (3 or 5) was coded as 3, two positions away (2 or 6) was coded as 2, and three positions away (1 or 7) was coded as 1. In the end, participants were asked to answer some demographic questions (spoken language at home, gender, and age).

Results

For ease of interpretability, in all studies we collapsed across the counterbalance factors in all analyses. See the SOM for analyses with these factors included.

Manipulation Check

People perceived Siri's gender-ambiguous voice ($M_{\text{ambiguous}} = 2.08$, $SD = 1.13$) as more ambiguous compared to Siri's female voice ($M_{\text{female}} = 1.44$, $SD = 0.77$; $t(300) = 12.55$, $p < .001$, $d = 0.72$, 95% CI [0.60, 0.85]) and Siri's male voice ($M_{\text{male}} = 1.36$, $SD = 0.69$; $t(299) = 8.56$, $p < .001$, $d = 0.52$, 95% CI [0.40, 0.64]).

Product Judgment

Most importantly, as predicted and depicted in Figure 1, a mixed ANOVA on product liking with *gender-ambiguity* and *gender of the gender-obvious voice* as factors revealed that there was a significant main effect of *gender-ambiguity* on product liking, $F(1, 599) = 58.41$, $p < .001$, $\eta_p^2 = .089$. A Bayesian repeated measures ANOVA strongly supported our hypothesis on product judgement, $BF_{10} = 4.25 \times 10^{10}$. The result of pairwise comparison tests revealed that people rated the product presented in a gender-ambiguous voice ($M_{\textit{male (ambiguous)}} = 4.76$, $SD = 1.23$; $M_{\textit{female (ambiguous)}} = 4.67$, $SD = 1.15$) less favorably than a similar option when it was presented in both a male voice ($M_{\textit{male}} = 5.04$, $SD = 1.21$; $F(1, 599) = 21.72$, $p < .001$, $\eta_p^2 = .035$; $BF_{10} = 3.2 \times 10^3$) and a female voice ($M_{\textit{female}} = 5.04$, $SD = 1.16$; $F(1, 599) = 37.81$, $p < .001$, $\eta_p^2 = .059$; $BF_{10} = 1.55 \times 10^6$) . The main effect of the *gender of the gender-obvious voice* ($F(1, 599) = .301$, $p = .58$, $BF_{10} = .072$) and the interaction ($F(1, 599) = 1.09$, $p = .296$, $BF_{10} = .032$) were not significant. See the SOM for additional details about the Bayesian analysis.

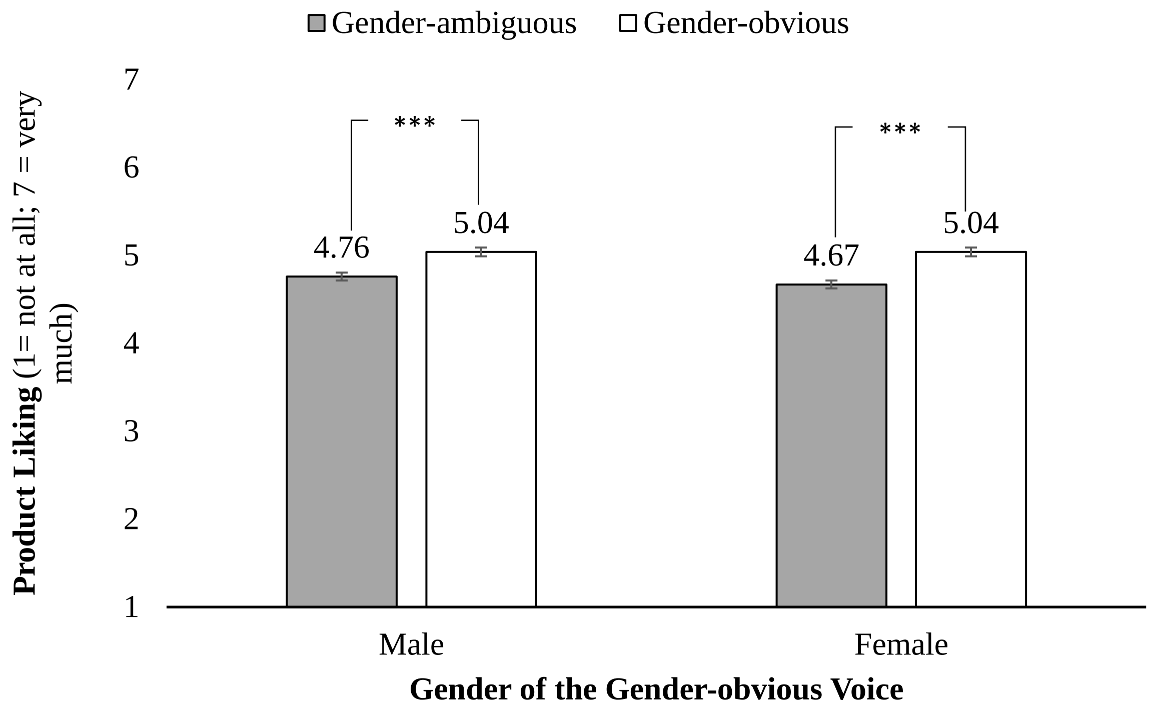


Fig. 1. Presenting product information in a gender-ambiguous voice (gray bars) resulted in less liking of the product when compared to both male (left bars) and female voices (right bars). Error bars represent standard errors. (*** $p < .001$).

Examining First Stimulus as Between-Participants Test. While not pre-registered, we tested the ratings of the first options presented to participants (before exposure to a second voice) as a between-participants test of the main hypothesis. A contrast test between voice gender conditions demonstrated that those who heard the information in a gender-ambiguous voice ($M_{\text{first (ambiguous)}} = 4.81$, $SD = 1.12$) evaluated it significantly less favorably than those who heard a gender-obvious voice ($M_{\text{first (gender-obvious)}} = 5.03$, $SD = 1.15$; $t(599) = 2.32$, $p = .021$, $d = 0.19$, 95% CI [0.03, 0.35]).

Social Disfluency

As expected, a paired-samples t-tests revealed that people found it harder to evaluate the gender of Siri's gender-ambiguous voice ($M_{ambiguous} = 2.53$, $SD = 1.77$) relative to Siri's gender-obvious voices ($M_{gender\text{-}obvious}$ = 1.52, SD = 1.14; t (600) = 15.01, $p < .001$, d = 0.61, 95% CI [0.52, 0.70]). As depicted in Figure 2, differences in social disfluency partially mediated the effect of gender-ambiguity on product judgment. A bootstrapped within-participants mediation analysis with 10,000 iterations (MEMORE; Montoya & Hayes, 2017) revealed that the decrease in product liking was explained by an increase in social disfluency (difficulty evaluating the gender of a voice) due to the voice gender ($b_{\text{total\_effect}}$ = - 0.32, SE =0.04, 95% CI [- 0.41, - 0.24], t(600) = - 7.64 $p < .001$; $b_{\text{direct\_effect}}$ = -24, SE =0.05, t(598) = - 4.71 $p < .001$, 95% CI [- 0.34, - 0.14]; $b_{\text{indirect\_effect}}$ = -0.08, SE =0.03, 95% CI [- 0.14, -0.03], p =.004). This result supports our proposal that social disfluency can explain negative reactions to gender-ambiguous voices.

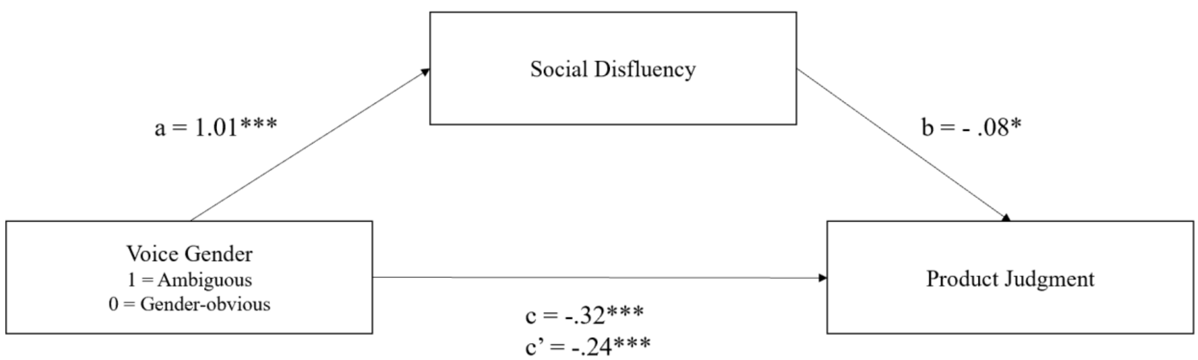


Fig. 2 – Social disfluency (difficulty evaluating the gender of each speaker) partially mediated the effect of the speaker's voice on product judgment. * $p < .05$; *** $p < .001$

Difficulty Processing the Information

In contrast to our expectation, a paired-samples t-test revealed that people perceived Siri's gender-ambiguous voice as harder to process ($M_{ambiguous}$ = 1.64, SD = 1.17) relative to Siri's gender-obvious voices ($M_{gender\text{-}obvious}$ = 1.43, SD = 1.00; t(600) = 5.93, $p < .001$, d = 0.24, 95% CI [0.16, 0.32]). The indirect effect of gender-ambiguity on liking judgment through difficulty processing the information was also significant using the MEMORE within-participants bootstrapping technique with 10,000 iterations ($b_{\text{indirect\_effect}}$ = - 0.09, *SE* = 0.02, 95% CI [- 0.12, - 0.05]). This result supports an alternate explanation that difficulty understanding gender-ambiguous voices may explain negative reactions to them, which we address in Study 2.

Study 2

Study 1 provided evidence that a gender-ambiguous voice can lower social fluency and spill-over to negatively influence ratings of products. However, one shortcoming may have been that the voices, although all in use on Siri, may have differed in more than gender-ambiguity (e.g., paralinguistic features: pace, prosody, pronunciation, etc.). Thus, in Study 2 we aimed to replicate Study 1 using one single voice where only the pitch was manipulated to create gender-ambiguity. Although gender is thought to be signaled in voice by more factors than just pitch (Pernet & Belin, 2012), it is the most important determinate (Oleszkiewicz et al., 2017). Using this design reduces the possibility that difficulty understanding the information can account for the results.

Method

Participants and Design

Based on the effect size observed in Study 1, a power analysis indicated that 198 participants would be required to reach a power of 80%. Adding a small margin, two hundred and seven participants from Mturk (48.8% female; age: *M* = 43.13 years; *SD* = 13.65) were recruited and paid to complete this pre-registered study (https://aspredicted.org/ZJH_SRC), all of whom were included in the analyses. We used a similar procedure to Study 1. However, we simplified the design to use only a female voice as a comparison (rather than both male and female, as before). To create the voices, as a base voice we began with the female voice of "Sophie," a digital voice used in text-to-speech programs. We raised the pitch to create a version of Sophie that sounds more feminine, and we lowered the pitch to create a version of Sophie to sound more gender-ambiguous. As in Study 1, participants listened to two brands of toothpaste, one described by a gender-ambiguous voice and the other described by a female voice. This results in a one-factor within-participants design testing reactions to *gender-ambiguity* (gender-ambiguous versus gender-obvious voices). *Stimulus order* and *voice-gender order* were counterbalanced as previously.

Procedure

After completing a CAPTCHA, participants heard two voices with different pitches. To create the gender-ambiguous voice, we manipulated the base voice of Sophie by lowering its pitch by one semitone to a level where its gender was perceived as ambiguous, by putting its pitch range between 145-175 Hz, the range defined by researchers as gender-ambiguous (Mullennix et al., 1995) without affecting the tempo and other voice characteristics. To create the female voice, we manipulated the base voice of Sophie by raising its pitch by one semitone.

Similar to the previous study, participants first heard audio information about one of the two brands in a female voice or a gender-ambiguous voice, depending on counterbalancing. Then, on a subsequent screen, they were asked to answer a question about product liking for the first toothpaste brand. Afterward, they heard about the second brand presented at different pitch levels. Subsequently, they were asked to answer the same question regarding their attitude toward the second brand. Next, participants were asked to guess the gender of the speakers as a manipulation check. In the end, they were asked to rate *social disfluency* and *difficulty processing the information* in the same way as the previous study before being asked to answer some demographic questions (spoken language, gender, and age).

Results

Manipulation Check

Participants perceived the gender of the voice with a lower pitch to be more ambiguous compared to the voice with a higher pitch ($M_{\text{lower-pitch}} = 2.51$, $SD = 1.20$; $M_{\text{higher-pitch}} = 1.62$, $SD = 0.86$, $t(206) = 11.00$, $p < .001$, $d = 0.76$, 95% CI [0.61, 0.92]), indicating a successful manipulation.

Product Judgment

A result of both classic and Bayesian paired-samples t-tests revealed a main effect of *gender-ambiguity* on product liking, $t(206) = 2.94$, $p = .004$, $d = .204$, 95% CI [0.07, 0.34]; $BF_{10} = 5.25$). People rated the product presented in a gender-ambiguous voice ($M_{ambiguous} = 5.20$, $SD = 1.29$) less favorably than the similar option presented in a female voice ($M_{female} = 5.40$, $SD = 1.22$), replicating our findings in Study 1.

Examining First Stimulus as Between-Participants Test. Similar to the previous study, while not pre-registered, we examined the ratings of the first options presented to participants as a between-participants test of the main hypothesis. Those who heard the first brand's information in the gender-ambiguous voice ($M_{\text{first (ambiguous)}} = 5.10$, $SD = 1.32$) liked it less than those who were informed about the first brand in the female voice ($M_{\text{first (female)}} = 5.42$, $SD = 1.17$), though only marginally so ($t(205) = 1.87$, $p = .062$, $d = .26$, 95% CI [-0.01, 0.53]).

Social Disfluency

As predicted, we observed a significant effect of gender-ambiguity on difficulty in understanding the gender of a narrator, *social disfluency*, $t(206) = 8.75$, $p < .001$, $d = 0.61$, 95% CI [0.46, 0.76]. When participants heard the product information in the gender-ambiguous voice ($M_{ambiguous} = 3.21$, $SD = 1.98$), they had more difficulty determining the gender of the speaker relative to the female voice ($M_{female} = 2.04$, $SD = 1.50$).

As in study 1, a within-participants bootstrapping analysis with 10,000 iterations revealed evidence for complete mediation (see Figure 3); the decrease in product liking was explained by an increase in social disfluency due to the voice gender manipulation (b_{total_effect} = - 0.20, *SE* =0.07, 99% CI [- 0.34, - 0.07], $t(206) = -2.94$ $p < .001$; b_{direct_effect} = -0.08, *SE* =0.03, 95% CI [- 0.24, 0.08], $p = 0.32$; $b_{indirect}$ = -0.12, *SE* =0.06, 99% CI [- 0.25, - 0.02], $p < .01$).

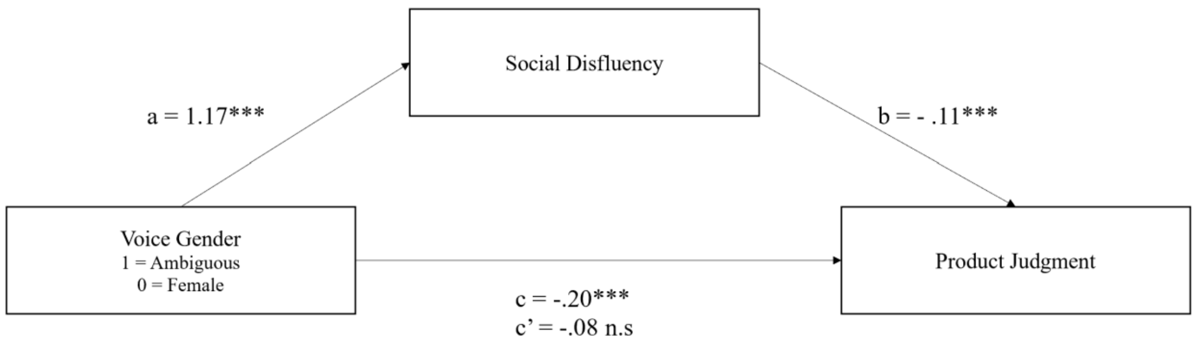


Fig. 3 – Social disfluency completely mediated the effect of voice gender on product liking. * $p < .05$; ** $p < .01$; *** $p < .001$.

Difficulty Processing the Information

All participants found processing the information to be equally easy (p =.81) when information was presented with the lower pitch voice ($M_{ambiguous}$ = 1.85, *SD* = 1.29) as when it was presented with higher pitch voice (M_{female} = 1.83, *SD* = 1.29). This suggests that manipulating the voice-gender did not affect difficulty processing the *information itself*, and differences regarding product liking are instead related to difficulty in processing *social information* (i.e., gender), what we have called *social disfluency*.

Study 3

In the previous studies, participants heard one each of a gender-ambiguous and gender-obvious voice describing a product in a single category. To ensure robustness, in study 3 we tested several voices and categories.

Method

Participants and Design

Because we were unsure if the effect size would be similar for novel voices and stimuli, sample size was determined based on the between-participants effect size observed in Study 2, targeting power of 97%. We recruited participants from Prolific to participate in a longitudinal study consisting of two sessions separated by a week. This longitudinal study employs a 2 (*gender-ambiguity*: gender-ambiguous vs. gender-obvious voices; within) × 4 (*product*: toothpaste, coffee, cordless drill, detergent; within) within-participants design. That is, participants rated each of the four products twice: once when narrated by a gender-ambiguous voice and once when narrated by a gender-obvious voice. We separated the two ratings by one week to decrease the salience of the repeated measures.

At Time 1, 901 participants located in the US and Canada were paid a small sum for completing this pre-registered study (https://aspredicted.org/XG6_NSP). Amongst them, 472 participants (49.5% female; age: M = 37.72 years; SD = 13.09) returned and completed our survey at Time 2 (52.38% retention rate). This retention rate was lower than we anticipated due to an outage on Qualtrics that made the survey at Time 2 unavailable for a brief period during the collection window. Following the preregistered stopping rule, data collection was terminated 48 hours after the commencement of Time 2 despite not reaching the target sample size of 600. All participants who completed the Time 2 study were included in the analysis.

Procedure

After completing a CAPTCHA bot check, participants were asked to rate their perception of different products presented in different voices. The design used 4 different base voices (2 of which were initially male and 2 female). The base voices had their pitch raised and lowered an equal amount to create the stimuli: a more gender-ambiguous version and a more gender-obvious version. That is, both versions of the voices were modified versions of the same base voice. All of the voices were digital voices used in various text-to-speech applications, and each base voice was associated with a particular product. For example, each toothpaste was described by a gender-ambiguous narrator and a gender-obvious narrator created from the same base voice. We used two different male voices from Google's text-to-speech service as the base voice for the first and the third product (i.e., toothpaste and drill), while we recorded the second and the third options (i.e., coffee and detergent) using, respectively, a female voice from Google and a voice named "Joanna" from a different text-to-speech service.

Each participant was asked to rate each product "how much do you like this [product]? (1 = not at all; 7 = very much)" before hearing about the next one. Unlike the previous studies, participants provided their ratings of the products on the same screen as the audio interface. As before, they could listen to the audio as often as they liked. This design reduces the plausibility of an alternate explanation for the previous results based on memory where gender-ambiguous voices are less memorable and thus rated lower.

There were four products used as stimuli presented in this order: toothpaste, coffee, cordless drills, and detergent. We measured evaluations of each product twice, with each measurement separated by one week. We counterbalanced the order of the voices such that at Time 1 some participants heard a gender-ambiguous voice (A) first, followed by a gender-obvious (G) one (GAGA) or vice-versa (AGAG). At Time 2, participants were in the reverse order condition as Time 1. Thus, each participant heard about the exact same product twice: once when narrated by a gender-ambiguous voice and once when narrated by a gender-obvious voice.

Results

Pretest

We conducted a pretest with 100 participants from Prolific. The results indicated that the manipulation was successful ($F(3, 294) = 122.98$, $p < .001$, η_p^2=.56), meaning that participants who heard the information in a gender-ambiguous voice evaluated the gender of the narrator to be significantly more ambiguous than those who heard the information in a gender-obvious voice, collapsed across all four voices. Similar results were obtained for each voice used in this study. See the SOM for the details of this analysis.

Product Judgment

Both classic and Bayesian repeated-measures ANOVAs were conducted with *gender-ambiguity* and *product* as within-participants factors to examine the effect of the gender ambiguous voices on product liking. Most importantly, as expected, the result revealed a significant main effect of *gender-ambiguity*, $F(1, 1413) = 167.76$, $p <.001$, $\eta_p^2 = .263$, $BF_{10} = 1.33 \times 10^{14}$, indicating overall less favorable product judgments for products described by a gender-ambiguous voices, replicating the past studies across different product stimuli and voices (see Figure 4).

The main effect of *product* was also significant ($F(3, 1413) = 22.91$, $p < .001$, $\eta_p^2 = .046$, $BF_{10} = 5.55 \times 10^{11}$), indicating that some products had more favorable ratings than others, inconsequential for our hypothesis. Finally, the interaction was also significant ($F(3, 1413) = 5.33$, $p =.001$, $\eta_p^2 = .011$, $BF_{10} = 10.43$), indicating that the magnitude of the effect of *gender-ambiguity* differed as a function of *product*, an effect about which we did not have a hypothesis.

Each contrast test between *gender-ambiguity* within *product* demonstrated that those who heard the information in a gender-ambiguous voice evaluated the product significantly less favorably than those who heard about it in a gender-obvious voice. For example, people rated the toothpaste less favorably when it was narrated in gender-ambiguous voice ($M_{\text{toothpaste (ambiguous)}}$ = 4.72, SD = 1.45) than when it was narrated in gender-obvious voice ($M_{\text{toothpaste (gender-obvious)}}$ = 5.12, SD = 1.23; F (1, 471) = 45.86, p <.001, η_p^2= .089, $BF_{10} = 1.45 \times 10^8$). The same pattern of result was found for all of the other products. The rating for the coffee was significantly lower when it was described by a gender-ambiguous narrator ($M_{\textit{coffee} \text{ (ambiguous)}}$ = 4.25, SD = 1.53) compared to when it was described by a gender-obvious one ($M_{\textit{coffee} \text{ (gender-obvious)}}$ = 4.69, SD = 1.44; F_{coffee} (1, 471) = 46.26, p <.001, η_p^2= .089, $BF_{10} = 1.74 \times 10^8$). The rating for the drill was significantly lower when it was described by a gender-ambiguous narrator ($M_{\textit{drill} \text{ (ambiguous)}}$ = 4.35, SD = 1.85) compared to when it was described by a gender-obvious one ($M_{\textit{drill} \text{ (gender-obvious)}}$ = 4.83, SD = 1.61; F_{drill} (1, 471) = 46.55, p <.001, η_p^2= .09, $BF_{10} = 1.98 \times 10^8$). The rating for the detergent was significantly lower when it was described by a gender-ambiguous narrator ($M_{\textit{detergent} \text{ (ambiguous)}}$ = 4.10, SD = 1.57) compared to when it was described by a gender-obvious one ($M_{\textit{detergent} \text{ (gender-obvious)}}$ = 4.82, SD = 1.39; $F_{detergent}$ (1, 471) = 105.76, p <.001, η_p^2= .183, $BF_{10} = 2.07 \times 10^{19}$).

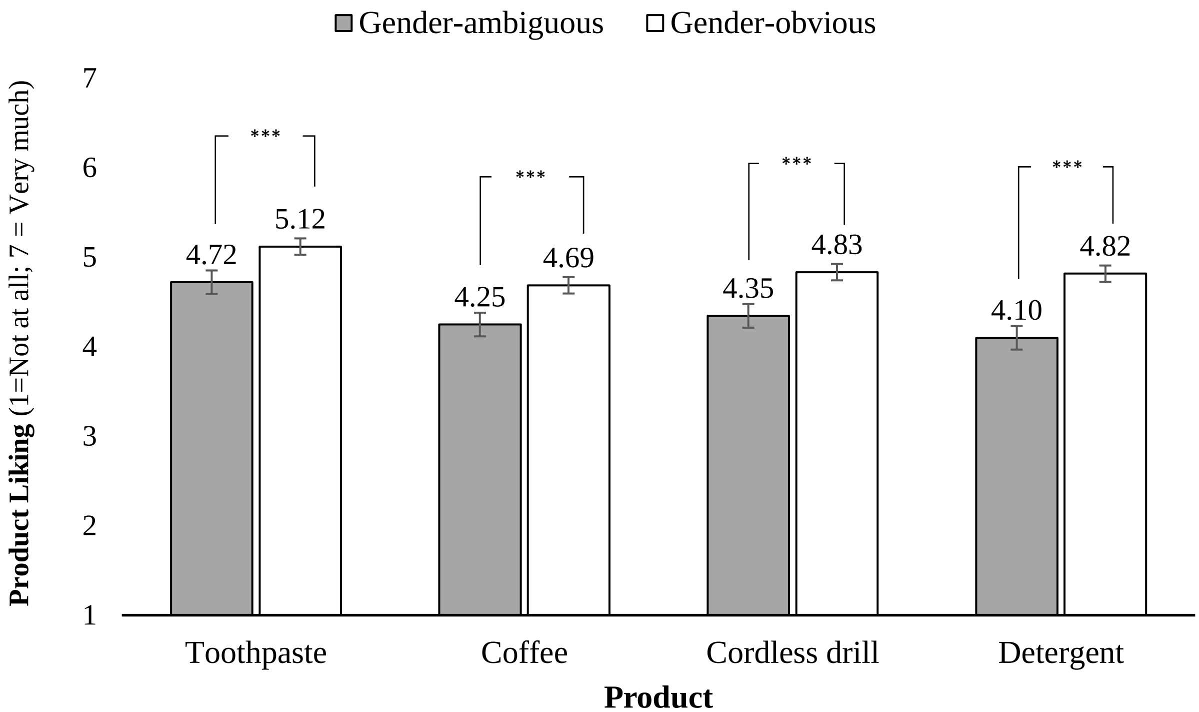


Fig. 4. Across different products, people liked a product described by a narrator with a gender-ambiguous voice (gray bars) less. Error bars represent standard errors. (*** $p < .001$)

Examining Time 1 as Between-Participants Test. As a between-participant test of our main hypothesis, we examined the data from Time 1 (only) including all 901 participants, regardless of whether or not they returned for Time 2. A mixed ANOVA was conducted with *gender-ambiguity order* (GAGA vs AGAG; between) and *product* (toothpaste, coffee, cordless drill, detergent; within) to examine the effect of gender-ambiguity on product liking. In this design, the first order group (GAGA) should rate products 1 and 3 as *better* than the other group (AGAG), as these products were narrated by a gender-obvious narrator. The same group should rate products 2 and 4 as *worse* than the other group. Thus, our main hypothesis is to observe a significant interaction.

As expected, the result revealed a significant interaction F (3, 2697) = 60.39, p <.001, η_p^2 = .063, BF_{10} = 4.94×10^{18}. This indicates that the liking of each product differed by *gender-ambiguity order*. There was also a significant main effect of *product* F (3, 2697) = 38.19, p <.001, η_p^2 = .041, indicating that some products were rated more favorably than others. There was no significant main effect of *gender-ambiguity order* F (1, 899) = 1.93, p =.166, η_p^2 = .00, BF_{10} =0.162, indicating that overall ratings did not differ between the between-participants order conditions.

To examine the interaction, we conducted a series of planned contrast tests for each product. The rating for the toothpaste was significantly lower when it was described by a gender-ambiguous narrator ($M_{\text{toothpaste (ambiguous)}}$ = 4.69, SD = 1.43) compared to when it was described by a gender-obvious one ($M_{\text{toothpaste (gender-obvious)}}$ = 5.03, SD = 1.28; F (1, 899) = 14.04, p <.001, η_p^2 = .02, BF_{10} = 71.03). The same was true for each additional product. Specifically, the rating for the coffee was significantly lower when it was described by a gender-ambiguous narrator ($M_{\textit{coffee}\text{ (ambiguous)}}$ = 4.10, SD = 1.56) compared to when it was described by a gender-obvious one ($M_{\textit{coffee}\text{ (gender-obvious)}}$ = 4.54, SD = 1.49; F (1, 899) = 19.01, p <.001, η_p^2 = .02, BF_{10} = 7.85×10^2). The rating for the drill was significantly lower when it was described by a gender-ambiguous narrator ($M_{\textit{drill}\text{ (ambiguous)}}$ = 4.24, SD = 1.90) compared to when it was described by a gender-obvious one ($M_{\textit{drill}\text{ (gender-obvious)}}$ = 4.77, SD = 1.64; F (1, 899) = 19.48, p <.001, η_p^2 = .02, BF_{10} = 9.89×10^2). The rating for the detergent was significantly lower when it was described by a gender-ambiguous narrator ($M_{\textit{detergent}\text{ (ambiguous)}}$ = 3.91, SD = 1.53) compared to when it was described by a gender-obvious one ($M_{\textit{detergent}\text{ (gender-obvious)}}$ = 4.74, SD = 1.38; F (1, 899) = 72.37, p <.001, η_p^2 = .08, BF_{10} = 6.2×10^{13}).

Study 4

Thus far, we have established that difficulty deciphering the gender of a voice can have negative consequences for product judgments across several voices and product categories, even for speakers equally easy to understand. We next investigated the psychological mechanism. Study 4 tested whether reducing *difficulty categorizing the gender* of a voice by naming it can reduce social disfluency. We reasoned that the gender-category of a speaker with a clearly-gendered name should be less ambiguous compared to one missing this information, and consequently categorization should be easier when such a name is provided (Huart et al., 2005). Thus, if difficulty categorizing causes social disfluency, giving a gender-ambiguous voice a gender-specific name should reduce the negative effect, compared to when this information is absent. Study 4 also used a between-participants design, ensuring methodological robustness.

Method

Participants and Design

An earlier version of this experiment suggested that this manipulation using names would not be an effective moderator of the effect of social disfluency. Thus, in this study we aimed to conduct a very highly-powered test in order to have better confidence that an observed null effect could not due to a lack of power. Accordingly, 1799 participants from Prolific (51.6% female; age: M = 36.49 years; SD = 12.75) located in the US and Canada were paid a small sum for completing this pre-registered study (https://aspredicted.org/MTD_P2F). No participants were excluded from the analyses. A toothpaste brand was used as a stimulus. For the voices, a male digital voice from Google's text-to-speech service was selected as a base voice. The pitch was raised to sound more gender-ambiguous and lowered and equal amount to sound more masculine. Participants were randomly assigned to conditions in a 2 (*gender-ambiguity*: gender-obvious vs. gender-ambiguous voice; between) × 2 (*named voices*: yes vs. no; between) between-participants design.

Procedure

After completing a CAPTCHA, participants were randomly assigned to conditions in which, unlike the previous studies, they heard only one voice. Half of the participants heard a description of a toothpaste brand narrated in the higher-pitched ambiguous version of the voice, while the other half heard it narrated in the lower-pitched male version of the voice.

To manipulate the ease of categorization, half of the participants were informed about the name of the speaker, David (which is a masculine name) through a text, "Please listen to some information presented by a voice assistant, David,..." The other condition did not learn about any names. If categorization contributes to the negative effect of hearing a gender-ambiguous voice, giving the voice assistant this gender-identifying name should moderate.

All participants reported the same measure of product liking used in our previous studies. We also asked a binary question "In this study, did the voice assistants have a name?" as a manipulation check of naming, and then asked participants to guess the gender of the speaker.

Furthermore, since the voices of homosexual people often fall outside of the gender normative pitch range (e.g., Baeck et al., 2011), it seems possible that listeners may judge a gender-ambiguous voice created by altering its pitch as sounding homosexual. To control the effect of prejudice, participants completed three explicit measures of *prejudice* against LGBTQ people adapted from Brigham (1993). They were asked their agreement with three items: "I would rather not have a LGBTQ person live in the same apartment building I live in," "I get very upset when I hear people make a prejudicial remark about members of the LGBTQ community," and "It would not bother me if my new roommate was a member of LGBTQ community" on a 5-point scale from strongly disagree to strongly agree.

Results

Manipulation Checks

Voice Gender Ambiguity and Categorization. A between-participants ANOVA on voice gender ambiguity with *gender-ambiguity* and *named voices* as between-participants factors revealed significant main effects of *named voices* ($F(1, 1795) = 32.45$, $p < .001$, $\eta_p^2 = .018$), and *gender ambiguity* ($F(1, 1795) = 654.13$, $p < .001$, $\eta_p^2 = .267$). More importantly, there was also a significant interaction, $F(1, 1795) = 27.14$, $p < .001$, $\eta_p^2 = .015$**.** While there was not a significant difference in ambiguity between gender-obvious male voices when they were called by a name or not ($F(1, 1795) = .121$, $p = .729$), participants perceived that the gender of the gender-ambiguous voice called by a name ($M_{\text{ambiguous (named)}} = 1.90$, $SD = 0.95$) to be less ambiguous compared to the gender-ambiguous voice called without a name ($M_{\text{ambiguous (unnamed)}} = 2.30$, $SD = 1.04$; $F(1, 1795) = 59.42$, $p < .001$, $\eta_p^2 = .032$), indicating that providing a name for the narrator reduced the gender-ambiguity for the gender-ambiguous narrator.

Naming. By condition, participants also correctly identified whether the voices were named ($X^2(1, N = 1799) = 1452.88$, $p < .001$, $V = .67$).

Product Judgment

As depicted in Figure 5, a between-participants ANOVA on product judgment with *gender-ambiguity* and *named voices* as between-participants factors revealed a significant main effect of *gender-ambiguity*, $F(1, 1795) = 40.18$, $p < .001$, $\eta_p^2 = .022$, $BF_{10} = 1.33\times10^7$, indicating overall less favorable product judgments for the product when it was described by a gender-ambiguous voice, replicating the previous studies. Moreover, the main effect of *named voices* was not significant (p = .73), indicating that providing a name for the narrator did not alter perceptions of the product. Finally, there was marginally significant two-way interaction of *gender-ambiguity* and *named voices*, F (1, 1795) = 2.80, p = .094, η_p^2 = .002, BF_{10} =0.06. However, this interaction was in the *opposite* direction we predicted. Replicating the past studies, in the no-name condition, we found that people rated the product less favorably when it was narrated in the gender-ambiguous voice ($M_{\text{ambiguous (unnamed)}}$ = 4.56, SD = 1.55) compared to the gender-obvious voice ($M_{\text{male (unnamed)}}$ = 4.89, SD = 1.35; F (1, 1795) = 10.85, p = .001, η_p^2 = .006; BF_{10} =17.43). Moreover, we observed a directionally *larger* effect in the same direction in the name condition. The toothpaste was liked less when the gender-ambiguous voice was called by a name ($M_{\text{ambiguous (named)}}$ = 4.42, SD = 1.59) compared to when the product was heard in gender-obvious voice, named David ($M_{\text{male (named)}}$ = 4.98, SD = 1.38; F(1, 1795) = 32.20, p < .001, η_p^2 = .02; BF_{10} =2.9×10^5). This result does not support the account explaining the effect based on categorization.

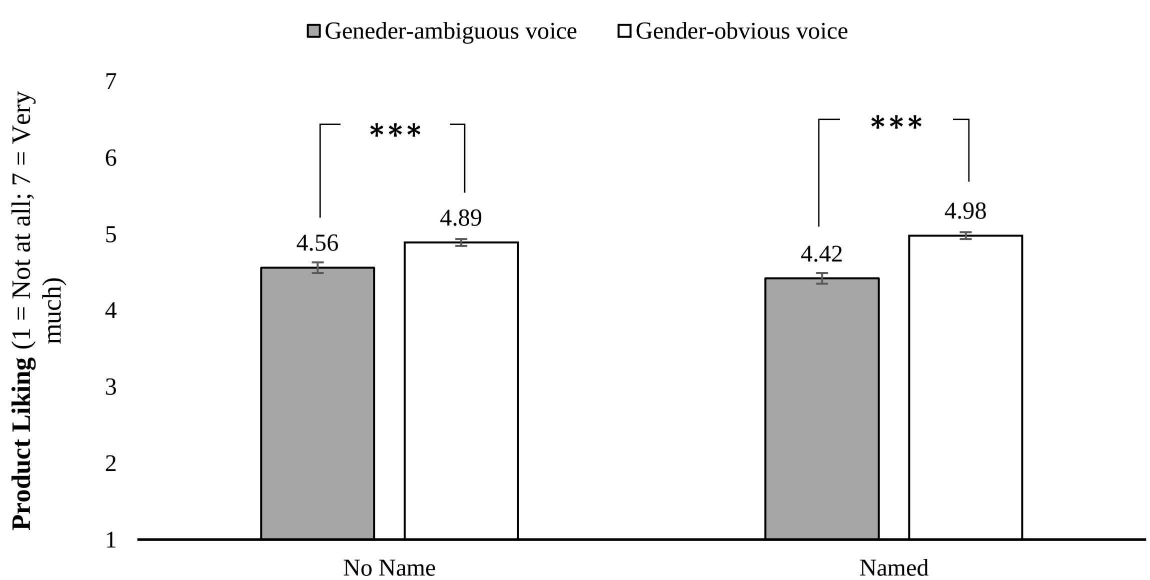


Fig. 5. Making categorizing the gender of the voices easier by assigning male names to each voice assistant (right bars) directionally amplified the negative effect of hearing a gender-

ambiguous voice, contrary to the expected pattern if difficulty categorizing were the cause of the negative effect. Error bars represent standard errors. (*** $p \leq .001$).

Prejudice Scale

The items of the scale were highly correlated ($\alpha = .74$), so we averaged them to form a single measure of prejudice. As expected, there was no effect of *named voices* (F (1,1795) = .024, $p = .88$, $\eta_p^2 = .000$, *gender-ambiguity* (F (1,1795) = .732, $p = .39$, $\eta_p^2 = .000$), or their interaction (F (1,1795) = .42, $p = .52$, $\eta_p^2 = .000$) on prejudice. Controlling *prejudice,* the result of a between-participants ANCOVA on product liking revealed a significant main effect of *gender-ambiguity*, $F(1, 1794) = 39.93$, $p <.001$, $\eta_p^2 = .022$, indicating overall less favorable judgments for products described by a gender-ambiguous voice, while the main effect of *named voices* was not significant ($p = .73$). The interaction was marginally significant ($F(1, 1794) = 2.85$, $p = .092$, $\eta_p^2 = .002$). This indicates that there is an adverse effect of voice gender-ambiguity on product liking above and beyond explicit prejudice. Moreover, the result does not show a significant effect of *prejudice* on product liking $F(1, 1794) = .83$, $p = .36$, $\eta_p^2 = .000$, suggesting that prejudice did not play a major role in the product evaluations.

Study 5

Study 4 failed to provide evidence supporting the role of categorization in reducing social fluency. However, another account based on familiarity suggests that if people are unfamiliar with the sound of a non-binary voice, they may experience less fluency compared to a more frequently-encountered gender-obvious voice. In study 5, we manipulated the amount of exposure to a gender-ambiguous voice, reasoning that increasing incidental exposure (by narrating the study’s instructions in the gender-ambiguous voice) should increase familiarity with the voice.

Method

Participants and Design

Because the design of this study was similar to that of Study 1, we aimed for a similar sample size. Six hundred and four Mturk participants (55.9% female; age: M = 43.04 years; SD = 12.47) located in the U.S. and Canada were paid a small sum for completing this pre-registered study (https://aspredicted.org/K9B_13X), all of whom were included in the analyses. Two hypothetical liquid hand soap brands were used as stimuli. For the voices, the female digital voice of Sophie with the pitch raised to sound more feminine and a version of Sophie with the pitch lowered to sound more gender-ambiguous were used, as in Study 2. Participants were randomly assigned to one of two between-participants conditions in a 2 (*instruction voice*: gender-ambiguous vs. female; between) × 2 (*product voice*: rating a product described by a female voice first and then rating a product described by a gender-ambiguous voice second; within) mixed design. *Stimulus order* was counterbalanced.

Procedure

All participants completed a CAPTCHA bot check at the beginning of the study. The procedure was similar to studies 1 and 2. However, in this study, the 35-second instructions for the task were spoken aloud to participants, whereas previously participants read the instructions to themselves. The voice that narrated the instructions varied by condition. Half of the participants heard them read in Sophie’s lower-pitched ambiguous voice, while the other half heard them read in Sophie’s higher-pitched female voice. Thus, half of the participants were randomly assigned to have additional exposure to the gender-ambiguous voice prior to the main task. After hearing the instructions in either a female voice or a gender-ambiguous voice, participants heard some information about two hand soaps, first in the female voice, then in the gender-ambiguous voice. We expected that hearing the study's instructions in the ambiguous voice should increase their familiarity with that voice, attenuating negative reactions due to unfamiliarity.

As the dependent measure, all participants completed the measures of product liking used in the previous studies. As a manipulation check participants guessed the gender of each voice as before. To measure *LGBTQ perception*, participants also answered, "To what extent do you think the voice of each speaker sounds like a person from the LGBTQ community?" (1 = Not at all, 7 = Very much).

Results

Manipulation Check

Participants perceived the gender of the voice with a lower pitch to be more ambiguous compared to the voice with a higher pitch ($M_{\text{lower-pitch}} = 2.44$, $SD = 1.15$; $M_{\text{higher-pitch}} = 1.85$, $SD = 0.97$, $t(603) = 12.42$, $p < .001$, $d = 0.51$, 95% CI [0.42, 0.59]), indicating a successful manipulation.

Product Judgment

Both classic and Bayesian repeated-measures ANOVAs on product liking with *instruction voice* and *product voice* as factors revealed a significant main effect of *product voice* ($F(1, 602) = 4.69$, $p = .031$, $\eta_p^2 = .01$, $BF_{10} = 16.97$) replicating prior studies, but no significant effect of *instruction voice* ($p = .78$). Most importantly, there was a significant two-way interaction $F(1, 602) = 16.62$, $p < .001$, $\eta_p^2 = .03$, $BF_{10} = 59.72$, as predicted. As depicted in Figure 6, when the instructions were read in a female voice, we observed the same pattern as in previous studies, as expected. Specifically, when the gender-ambiguous voice described the second product, it was liked less ($M_{\text{ambiguous}} = 5.12$, $SD = 1.28$) compared to the first product, which was described by a voice that sounded female (M_{female} = 5.35, *SD* = 1.20; $F(1, 602)$ = 19.68, $p < .001$, η_p^2=.032). However, when the instructions were read in a gender-ambiguous voice, the difference was completely eliminated, with no difference between the second product ($M_{\text{ambiguous}}$ = 5.24, *SD* = 1.32) and the first (M_{female} = 5.17, *SD* = 1.24; $F < 1$). This provides evidence for the role of disfluency due to low familiarity as the underlying mechanism.

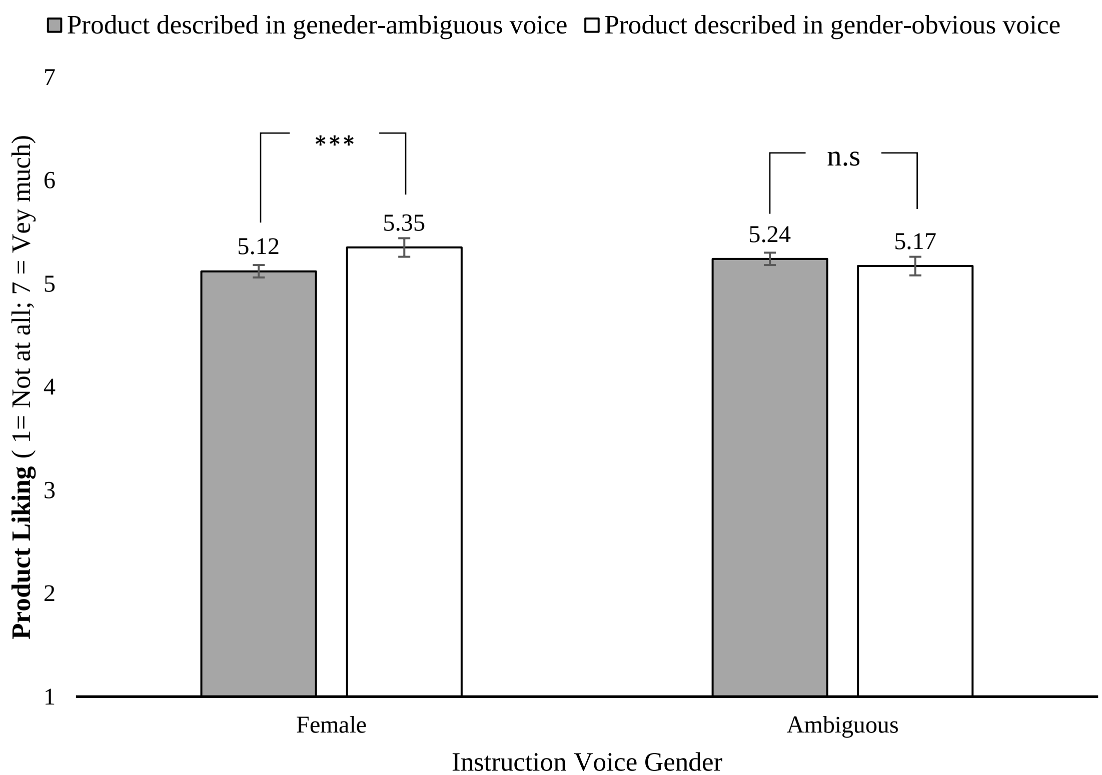


Fig. 6. Hearing the instructions for the study read in the gender-ambiguous voice attenuated the negative effect of hearing a gender-ambiguous voice. A negative effect was present only when the instructions were read by a female narrator (left bars), replicating previous studies. When familiarity was increased by narrating the instructions in a gender-ambiguous voice, the negative effect was eliminated (right bars). Error bars represent standard errors. (*** $p < .001$).

LGBTQ Perception

While the overall level of prejudice against the LGBTQ community should not have varied by condition due to random assignment, if with greater exposure, the gender-ambiguous voice seemed less like a member of the LGBTQ community, the two groups still may have exhibited different levels of bias. To investigate this possibility, we analyzed the question about the extent to which each voice sounded like a member of the LGBTQ community.

As depicted in Figure 7, the results were not consistent with a prejudice interpretation. While perceptions about group membership did differ between conditions, as indicated by a significant interaction ($F(1, 602) = 4.62$, $p = .032$, $\eta_p^2 = .008$), participants exhibited significant differences between the gender-ambiguous voice and the female voice in LGBTQ community perception for both instruction-voice conditions $F(1, 602) = 179.42$, $p < .001$, $\eta_p^2 = .23$. Specifically, in both the female instruction condition ($M_{ambiguous\ (instruction\text{-}female)} = 3.83$, $SD = 1.72$; $M_{female\ (instruction\text{-}female)} = 2.82$, $SD = 1.53$; $F(1,602) = 122.03$, $p <.001$, $\eta_p^2 = .17$) and the ambiguous instruction condition ($M_{ambiguous\ (instruction\text{-}ambiguous)} = 3.66$, $SD = 1.69$; $M_{female\ (instruction\text{-}ambiguous)} = 2.93$, $SD = 1.46$; $F(1,602) = 62.00$, $p <.001$, $\eta_p^2 = .09$), the female voice was significantly less likely to be perceived as a member of the LGBTQ community compared to the gender-ambiguous voice. Thus, prejudice could have affected *both* conditions and is inconsistent with the complete attenuation observed in liking. The main effect of instruction voice was not significant (F $(1,602) = .063$, $p = .80$).

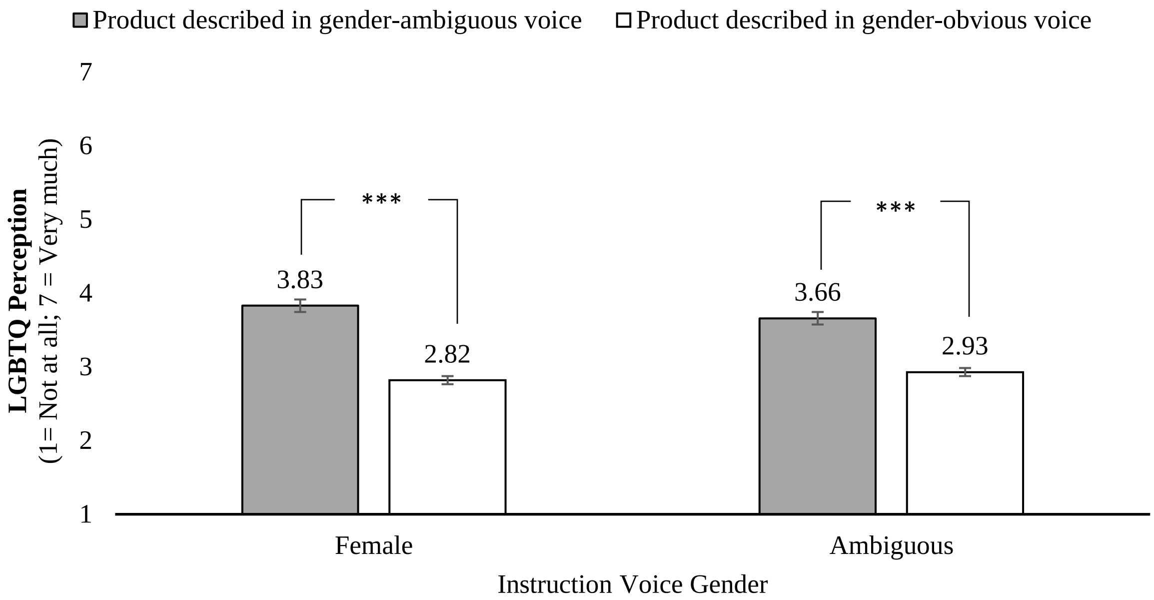


Fig. 7. Regardless of hearing the instructions for the study read in either the gender-ambiguous voice (right bars) or the female voice (left bars), people perceived the gender-ambiguous voice as more likely to be a member of the LGBTQ community compared to the female voice. Error bars represent standard errors. (*** $p < .001$).

Replication

Because hand soap may be considered a feminine product category, we replicated this study using voices from Siri, comparing the gender-ambiguous voice of Siri against a *male* Siri voice. The results were consistent and are reported in the SOM. This suggests that congruency between the product category and the voice in terms of gender (Nass et al., 1997), cannot explain negative effect of voice gender-ambiguity. This explanation also cannot explain the observed interaction.

General Discussion

Across five experiments, people rated products described by narrators with gender-ambiguous voices less favorably than when they were described by clearly male or female narrators. Gender-ambiguous voices created social disfluency–an uneasy feeling related to difficulty understanding the gender of the speaker that spilled over to affect evaluations of the products being described. The effects were eliminated by increasing familiarity with the voices, supporting the idea that exposure to these voices is a positive development.

While we have explored how social disfluency affects product judgments, the implications may be broader. For instance, would a news story reported by a gender-ambiguous speaker be more likely to be considered “fake news?” Could a similar effect occur for other socially unfamiliar voices? Indeed, it seems plausible that intuition about social disfluency could have led managers to hire radio personalities that sound highly uniform (Wolchover, 2011), potentially creating bias against people who do not sound familiar to audiences.

The idea of social disfluency may also extend beyond voices (Lick & Johnson, 2015). For example, related work in the visual domain has investigated social perceptions of gender-ambiguous faces, finding negative effects only after a gender-categorization task (Owen et al., 2016; Winkielman et al., 2015; Yamada et al., 2013). Though the current effects occurred without an explicit categorization task and did not seem to be driven by difficulty categorizing, we did observe that drawing attention to the gender (by providing a gender-identifying name) made the negative effect directionally stronger. Why does highlighting gender amplify negative reactions? This seemingly-consistent finding is particularly unusual, as drawing someone’s attention to the source of disfluency can attenuate its effect (Winkielman et al., 1998). Future researchers may wish to investigate further. Understanding the difference (or similarity) between visual and auditory person-perception may also be an interesting avenue for future research.

Relatedly, we have investigated artificial human voices. Scientists studying artificial visual representations of humans have observed a phenomenon known as the “uncanny valley,” where observers describe them as eerie. They have explained this effect in terms of both difficulty categorizing (e.g. Yamada et al., 2013) and low familiarity with existing categories (e.g. Diel & Lewis, 2022). Future researchers may wish to investigate an uncanny valley for speech. Since speech is a strong indicator of humanness (Schroeder & Epley, 2016), as (until recently) humans had a monopoly on speech, it may be that speech is automatically processed as human, and only particularly unfamiliar deviations from typical speech begin to engender disfluency.

This discussion also highlights an important limitation. While we expect similar effects for gender-ambiguity for actual human voices, we did not test any here. Our results are also limited to the populations we tested. Indeed, our theory predicts smaller effects among people more familiar with such voices.

Our results contribute to scholarly work on fluency (Alter & Oppenheimer, 2009) and social categorization (e.g., Bernstein et al., 2007; Hugenberg & Bodenhausen, 2004; Johnson et al., 2015; Winkielman, et al., 2015), providing insight into how people react to gender-ambiguity in the context of voice communication. We also provide an answer to the important question which motivated this research. While negative reactions to gender-ambiguous voices may indeed occur initially, the findings provide hope that hearing these voices more frequently can be the remedy.

APPENDIX A

Gender-Ambiguous Voices and Social Disfluency

APPENDIX A1. ADDITIONAL ANALYSIS ON STUDIES

For ease of interpretability, in Studies 1, and 2, we collapsed across the counterbalance factors in all analyses. Here, the details of our main findings are available following our preregistered plans for analysis, which involves considering these factors.

Study 1 – Main Findings

Method

Six hundred and two Mturk participants rated two hypothetical brands of toothpaste in a 2 (*rating order*: first toothpaste and second toothpaste; within) × 2 (*gender of the gender-obvious voice*: female vs. male; between) mixed design. *Brand order* (CleanMint first vs. MintyClean first; within) and *voice gender order* (ambiguous voice first vs. gender-obvious voice first; between) were counterbalanced.

Result

Product Judgment

The result showed that there is no *brand order* effect ($p = .163$) or interactions of it with other factors ($ps > .158$) on liking judgement. Combining data on *brand order*, the result of a mixed ANOVA on product liking with *rating order*, *voice gender order*, and *gender of the gender-obvious voice* as factors revealed that there was a significant main effect of *rating order*, about which we did not have a hypothesis, $F(1, 597) = 4.10$, $p = .043$, $\eta_p^2 = .007$. Most importantly, and as predicted, there was a significant two-way interaction between *rating order* and *voice gender order*, $F(1, 597) = 58.69$, $p < .001$, $\eta_p^2 = .09$. There were no differences between product ratings as a function of the gender of the gender-obvious voice (i.e., male vs.

female voices) ($F(1, 597) = .032$, $p = .86$, $\eta_p^2 = .000$). Moreover, there was not a significant three-way interaction between the factors ($F(1, 597) = 1.13$, $p = .29$, $\eta_p^2 = .002$).

We explored the interaction using planned contrast tests. Consistent with our hypothesis, collapsing data across the *gender of the gender-obvious voices*, when the ambiguous voice was first, participants rated the first option less favorably than the second option, which was described in a gendered voice ($M_{first\ (ambiguous)} = 4.81$, $SD = 1.12$; $M_{second\ (gender\text{-}obvious)} = 5.05$, $SD = 1.22$; $F(1, 597) = 15.91$, $p < .001$, $\eta_p^2 = .026$) and vice versa ($M_{first\ (gender\text{-}obvious)} = 5.03$, $SD = 1.15$; $M_{second\ (ambiguous)} = 4.62$, $SD = 1.25$; $F(1, 597) = 46.82$, $p < .001$, $\eta_p^2 = .073$). This provides evidence for a negative effect of gender-ambiguity.

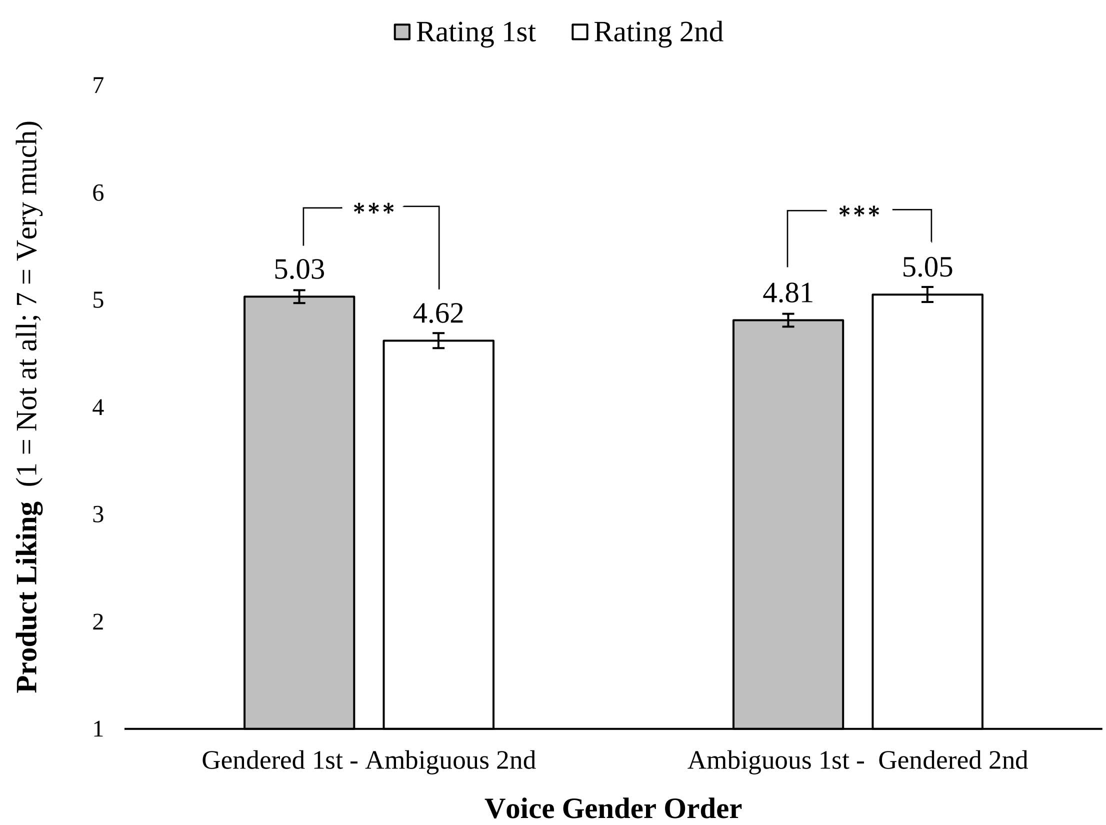


Fig. 8. Presenting product information in a gender-ambiguous voice (bars labeled "Ambiguous") resulted in less liking of the product compared to male and female voices combined (labeled "Gendered"). Error bars represent standard errors. (*** $p < .001$).

Furthermore, as seen in Figure 9, the results of pairwise comparison tests within the *gender of the gender-obvious voice* factor revealed that, when comparing to a female voice, when the ambiguous voice was first, participants rated the first option less favorably than the second option, which was described in a female voice ($M_{first\ (ambiguous)} = 4.78$, $SD = 1.13$; $M_{second\ (female)} = 5.05$, $SD = 1.09$; $F(1, 597) = 29.62$, $p < .001$, $\eta_p^2 = .05$) and vice versa ($M_{first\ (female)} = 5.03$, $SD = 1.09$; $M_{second\ (ambiguous)} = 4.56$, $SD = 1.17$; $F(1, 597) = 10.76$, $p = .001$, $\eta_p^2 = .02$). We found the same pattern of result comparing the gender-ambiguous with the male voice. The results revealed that when the ambiguous voice was first, participants rated the first option less favorably than the second option, which was described in a male voice ($M_{first\ (ambiguous)} = 4.85$, $SD = 1.12$; $M_{second\ (male)} = 5.05$, $SD = 1.19$; $F(1, 597) = 5.60$, $p = .018$, $\eta_p^2 = .009$) and vice versa ($M_{first\ (male)} = 5.03$, $SD = 1.22$; $M_{second\ (ambiguous)} = 4.68$, $SD = 1.32$; $F(1, 597) = 17.90$, $p < .001$, $\eta_p^2 = .029$).

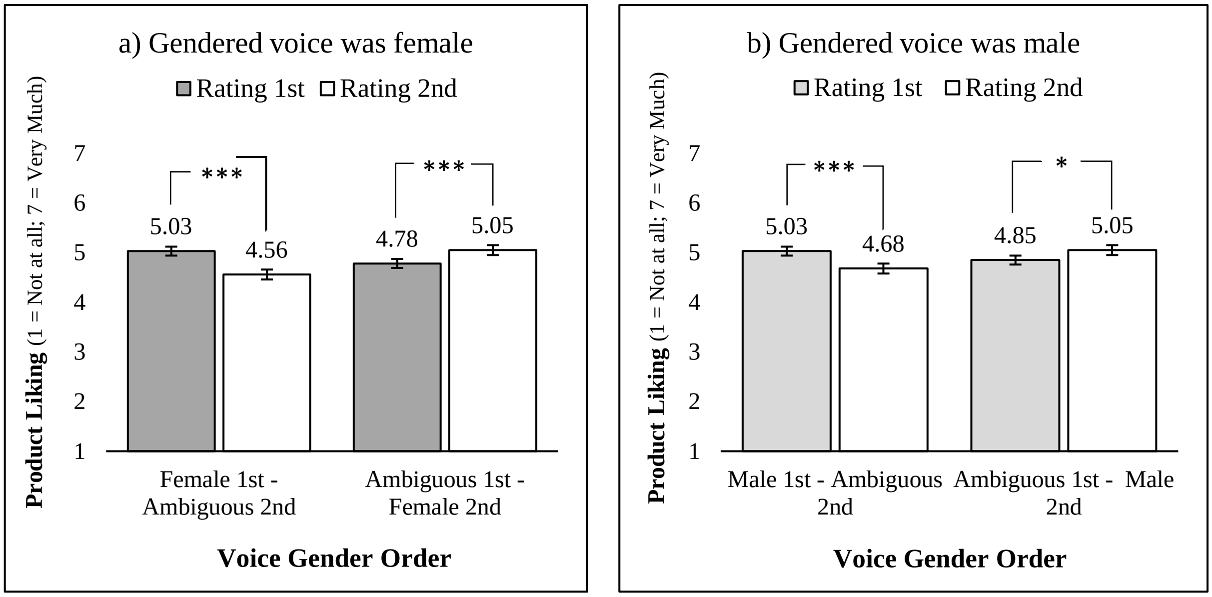


Fig. 9. Regardless of when compared to either a female voice (panel a) or a male voice (panel b), people liked a product with a gender-ambiguous voice less. Error bars represent standard errors. (***p < .001; * p < .05)

Social Disfluency

As expected, there was no *brand order* effect (p = .41). Collapsing data across the brands, we observed a significant two way interaction between *voice gender order* and *rating order* (F(1, 597) = 228.89, p < .001, η_p^2 = .28) and a three-way interaction between them and the *gender of the gender-obvious voice* on our proposed mediator, social disfluency, (F(1, 597) = 9.40, p = .002, η_p^2 = .02), while there was no two way interaction between the *gender of the gender-obvious voice* and *rating order* (p = .40), and the rating order main effect on social disfluency was marginally significant (F(1, 597) = 3.34, p = .07, η_p^2 = .006).

Most importantly, we found that there was no significant interaction of *voice gender order* and *gender of the gender-obvious voice* (p = .93) on social disfluency. The results of pairwise comparison tests revealed that regardless of whether the gender-obvious voice is a female voice ($M_{first\ (female)}$= 1.80, SD = 1.35), or a male voice ($M_{first\ (male)}$= 1.55, SD = 1.13) presented first, people found it difficult to understand the gender of a narrator with the gender-ambiguous voice came second (First female: $M_{second\ (ambiguous)}$ = 3.07, SD = 1.85, $F(1, 597)$ = 91.17, $p < .001$, $\eta_p^2 = .132$; First male: $M_{second\ (ambiguous)}$ = 2.53, SD = 1.82, $F(1, 597)$ = 54.59, $p < .001$, $\eta_p^2 = .084$). We found the same significant pattern of result when the gender-ambiguous voice came first. Regardless of whether the gender-obvious voice is a female voice ($M_{second\ (female)}$= 1.43, SD = 1.08), or a male voice ($M_{second\ (male)}$= 1.31, SD = 0.90) presented second, people found it difficult to understand the gender of a narrator with the gender-ambiguous voice came first (Second female: $M_{first\ (ambiguous)}$ = 2.57, SD = 1.74, $F(1, 597)$ = 74.96, $p < .001$, $\eta_p^2 = .112$; Second male: $M_{first\ (ambiguous)}$ = 1.93, SD = 1.49, $F(1, 597)$ = 21.84, $p < .001$, $\eta_p^2 = .035$).

Study 2 – Main Findings

Method

Two hundred and seven Mturk participants rated two hypothetical brands of toothpaste in a 2 × 2 mixed design with *rating order* (first toothpaste and second toothpaste) as a within factor and *voice gender order* (female first vs. ambiguous first) as a between factor. *Brand order* was counterbalanced. Participants were randomly assigned to one of the two conditions where they first heard audio information about one of the two brands in a female voice or a gender-ambiguous voice. Then, they were asked to answer a question about product liking for the first toothpaste brand. Afterward, they heard about the second brand presented at different pitch levels. They were asked to answer the same question regarding their attitude toward the second brand. Given this design, we pre-registered a prediction for a 2-way interaction of *rating order* and the *voice gender order* on product judgments such that people rate the product presented with a gender-ambiguous voice less favorably than the product presented in a female voice.

Results

Manipulation Check

The voice gender manipulation was effective ($F(1, 205) = 124.04$, $p < .001$, $\eta_p^2 = .38$), as indicated by a significant two-way interaction of *voice gender order* and *rating order* on gender ambiguity with a mixed ANOVA. Collapsing across *voice gender order*, participants perceived the gender of the voice with a lower pitch to be more ambiguous compared to the voice with a higher pitch ($M_{lower\text{-}pitch} = 2.51$, $SD = 1.20$; $M_{higher\text{-}pitch} = 1.62$, $SD = 0.86$, $t(206) = 11.00$, $p < .001$, $d = 0.85$).

Product Judgment

Following our preregistration, the result of a mixed ANOVA showed that there is no *brand order* effect (p = .287) or interaction of it with *voice-gender order* (p = .976) on product liking. Combining data on *brand order*, the result on product liking with *rating order* and *voice-gender order* as factors revealed a significant two-way interaction ($F(1, 205) = 8.54$, $p = .004$, $\eta_p^2 = .04$), as predicted. No other effects were significant (ps >.1). The result of pairwise comparison tests showed that when the gender-ambiguous voice was first, participants liked the first option ($M_{first\ (ambiguous)} = 5.10$, $SD = 1.23$) less than the second option heard in a female voice ($M_{second\ (female)} = 5.38$, $SD = 1.23$; $F(1, 205) = 8.70$, $p = .004$, $\eta_p^2 = .041$). In contrast, when the gender-ambiguous voice was second, participants liked the first option (presented in a female voice) ($M_{first\ (female)} = 5.42$, $SD = 1.25$) more than the second option ($M_{second\ (ambiguous)} = 5.30$, $SD = 1.25$), though the contrast was not significant ($F(1, 205) = 1.43$, $p = .23$, $\eta_p^2 = .007$).

Social Disfluency

As expected, there was no *brand order* effect (p = .16). Collapsing data across the brands, we observed a significant two-way interaction between *voice gender order* and *rating order* on our proposed mediator, social disfluency $F(1, 205) = 78.23$, $p < .001$, $\eta_p^2 = .276$, as predicted. When participants heard the product information in the ambiguous voice, they had more difficulty determining the gender of the speaker relative to the female voice, both when the gender-ambiguous voice came first ($M_{first\ (ambiguous)} = 2.81$, $SD = 1.88$; $M_{second\ (female)} = 1.90$, $SD = 1.35$; $F(1, 205) = 23.62$, $p < .001$, $\eta_p^2 = .103$) and second ($M_{first\ (female)} = 2.18$, $SD = 1.63$, $M_{second\ (ambiguous)} = 3.62$, $SD = 2.00$; $F(1, 205) = 58.20$, $p < .001$, $\eta_p^2 = .221$).

Study 3 – Pretest of Voices from Study 3

The goals of this pretest were twofold. First, we aimed to verify that our manipulation of gender-ambiguity was effective. Second, we tested if manipulating the pitch of the voice also manipulated the perceived age of the speaker. We were concerned that older-sounding voices might be considered more competent. If gender-ambiguous voices always sound younger, this may be an alternate explanation for the pattern of results we have observed.

Method

In this pretest, one-hundred participants from Prolific (67% female; age: M = 35.31 years; SD = 12.14) located in the US and Canada were recruited and paid a small sum for completing this pre-test, all of whom were included in the analyses. Following the design used in Study 3, we used 4 different base voices (2 of which were initially male and 2 female). The base voices had their pitch raised and lowered an equal amount to create the stimuli: a more gender-ambiguous version and a more gender-obvious version. That is, both versions of the voices were modified versions of the same base voice. All of the voices were digital voices used in various text-to-speech applications, and each base voice was associated with a particular product. Each product was described by a gender-ambiguous narrator and a gender-obvious narrator created from the same base voice. Following Study 3, there were four products used as stimuli presented in this order: toothpaste, coffee, cordless drills, and detergent. Participants were randomly assigned to one of two between-participants conditions where they heard about the products starting with a gender-ambiguous voice (A) followed by a gender-obvious (G) one (GAGA) or vice-versa (AGAG). After hearing about each product description, they were asked to guess the gender and age of the voice.

To measure the perception of the voice gender, we asked participants rate the speaker's voice gender scaled from "1 = clearly male", "4 = ambiguous", and "7 = clearly female." Then, we recoded the response distance from the mid-point as a measure of *voice-gender ambiguity* (1 = Highly clear; 4 = Highly ambiguous). That is, a response at the midpoint of the gender question (4) was coded as 4, one position away (3 or 5) was coded 3, two positions away (2 or 6) were coded as 2, and three positions away (1 or 7) was coded a 1.

To measure the perception about each narrator's age, we asked them to answer "what is your best guess about the age of the speaker? (1= quite young, 7 = quite old)."

Result

Voice gender ambiguity

As depicted in Figure 10, a mixed ANOVA was conducted with *gender-ambiguity order* (GAGA vs AGAG; between) and *product* (toothpaste, coffee, cordless drill, detergent; within) as factors to examine the effect on voice gender ambiguity. The result showed a significant interaction between *product* and *gender-ambiguity order* (F (3,294) = 122.98, p < .001, η_p^2=.56), as expected. There was also a significant main effect of *product* (F (3,294) = 6.58, p < .001, η_p^2=.06). Since each product was narrated by a different voice, the most likely explanation for this effect is that perceptions of gender-ambiguity differed as a function of the narrator. The main effect of *gender-ambiguity order* was not significant (F (1,98) = 2.35, p = .128, η_p^2=.023).

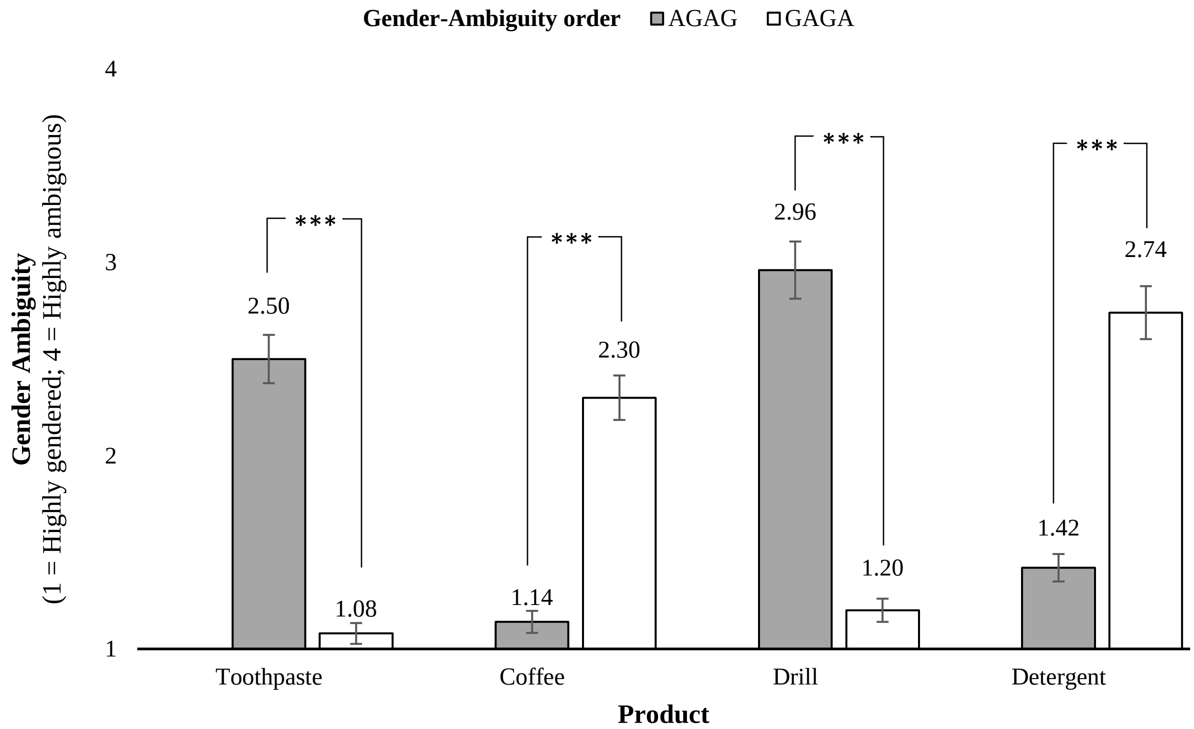


Fig. 10. Perception about the gender of different voices presenting products in a gender-obvious voice or a gender-ambiguous voice. The gray bars represent AGAG order where toothpaste was heard first in a gender-ambiguous voice (A) followed by a gender-obvious (G) one. The white bars represent the GAGA order where the toothpaste was heard in a gender-obvious voice first (G) followed by a gender-ambiguous voice (A) (*** $p < .001$).

The result of the contrast tests within each *product* demonstrated that those who heard the information in a gender-ambiguous voice evaluated the gender of the narrator to be significantly more ambiguous than the corresponding gender-obvious voice. For example, people rated the narrator of toothpaste more ambiguous when it was described in gender-ambiguous voice ($M_{\text{toothpaste (ambiguous)}} = 2.50$, $SD = 0.97$) than when it was narrated in the male voice ($M_{\text{toothpaste (male)}} = 1.08$, $SD = 0.44$; F (1, 98) = 87.93, p <.001, $\eta_p^2 = .47$). The same pattern of result was found for the other voices. The rating of gender ambiguity for the narrator of coffee was significantly higher when it was described by a gender-ambiguous voice ($M_{\text{coffee (ambiguous)}}$ = 2.30, *SD* = 1.16) compared to when it was described in a female voice ($M_{\text{coffee (female)}}$ = 1.14, *SD* = 0.45; F (1, 98) = 33.64, p <.001, η_p^2= .31). The rating of gender ambiguity for the narrator of drill was significantly higher when it was described by a gender-ambiguous voice ($M_{\text{drill (ambiguous)}}$ = 2.96, *SD* = 0.90) compared to when it was described by a male voice ($M_{\text{drill (male)}}$ = 1.20, *SD* = 0.45; F (1, 98) = 152.03, p <.001, η_p^2= .61). The rating for the speaker of the detergent was significantly higher described in a gender-ambiguous voice ($M_{\text{detergent (ambiguous)}}$ = 2.74, *SD* = 1.24) compared to when it was described by a female-sounding voice ($M_{\text{detergent (female)}}$ = 1.42, *SD* = 0.78; F (1, 98) = 40.35, p <.001, η_p^2= .29).

Voice age

The result of a mixed ANOVA with *gender-ambiguity order* as a between factor and *product* as a within factor showed a main effect of *gender-ambiguity order* (F (1,98) = 7.13, p = .01, η_p^2=.07) on perception about the age of the narrator. This is because three out of four voices in the GAGA condition sounded older than in the AGAG condition. The main effect of *product* was also significant (F (3,294) = 17.70, p < .001, η_p^2=.15), indicating that the perceived age depended on the voice of the narrator. However, these main effects were qualified by a significant interaction (F (3,294) = 25.26, p < .001, η_p^2=.21). This indicates that the pattern of age with respect to the order condition differed as a function of product/narrator (See Figure 11).

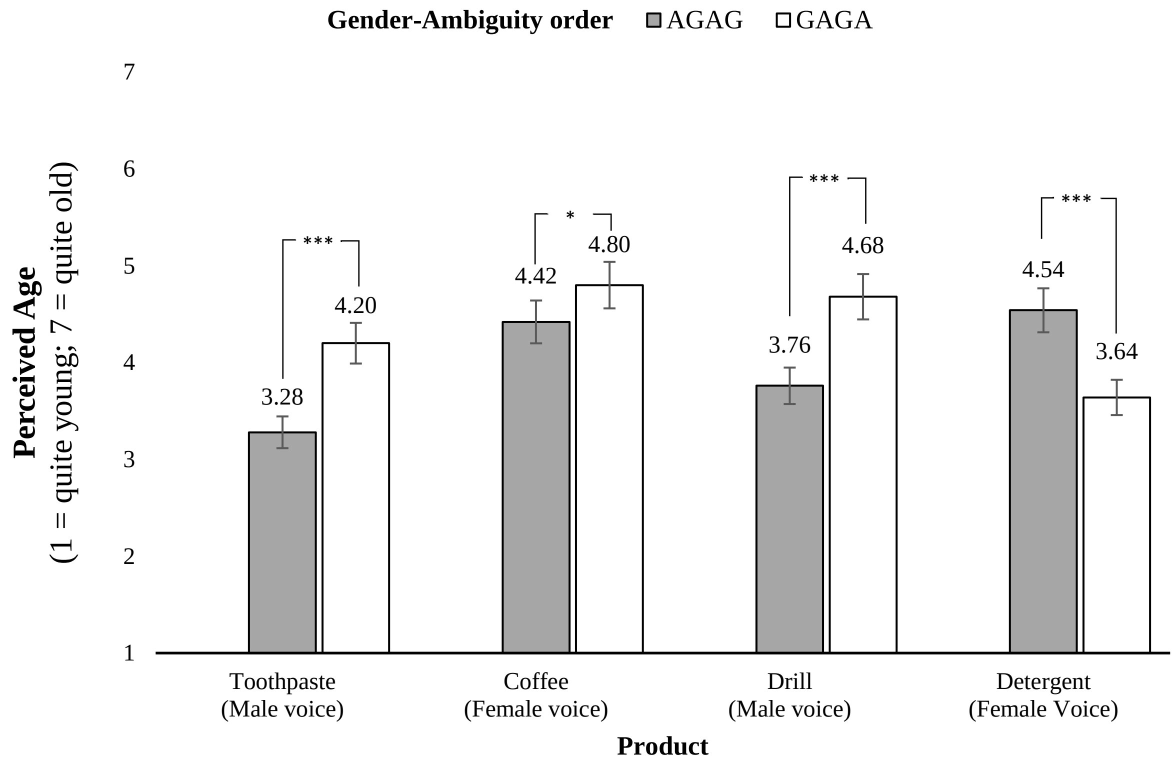


Fig. 11. Perceived age of different voices. The gray bars represent AGAG order where toothpaste was heard first in a gender-ambiguous voice (A) followed by a gender-obvious (G) one. The white bars represent the GAGA order where the toothpaste was heard in a gender-obvious voice first (G) followed by a gender-ambiguous voice (A) (*** $p < .001$; * $p < .05$).

Descriptively, people judged the age of a narrator of toothpaste to be older when it was described in male voice ($M_{\text{toothpaste (male)}} = 4.20$, $SD = 0.83$) than when it was narrated in a gender-ambiguous voice ($M_{\text{toothpaste (ambiguous)}} = 3.28$, $SD = 0.99$; F (1, 98) = 25.26, p <.001, η_p^2= .205). The same pattern of result was found for the other male voice describing the drill. The rating of age for the narrator of drill with a male voice was significantly older ($M_{\text{drill (male)}} = 4.68$, $SD = 1.08$) compared to when the speaker presented the drill with the gender-ambiguous version of the base voice ($M_{\text{drill (ambiguous)}} = 3.76$, $SD = 1.20$; F (1, 98) = 16.20, p <.001, η_p^2= .142).

For the female base voices, the pattern was mixed. People rated the age of a narrator describing coffee to be older when it was described in a gender-ambiguous voice ($M_{\text{coffee (ambiguous)}}$ = 4.80, *SD* = 0.90) than when it was narrated in a female voice ($M_{\text{coffee (female)}}$ = 4.42, *SD* = 0.73; F (1, 98) = 5.35, p =.023, η_p^2= .05). However, this pattern of result was not found for the other female voice. The rating of age for the narrator of detergent with a female voice was significantly older ($M_{\text{detergent (female)}}$ = 4.54, *SD* = 0.76) compared to when the speaker presented the detergent with the gender-ambiguous version of the base voice ($M_{\text{detergent (ambiguous)}}$ = 3.64, *SD* = 1.10; F (1, 98) = 22.57, p <.001, η_p^2= .19).

Because we obtained negative effects on product ratings for gender-ambiguity for all of these voices (including for coffee when the gender-ambiguous voice sounded older; see study 3 in the main text), but we did not find a consistent pattern with respect to age, it seems highly unlikely that the alternate explanation relating age and competence is the best explanation for the results.

APPENDIX A2. BAYESIAN ANALYSIS RESULTS

Study 1 – Bayesian repeated measures ANOVA

We estimated Bayes factors (BFs) to quantify the strength of evidence for effects with one degree of freedom. BFs allowed us to assess evidence in favor of the alternative over the null hypothesis (BF_{10}) or of insensitivity toward hypotheses. Following our main design in Study 1, we conducted a Bayesian repeated measures ANOVA using JASP software (JASP Team, 2020). Following our prediction, we expected that people would evaluate a product presented in gender-ambiguous voice less favorably than a product presented in a gender-obvious voice (H_1). The result of the Bayesian analysis on product liking with *gender-ambiguity* and *gender of the gender-obvious voice* as factors revealed that there was a significant main effect of *gender-ambiguity* which revealed *extreme* support for our hypothesis, $BF_{10} = 4.25\times10^{10}$ (see Table. 1). Although Bayes factors provide a continuous measure of evidence, Bayesian statisticians often used criterion values to facilitate decision-making. Following Lee and Wagenmakers (2013), Bayes factors (BF_{10}) > 3 (or 10, 30, or 100 respectively) are usually providing substantial (or strong, very strong, or extreme respectively) evidence for the alternative over the null hypothesis (See Table 2).

Table 1

Bayesian analysis of effects (Liking)

Effects	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF_{incl}
Gender-Ambiguity	0.600	0.400	1.000	1.567×10^{-11}	$4.254\times10^{+10}$
Gender of the gender-obvious voice	0.600	0.400	0.137	0.863	0.106
Gender-Ambiguity ∗ Gender of the gender-obvious voice	0.200	0.800	0.018	0.982	0.073

Table 2

A heuristic classification scheme for Bayes factors BF_{10} (Lee & Wagenmakers, 2013, p. 105)

Bayes factor	Evidence category
> 100	Extreme evidence for H1

Bayes factor	Evidence category
30 - 100	Very strong evidence for H1
10 - 30	Strong evidence for H1
3 - 10	Moderate evidence for H1
1 - 3	Anecdotal evidence for H1
1	No evidence
1/3 - 1	Anecdotal evidence for H0
1/10 - 1/3	Moderate evidence for H0
1/30 - 1/10	Strong evidence for H0
1/100 - 1/30	Very strong evidence for H0
< 1/100	Extreme evidence for H0

As shown in Table 3, the result of pairwise comparison Bayesian tests revealed strong evidence in favor of our hypothesis in regard with rating less favorably a product presented in a gender ambiguous voice compared with when the same option presented either in a female (BF_{10} = 1.54×10^6) or a male voice (BF_{10} = 3.22×10^3).

Table 3

Bayesian analysis of effects within gender-obvious voice categories (Liking)

Measure 1		Measure 2	BF_{10}	error %	Median	95% credible interval
Gender-ambiguous voice	-	Female voice	$1.545\times10^{+6}$	1.230×10^{-12}	- 0.344	[-0.460, -0.228]
Gender-ambiguous voice	-	Male voice	$3.22\times10^{+3}$	9.263×10^{-6}	- 0.272	[-0.387, -0.157]

To estimate the magnitude of the effect, we obtained a posterior distribution and credible interval per product (see Figure 12). Comparing the gender-ambiguous voice with a female voice, we obtained the median of − 0.344 with 95% CI [− .460, − 0.228] for the standardized effect size δ. This means that, under the assumption that the effect exists and given the model we specified, we can be 95% certain that the true value of δ lies between − .460, − 0.228. We obtained the same pattern of result in regard with the male voice.

b) Gendered voice was female

b) Gendered voice was male

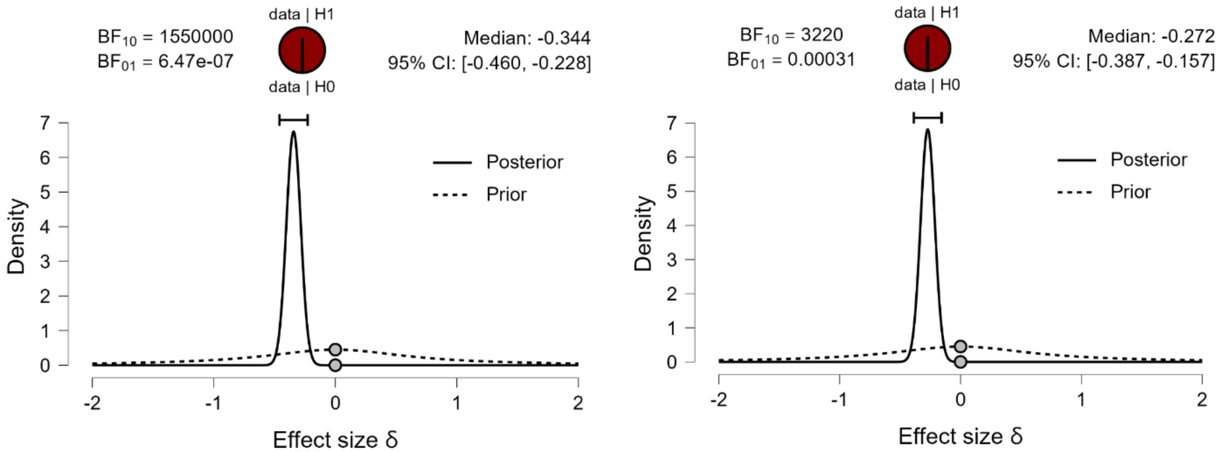


Fig. 12. Prior and posterior distribution per gender-obvious voices.

Study 2 – Bayesian paired-samples t-test.

Following our main design in Study 2, we conducted a Bayesian paired-samples t-test in JASP to quantify the evidence in favor of our specific hypotheses using a default zero-centered Cauchy prior with a scale of 0.707 (Morey & Rouder, 2011). As we expected, the Bayesian analysis on product liking with *gender-ambiguity* and *gender of the gender-obvious voice* as factors revealed *moderate* evidence in favor of the main effect of *gender-ambiguity*, BF_{10} = 5.025 (see Table. 4).

Table 4

Bayesian analysis of effects (Liking)

Measure 1		Measure 2	BF_{10}	error %	Median	95% credible interval
Gender-ambiguous voice	-	Female voice	5.025	0.005	- 0.201	[-0.338, -0.064]

To estimate the magnitude of the effect, we obtained a posterior distribution and credible interval per product (see Figure 13). We obtained the median of − 0.201 with 95% CI [− .338, − 0.064] for the standardized effect size δ. This means that, under the assumption that the effect exists and given the model we specified, we can be 95% certain that the true value of δ lies between − .338, − 0.064.

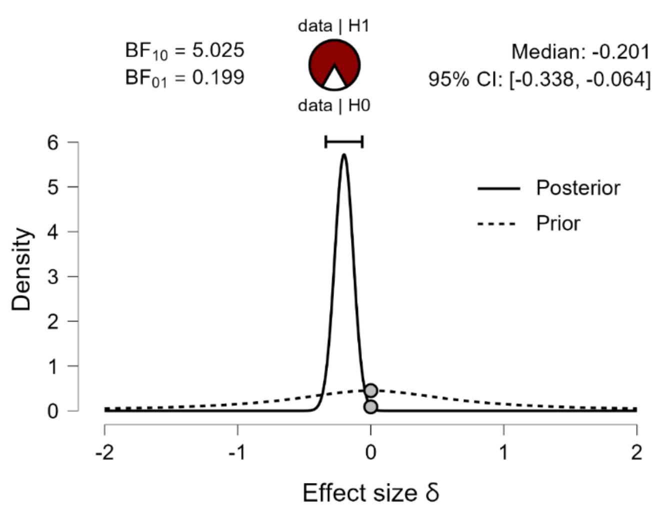


Fig. 13. Prior and posterior distribution

Study 3 – Bayesian repeated measures ANOVA

As shown in Table 5, the result of a Bayesian repeated-measures ANOVA test on product liking with *gender-ambiguity* and *product* as within-participants factors revealed *extreme* evidence in favor of the main effects *of gender-ambiguity* (BF_{10} =1.33×10^{14}), supporting our main hypothesis regarding less favorable product judgments for the product when it was described by a gender-ambiguous voice. Moreover, we found very *strong* evidence for both the main effect of *product* (BF_{10} =5.55×10^{11}), and the interaction of *gender-ambiguity* and *product* (BF_{10} =10.43).

Table 4

Bayesian analysis of effects (Liking)

Effects	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF_{incl}
Product	0.600	0.400	1.000	1.202×10^{-12}	$5.546\times10^{+11}$
Gender Ambiguity	0.600	0.400	1.000	4.996×10^{-15}	$1.334\times10^{+14}$
Gender Ambiguity ✻ product	0.200	0.800	0.723	0.277	10.433

The result of separate Bayesian repeated-samples t-test for each product revealed *extreme* evidence in favor of our hypothesis (See Table 6).

Table 6

Bayesian analysis of effects for each product (Liking)

Measure 1		Measure 2	BF_{10}	error %	Median	95% credible interval
gender-ambiguous	-	gender-obvious				
Toothpaste	-	Toothpaste	$1.446\times10^{+8}$	1.420×10^{-14}	- 0.310	[-0.402, -0.217]
Coffee	-	Coffee	$1.739\times10^{+8}$	1.178×10^{-14}	- 0.311	[-0.403, -0.219]
Drill	-	Drill	$1.982\times10^{+8}$	1.030×10^{-14}	- 0.312	[-0.404, -0.220]
Detergent	-	Detergent	$2.072\times10^{+19}$	5.022×10^{-26}	- 0.471	[-0.566, -0.376]

To estimate the magnitude of the effect, we obtained a posterior distribution and credible interval per product (see Figure 14). For Toothpaste, as an example, we obtained the median of − 0.31 with 95% CI [− .402, − 0.217] for the standardized effect size δ. This means that, under the assumption that the effect exists and given the model we specified, we can be 95% certain that the true value of δ lies between − .402, − 0.217. Further results provided with Table 6 showed credible intervals obtained from posterior distribution for other products. In conclusion, there is strong evidence for the presence of the effect, and small uncertainty about its size.

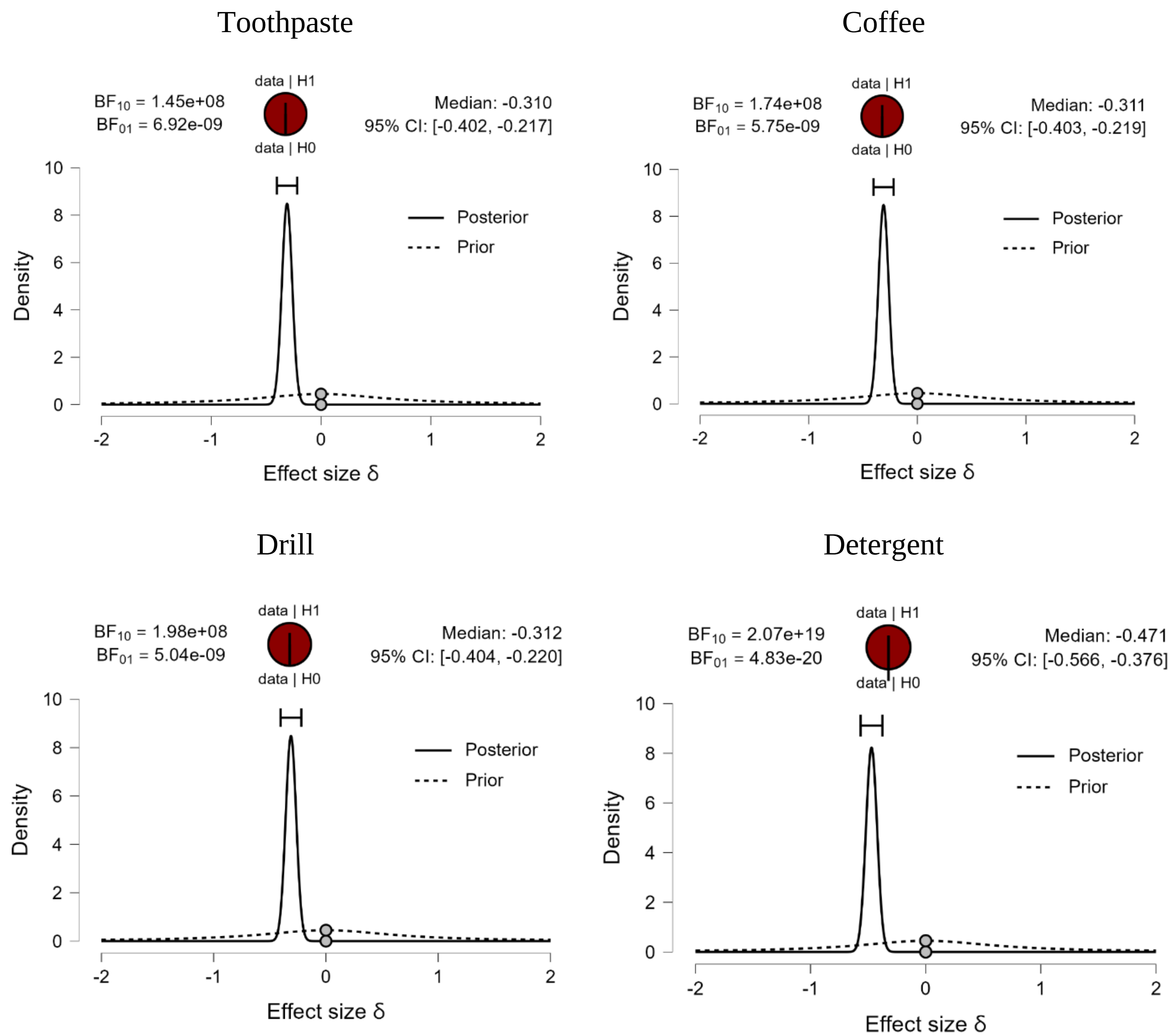


Fig. 14. Prior and posterior distribution for each product

Study 4 – Bayesian repeated measures ANOVA

As shown in Table 7, the result of a Bayesian ANOVA test on product liking with *gender-ambiguity* and *named voices* as between-participants factors revealed *extreme* evidence in favor of our main effect *of gender-ambiguity* (BF_{10} =1.33×10^7), less favorable product judgments for the product when it was described by a gender-ambiguous voice. Moreover, we found very *strong* evidence in favor of the null hypothesis regarding both the main effect of *named voices* (BF_{10} =0.05), and the two-way interaction of *gender-ambiguity* and *named voices* (BF_{10} =0.06).

Table 7

Bayesian analysis of effects (Liking)

Effects	**P(incl)**	**P(excl)**	**P(incl|data)**	**P(excl|data)**	**BF_{incl}**
Gender Ambiguity	0.600	0.400	1.000	5.020×10^{-8}	1.328×10^{+7}
Named Voices	0.600	0.400	0.068	0.932	0.049
Gender Ambiguity ✻ Named Voice	0.200	0.800	0.015	0.985	0.060

Running the Bayesian analysis separately for no-name and named conditions, respectively, the results revealed *strong* evidence in favor of our main hypothesis (BF_{10} =17.43), and *extreme* evidence against social categorization account (BF_{10} =2.9×10^5).

Table 8

Bayesian analysis of effects (Liking)

	Effects	**P(incl)**	**P(excl)**	**P(incl|data)**	**P(excl|data)**	**BF_{incl}**
No name	Gender Ambiguity	0.500	0.500	0.946	0.054	17.432
Named voices				1.000	3.37×10^{-6}	296861.911

Study 5 – Bayesian repeated measures ANOVA

Following our main design in Study 5, as shown in Table 11, the result of Bayesian repeated-measures ANOVA on product liking with *instruction voice* and *product voice* as factors demonstrated *strong* evidence in favor of the interaction effect of *product voice* and *instruction voice*, $BF_{10} = 59.72$. Moreover, replicating our previous findings regarding the main effect, the result supports our hypothesis *strongly* showing that people have negative judgment toward gender-ambiguous voice compared to the gender-obvious voice ($BF_{10} = 16.97$).

Table 9

Bayesian analysis of effects (Liking)

Effects	**P(incl)**	**P(excl)**	**P(incl|data)**	**P(excl|data)**	**BF_{incl}**
Product voice	0.600	0.400	0.962	0.038	16.969
Instruction voice	0.600	0.400	0.946	0.054	11.697
Product voice ✻ Instruction voice	0.200	0.800	0.937	0.063	59.72

APPENDIX A3. SUPPLEMENTAL STUDIES

Study S1. Replication of Study 5 using Siri's voices.

Method

Participants and design

Three hundred MTurk users (52.7% female; age: $M = 40.04$ years; $SD = 11.34$) located in the U.S. and Canada were paid a small sum for completing this pre-registered study (https://aspredicted.org/GPB_HCB), all of whom were included in the analyses. Two hypothetical liquid hand soap brands were used as stimuli, presented in either a male Siri voice or an ambiguous Siri voice (Table 2). Participants were randomly assigned to one of four between-participants conditions in a 2 (*instruction voice gender*: ambiguous vs. male; between) × 2 (*product voice gender order*: gender-ambiguous first vs. male first; between) × 2 (*rating order*: first and second; within) mixed design. Brand order was counterbalanced.

Procedure

The procedure was similar to Studies 2 and 3. However, in this study, the instructions for the task were spoken aloud to participants, whereas previously participants read the instructions to themselves. The voice that narrated the instructions varied by condition. Half of the participants heard them read in Siri's gender-ambiguous voice, while the other half heard them read in Siri's male voice. Thus, half of the participants were randomly assigned to have additional exposure to the gender-ambiguous voice prior to the main task. This should increase their familiarity with the voice.

As the dependent measure, all participants filled out the measure of product liking used in the previous studies.

Table 10

Stimuli for Experiment S1. Both Brands' Information Was the Same, and All Participants Heard Both Brands Information.

Cleansoft brand Information	Softsoap brand Information
The new liquid soap brand name is CleanSoft. This Antibacterial Hand Wash effectively washes away dirt and germs while nourishing your skin with 100% gentle cleansers with no Paraben. It is also formulated without phthalates. It is enriched with moisturizer and pH balanced for your skin. This hand wash contains a combination of 5 skin nutrients and moisturizers that work with skin to maintain hand softness and create new moisturizers for lasting nourishment.	The new liquid soap brand name is SoftSoap. It is made with a 5-moisturizer blend and protects from skin dryness. This Antibacterial Hand Wash eliminates almost 100% of bacteria found in the household in seconds. It gives your hands lasting nourishment with every pump. It is Paraben free and formulated without phthalates. This gentle cleanser offers a rich, indulgent texture that provides instant softness and nourishment that lasts for hours.

Result

A mixed ANOVA on product judgment with *instruction voice gender, product voice gender,* and *rating order* revealed a significant three-way interaction F (1, 296) = 3.91, p = 0.049 η_p^2 = 0.013. When the instructions were read in a male voice, we observed the same pattern as in previous studies, as expected. Specifically, when the gender-ambiguous voice described the first product, it was liked less ($M_{\text{first (ambiguous)}}$ = 5.31, SD = 1.26) compared to the second voice, which sounded male ($M_{\text{second (male)}}$ = 5.41, SD = 1.30) and vice versa ($M_{\text{first (male)}}$ = 5.39, SD = 1.32 vs. $M_{\text{second (ambiguous)}}$ = 5.18, SD = 1.25). However, when the instructions were read in an ambiguous voice, giving participants more exposure and familiarity, they rated the option presented in an ambiguous gender voice directionally *more* favorably $M_{\text{first (ambiguous)}}$ = 5.60, SD = 1.24 vs. M

$_{\text{second (male)}}$ = 5.51, *SD* = 1.21; $M_{\text{first (male)}}$ = 5.21, *SD* = 1.35 vs. $M_{\text{second (ambiguous)}}$ = 5.25, *SD* = 1.40), though not significantly so. In addition, there was a significant interaction between rating order and the brand order counterbalance factor F (1, 292) = 4.08, p = .044, η_p^2 = .014 when we tested a model which included that factor.

REFRENCES

Alter, A. L., & Oppenheimer, D. M. (2008). Easy on the mind, easy on the wallet: The roles of familiarity and processing fluency in valuation judgments. *Psychonomic Bulletin & Review*, *15*(5), 985–990. https://doi.org/10.3758/PBR.15.5.985

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Baeck, H., Corthals, P., & Borsel, J. van. (2011). Pitch characteristics of homosexual males. *Journal of Voice*, 25(5). https://doi.org/10.1016/j.jvoice.2010.10.019

Birtel, M. D., & Crisp, R. J. (2012). "Treating" prejudice: An exposure-therapy approach to reducing negative reactions toward stigmatized groups. *Psychological Science*, 23(11), 1379-1386. https://doi.org/10.1177/0956797612443838

Cartei, V., Oakhill, J., Garnham, A., Banerjee, R., & Reby, D. (2020). "This is what a mechanic sounds like": Children's vocal control reveals implicit occupational stereotypes. *Psychological Science, 31*(8), 957-967. https://doi.org/10.1177/0956797620929297

Diel, A., & Lewis, M. (2022). The deviation-from-familiarity effect: Expertise increases uncanniness of deviating exemplars. *PLOS ONE*, 17(9), e0273861. https://doi.org/10.1371/journal.pone.0273861

Eberhardt, J. L., Dasgupta, N., & Banaszynski, T. L. (2003). Believing is seeing: The effects of racial labels and implicit beliefs on face perception. *Personality and Social Psychology Bulletin*, 29(3), 360-370. https://doi.org/10.1177/0146167202250215

Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology*, 42(9), 2213-2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

Huart, J., Corneille, O., & Becquart, E. (2005). Face-based categorization, context-based categorization, and distortions in the recollection of gender ambiguous faces. *Journal of Experimental Social Psychology*, *41*(6), 598-608. https://doi.org/10.1016/j.jesp.2004.10.007

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in Social Categorization: The Role of Prejudice and Facial Affect in Race Categorization. *Psychological Science*, 15(5), 342–345. https://doi.org/10.1111/j.0956-7976.2004.00680.x

Iyer, E. S., & Debevec, K. (1986). Gender stereotyping of products: are products like people?. In *Proceedings of the 1986 Academy of Marketing Science (AMS) Annual Conference 2014* (pp. 40-45). Springer. https://doi.org/10.1007/978-3-319-11101-8_9

Johnson, K. L., Lick, D. J., & Carpinella, C. M. (2015). Emergent research in social vision: An integrated approach to the determinants and consequences of social categorization. *Social and Personality Psychology Compass*, *9*(1), 15–30. https://doi.org/10.1111/spc3.12147

Ko, S. J., Judd, C. M., & Blair, I. V. (2006). What the voice reveals: Within-and between-category stereotyping on the basis of voice. *Personality and Social Psychology Bulletin*, *32*(6), 806-819. https://doi.org/10.1177/0146167206286627

Kruger, J., Epley, N., Parker, J., & Ng, Z.-W. (2005). Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology*, *89*(6), 925–936. https://doi.org/10.1037/0022-3514.89.6.925

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Lev-Ari, S., & Keysar, B. (2010). Why don’t we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Lick, D. J., & Johnson, K. L. (2013). Fluency of visual processing explains prejudiced evaluations following categorization of concealable identities. *Journal of Experimental Social Psychology*, *49*(3), 419–425. https://doi.org/10.1016/j.jesp.2013.01.001

Lick, D. J., & Johnson, K. L. (2015). The interpersonal consequences of processing ease: Fluency as a metacognitive foundation for prejudice. *Current Directions in Psychological Science*, 24(2), 143–148. https://doi.org/10.1177/0963721414558116

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6-27. https://doi.org/10.1037/met0000086

Morey R. D., Rouder J. N. (2011). Bayes factor approaches testing interval null hypotheses. *Psychological Methods*, 16(4), 406–419.

Mullennix, J. W., Johnson, K. A., Topcu-Durgun, M., & Farnsworth, L. M. (1995). The perceptual representation of voice gender. *The Journal of the Acoustical Society of America*, *98*(6), 3080–3095. https://doi.org/10.1121/1.413832

Nass, C., & Brave, S. (2005). *Wired for Speech*. MIT Press.

Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, *27*(10), 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

Newman, E. J., & Schwarz, N. (2018). Good sound, good research: How audio quality influences perceptions of the research and researcher. *Science Communication*, *40*(2), 246–257. https://doi.org/10.1177/1075547018759345

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin and Review*, 24(3), 856-862. https://doi.org/10.3758/s13423-016-1146-y

Owen, H. E., Halberstadt, J., Carr, E. W., & Winkielman, P. (2016). Johnny Depp, reconsidered: How category-relative processing fluency determines the appeal of gender ambiguity. *PLOS ONE*, *11*(2), e0146328. https://doi.org/10.1371/journal.pone.0146328

Peracchio, L. A., & Tybout, A. M. (1996). The moderating role of prior knowledge in schema-based product evaluation. *Journal of Consumer Research*, *23*(3), 177–192. https://doi.org/10.1086/209476

Pernet, C. R., & Belin, P. (2012). The role of pitch and timbre in voice gender categorization. *Frontiers in Psychology*, *3*, 23. https://doi.org/10.3389/fpsyg.2012.00023

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?. *Personality and Social Psychology Review*, *8*(4), 364-382. https://doi.org/10.1207/s15327957pspr0804_3

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*(1), 45-48. https://doi.org/10.1111/1467-9280.00008

Reis, H. T., Maniaci, M. R., Caprariello, P. A., Eastwick, P. W., & Finkel, E. J. (2011). Familiarity does indeed promote attraction in live interaction. *Journal of Personality and Social Psychology*, *101*(3), 557-570. https://doi.org/10.1037/a0022885

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science*, *26*(6), 877–891. https://doi.org/10.1177/0956797615572906

Schroeder, J., & Epley, N. (2016). Mistaking minds and machines: How speech affects dehumanization and anthropomorphism. *Journal of Experimental Psychology: General*, *145*(11), 1427–1437. https://doi.org/10.1037/xge0000214

Schwarz, K., Fontanari, A. M. V., Costa, A. B., Soll, B. M. B., da Silva, D. C., de Sá Villas-Bôas, A. P., Cielo, C. A., Bastilha, G. R., Ribeiro, V. V., Dorfman, M. E. K. Y., & Lobato, M. I. R. (2018). Perceptual-Auditory and Acoustical Analysis of the Voices of Transgender Women. *Journal of Voice*, *32*(5), 602–608. https://doi.org/10.1016/j.jvoice.2017.07.003

Song, H., & Schwarz, N. (2009). If it's difficult to pronounce, it must be risky: Fluency, familiarity, and risk perception. *Psychological Science*, *20*(2), 135–138. https://doi.org/10.1111/j.1467-9280.2009.02267.x

Sujan, M. (1985). Consumer knowledge: Effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research*, *12*(1), 31–46. https://doi.org/10.1086/209033.

Team, J. A. S. P. (2023). JASP (Version 0.17. 2)[Computer software].

West, M., Kraut, R., & Chew, H. E. (2019). *I'd blush if I could: Closing gender divides in digital skills through education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000367416

Wetzels, R., Raaijmakers, J. G., Jakab, E., & Wagenmakers, E. J. (2009). How to quantify support for and against the null hypothesis: A flexible WinBUGS implementation of a default Bayesian t test. P*sychonomic bulletin & review*, 16, 752-760.

Winkielman, P., Olszanowski, M. & Gola, M. (2015), Faces in-between: Evaluations reflect the interplay of facial features and task-dependent fluency. *Emotion*, 15(2), 232–42. https://doi.org/10.1037/emo0000036

Winkielman, P., Schwarz, N., & Belli, R. F. (1998). The role of ease of retrieval and attribution in memory judgments: Judging your memory as worse despite recalling more events. *Psychological Science*, 9(2), 124–126. https://doi.org/10.1111/1467-9280.00022

Wolchover, N. (2011). Why do news anchors all talk the same? *NBC News*. https://www.nbcnews.com/id/wbna44740700

Yamada, Y., Kawabe, T., & Ihaya, K. (2013). Categorization difficulty is associated with negative evaluation in the "uncanny valley" phenomenon: Categorization and evaluation. *Japanese Psychological Research*, 55(1), 20–32. https://doi.org/10.1111/j.1468-5884.2012.00538.x

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9* (2p2), 1–27. https://doi.org/10.1037/h0025848

CHAPTER 2

DISFLUENCY INCREASES RELIANCE ON HEURISTIC CUES IN CONSUMER CHOICE

ABSTRACT

This research provides evidence that when information is disfluent (difficult to process), consumers tend to simplify their decision-making process by relying on heuristic product information, rather than meticulously processing all available information. In other words, they rely more heavily on factors like brands, countries of origin, customer review ratings, or recommendations. This occurs because consumers conflate the feeling of disfluency with the difficulty of choosing. This effect of disfluency is especially strong for quick decisions, but also can influence more effortful ones. The findings challenge previous research on disfluency and dual processing modes and add to our understanding of brands and recommendations.

Keywords: processing disfluency, heuristics, brand, recommendation.

Marketing materials, including advertisements, apps, and websites, often present challenges to consumers due to factors such as unconventional fonts, distracting background images, and low image resolution. These elements contribute to disfluency–feelings of difficulty experienced when processing information. Understanding how disfluency influences consumer choices is a critical question. When consumers encounter disfluent information, they may hesitate to make a selection from the available options (Novemsky et al. 2007). However, if they do make a choice, disfluency may also impact which option they ultimately select, the focus of this research.

Consumers make decisions that rely primarily on two distinct strategies (Wright 1975). While some decisions are processed carefully and in-detail, others are made quickly and with simplified information. This type of dual-process framework predicts qualitatively different judgments depending on which reasoning strategy is used (Amit et al. 2021). Various dual-processing models differentiate *heuristic*, simple inferential, peripheral, or intuitive processing (i.e. system I) from *systematic*, detailed, central, or deliberate processing (i.e. system II) (Amit et al. 2021; Maheswaran and Levy 1990; Kahneman and Frederick 2002; Chaiken 1980, 1987; Petty and Cacioppo 1986, Chaiken and Maheswaran 1994). While heuristic processing utilizes easily-processed information and simplification strategies, systematic processing entails relatively thorough and analytic scrutiny of judgment-relevant information. We wondered if the ease or difficulty of processing information can affect which processing style a consumer will employ when making judgments and choices.

Past researchers have found that disfluency can lead to *systematic* reasoning during problem-solving tasks (e.g. Alter, Oppenheimer, Epley, and Eyre 2007; Alter 2013; Song and Schwarz 2008). They reasoned that disfluency serves as a metacognitive cue that something is amiss. Therefore, more careful attention should be paid, in order to understand the source of the uneasy feeling. This heightened attention should lead to better performance. However, other scholars have recently cast doubt on these findings, arguing that disfluency *does not* lead to analytical/deliberate processing (e.g. Meyer, Frederick, Burnham, et al. 2015; Thompson et al. 2013). While our research also questions whether disfluency leads to systematic processing, we go one step further to show that, in a consumer choice context, disfluency can serve as a metacognitive signal that choosing may be difficult. In turn, this can lead consumers to simplify the decision-making process by using heuristic product information rather than fully and carefully processing all the available information. That is, they may rely more heavily on heuristic-based attributes, such as a brand name (Lee and Shavitt 2009; Maheswaran et al. 1992; Gunasti and Ross Jr 2010; Yan and Duclos 2013), a country of origin (Maheswaran 1994; Chang 2004; Hong and Wyer Jr 1989) or a recommendation (e.g. Senecal and Nantel 2004; Mariadassou et al. 2023; Munz and Morwitz 2019; Dellaert and Häubl 2012).

This research contributes to a scholarly understanding of disfluency and dual processing models. We present seven pre-registered studies in which we manipulate disfluency to show how disfluency can lead consumers to rely on heuristic cues. We contribute to understanding when heuristic cues such as a brand name, country of origin, review score, or recommendation will be used by consumers. In addition to demonstrating the mediating role of heuristic reasoning on the effect of disfluency on a choice share, we provide additional evidence for our theory by examining the moderating effects of attribution and processing style. We further show this effect in a choice setting with actual marketplace consequences.

THEORETICAL BACKGROUND

In this section, we draw upon research on disfluency (e.g. Novemsky et al. 2007; Alter et al. 2007; Schwarz et al. 2021), processing styles (e.g. Amit et al. 2021; Kahneman and Frederick 2002; Chaiken 1980, 1987; Chaiken and Maheswaran 1994; Maheswaran and Levy 1990), and the relationship between the two to derive our prediction that disfluency increases the use of heuristic cues rather than systematic cues in a consumer choice context.

Processing Disfluency

Processing disfluency refers to the subjective feeling of difficulty experienced when performing a cognitive task like decision-making (Novemsky et al. 2007; Schwarz 2004). Numerous aspects of the way a target object is presented can contribute to processing disfluency, which in turn can have an impact on how well the target object is perceived and hence evaluated.

A substantial body of research has documented the impact of disfluency on product judgment and evaluation, focusing on different consumption domains (e.g. Pocheptsova et al. 2010), brand evaluation (e.g. Mohsenin and Munz 2023; Lee and Labroo 2004; Lee and Shavitt 2009; Labroo et al. 2009), price perception (e.g. Janiszewski and Meyvis 2001; Thomas and Morwitz 2009; Wadhwa and Zhang 2015), and choice preferences (e.g. Novemsky et al. 2007; Dhar 1997; Walter et al. 2020). Consistently, experiencing disfluency leads to less favorable judgments and evaluations (see Schwarz et al. 2021) and can lead to deferral in a choice context (Novemsky et al. 2007). These effects are said to be due to consumers misattributing the negative feeling of difficulty processing information about a product to being a reaction to the product itself.

While researchers have long documented the role of disfluency in consumers' judgments and product evaluations, less attention has been paid to the question of *how* disfluency affects information processing in choice-making settings. As we alluded to earlier, most existing research suggests that disfluency leads to more systematic or analytical processing (Alter et al. 2007; Diemand-Yauman et al. 2011; Song and Schwarz 2008), though recently this idea has come into question (Meyer et al. 2015; Thompson et al. 2013; Sirota, Theodoropoulou and Juanchich 2021a). A helpful framework for understanding how fluency might affect cognitive processing involves dual systems of processing, namely heuristic, peripheral or System I processing and systematic, central, or System II processing (Maheswaran and Levy 1990; Kahneman and Frederick 2002; Chaiken 1980, 1987; Chaiken and Maheswaran 1994; Petty and Cacioppo 1986).

Dual Processing Systems: Heuristic and Systematic

A *dual-system reasoning framework* (Kahneman and Frederick 2002; Chaiken 1980, 1987; Chaiken and Maheswaran 1994; Chaiken and Ledgerwood 2011; Petty and Cacioppo 1986) could provide a basis for predicting how disfluency influences information processing. Information processing is often conceived as falling along a continuum between more heuristic or peripheral versus more systematic or central types of processing (e.g. Chaiken 1980; Petty and Cacioppo 1984), where heuristic processing is a less effortful type of decision making (e.g. Chaiken and Maheswaran 1994). We will use the labels heuristic and systematic processing hereafter. Decision makers relying on heuristic processing are more likely to make inferences using only the information directly available to them (e.g. Bodenhausen and Wyer 1985;

Schwarz 1998), often leading to less accurate judgments (e.g. Payne, Bettman, and Johnson 1988). Heuristic processing involves using relatively salient and easily comprehended cues that activate well-learned judgmental rules or "heuristics" as mental shortcuts. Systematic processing, on the other hand, involves attempting to thoroughly understand any available information through careful attention, deep thinking, intensive reasoning, and comprehensive and analytical scrutiny of judgment (Maheswaran, Mackie, and Chaiken 1992). Some scholars also conceptualize systematic thinking as serving to monitor and update intuitive processing (e.g. Simmons and Nelson 2006).

Since systematic processing occurs when people possess both high motivation and high ability, when lacking them, people often utilize heuristic cues (e.g. Chen, Duckworth, and Chaiken 1999; Ratneshwar and Chaiken 1991) as a mental shortcut to simplify their decision task. Maheswaran (1994) has found that consumers apply stereotypes based on country of origin under conditions of low ability and argument ambiguity. On this line of reasoning, Chang (2004) showed that relying on country-of-origin cues to infer product quality and form product evaluations is rooted in heuristic processing, in line with research showing that brand name acts as a heuristic cue (Maheswaran et al. 1992).

While low motivation and low ability have been shown to influence a consumers' use of a heuristic processing style, here, we explore how disfluency can also determine how people process information. Thus, in this research, in a consumer choice context, we intend to answer the question, "does disfluency affect which processing style a consumer will employ when making judgments and choices?"

Disfluency and Processing Systems

Researchers have already examined this question of how disfluency affects processing style. Some results suggests that disfluency leads to systematic processing, while others find no relationship. Finally, while not investigating this question directly, some evidence suggests that the opposite may be possible; that disfluency could lead to relying on heuristics. We next consider each of these in turn.

Disfluency increases systematic processing. Alter et al. (2007) provided preliminary evidence that people take a systematic, analytical approach when they experience disfluency. They showed these effects by documenting improved performance on the Cognitive Reflection Test (CRT) and a six-item syllogistic reasoning task. The CRT is a test in which simple math problems seem to have an intuitive answer that proves incorrect upon more careful consideration (Frederick 2005). A syllogism is a kind of logical reasoning about relationships. In both problem-solving tasks, the authors found participants in a disfluent-font condition answered a greater proportion of the questions correctly than those in a fluent-font condition. They also provided evidence that when information in a product review was presented in a more difficult-to-read font, people were more likely to adopt a systematic approach to information processing. That is, their favorability ratings were more heavily influenced by the quality of the arguments than by heuristic cues such as the attractiveness of the communicator.

Moreover, Song and Schwarz (2008) found that reduced processing fluency attenuates the Moses illusion. When participants were asked, "*How many animals of each kind did Moses take aboard the ark?*" 88% responded erroneously with two. The correct answer is "none" because Moses had nothing to do with *Noah's* ark. Thus, the Moses illusion is the failure to notice the swapped names. They found that the error rate dropped from 88% to 53% when the question was posed in a disfluent font.

Further supporting the idea that disfluency leads people to pay more careful attention, Diemand-Yauman et al. (2011) found that reading information written in a difficult-to-read font increased the accuracy of recall by 14%. Simmons and Nelson (2006) also found that presenting information in a hard-to-read font led gamblers to bet against their favorite more often, as with more careful attention, they more carefully considered the point spread. Similarly, Walter and his colleagues (2020) found that an unsystematic spatial arrangement of a set of products resulted in perceptual disfluency. This, in turn, promoted the choice of unfamiliar products, as consumers were motivated to explore a greater number of alternatives in the assortment.

Disfluency does not increase systematic processing. While these authors have argued that the results arose because people spent more time and effort processing information when experiencing disfluency, Meyer and his colleagues (2015) found no evidence that disfluency leads to analytical reasoning or systematic processing. Their aggregated data from 16 experiments failed to provide evidence that disfluent fonts affect performance on the CRT or the syllogism task. In addition, Meyer et al. (2015) also examined potential moderating variables, including cognitive capacity, presentation format, and experimental setting. Still, they found no benefits for a disfluent font under any conditions.

Thompson and her colleagues (2013) also challenged the association between perceptual fluency and analytical or systematic thinking, finding no relationship. They examined the impact of the disfluency on a wide range of analytical processing tasks, including rethinking (i.e. probability of changing answers), processing time, accuracy (i.e. probability of correct answers), answer fluency (i.e. how easily and quickly a response comes to mind), and Feelings of Rightness rating (FOR; Thompson et al. 2011). They found that disfluency did not affect FOR or any other analytical engagement indicators (i.e. response time and answer changes). They also conducted a series of experiments to examine the beneficial effects of perceptual disfluency on deductive reasoning, including the CRT and the two syllogism reasoning tasks. While they failed to observe any evidence that creating disfluent processing conditions produced more correct solutions generally, they did find that the perceptual disfluency prompted additional processing, and, in turn, accuracy on the CRT among people with high cognitive ability.

More recently, Sirota and his colleagues (2021a) examined the impact of perceptual disfluency on cognitive reflection by using both verbal (Sirota et al. 2021b) and standard numerical CRT. They also failed to find evidence supporting the hypothesis that a disfluent font increases systematic thinking.

Disfluency increases heuristic processing. When examining the effects of disfluency on consumer choices, Novemsky and his colleagues (2007) found that greater disfluency during decision making led to either a compromise effect or choice deferral. While the authors did not interpret these results as aligning with a particular processing style (heuristic vs. systematic), we note that compromise and deferral could be considered ways to simplify a difficult decision (Dhar and Simonson 2003). Thus, their results could be considered evidence that disfluency may encourage consumers to process information more heuristically, as we have proposed. While our findings align with those of Novemsky et al, including replicating the main finding regarding deferral, we situate their findings as part of a broader framework. By directly considering heuristic versus systematic processing styles, we can additionally make predictions about consumers' use of *heuristic cues* as a function of disfluency. From this perspective, heuristic cues such as a brand name or country of origin can have more influence over a choice under disfluent processing conditions compared fluent conditions, with broad theoretical and managerial implications.

Other sources of decision difficulty and dual processing

In research on disfluency, the difficulty of the task itself is held constant. Consumers face the same set of choice options, encounter the same information about them, and are given ample time to consider them. Disfluency does not derive from the difficulty of the choice task itself, but rather from ancillary factors such as the font in which the information is presented. However, difficulty can also arise from the decision itself. For example, choosing from a large set of options can be more difficult than choosing among fewer (Iyengar and Lepper 2000), and choosing can also be more difficult when consumers must make tradeoffs, such as when they must pay a higher price for higher quality (Bettman et al. 1993). In these cases, choosers often simplify difficult decisions by adopting a more heuristic processing style. For example, when faced with many options, consumers may be less likely to make a purchase (Iyengar and Lepper, 2000; Gao and Simonson 2016). Similarly, consumers may defer more often when making a difficult decision requiring tradeoffs, especially when under time pressure (Dhar and Nowlis 1999). Time pressure can also lead to social mimicry, with consumers more likely to interpret the choices of others as acceptable defaults, rather than carefully processing the options. (Huh et al. 2011).

Situational differences can also play a role in determining the depth of processing in which a consumer will engage. For example, use of heuristics increased under low processing effort (e.g. Allard et al. 2019; Chaiken et al. 1989; Kaplan et al. 1993), but when processing was effortful, there was a greater likelihood that consumers would process information systematically (Suri and Monroe 2003). These findings show that when consumers do not have available resources to carefully process information, they tend to rely on heuristic processing strategies.

Along these lines, Pocheptsova et al. (2009) showed that people rely more heavily on simpler, more intuitive styles of thinking in their decisions when they are depleted. Moreover, Mukherjee and Lee (2016) have shown a positive effect of scarcity appeals on product evaluation when cognitive load is high. Considering scarcity to be a heuristic cue signaling value and premium quality (e.g. Brannon and Brock 2001; Inman et al. 1997), their findings could show that the more favorable evaluation of a scarce option roots in heuristic reasoning.

Furthermore, compared to when information is presented as written text, difficulty comparing options when they are spoken aloud can lead to choosing a recommended option more frequently (Mariadassou et al. 2023; Munz and Morwitz 2019) and to deferring choice at a higher rate (Munz and Morwitz 2019). These findings align with the idea that the fleeting nature of auditory information intensifies cognitive difficulty (Singh et al. 2012). If accepting a recommendation is a heuristic means of coping with the difficulty of the decision-making task, these results provide further evidence of cognitive difficulty leading to heuristic information processing.

Altogether, these findings are consistent with our notion that when faced with a difficult task, consumers will be more likely to use heuristic simplifying strategies.

CURRENT RESEARCH

We propose that within a consumer choice context, disfluency may lead individuals to rely on heuristic cues, such as brand names or recommendations, as a means to simplify their decision-making process. We predict that disfluency will lead consumers to use a heuristic processing style, which will in turn cause them to focus on heuristic product attributes as the basis for their choice. Importantly we predict that this will occur because consumers will mistake the feeling of disfluency as being about the difficulty of the choice (Novemsky et al. 2007).

Our research emphasizes metacognitive experiences in decision-making, contrasting with prior studies that induced difficulty by altering decision content or introducing factors like time constraints or reduced cognitive capacity. We manipulate disfluency by external variables (fonts, background images, image resolution) while holding traditional factors known to influence difficulty constant. We expect disfluency to mainly impact consumer choices when the disfluency is attributed to the choice. If consumers are made aware that the disfluency is from an irrelevant source, its impact should diminish. Hence, reliance on a heuristic cue should increase when disfluency is linked to perceived decision difficulty.

Our theoretical framework is grounded in the concept of heuristic-systematic dual-processing model. We posit that the impact of disfluency on consumer choice decisions is mediated through consumers' decision-making styles, wherein disfluency enhances heuristic reasoning and increases reliance on heuristic product information.

We describe seven pre-registered studies testing whether disfluency leads consumers to use heuristic cues as a means to simplify a difficult-seeming task (Table 1). The studies measure purchase decisions in both a forced-choice setting and when consumers can defer choice, and test for effects across several heuristic cues (e.g. a brand name, country-of-origin, and recommendation). Furthermore, effects are demonstrated in both hypothetical and consequential choice settings, and in choices between several options. Disfluency was manipulated using the most commonly used manipulation in the literature (i.e. font type), as well methods commonly encountered in the marketplace (i.e. text appearing on a background image and poor image resolution). In addition, the studies supported our conceptualization through both moderation (e.g. attribution, processing effort) and mediation. The effect is robust either in memory-based (i.e. sequentially) or stimulus-based (i.e. simultaneously) evaluations.

All of the studies followed preregistered plans for data collection and analysis. For all studies, supplemental material, data, pre-registrations, and survey materials are posted on ResearchBox (https://researchbox.org/686). To ensure participants' rights were protected, an Institutional Review Board (IRB) approved the experiments.

TABLE 11

SUMMARY OF STUDIES AND RESULTS

	Stimuli	Heuristic cue	Disfluency Manipulation	Dependent Variable (Mediator)	Moderator	Disfluent	Fluent	Result and Effect
Study 1 (N = 602)	Blender	Brand name / Country of origin	Font type [S]	1) Choice share of a better brand	-	33%	23.5%	Wald $\chi^2(1) = 6.64$ $p = .01$, $OR = 1.60$
				2) Choice deferral		15.3%	6.6%	Wald $\chi^2(1) = 11.09$, $p = .001$, $OR = 2.55$
Study 2 (N = 1807)	Coffee grinder [IC, R]	Brand name	Font type [Sq]	Choice share of a better brand	-	64.4.9%	54.2%	Wald $\chi^2(1) = 19.42$, $p < .001$, $OR = 1.53$
Study 3 (N = 600)	Air fryer [R]	Brand name	Background image [Sq]	Choice share of a better brand	Attribution absent	61.2%	47.7%	Wald $\chi^2(1) = 5.46$, $p = .019$, $OR = 1.74$
					Attribution present	50.3%	49%	Wald $\chi^2(1) = .053$, $p = .82$, $OR = 1.05$
Study 4 (N = 600)	Toaster	Recommendation	Font type [S]	Choice share of a recommended option	-	24.9%	14.4%	Wald $\chi^2(1) = 10.32$, $p = .001$, $OR = 1.97$
				(Heuristic-systematic attributes listed as reasons for choice)		M =.14, *SD* = .98	M =.38, *SD* =.96	$b_{\text{indirect\_effect}} = 0.14$, $SE = 0.06$, 95% CI [0.05, 0.27]
Study 5 (N = 600)	Blender [R]	Brand name	Background Image [Sq]	Choice share of a better brand	-	65.6%	42.2%	Wald $\chi^2(1) = 32.30$, $p < .001$, $OR = 2.61$
				(Intuitive vs. systematic thinking style scale)		M = 4.16, *SD* = 1.81	M = 3.68, *SD* = 1.85	$b_{\textit{indirect\_effect}} = -0.21$, $SE = 0.07$, 99% CI [0.04, 0.42]
Study 6 (N = 602)	Printer [R]	Brand name	Background Image [Sq]	Choice share of a better brand	Fast, low effort decision	77%	48%	Wald $\chi2(1) = 25.94$, $p < .001$, $OR = 3.62$
					Slow, high effort decision	56.3%	42.3%	Wald $\chi2(1) = 5.85$, $p = .016$, $OR = 1.76$
Study 7 (N = 1202)	Blender [R]	Brand name	Image Resolution [Sq]	Choice share of a better brand	-	41.3%	32.8%	Wald $\chi^2(1) = 9.14$, $p = .002$, $OR = 1.44$

[IC] Incentive-Compatible choice setting.
[R] Real brand names.
[S] Simultaneous evaluation.
[Sq] Sequential evaluation.

STUDY 1: DEFERRAL AND CHOICE SHARE

The purpose of study 1 was to provide evidence that disfluency can increase consumer use of simplification strategies; either choosing not to choose (deferral) or choosing an option with a better heuristic cue. We predicted that when information about options is presented in a disfluent font, consumers would choose based on heuristic cues, preferring better brands (Maheswaran et al. 1992) and more favorable countries-of-origin (Maheswaran 1994; Chang 2004) than they would if the information were instead presented in a more fluent font. We have also argued that deferring choice may be a simplification strategy, reasoning that choosing not to choose can be a way for consumers to avoid a difficult-seeming decision task. Thus, we also predict that consumers will defer choice at a higher rate when the font is disfluent (vs. fluent), replicating past results while putting them in a broader framework (Novemsky et al. 2007). In all studies, participants received identical descriptions of a set of choice options where one dominated on systematic cues and the other dominated on heuristic cues. In this study, we used the most common disfluency manipulation in the existing literature, print font (e.g. Alter et al. 2007; Novemsky et al. 2007; Thompson and Ince 2013; Sela and Berger 2012; Lee and Labroo 2004; Song and Schwarz 2008). This manipulation is said to hold the difficulty of the actual task constant while manipulating the subjective feeling of difficulty (Novemsky et al. 2007).

Method

Six hundred and two American and Canadian participants (M_{age} = 39.15 *SD* = 11.32; 48.8% Female) from Amazon Mechanical Turk (Mturk) were paid a small sum for completing this experiment. All participants who completed the experiment were included in the analysis. As with many consumer purchase situations, participants could choose between the options or defer the choice (choose none). This study was a two-cell between-subjects design manipulating disfluency. Following the existing literature, we manipulated disfluency by presenting the two options to half of the participants with the descriptions presented in an easy-to-read standard font (Arial font, see web appendix), while the others saw the same descriptions in a hard-to-read font (Amarix SC, cursive font, web appendix).

The stimuli for this experiment were two blenders. Participants were asked to imagine that they needed to buy a new blender to make smoothies and juices. One of the blenders was superior based on systematic cues (i.e. it had higher wattage, more capacity, and blending functions), while the other blender had a better brand name (Hamilton Beach vs. iTOChin), a heuristic cue (Maheswaran et al. 1992).

Brand name and country of origin as heuristic cues. Participants chose between options that varied in brand familiarity. There was a well-known brand in the US market (Hamilton beach) and an unknown foreign brand not available in the participants' market (iTOChin). We selected these brands as a strong initial test of our hypothesis. That is, both brand (Maheswaran et al. 1992) and country-of-origin (Maheswaran 1994; Chang 2004) can be heuristic cues. Since iTOChin strongly suggests that the product may be Chinese, it might serve as a negative heuristic cue to some participants (Maheswaran 1994). To verify, after the dependent variable, participants responded to items that assessed the familiarity of each brand (Lee and Shavitt 2009; Maheswaran et al. 1992) as a manipulation check. Participants were asked to rate "How familiar were you with the survey blender brands?" (1=Not at all familiar; 7 = Extremely familiar). Some information, such as price, products' material, and weight of the options kept constant between the two options. All participants saw both options presented in a top-and-bottom format on the same page. The presentation order of brands was counterbalanced.

The primary dependent variable was the response to the question, "If you were going to buy a blender, which one would you like to choose?" Participants could select one of three options: Blender A, Blender B, or "For now, I would not buy either one of these." For a manipulation check of font disfluency, participants were asked to rate "How difficult was it to read the information?" (1= Not at all difficult, 7=Very difficult). For a manipulation check of brand names, we asked participants to answer the brand familiarity question adapted from Lee and Shavitt (2009), and Maheswaran et al. (1992) explained earlier.

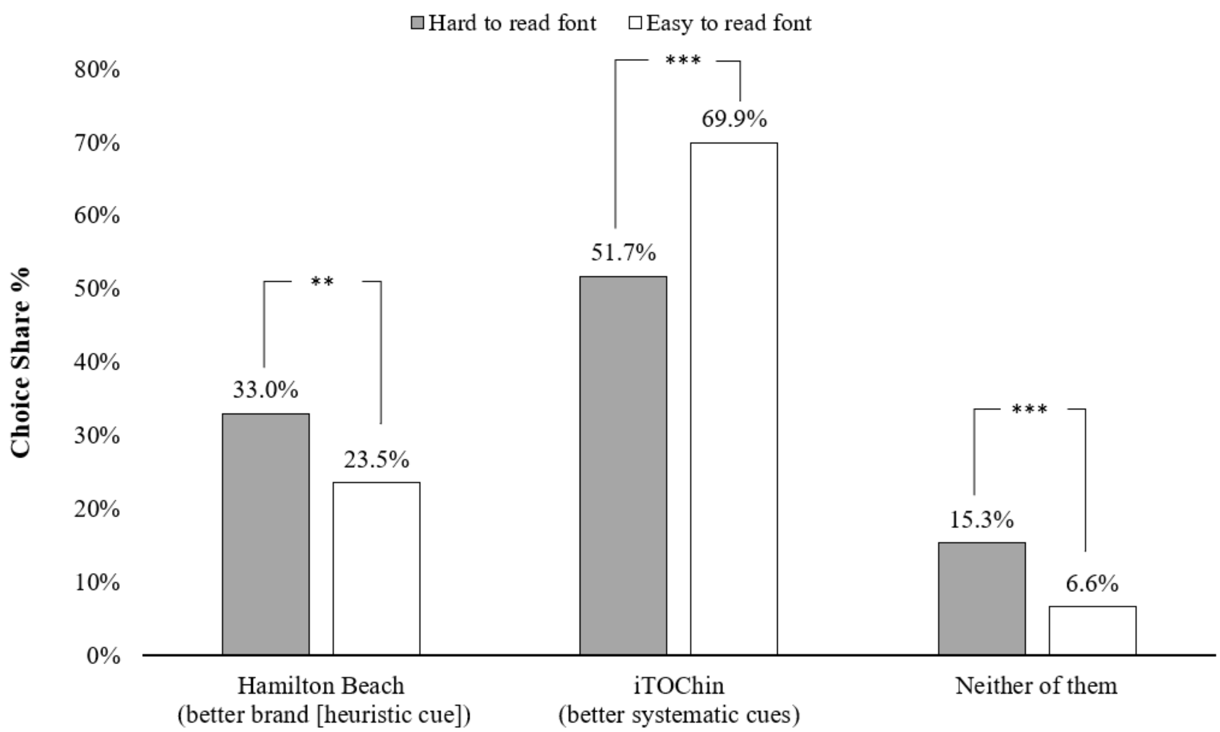
Results

Manipulation checks. The manipulation of fluency was successful. Participants rated the information presented in Amarix SC, cursive font as more difficult to read than the easy-to-read Arial font ($M_{hard\text{-}to\text{-}read\text{-}font} = 4.83$, $SD = 1.91$ vs. $M_{easy\text{-}to\text{-}read\text{-}font} = 1.78$, $SD = 1.59$, $t(600) = 23.74$ $p < .001$, $d = 1.93$). The manipulation of the brand name was also successful. A paired-sample t-test on brand name familiarity revealed that participants were more familiar with *Hamilton beach* (vs. *iTOChin*) brand name ($M_{Hamilton\_beach} = 5.12$, $SD = 2.00$ vs. $M_{iTOChin}$ 1.47, $SD = 1.16$, $t(601) = 39.48$, $p < .001$, $d = 1.61$).

Dependent variable. Following our preregistration, the result showed that there was no significant brand order effect ($p = .083$) or interaction of order with the font type ($p = .702$) on choice share. Thus, we collapsed across counterbalancing orders. We tested for overall differences in choices first. To do so, we ran a chi-square test with disfluency (hard-to-read, easy-to-read font) as the independent variable and choice share (Hamilton beach = 1; iTOChin = 2; and deferral = 3) as the dependent variable, revealing significant differences ($\chi^2(2) = 23.42$, $p < .001$, $V = .197$, see Figure 15).

FIGURE 15

DISFLUENCY LEADS CONSUMERS TO SIMPLIFY THE CHOICE BY CHOOSING A BETTER BRAND (LEFT TWO BARS) OR DEFERRING CHOICE (RIGHT TWO BARS).



NOTE. — ***$p \leq .001$; **$p \leq .01$

Next, we recoded the DV to be binary (*did the participant choose the well-known brand or not?*) and analyzed the data with a chi-square test. This is the primary test of the prediction that reading choice options in a hard-to-read font leads people to choose a well-known brand more often. The results supported this prediction. Whereas 23.5% of participants who received the easy-to-read font chose the Hamilton beach, 33% of those who received the hard-to-read font did so (Wald $\chi^2(1) = 6.64$, $p = .01$, $OR = 1.602$, 95% CI [1.12, 2.29]).

Finally, we recoded the original DV to be binary (*did the participant defer or not?*) and analyzed the data with a chi-square test. This is the primary test of the prediction that people are more likely to defer in the disfluent condition (Novemsky et al. 2007). The results supported this prediction. Whereas 6.6% of participants who received the easy-to-read font deferred, 15.3% of those who received the hard-to-read font did so (Wald $\chi^2(1) = 11.09$, $p = .001$, $OR = 2.55$, 95% CI [1.47, 4.43]) replicating previously reported findings (Novemsky et al. 2007).

Memory Checks. For exploratory purposes, to check whether participants differed in whether or not they recalled the systematic product information by condition, we asked, “Which blender has only 5 blending functions?” (Hamilton Beach or iTOChin), and “What was the wattage of Blender A/B [iTOChin]? (700w, or 900w).” Participants in the easy-to-read condition answered correctly more often than participants in the hard-to-read font condition for both items (blending function: 89.4% vs. 84% correct; ($\chi^2(1) = 3.81$, $p = .051$); blender power in wattage: 91.7% vs. 79.3% correct; ($\chi^2(1) = 18.83$, $p < .001$). This may indicate a shift in attention, consistent with our theory that disfluency leads to a focus on heuristic cues (and less focus on systematic cues), or it may indicate that participants in the hard-to-read font condition simply did not read. To help disentangle these possibilities, we analyzed the responses of 76.4% of participants ($N = 460$) who responded correctly to both memory-check questions, reasoning that participants must first read the information in order to remember it. The same pattern of results emerged as presented earlier. Specifically, consumers chose the better brand more often when it was difficult to read the information (Wald $\chi^2(1) = 5.13$, $p = .023$, $OR = 1.63$, 95% CI [1.07, 2.49]), and they were also more likely to defer choice (Wald $\chi^2(1) = 6.09$, $p = .014$, $OR = 2.25$, 95% CI [1.18, 4.28]). Thus, the overall results are unlikely to be explained by participants failing to read.

Reading Time. We also explored differences in reading time. Participants in the hard-to-read font condition spent on average around two seconds more time reading the information, although the time difference was not significant ($M_{\text{hard-to-read-font}} = 33.88$ seconds, $M_{\text{easy-to-read-font}} = 31.74$ seconds; $p = .68$); providing further evidence that participants seeing the hard-to-read font did not simply skip the information.

Discussion

Study 1 provided evidence for our hypothesis that disfluency leads to choice simplification in general and, specifically, using heuristic cues to make a choice. Consumers were more likely to choose the option with a better brand name when the information was presented in a hard-to-read font. The result is consistent with the idea that consumers process a brand name (Gunasti and Ross. Jr 2010; Maheswaran et al. 1992) or a country-of-origin (Maheswaran 1994; Chang 2004; Hong and Wyer Jr 1989) as a heuristic cue. Moreover, we replicated past findings that consumers defer at a higher rate under disfluent processing conditions (Novemsky et al 2007), interpreting them in a broader framework of heuristic versus systematic processing. Both findings support our line of reasoning regarding the role of using a heuristic processing style to simplify a decision task when it feels disfluent. However, since the effect on heuristic cues is novel to this research, we focus on it in the remaining studies.

STUDY 2: INCENTIVE-COMPATIBLE CHOICE

In Study 1, we showed that disfluency can lead consumers to simplify a difficult-seeming decision by either deferring choice or choosing based on brand, a heuristic cue. However, the hypothetical nature of the choice may be considered a limitation. This study examined how disfluency can lead to heuristic reasoning in an incentive-compatible choice setting. Participants were entered into a lottery (adapted from Song and Sela 2023) for a chance to win their chosen outcome, giving them an incentive to choose well. In this study, we provided participants with true information about two real-world coffee grinders. The coffee grinders differed in their systematic and heuristic attributes. We expected participants to use the brand names as heuristic cues, choosing the more familiar brand of coffee grinder more often when it was hard to read information.

Method

Eighteen hundred and seven American consumers recruited through Prolific participated in the study for a small monetary payment (M_{age} = 40.51, *SD* = 14.11; 5.4% Female). There were no data exclusions. In this pre-registered study, we used a two-cell between-participants experimental design manipulating disfluency using the same fonts as in Study 1. Before being assigned to one of the font conditions, participants were truthfully told that in addition to their monetary reward, one randomly selected person would receive their chosen option, to make the decision incentive-compatible[1]. We selected two real-world brands and their information to present to participants. While one of the coffee grinders had a well-known, familiar brand (a heuristic cue), "Cuisinart," the other one, "Wirsh" had better systematic features regarding power and capacity.

In Study 1, we observed the predicted effect using a stimulus-based (i.e. simultaneous) evaluation format. To show that the effect is robust in memory-based evaluation format, all participants saw both options presented sequentially on different screens. The presentation order was counterbalanced.

The dependent measure was the choice of coffee grinders from a binary choice set. After reading the options' information, participants were asked to make a choice by asking, "If you were going to buy one of them, which one would you like to choose? (Cuisinart coffee grinder vs. Wirsh coffee grinder)." After responding to the dependent measure, participants answered a series of questions. To check our disfluency manipulation, we asked participants, "How difficult was it to read the information?" (1= Not at all difficult; 7=Very difficult). For a manipulation check, we asked participants to answer the brand familiarity question, "How familiar were you with the survey coffee grinders brands? (1 = Not at all familiar; 7 = Extremely familiar)."

Results

Manipulation checks. The manipulation of fluency was successful. As expected, the hard-to-read font was rated significantly more difficult to read than the easy-to-read font ($M_{hard\text{-}to\text{-}read\text{-}}$

[1] Privacy protections on Prolific prevent mailing actual products, so the monetary value was bonused to the winning participant instead.

$_{font}$ = 5.46, *SD* = 1.65 vs. $M_{easy\text{-}to\text{-}read\text{-}font}$ = 2.32, *SD* = 1.59, *t* (1805) = 41.21, *p* < .001, *d* = 1.94). Moreover, the manipulation of brand familiarity was also successful. A paired-sample t-test on brand name perception revealed greater familiarity with *Cuisinart* than *Wirsh* coffee grinder ($M_{wellknown\_option}$ = 4.82, *SD* = 2.09 vs. $M_{lessknown\_option}$ = 1.58, *SD* = 1.22, *t* (1806) = 50.01, *p* < .001, *d* = 1.36).

Choice share. We ran a logistic regression with font type as the independent variable and the well-known coffee grinder choice share (Cuisinart = 1; Wirsh = 0) as the dependent variable[2]. We predicted that when the descriptions of the options were presented in a hard-to-read font, people would choose the more familiar brand name more often. The result of the logistic regression supported our prediction. Amongst participants who saw the products' information in a hard-to-read font, 64.4 % chose the *Cuisinart* coffee grinder, whereas only 54.2% of those who saw the information in easy-to-read font did so (Wald $\chi^2(1) = 19.42$, *p* < .001, *OR* = 1.53, 95% CI [1.27, 1.85]), which replicates the result of Study 1 and supports our main hypothesis. Hence, this study showed that the effect can manifest in an incentive-compatible choice setting.

STUDY 3: METACOGNITIVE NATURE OF THE EFFECT

The purpose of Study 3 was threefold. First, this study sought to replicate our previous findings. Second, this study aimed to examine another manipulation of disfluency which is common in the marketplace. Frequently, retailers' web sites contain photos of products being

---

[2] There was a significant order effect such that the first coffee grinder was chosen more often. However, there was no interaction with the font manipulation, so we presented the results collapsed across order.

used by consumers, and often information about the product is written on these images (see web appendix). We argue that the written information with these photos in the background can sometimes be difficult to read in the same way that a font can sometimes be difficult to read. Thus, this type of product presentation can increase the disfluency, which could lead to choosing based on heuristic cues. Third, to test if the effects we have observed are due to a metacognitive process, we employed an attribution paradigm, following existing research (Novemsky et al. 2007; Lee and Shavitt 2009; Pocheptsova et al. 2010; Sela and Berger 2012). In a metacognitive explanation for disfluency-based effects, a consumer is thought to misattribute the feeling of difficulty of reading the information to be about difficulty choosing between the options. Thus, if the experimenter makes a participant aware that the feeling of difficulty is instead due to the manipulation, the effects on choices should be diminished, as they should be less likely to use a simplifying strategy if the choice does not seem difficult. We use this paradigm in the current study.

In sum, we expected that people would choose a well-known brand at a higher rate when disfluency was high. However, we expected the effect to be reduced or eliminated when participants correctly attributed the difficulty to the background image. That is, this attribution should undermine (or explain away) the informational value of their difficulty experience (Novemsky et al. 2007; Lee and Shavitt 2009; Pocheptsova et al. 2010), resulting in lower rates of heuristic reasoning that are similar to the fluent condition.

Method

In this pre-registered study, we obtained data for 600 American and Canadian participants who completed the study on desktop computers (M_{age} = 40.96, SD = 12.05; 45.3% Female) from Mturk in exchange for a small payment. All participants were included in the analysis. We used two air fryer brands as stimuli in this study. This study followed a 2 (Disfluency: low [no background image] vs. high [a background image]) by 2 (Attribution: present vs. absent) between-participants design. To manipulate disfluency, we used a presentation format that closely matched real product presentations on Amazon.com. We presented the air fryers to half of the participants with the textual information presented over a background image, while the others saw the same information about the same air fryers without a photo background. In the disfluent condition, we used two different background images for each option, counterbalancing the order.

We crossed the background image manipulation with a manipulation aimed at redirecting the attribution of difficulty arising from the background image as a source of disfluency. Half of the participants in each image condition received an additional sentence that stated, "This information presented next is hard to read because of the background image." Such information was not provided for the other half.

Participants were asked to imagine that they needed to buy a new air fryer. Similar to the previous studies, one of the options was superior on its systematic cues (i.e. cooking functions, wattage power, oil reduction, capacity, and voice control), while the other choice had only a better brand name, a heuristic cue. We selected two existing air fryer brands on the market, Ninja, as an established brand, and *Cosori*, as a less established brand. Other information, such as the position of information on the backgrounds, air fryers' price, color, and material kept constant between the two options. All participants saw both options presented sequentially on different screens (see web appendix). The presentation order was counterbalanced.

The dependent measure was the binary response to the question, “If you were going to buy one of them, which one would you choose?” (Ninja or Cosori). After responding to the dependent measure, participants answered a series of questions. To check whether our brand name and disfluency manipulations were effective, we asked the same questions as in the previous studies. Near the end, we included two memory check questions to assess the extent to which participants considered systematic cues by asking, “What was the number of Ninja air fryer's cooking functions?” (5 on-touch cooking functions or 8 on-touch cooking functions), “Which air fryer did give you a voice-activated cooking control?” (Ninja or Cosori).

Results

Manipulation checks. Participants rated the information presented with a background image as more difficult to read than without a background image ($M_{background\_image} = 5.28$, $SD = 1.76$ vs. $M_{white\_background} = 3.51$, $SD = 1.94$, $t\ (598) = 11.66$, $p < .001$, $d = .96$) regardless of whether they were informed about the source of disfluency or not ($F\ (1, 596) = .23$, $p = .63$). Moreover, the manipulation of brand familiarity was also successful. A paired-sample t-test on brand name familiarity revealed more familiarity with the Ninja (vs. Cosori) brand name ($M_{ninja} = 4.49$, $SD = 2.19$ vs. $M_{cosori} = 2.35$, $SD = 1.90$, $t\ (599) = 22.10$, $p < .001$, $d = 0.90$).

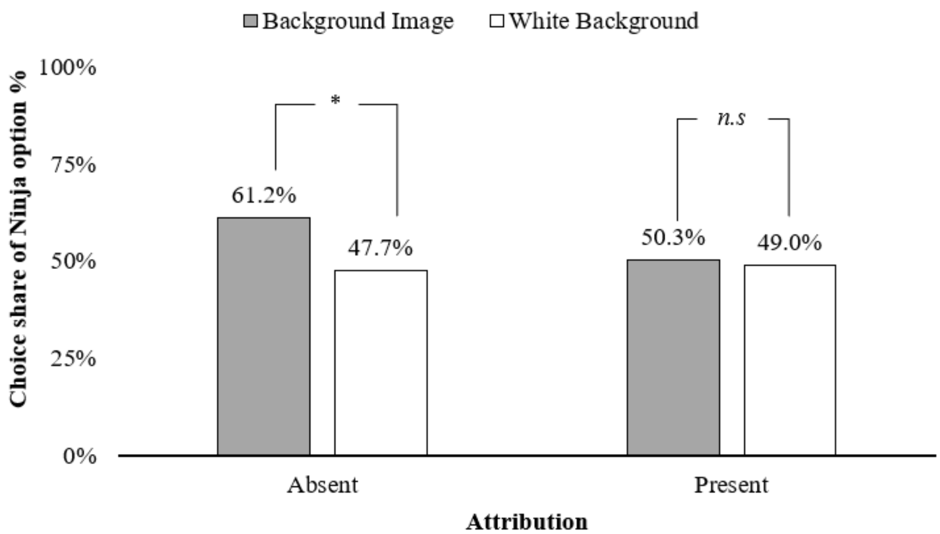
Choice Share. Following our preregistration, the result of a logistic regression showed that there is no significant brand order effect ($p = .17$), effect of different backgrounds ($p = .57$), or their interactions with source attribution (p = .34) and disfluency (p = .26) on choice share. Thus, we collapsed the data across the two counterbalancing factors.

We expected that in the absence of an instruction that mentions the source of disfluency, participants would misinterpret the disfluency caused by the background image as reflecting choice difficulty, and that this would result in an increased incidence of heuristic reasoning. The results supported the predictions. Without a warning about the difficulty of reading the information, participants who saw product information on a background image chose the *Ninja* air fryer 61.2% of the time, whereas 47.7% of those who saw the information without a background image did so ($\chi^2(1) = 5.46$, $p = .019$, $OR = 1.74$, 95% CI [1.09, 2.75], See Figure 16). In contrast, when the instructions mentioned that the background image was the source of disfluency, there was no difference in the rate of choosing the Ninja air fryer among those receiving the information without a background image ($P_{white\_background}$ = 49%) and those receiving the background image ($P_{background\_image}$ = 50.3%, $\chi^2(1) = .053$, $p = .82$).

When the product information was presented with a background image, participants chose the Ninja air fryer more often when they were not aware of the source of disfluency ($P_{attribution\text{-}absent}$ = 61.2% vs. $P_{attribution\text{-}present}$ = 50.3%, $\chi^2(1) = 3.59$, $p = .058$), which supported the notion that the observed effect is metacognitive in nature.

FIGURE 16

DISFLUENCY LEADS TO HEURISTIC INFORMATION PROCESSING (LEFT BARS) EXCEPT WHEN THE BACKGROUND IMAGE IS IDENTIFIED AS THE SOURCE OF DISFLUENCY (RIGHT BARS)



NOTE. — *p < .05

Memory Checks. There were no significant differences in the performance of participants when asked to recall information about the voice-activated cooking function in relation to attribution conditions (p = .94), background image conditions (p = .17), or their interaction (p = .99). Similarly, no significant effect of attribution conditions (p = .48), background image conditions (p = .38), or their interaction (p = .73) were observed on participants' ability to recall cooking functions of air fryers. While differing from the results of Study1, this pattern is also consistent with the idea that the results are not due to participants' failure to read in the difficult-to-read condition, as both groups showed similar recall.

Discussion

Studies 1 and 2 showed that people choose using heuristics when it is hard to process information. Study 3 further provided evidence that the effect may be metacognitive in nature. That is, people tend to misattribute their feelings of disfluency to the difficulty of choosing, and thus they rely on heuristics to cope with the difficulty. When they do not attribute the feelings of difficulty to the choice task, they no longer use a simplifying strategy. This suggests that the effect is not cognitive (e.g. participants cannot read the information), but is rather metacognitive, requiring an inference about the source of difficulty for the effect to come about.

STUDY 4: SYSTEMATIC-HEURISTIC CUE UTILIZATION AS A MEDIATOR

Study 4 was designed to extend the findings of previous studies in two ways. First, we tested for a mediating process. We expected that reading a difficult-to-read font would lead to a larger difference in heuristic vs. systematic cue utilization compared to an easy-to-read font. We predicted that this difference in use of cues would mediate the effect on the choice described above. Second, Study 4 tested a recommendation as a heuristic cue, rather than a brand name as previously.

Similar to Studies 1 and 2, we manipulated disfluency using fonts. We hypothesized that consumers would choose the recommended option more often when the font was harder to read. To measure the mediator, we asked participants to choose amongst a list of attributes as reasons for their choice. With some of the cues being heuristic and some being systematic, we could measure their relative use during the task.

Method

In this pre-registered study, we recruited 600 American and Canadian participants (M_{age} = 41.85, *SD* = 12.18; 44.8% Female) from Mturk in exchange for a small payment. All participants who completed the experiment were included in the analysis. Like Studies 1 and 2, participants were randomly assigned to one of two between-participant conditions: hard-to-read and easy-to-read fonts.

Participants were asked to imagine that they needed to buy a new toaster. They were presented with descriptions of two toasters from which they were asked to choose one. One of the toasters was superior on systematic cues (i.e. it has higher power, more toast-shade settings), while the other toaster was superior on a heuristic cue, a recommendation from Amazon (Amazon's Choice). Both toasters have the same price and the number of toaster slots. The only difference between conditions was the font. Participants saw information about both alternatives simultaneously with the order of the toasters counterbalanced. To counterbalance the order, we randomly assigned participants to see either the recommended toaster as "Toaster A" or the non-recommended as "Toaster A" in the top position (first). Both options were shown on the same screen simultaneously.

Similar to our previous studies, the primary dependent variable was the choice share between the two options. Participants answered, "If you were going to buy a toaster, which one would you like to choose?" (Toaster A or Toaster B). To check of the font disfluency manipulation, participants were asked to rate "How difficult was it to read the information?" (1= Not at all difficult, 7=Very difficult).

We then asked participants to select reasons why they chose the toaster from a list of reasons we provided. We expected that people would choose more heuristic cues as reasons to buy an option if it was hard to process the information. To test this prediction, following Chaiken and Eagly (1983), we categorized reasons such as those recommended by Amazon (i.e. a credible source) and customer review ratings (i.e. other people's opinion) as heuristic cues (Chaiken and Maheswaran 1994), while the other two reasons, the number of toast shade settings and toaster power, were categorized as systematic cues. We asked, “Please select the reason(s) why you chose one of the toasters. (You can select more than one) (Recommended by Amazon, Number of toast shade settings, Toaster Power, and Consumers’ review ratings).” To form a single relative reasoning score as a potential mediator of this effect for each participant, we subtracted the number of heuristic cues from the number of systematic cues. Thus, higher scores would indicate that more systematic cues were used to make the choice, and lower scores would indicate that more heuristic cues were used to make the choice.

Similar to the previous studies, we included three memory check questions to see to what extent participants recall heuristic vs. systematic cues while processing the information by asking, “Which toaster was recommended in this survey?” (Toaster A or Toaster B), “In reviews, how many stars did Toaster A/B receive?” (4.2 vs. 4.3) and “How much electricity (Power) does Toaster A/B use?” (900w vs. 1200w).

Results

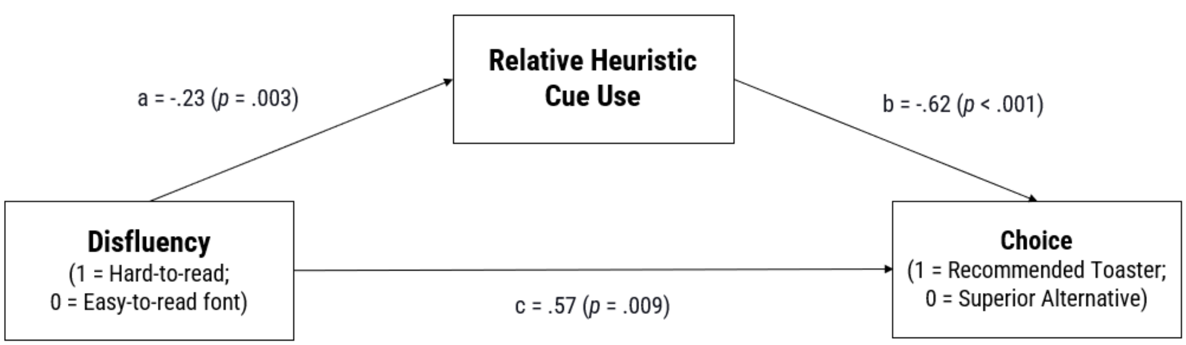
Manipulation check. Participants rated the information presented in the hard-to-read *Amarix SC, cursive* font as more difficult to read than the easy-to-read font ($M_{hard\text{-}to\text{-}read\text{-}font}$ = 4.67, SD = 1.83 vs. $M_{easy\text{-}to\text{-}read\text{-}font}$ = 1.72, SD = 1.08, $t(598)$ = 24.04 $p < .001$, d = 1.96).

Choice Share. There was no effect of counterbalance order (p = .54), so we collapsed the data, as preregistered. The result of a logistic regression with font type (1= hard-to-read font vs. 0 = easy-to-read font) as the independent variable and the recommended toaster share (Recommended toaster = 1; Superior toaster = 0) as the dependent variable revealed the main effect of disfluency (Wald $\chi^2(1) = 10.32$, $p = .001$, $OR = 1.98$, 95% CI [1.30, 2.99]), replicating our findings in the previous studies. Whereas 24.9% of participants who read the hard-to-read font chose the recommended option, 14.4% of those who received the easy-to-read font did so, replicating the previous studies except with a recommendation as a heuristic cue.

Mediation (Reasons to buy). As depicted in Figure 17, those who experienced disfluency showed lower differences between systematic and heuristic reasoning ($M_{hard\text{-}to\text{-}read\text{-}font} = .14$, $SD = .98$) than did those who experienced fluency ($M_{easy\text{-}to\text{-}read\text{-}font} = .38$, $SD = .96$, $t(598) = 2.95$, $p = .003$ $d = 0.24$). These differences in reasoning mediated the choice share of the recommended option. That is, the indirect effect of disfluency on the choice share through reasons to buy was significant using a bootstrapping technique (Preacher and Hayes 2004) with 10,000 iterations ($b_{indirect\_effect} = 0.14$, $SE = 0.06$, 95% CI [0.05, 0.27]).

FIGURE 17

HEURISTIC CUE USE MEDIATES THE EFFECT OF DISFLUENCY ON THE CHOICE SHARE



Memory Checks. At the end of the study, we included three questions to examine the extent to which participants recalled the product information. Memory for the heuristic cues (recommendation and review rating) did not differ across the font conditions respectively ($M_{hard\text{-}to\text{-}read\ font} = 81.6\%$ vs. $M_{easy\text{-}to\text{-}read\ font} = 81.6\%$, $\chi^2(1) = .13$, $p = .72$; $M_{hard\text{-}to\text{-}read\ font} = 90.6\%$ vs. M

easy-to-read font = 87.4%, Wald $\chi^2(1) = 1.62$, $p = .203$). On the other hand, when recalling the systematic cue (the power of the inferior toaster), participants had better recall when it is easier to read information (Wald $\chi^2(1) = 7.03$, $p = .008$, $OR = 2.25$, 95% CI [1.24, 4.11]). As in Study 1, these findings support our notion that participants who process information more fluently will focus more attention on systematic cues, and thus be more likely to remember them, though not pre-registered.

As in Study 1, we explored whether the main effect persisted among only participants who recalled the details correctly. Among the 69.1% of participants who responded correctly to all memory-check questions ($N = 415$), the pattern of results did not change. That is, participants chose the recommended option at a higher rate ($P_{\textit{hard-to-read-font}}$ = 23.4% vs. $P_{\textit{easy-to-read-font}}$ =15.2%) when it was hard to read the information (Wald $\chi^2(1) = 4.40$, $p = .036$, $OR = 1.70$, 95% CI [1.036, 2.79]). This is further evidence that when examining only participants who actually read the information, the effect persists.

Reading Time. We did not pre-register a hypothesis for reading time. Participants in the hard-to-read font condition spent on average around eight seconds more time reading the information, although the time difference was only marginally significant ($M_{\textit{hard-to-read-font}}$ = 39.44 seconds, $SD = 68.88$ vs. $M_{\textit{easy-to-read-font}}$ = 31.63 seconds, $SD = 23.32$; $F(1, 598) = 3.45$, $p = .064$). This is further suggestive evidence that participants in hard-to-read font condition read the information.

STUDY 5: SUBJECTIVE THINKING STYLE MEDIATION EFFECT

In Study 5, we demonstrated the mediating process by measuring self-reported thinking style, adapted from Pocheptsova et al. (2010). Instead of selecting the reasons for choosing what they did (as in Study 4), we asked participants to report the extent to which their processing style was systematic or intuitive. We predicted that participants in the disfluent condition would choose the better brand more often, and that these effects would be mediated by increased reliance on an intuitive processing style.

Method

In this pre-registered study, we obtained data for 600 American participants using desktop computers (M_{age} = 41.23, *SD* = 13.78; 42% Female) from Prolific in exchange for a small payment. All participants who completed the experiment were included in the analysis. The procedure of this study largely followed the pattern of Study 3. As in Study 3, to manipulate disfluency we presented product information with or without a background image to mimic the presentation of real products on Amazon.com. Participants were asked to imagine that they needed to buy a new blender. We used brand name as a heuristic cue. For realism, in this study, we use actual blender brands, Ninja as an established brand and Shardor as a less established brand. Other information, such as the position of information on the backgrounds, blender's price, color, and material kept constant between the two options. All participants saw both options presented sequentially on different screens. The order was counterbalanced.

The dependent measure was the binary response to the question, "If you were going to buy one of them, which one would you choose?" (Ninja or Shardor). Next, to assess the mediator—thinking style—participants reported how they evaluated the options using a 7-point bipolar scale (1 = Very systematically; 7 = Very intuitively; adapted from Pocheptsova et al. 2010). In addition, participants answered a question about brand familiarity and a manipulation check of disfluency asked in the previous studies.

Results

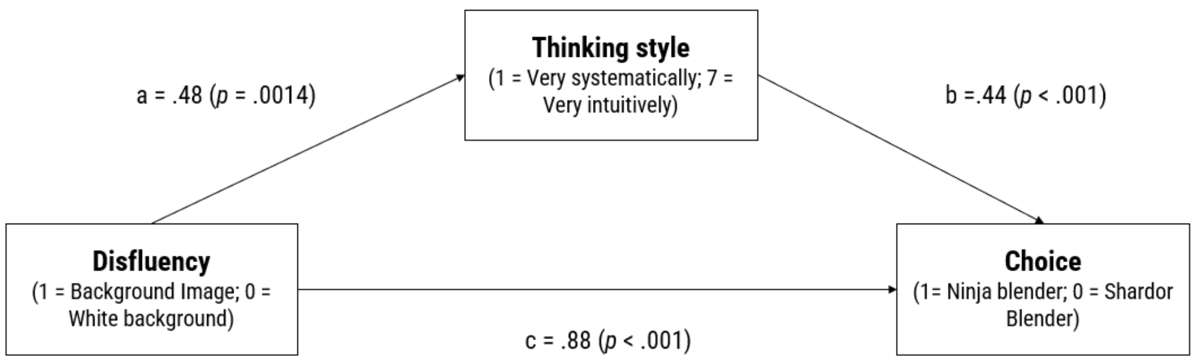
Manipulation checks. Participants rated the information presented with a background image as more difficult to read than without a background image ($M_{background\_image} = 4.54$, $SD = 1.95$ vs. $M_{white\_background} = 2.72$, $SD = 1.71$, $t(598) = 12.11$, $p < .001$, $d = .99$). In addition, they were more familiar with *Ninja* (vs. *Shardor*) brand name ($M_{Ninja} = 5.08$, $SD = 1.94$ vs. $M_{Shardor} = 1.54$, $SD = 1.11$, $t(599) = 40.84$, $p < .001$, $d = 1.67$).

Choice share. There was no significant brand order effect ($p = .195$) or interaction of order with the background image ($p = .411$) on choice share. Following our preregistration, combining data on brand order, the result of a logistic regression with visual disfluency (background image, white background) as the independent variables and Ninja choice share (Ninja = 1; Shardor = 0) as the dependent variable revealed the predicted effect of disfluency (Wald $\chi^2(1) = 32.30$, $p < .001$, $OR = 2.61$, 95% CI [1.87, 3.63]). Amongst participants who saw products' information on a white background, 42.2 % chose Ninja blender, whereas 65.6% of those who saw the information with a background image did so, replicating the previous studies.

Mediation (Thinking style). We proposed that disfluency would lead to intuitive processing. As depicted in Figure 18, those who experienced disfluency showed higher intuitive processing ($M_{background\_image}$ = 4.16, *SD* = 1.81) than did those who experienced fluency ($M_{white\_background}$ = 3.68, *SD* = 1.85, t (598) = 3.21, p = .001, d = 0.26). These differences in processing information mediated the choice. That is, the indirect effect of disfluency on the choice share through *thinking style* was significant using a bootstrapping technique (Preacher and Hayes 2004) with 10,000 iterations ($b_{indirect\_effect}$ = - 0.21, *SE* = 0.07, 99% CI [0.04, 0.42]).

FIGURE 18

INTUITIVE PROCESSING OF INFORMATION INCREASES THE CHOICE SHARE OF A BETTER BRAND NAME OPTION AND MEDIATES THE EFFECT OF DISFLUENCY ON CHOICE SHARE



STUDY 6: PROCESSING STYLE AS A MODERATOR

In Studies 4 and 5, we provided evidence that consumers process information less systematically when they experience disfluency, and that this in turn leads them to base their decisions on heuristic product information. In this study, instead of measuring processing style, we manipulated it through participants' effort invested in the decision making. Previously, it has been shown that decision makers process more heuristically when there are time constraints on them (e.g. Allard et al. 2019; Chaiken et al. 1989; Kaplan et al. 1993; Dhar and Nowlis 1999). In contrast, slow decision makers are more likely to process information systematically (Suri and Monroe 2003; Allard et al. 2019). As such, in this study, we manipulated processing style by creating a belief that quick decisions would lead to the best outcomes for some participants (heuristic processing), while we encouraged others to make slow, effortful decisions (systematic processing). When crossed with a manipulation of disfluency, we expected that this manipulation would have a moderating effect on the relationship between disfluency and choice share of a well-known option.

Specifically, since both fast decision-making and disfluency can lead consumers to use a heuristic processing style on their own, when the two are paired, the likelihood of using heuristics should be even higher. In contrast, instructing participants to make slower decisions could attenuate the tendency to process information heuristically that disfluency would otherwise cause. Thus, the effect of disfluency on choice should be smaller for slow, high-effort decisions than for fast, low-effort decisions.

Method

Six hundred and two American and Canadian online panel participants (M_{age} = 41.55, *SD* = 12.73; 53.5% Female) recruited through Amazon Mturk took part in this pre-registered experiment in exchange for payment. This study used a 2 (Disfluency: low [no background image] vs. high [a background image]) by 2 (Processing style: more heuristic processing [fast and effortless] vs. less heuristic processing [slow and effortful]) between-participants design. The dependent variable of interest was the proportion of participants choosing the printer with a better brand name. All participants who completed the experiment were included in the analysis. The survey could only be completed on a desktop computer.

Before participants made their choice, they were randomly assigned to one of the four conditions. We manipulated disfluency, similar to Studies 3 and 5, by presenting the two printers to half of the participants with the textual information presented over a background image,[3] while the others saw the same information about the same printers without a photo background.

We manipulated processing style by manipulating participants' effort invested in the decision making (Inbar et al. 2011; Allard et al. 2019). As depicted in the web appendix, half of the participants saw information reporting that "fast-is-accurate" when it comes to making decisions. They read that psychological research showed that "people often make better choices when they decide quickly and effortlessly. A quick response often leads to making the best choice possible." The text concluded with instructions to "quickly consider the options and immediately make your selection." This manipulation creates a sense of time-pressure in the "fast-is-accurate" condition without placing any actual constraint on the amount of time participants could process the stimuli.

The other half of participants, in the "slow-is-accurate" belief condition, saw information that "recent psychological research has shown that making better choices often takes more time and effort. Taking your time often leads to making the best choice possible." The text concluded with the instruction to "fully consider the available options before making your selection."

Participants were instructed to imagine themselves having a printer which stopped working suddenly. They were told that since their printer model was non-repairable, they needed to buy a new one, and they found two laser printers with the same price on the market; both met

---

[3] Since the order effect of using different background images was not significant in Study 2 (p = .28), in this study we did not counterbalance the background images.

their needs. Similar to previous studies, one of the printers was better on systematic cues (i.e. larger paper tray size and printing output, higher printing speed, voice printing feature, and more connectivity options). In contrast, the other printer had a better brand name as a heuristic cue. For realism, in this study, we use actual printer brands, EPSON as an established brand, and PANTUM as a less established brand. Other information, such as price and the printer color, was held constant between the two options. All participants saw both options presented sequentially on different screens. The presentation order was counterbalanced.

Checking for brand name and disfluency manipulations, we asked the same questions used in the previous studies. As a check of our processing speed manipulation, we used both subjective and objective measurements, following the existing literature. Regarding the subjective check (Dhar and Nowlis 1999), participants are requested to answer two questions: 1) "How much time pressure did you feel when making your choice?" (1 = No pressure; 7 = Very much pressure); 2) "How fast did you need to make your decisions?" (1 = Not at all fast; 7 = Very fast). To check objective measurement, we measured the time spent on the pages to read the information (Allard et al. 2019).

Results

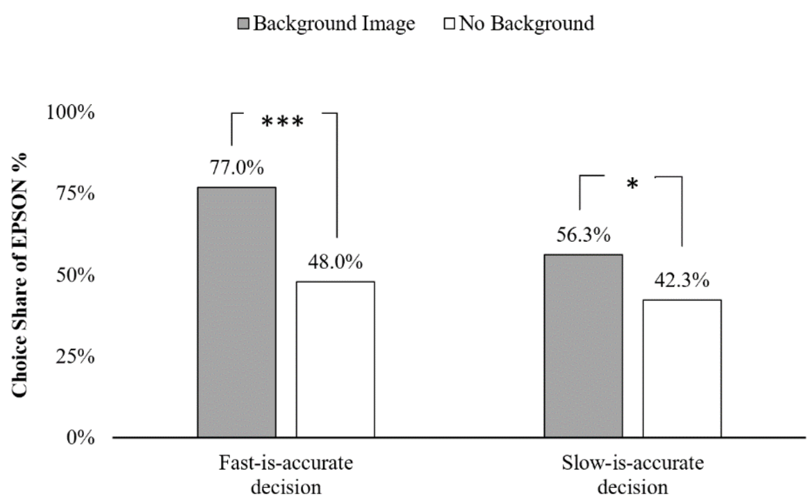
Manipulation checks. Participants rated the information presented with a background image as more difficult to read than without a background image ($M_{background\_image} = 4.03$, $SD = 1.91$ vs. $M_{white\_background} = 2.81$, $SD = 1.75$, $t(600) = 8.15$ $p < .001$, $d = 0.66$). The manipulation of brand name familiarity was also successful. A paired-sample t-test on brand name familiarity revealed that more familiarity with EPSON (vs. PANTUM) brand name ($M_{Epson} = 5.33$, $SD =$

1.92 vs. M_{Pantum} 1.45, SD = .98, t (601) = 43.39, p < .001, d = 1.77). Furthermore, participants overall took longer to reach a decision in the slow-is-accurate condition ($M_{\textit{slow-is-accurate}}$ = 45.69, SD = 38.03) compared with the fast-is-accurate condition ($M_{\textit{fast-is-accurate}}$ = 15.78, SD = 13.64; t (600) = 12.86, p < .001, d = 1.05). Moreover, participants felt more time pressure (α = 0.89) in the fast-is-accurate condition ($M_{\textit{slow-is-accurate}}$ = 5.33, SD = 1.32; $M_{\textit{fast-is-accurate}}$ = 2.17, SD = 1.16; t (600) = 31.19, p < .001, d = 2.54). These results support the validity of our processing speed manipulation.

Choice Share. We predicted that processing style would moderate the effect of disfluency on heuristic reasoning (i.e. choosing the EPSON printer more often). We found that there was no brand order effect (p = .29) or interaction of order with the processing style and disfluency (ps > .345) on choice share. Combining data on brand order in accordance with our preregistered plan for analysis, we conducted a logistic regression with EPSON printer choice share as the dependent variable, and disfluency (with a background image vs. no background image), processing style (slow-is-accurate vs. fast-is-accurate), and their interaction as predictors. Replicating our previous findings, we found a main effect of disfluency (Wald $\chi^2(1)$ = 28.25, p < .001, OR = 2.47, 95% CI = [1.77, 3.45]), such that 66.7% of participants chose the EPSON printer when its information presented with a background image (vs. 45.15% when it was presented without a background image). In addition, we found a main effect of processing style (Wald $\chi^2(1)$ = 11.16, p < .001, OR = 0.57, 95% CI [0.41, 0.79]) on the choice share of the EPSON printer, as expected based on past literature. Most importantly, we found a two-way interaction between processing style and disfluency in the choice share of the better brand (Wald $\chi^2(1)$ = 4.42, p = .036, OR = 2.06, 95% CI [1.05, 4.04], See Figure 19). Consistent with our theorizing, while 77% of participants chose the EPSON printer in the fast-is-accurate condition when it was hard to process the information, only 48% of them did so when it was easy to process the information (Wald $\chi^2(1) = 25.94$, $p < .001$, $OR = 3.62$, 95% CI [2.21, 5.94]). However, this effect was attenuated when participants processed the information slowly and effortfully (Wald $\chi^2(1) = 5.85$, $p = .016$, $OR = 1.76$, 95% CI [1.11, 2.78]).

FIGURE 19

PROCESSING STYLE INCREASED THE MAGNITUDE OF THE DIFFERENCE BETWEEN FLUENT AND DISFLUENT PRESENTATION.



NOTE. — ***$p < .001$; * $p < .05$

Memory Checks. To explore differences in product information recall by condition, we asked participants two memory-check questions: “Which printer offered voice-activated printing?” (EPSON or PANTUM) and “What was the tray capacity of Printer A/B?” (150 or 250). Memory for the former cue (voice assistant) differed across both the disfluency conditions ($P_{background\_image} = 74.6\%$ vs. $P_{white\text{-}background} = 89.3\%$, $\chi^2(1) = 10.90$, $p = .001$) and the processing style conditions ($P_{fast\text{-}is\text{-}accurate} = 77.2\%$ vs. $P_{slow\text{-}is\text{-}accurate} = 87.4\%$, $\chi^2(1) = 4.73$, $p = .03$), while the interaction of them was not significant (*p* = .76). On the other hand, when recalling the capacity of the printers, there were no significant main effect of disfluency (*p* = .96) and its interaction with processing style (*p* = .50), while there was a marginal significant main effect of processing style ($\chi^2(1) = 3.72$, $p = .054$). We analyzed responses from the 71.9% of participants (N = 433) who answered both questions correctly. Results aligned with our theory, showing main effects of disfluency (Wald $\chi^2(1) = 16.11$, $p < .001$, *OR* = 3.37, 95% CI = [1.86, 6.10]) and processing style (Wald $\chi^2(1) = 8.96$, $p = .003$, *OR* = 2.99, 95% CI [1.37, 4.46]). Crucially, we also observed an interaction effect of both factors on the choice share of the EPSON printer (Wald $\chi^2(1) = 4.24$, $p = .039$, *OR* = 0.44, 95% CI [0.20, 0.96]), controlling for product information recall.

Discussion

In this study, we aimed to manipulate processing style through a manipulation of processing speed. When crossed with a manipulation of disfluency (presenting text over a background-image or not), we expected an interaction. If disfluency leads to heuristic processing, as we have hypothesized, its effect should be strongest under conditions that further facilitate heuristic processing (such as when consumers make quick, effortless decisions). In contrast, the effect of disfluency may be attenuated when there is a countervailing factor, in this case an instruction to process the information slowly and carefully. The results were consistent with our predictions and further supported our theoretical account as to the mechanism underlying the effect of disfluency on consumer choices.

STUDY 7: A FOUR-CHOICE SETTING TO ADDRESS MANAGERIAL RELEVENCE

In the previous studies, we obtained the results using both actual and hypothetical choices among two options. Here, we further explored the implications of this research by testing the main effect in a choice setting that more closely mimicked real-world purchase decisions by presenting consumers with more than two options. Moreover, we used image resolution to manipulate disfluency, as low resolution for product images is common in the marketplace (see web appendix).

Method

In this pre-registered study, we recruited 1,202 American and Canadian Mturk users (M_{age} = 42.68, *SD* = 13.25; 55.7% Female) in exchange for payment. Participants were randomly assigned to two between-participants conditions which manipulated fluency. To closely mimic the real-world environment, we manipulated the image quality by lowering the images resolution in the disfluent condition. While participants in the fluent condition were faced with four real-market blenders (i.e. Hamilton beach, Kogan, Amzchef, and ICUCINA) presented with clear images, the other half in disfluent condition found the same blenders shown with poor image resolution, holding constant the image size (see web appendix).

The task and instructions were similar to those in previous studies, in which subjects were asked to make a choice between alternatives. The main difference was that people were asked to make a choice between four alternatives instead of two. Following the previous studies participants were informed about four blenders. Three of the blenders were better on systematic cues (i.e. more capacity, more power wattage, and more numbers of speed). In contrast, the other blender had only a better brand name, which could be used as a heuristic cue. Other information, such as price was kept constant between the blenders to control any price-based inferences. All presented information was real-market information collected through the blender brands' websites. All participants saw all four options presented sequentially on different screens. We randomly rotated the order of product displays across participants to avoid primacy or recency effects.

The dependent variable was the choice between the four blenders. To check the disfluency and brand familiarity manipulations, we asked the same questions as previously.

Results

Manipulation checks. Participants rated the information presented in low-resolution as more difficult to read than the high-resolution images ($M_{high\text{-}image\text{-}quality} = 2.68$, $SD = 1.70$ vs. $M_{low\text{-}image\text{-}quality} = 4.46$, $SD = 1.84$, $t(1200) = 17.40$, $p < .001$, $d = 1.00$). The manipulation of brand name familiarity was also successful. A repeated measure ANOVA on brand name familiarity revealed that more positive ratings were obtained when the product featured *Hamilton beach* ($M_{Hamilton\ beach} = 5.41$) (vs. *others*) brand name ($M_{Kogan} = 1.31$, $M_{Amzchef} = 1.47$, $M_{ICUCINA} = 1.26$, $F(3, 1199) = 1230.42$, $p < .001$, $\eta_p^2 = .76$).

Choice Share. Following our preregistration, the result of logistic regression with product image quality (low-resolution vs. high-resolution) as the independent variable and the Hamilton beach blender choice share (*Hamilton beach = 1*; *other blenders = 0*) as the dependent variable revealed a significant effect of disfluency (Wald $\chi^2(1) = 9.14$, $p = .002$, $OR = 1.44$, 95% CI [1.14, 1.82]), replicating our findings in the previous studies. Whereas 41.3% of participants who saw the products' information in low-quality images chose the Hamilton beach option, 32.8% of those who saw the high-quality image information did so. No differences were observed between the choice share of the other blenders (see Figure 20).

FIGURE 20

LOW IMAGE-QUALITY INCREASED THE LIKELIHOOD OF CHOOSING A BLENDER WITH A BETTER BRAND NAME, HAMILTON BEACH.

NOTE. — $^{**}p < .01$

Discussion

This study focused on the issue of managerial relevance, specifically exploring decision-making scenarios involving more than two options, which closely resemble real-world conditions. The findings of this study further support the significance of heuristic reasoning when it is hard to process information, replicating our previous findings.

GENERAL DISCUSSION

While past researchers have found that disfluency activates an analytical or systematic thinking style (e.g. Alter et al. 2007; Alter 2013; Song and Schwarz 2008; Walter et al. 2020; Simmons and Nelson 2006), other scholars have recently questioned their findings (e.g. Meyer et al. 2015; Thompson et al. 2013). We go one step further to show that when disfluency signals that choosing is difficult, people simplify the decision-making process by using heuristics rather than relying on systematic features which need more computational processing. This means consumers rely more heavily on heuristic attributes, such as a brand name (e.g. Lee and Shavitt 2009; Maheswaran et al. 1992; Gunasti and Ross Jr 2010), country of origin (Maheswaran 1994; Chang 2004; Hong and Wyer Jr 1989), or a recommendation (e.g. Senecal and Nantel 2004; Mariadassou et al. 2023; Munz and Morwitz 2019; Dellaert and Häubl 2012). This idea is supported by evidence from seven pre-registered studies.

Study 1 provided initial evidence for our hypothesis, showing that consumers were more likely to choose an option with a better brand name and avoid choosing an option from a less-desirable country of origin when the information was difficult to read. The subsequent studies further supported the main hypothesis, showing that disfluency led to increased reliance on heuristic cues in an incentive-compatible choice setting (Study 2) and when presented with textual product information over a distracting background image (Study 3). Additionally, Study 3 revealed that participants were less likely to rely on heuristic cues when they were aware of the task's difficulty, suggesting that without this awareness, consumers mistake the feeling of disfluency with the difficulty of choosing, leading them to simplify the task by processing the information heuristically. Studies 4 and 5 explored the mediation of the effect through heuristic versus systematic reasoning. Study 6 instead manipulated consumers' processing style. The findings showed that individuals who believed that faster decision-making led to better outcomes had a higher tendency to rely on heuristic processes in the face of disfluency. Lastly, Study 7 aimed to align the research with real-world marketplace practices by increasing the number of options and varying image resolution as another form of disfluency. The results replicated the previous findings, indicating that individuals tend to rely more on heuristic cues when faced with information processing difficulties.

Overall, these studies provide evidence for the hypothesis that disfluency increases reliance on heuristic cues, simplifying the decision-making process.

Theoretical Implications

First, our studies make a significant contribution to the existing literature on processing disfluency (e.g. Lee and Labroo 2004; Labroo et al. 2009; Dhar 1997; Walter et al. 2020; Dhar and Simonson 2003; Novemsky et al. 2007) by examining which option will be chosen. Within the consumer choice context, our research builds upon the work of Novemsky and his colleagues (2007) in two ways. First, while their focus was on how the perception of difficulty influences the decision to make a purchase (or not), we investigate which option consumers are more likely to choose, utilizing a forced choice setting. Second, we extend their findings by providing insights into why individuals tend to defer their decisions when faced with difficulty, aligning with our perspective that compromise and deferral can be viewed as heuristics used to simplify complex decisions.

Moreover, we provide direct evidence that disfluency encourages consumers to process information more heuristically, supporting and extending current findings (e.g. Huh et al. 2011; Mukherjee and Lee 2016; Mariadassou et al. 2023; Munz and Morwitz 2019). Previous studies altered various aspects of choice tasks to increase their difficulty (i.e. increasing cognitive load, or increasing the complexity of choice sets with more options and alternatives). In contrast, we examined disfluency by manipulating ancillary aspects of the task (font, background image, resolution) while keeping the choice itself the same. The result of Study 3 provides further support for a metacognitive interpretation. Specifically, the effects relied on misinterpreting the difficulty due to disfluency to be about difficulty choosing.

In contrast to existing findings regarding the relationship between disfluency and deliberate processing (i.e. Alter et al. 2007; Song and Schwarz 2008; Walter et al. 2020), we found that disfluency led consumers to choose options with superior heuristic cues. Why do the present results appear to be *opposite* to the earlier findings?

One explanation may be may due to different norms regarding statistical power between the earlier results and the more recent ones. Indeed, others have already questioned the earlier findings based on evidence from more powerful experimental tests (e.g. Meyer et al. 2015). Yet, even the more recent, better-powered studies did not find an increase in the use of heuristics, instead finding no differences between fluency conditions.

Another explanation for why the present results diverge could be that earlier studies and the replication attempts used mathematical problem-solving (Alter et al. 2007, Meyer et al. 2015) or other similar problem-solving tasks (Meyer et al. 2015; Thompson et al. 2013; Sirota 2021a, b). In contrast, our studies involved a comparative consumer choice context. It may be that a person's response to disfluency differs as a function of the task. We have argued that the effect of disfluency on choice is predicated on a consumer inferring that the choice is difficult as a result of the disfluency, a process necessarily absent when a consumer is not making a choice (as when solving a math problem). This idea is corroborated by the results of Study 3, which showed that without this inference, there was no difference between fluency conditions.

Math problems and some other types of problem-solving tasks also may have a different structure than consumer choice problems. For the former types of problems, there is a “correct” answer that cannot be approximated by using a heuristic. That is, a nearly correct answer to a math problem is an incorrect answer, while a nearly ideal solution to a consumer choice is probably good enough. Thus, it seems plausible that heuristic solving strategies should be more likely when a merely satisfactory solution will suffice.

Second, this research contributes to the literature on dual processing systems (e.g. Kahneman and Frederick 2002; Chaiken 1980, 1987; Chaiken and Maheswaran 1994; Petty and Cacioppo 1986). While the previous findings showed people are more likely to use heuristic processing under low motivation or low ability to process information, the results reveal that in the context of consumer decision-making, disfluency can also lead to using heuristic reasoning rather than evaluating options systematically. We showed direct evidence for this process in Studies 4, and 5. We documented differences in decision-making strategies of consumers when facing disfluency. That is, when it comes to subjective feeling of difficulty in making a choice, people evaluate options more intuitively by relying more on heuristics like brand names or recommendations. In contrast, alternatives are evaluated in more details by relying less on heuristic cues but more one systematic features when products information is easy to read. This is because people tend to simplify their decision-making process to cope with difficulties they feel while making a purchase decision.

We also contribute to understanding the situational factors in which brands can be more or less impactful on consumer decision making (Walter et al. 2020; Lee and Shavitt 2009). Some past research has suggested that perceptions of well-established brands can be hurt when disfluency signals to a customer that he or she may not understand the brand very well (Lee and Shavitt 2009). Disfluency can also lead to consumers choosing an unfamiliar brand more often due to exploratory search inclination (Walter et al. 2020). In contrast, we document an advantage for established brands when the choice environment leads to disfluent information processing. Specifically, it can increase the likelihood of a well-established brand being chosen.

We extended the current body of research on recommendation acceptance in the context of voice interfaces, building upon prior work (e.g. Mariadassou et al. 2023; Munz and Morwitz 2019). Previous studies have indicated a higher likelihood of people adhering to auditory recommendations compared to visual ones. This difference has been attributed to challenges in processing auditory information versus text-based information (Munz and Morwitz 2019), as well as a heightened sense of urgency (Mariadassou et al. 2023). To contribute to these findings, we delved into the intricacies of dual processing models. In Study 4, our research revealed that individuals more frequently opt for recommended choices, often driven by heuristic reasoning as a means to navigate challenging decision tasks.

Alternate Explanation: Selective Reading

It is possible that when information is hard to read, people simply may read less of it. That is, rather than disfluency leading people to make an inference about the difficulty of the *choice*, people may instead try to avoid the difficult *task* of reading the product information and selectively read only the information they deem most important (e.g. a brand name). This alternate explanation is in line with our hypothesis about simplification, but is subtly different in its proposed mechanism. While perfectly plausible and possibly contributing to the observed effects, it is unlikely to fully account for the data.

In Study 3, the effect disappeared when participants were merely informed that reading the information may be harder due to the background image. This pattern of results is consistent with a metacognitive interpretation (Novemsky et al. 2007) but is less easily explainable with this alternate view in mind. That is, the effect only occurred when participants misattributed the negative feeling of disfluency to the choice task itself. Furthermore, differences in memory for product information and reading time did not consistently differ between fluency conditions, suggesting that participants in the disfluent conditions were not simply skipping the information. More to the point, when examining only the subset of participants who correctly recalled all of the product details, the effects of disfluency on choice remained significant, suggesting that the effect occurs even when participants read the information at similar rates between the conditions.

Practical and Managerial Implications

It is noteworthy that just by presenting information in a hard-to-read font, adding a background image to product information, or presenting information in an image with degraded resolution consumers' decisions about the options shifted to favor a better brand name or a recommended product. Our findings align with the idea that brands play a vital role in projecting a positive image, signaling distinct characteristics or overall product quality to consumers, and impacting their purchasing decisions (Gunasti et al. 2010).

Our research also highlights how subtle differences in the presentation of options, particularly in e-commerce settings, can lead to unexpectedly disparate effects between well-established and less-established brands. For example, imagine that a retail web site decides to apply an algorithm to the images on its site that lowers their resolution as a way to save server space. On its face, this change appears neutral; all brands will have their image resolution reduced, and so all brands should be equally affected. However, our research suggests that this action could disproportionately favor well-established brands. Thus, retailers and less-established brands should be mindful of such changes.

Along these lines, the research also highlights how marketing managers for less-established brands should critically focus on ensuring fluency in their marketing materials. That is, we do not advocate for well-established brands to create disfluency in hopes of gaining market share. Instead, we hope to encourage less-established brands to improve the fluency in their marketing materials to avoid any unnecessary disadvantage. This provides them with benefits while additionally improving the user experience for the customer.

Conclusion

When information about choice options is difficult to process, consumers are more likely to choose a product that excels on a heuristic attribute such as a brand name, country of origin, or recommendation. This occurs because they mistake the difficulty of processing the information about the products with difficulty choosing between the options. Thus, they try to simplify the difficult-seeming choice by using heuristics. The results are counter to existing literature on the relationship between disfluency and processing style and have large practical implications.

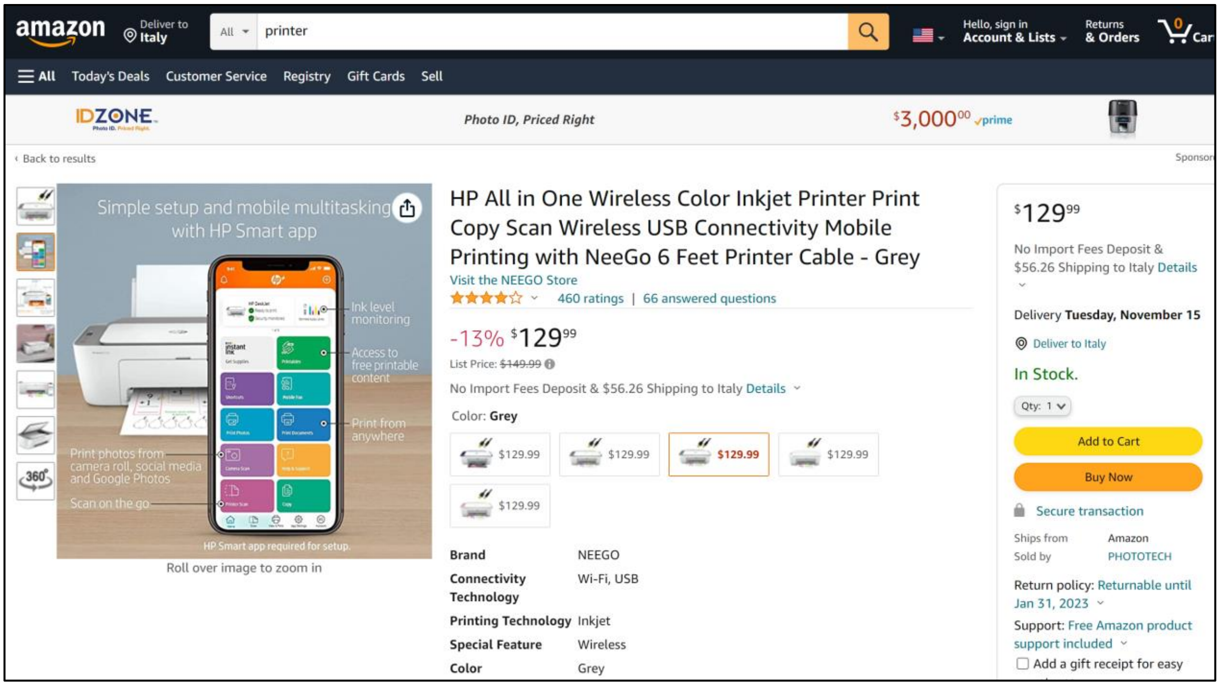
APPENDIX B

Disfluency Increases Reliance on Heuristic Cues in Consumer Choice

REAL-WORLD TYPE OF DISFLUENCY IN CONSUMER CHOICE SETTING

FIGURE 21

TEXT APPEARING ON A BACKGROUND IMAGE USED IN STUDIES 3, 5 AND 6.



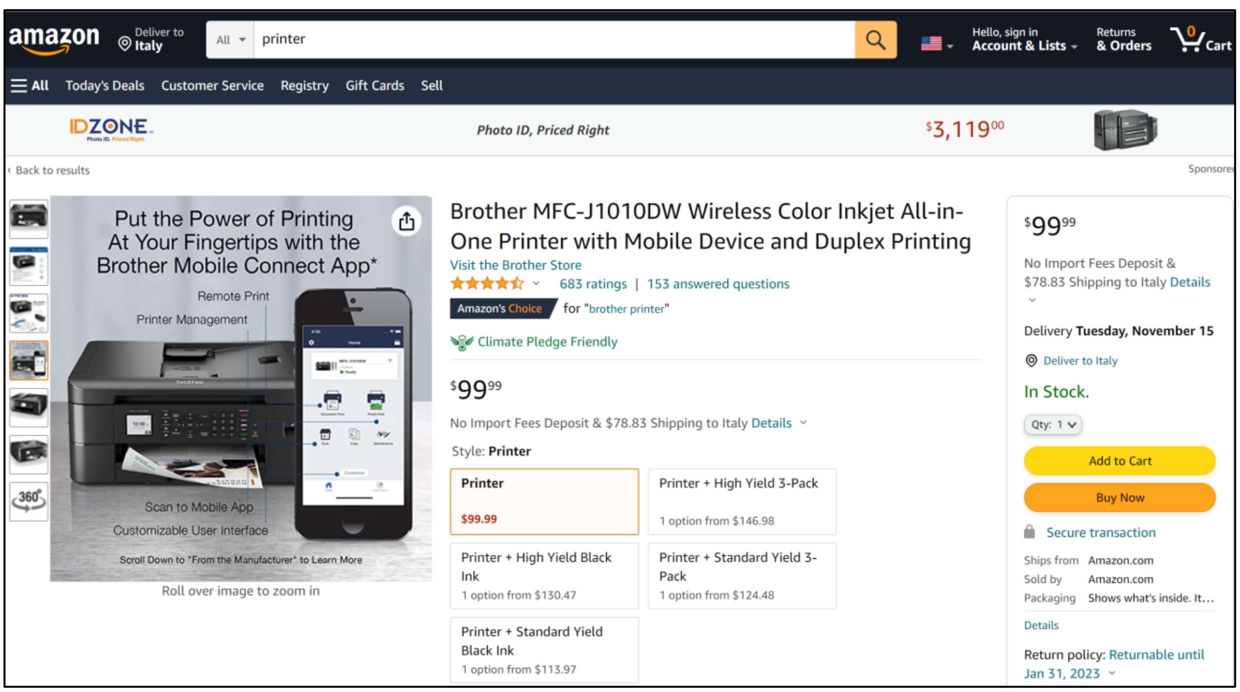


FIGURE 22

AMAZON CHANGES THE IMAGE QUALITY OF PRODUCTS TO MAKE WEBSITE LOADS FASTER USED IN STUDY 7. THE QUALITY OF THE PRITER IN THE LEFT-HAND SIDE IS REDUCED COMPARED TO ANOTHER OPTION IN THE RIGHT-HAND SIDE FROM THE SAME CATEGORY PRESENTED AT HIGHER RESOLUTION

DISFLUENCY MANIPULATION – STUDY 1 (FONT TYPE)

FIGURE 23

STIMULI FOR STUDY 1. BLENDERS A AND B IN AN EASY-TO-READ FONT CONDITION.

Blender A

Brand	iTOChin
Material	Glass
Wattage	900 watts
Item Weight	2.5 pounds

About this Item:

- 2 jars Included for versatile blending: a 40 oz. glass blender jar and 20 oz. travel jar.
- 7 blending functions include mix, puree, chop, milkshake, smoothie, icy drink and crush ice, plus pulse.
- Ice-crushing stainless steel blades

Price: $54.99

Blender B

Brand	Hamilton Beach
Material	Glass
Wattage	700 watts
Item Weight	2.5 pounds

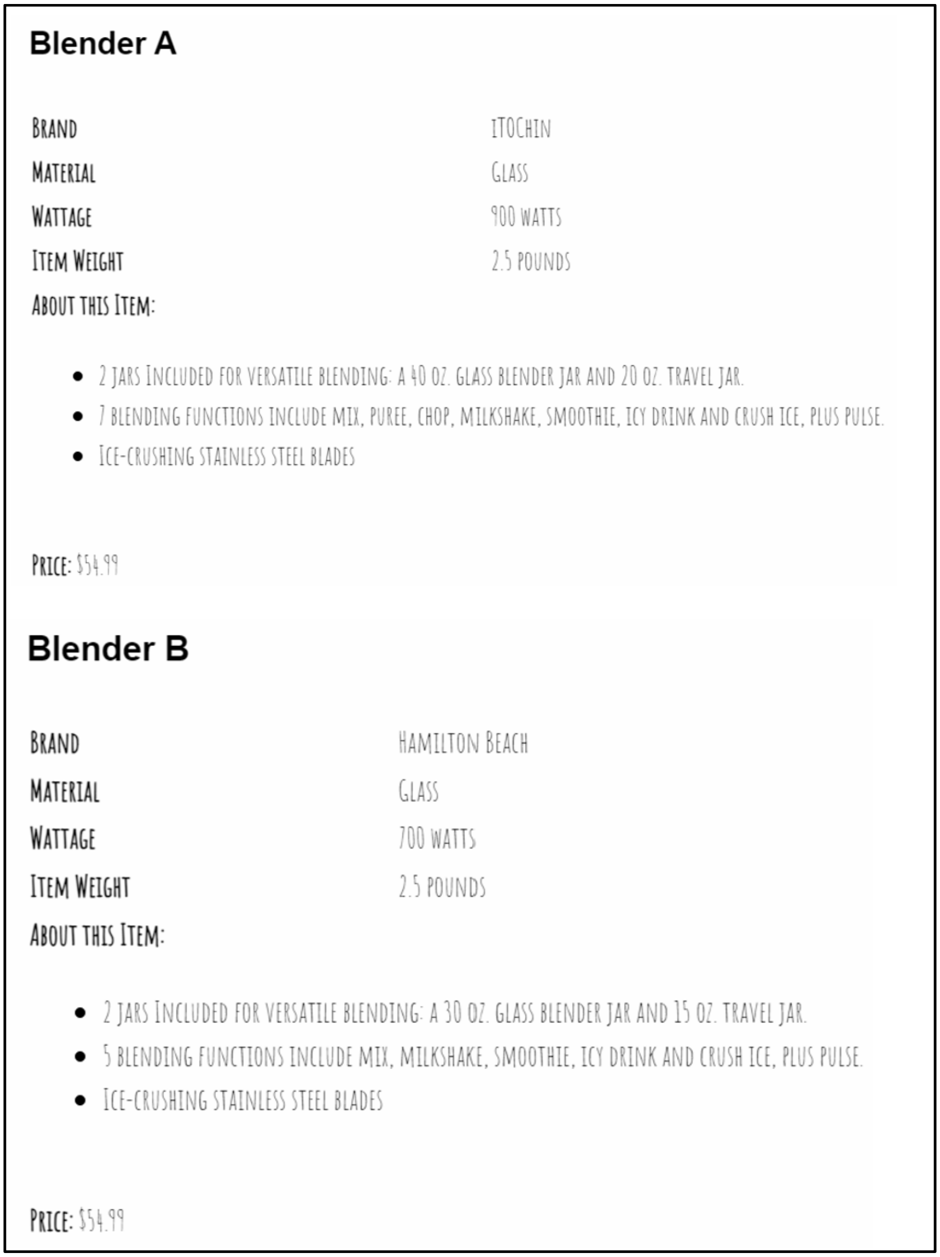
About this Item:

- 2 jars Included for versatile blending: a 30 oz. glass blender jar and 15 oz. travel jar.
- 5 blending functions include mix, milkshake, smoothie, icy drink and crush ice, plus pulse.
- Ice-crushing stainless steel blades

Price: $54.99

FIGURE 24

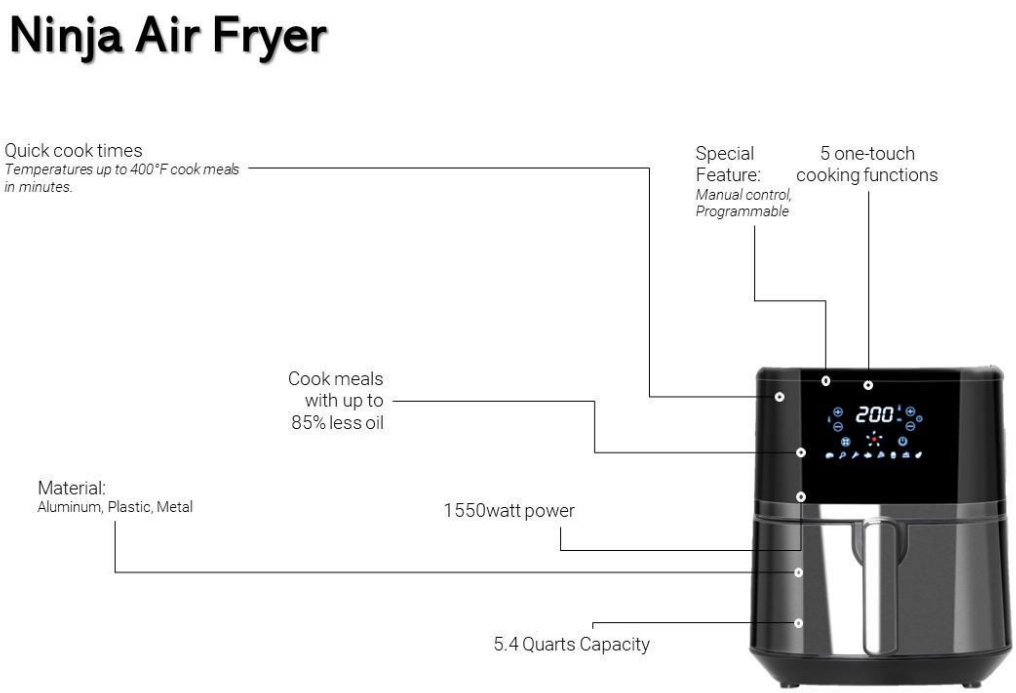
STIMULI FOR STUDY 1. BLENDER A AND B IN A HARD-TO-READ-FONT CONDITION.



DISFLUENCY MANIPULATION – STUDY 3 (BACKGROUND IMAGE)

FIGURE 25

STIMULI FOR STUDY 3. AIR FRYERS A IN A FLUENT CONDITION ON A WHITE BACKGROUND.



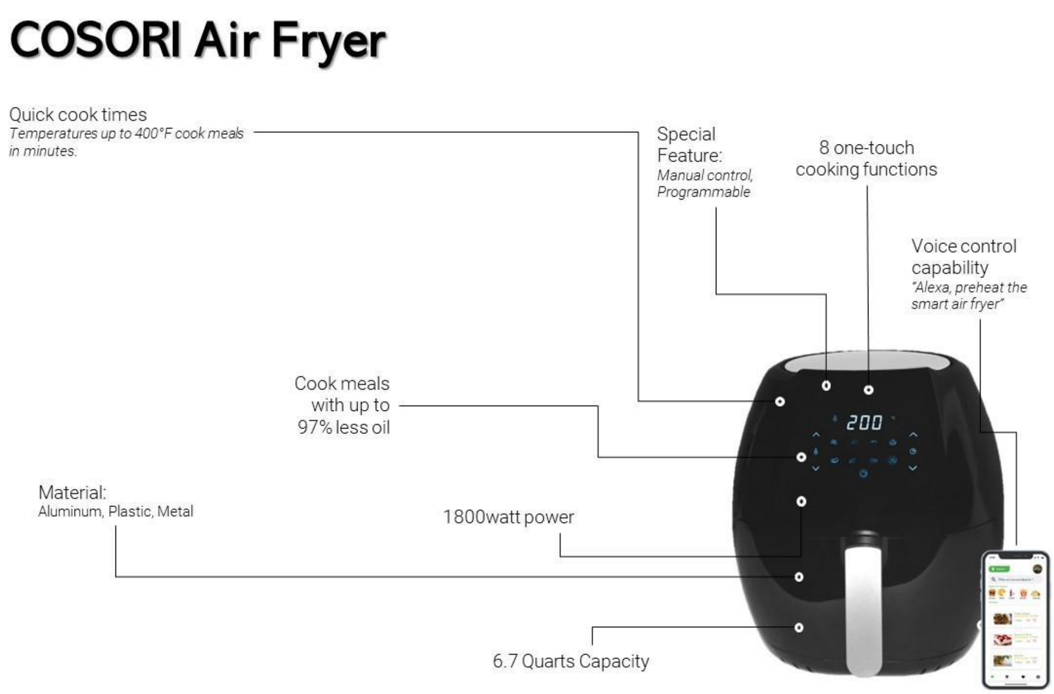
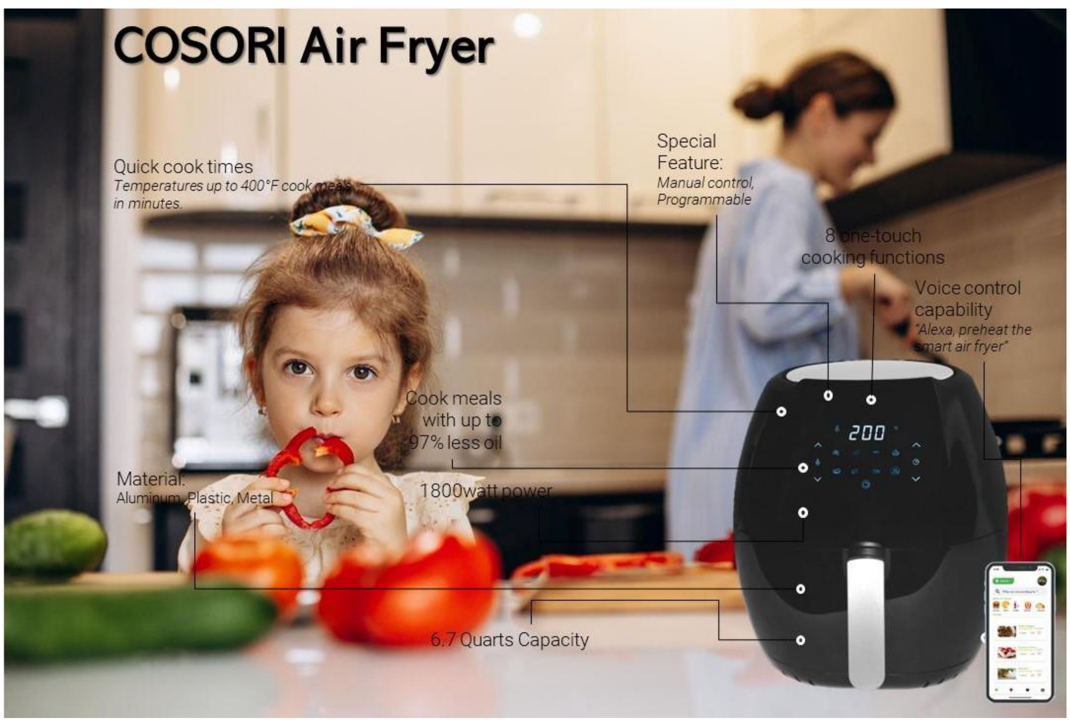


FIGURE 26

STIMULI FOR STUDY 3. AIR FRYERS IN DISFLUENT CONDITION WITH DIFFERENT BACKGROUND IMAGES. WE COUNTERBALANCED THE ORDER OF BACKGROUND IMAGES.



PROCESSING STYLE MANIPULATION

(STUDY 6; ADAPTED FROM ALLARD, HARDISTY, AND GRIFFIN 2019)

Slow-is-accurate decision Condition

Recent psychological research has shown that "**making better choices often takes more time and effort. Taking your time often leads to making the best choice possible.**" Therefore, for the next question, a slow response time will be critical.

We ask you to take as much time as needed to respond and think carefully about your choice. Fully consider the available options before making your choice.

(please click below before proceeding to the selection task)

I will make a SLOW decision

○

Fast-is-accurate decision Condition

Recent psychological research has shown that "**people often make better choices when they decide quickly and effortlessly. A quick response often leads to making the best choice possible.**" Therefore, for the next question, a quick response time will be critical.

We ask you to take as little time as needed to respond and think quickly about your choice. Quickly consider the available options and immediately make your choice.

(please click below before proceeding to the selection task)

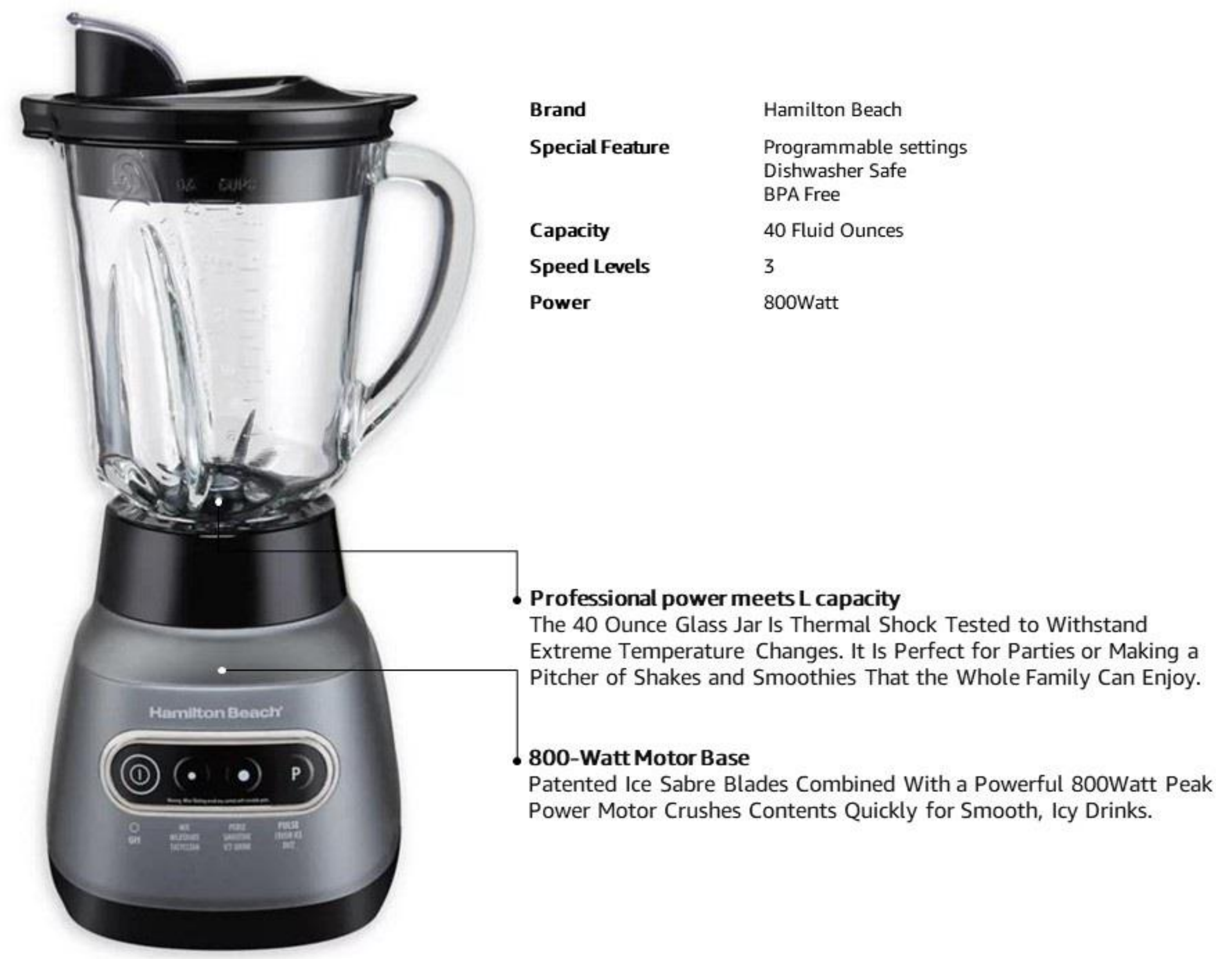
I will make a FAST decision

○

DISFLUENCY MANIPULATION – STUDY 7 (IMAGE QUALITY)

FIGURE 27

STIMULI FOR STUDY 7. BLENDERS INFORMATION IN HIGH RESOLUTION CONDITION PRESENTED ON DIFFERENT SCREENS RANDOMLY



Brand	Amzchef
Special Feature	3 Programmable settings BPA Free Dishwasher Safe
Capacity	68 Fluid Ounces
Speed Levels	3
Power	1800Watt

- **Self-Cleaning Function**
  Easy to clean with dish soap and water, use pulse function, your professional high-speed blender could clean itself in 30 to 60 seconds.

- **Smart Panel Control**
  Different settings give you the control you need to get the perfect consistency every time.

- **1800-Watt Motor Base**
  Equipped with 1800 watts 26000 RPM high powered motor, the AMZCHEF Countertop Blender creates smoothies, ice drinks, shakes, and more with one touch of the button.

Brand	ICUCINA
Special Feature	3 Programmable settings BPA Free Dishwasher Safe
Capacity	64 Fluid Ounces
Speed Levels	5
Power	1600Watt

- **Easy to Clean**
  Easy as 1,2,3! Add warm water, a touch of soup, turn the dial for 2 seconds and have your countertop blender clean itself, no disassembly required! If needed, dishwasher safe.

- **1600-Watt Motor Base**
  1600watt motor spinning at 30,000 rotations per minute blends the toughest ingredients from frozen fruits to ice cubes to the thick powder in seconds with no wear and tear.

Brand	Kogan
Special Feature	BPA Free Dishwasher Safe
Capacity	68 Fluid Ounces
Speed Levels	10
Power	2000Watt

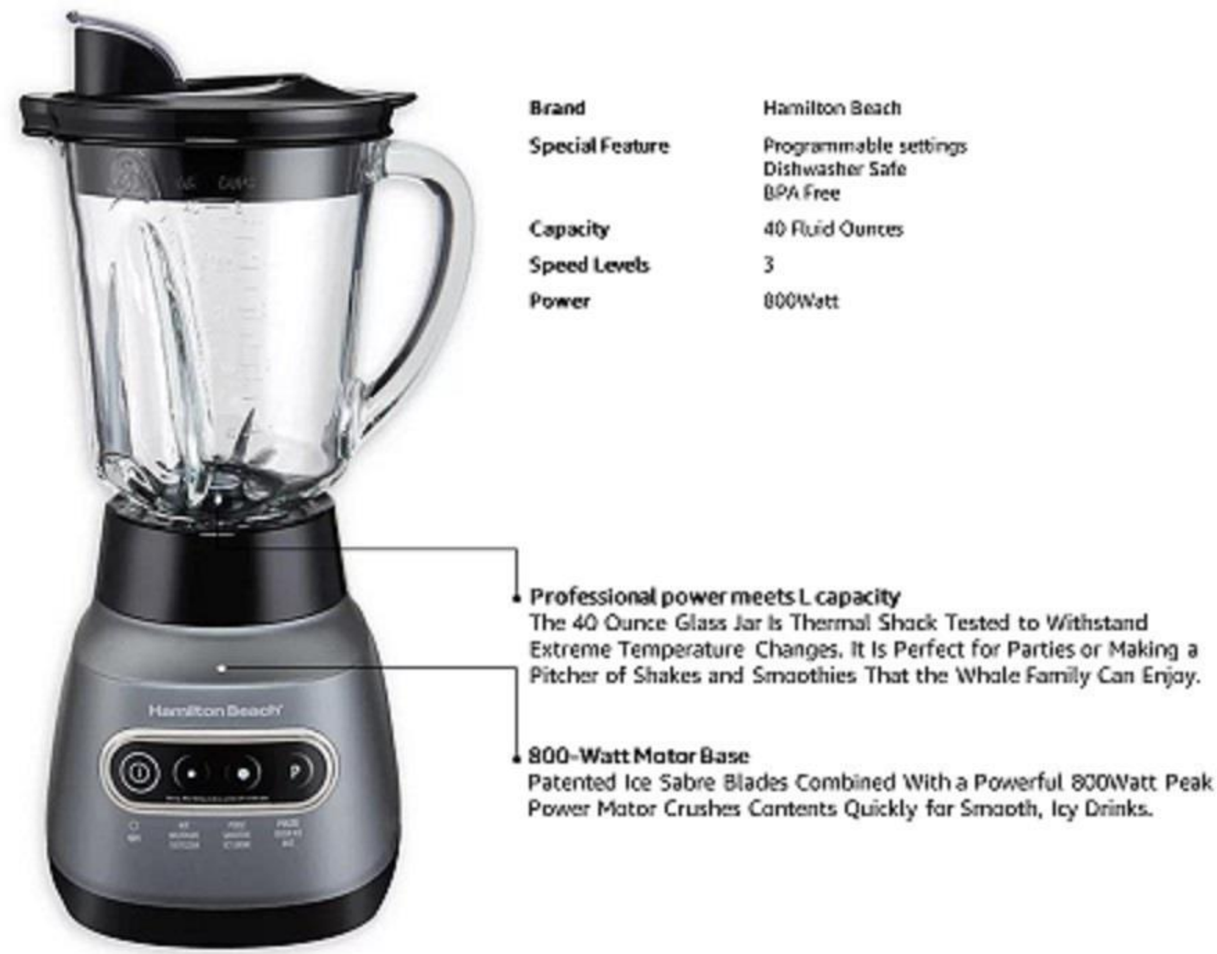
- **6 Stainless Steel Blade**
  Heavy duty stainless steel blades crushed ice and vegetables efficiently, creating shakes and smoothies in second.

- **Variable Speed Control & Timer**
  you're in control—adjust the speed at any time during the blend to achieve a variety of textures, from silky smoothies to chunky salsa.

- **2000-Watt Motor Base**
  The powerful 2200W professional blender with up to 32000 RPM can break the cell walls of both soft and hard ingredients in seconds. The upgraded 6 lengthened sharp blades makes blending range wider and more efficient.

FIGURE 28

STIMULI FOR STUDY 7. BLENDERS INFORMATION IN LOW RESOLUTION CONDITION PRESENTED ON DIFFERENT SCREENS RANDOMLY




Brand
Amzchef
Special Feature
3 Programmable settings
BPA Free
Dishwasher Safe
Capacity
68 Fluid Ounces
Speed Levels
3
Power
1800Watt
Self-Cleaning Function
Easy to clean with dish soap and water, use pulse function, your professional high-speed blender could clean itself in 30 to 60 seconds.
Smart Panel Control
Different settings give you the control you need to get the perfect consistency every time.
1800-Watt Motor Base
Equipped with 1800 watts 26000 RPM high powered motor, the AMZCHEF Countertop Blender creates smoothies, ice drinks, shakes, and more with one touch of the button.
amzchef


Brand
ICUCINA
Special Feature
3 Programmable settings
BPA Free
Dishwasher Safe
Capacity
64 Fluid Ounces
Speed Levels
5
Power
1600Watt
Easy to Clean
Easy as 1,2,3! Add warm water, a touch of soup, turn the dial for 2 seconds and have your countertop blender clean itself, no disassembly required! If needed, dishwasher safe.
1600-Watt Motor Base
1600watt motor spinning at 30,000 rotations per minute blends the toughest ingredients from frozen fruits to ice cubes to the thick powder in seconds with no wear and tear.


Brand
Kogan
Special Feature
BPA Free
Dishwasher Safe
Capacity
68 Fluid Ounces
Speed Levels
10
Power
2000Watt
6 Stainless Steel Blade
Heavy duty stainless steel blades crushed ice and vegetables efficiently, creating shakes and smoothies in second.
Variable Speed Control & Timer
you're in control—adjust the speed at any time during the blend to achieve a variety of textures, from silky smoothies to chunky salsa.
2000-Watt Motor Base
The powerful 2200W professional blender with up to 32000 RPM can break the cell walls of both soft and hard ingredients in seconds. The upgraded 6 lengthened sharp blades makes blending range wider and more efficient.

REFERENCES

Allard, Thomas, David J. Hardisty and Dale Griffin (2019), "When "More" Seems Like Less: Differential Price Framing Increases the Choice Share of Higher-Priced Options," *Journal of Marketing Research,* 56 (5), 826-41.

Alter, Adam L. (2013), "The Benefits of Cognitive Disfluency," *Current Directions in Psychological Science*, 22 (6), 437-42.

Alter, Adam L., Daniel M. Oppenheimer, Nicholas Epley, and Rebecca N. Eyre (2007), "Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning," *Journal of Experimental Psychology: General*, 136 (4), 569-76.

Amit, Adi, Sari Mentser, Sharon Arieli, and Niva Porzycki (2021), "Distinguishing Deliberate from Systematic Thinking," *Journal of Personality and Social Psychology*, 120 (3), 765-88.

Bettman, James R., Eric J. Johnson, Mary F. Luce, and John W. Payne (1993), "Correlation, Conflict, and Choice," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 931–51.

Bodenhausen, Galen V., and Robert S. Wyer (1985), "Effects of Stereotypes in Decision Making and Information-Processing Strategies," *Journal of Personality and Social Psychology,* 48 (2), 267-82.

Brannon, Laura A., and Timothy C. Brock (2001), "Scarcity Claims Elicit Extreme Responding to Persuasive Messages: Role of Cognitive Elaboration," *Personality and Social Psychology Bulletin*, 27 (3), 365-75.

Chaiken, Shelly (1980), “Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion,” *Journal of Personality and Social Psychology*, 39 (5), 752-66.

Chaiken, Shelly (1987), “The Heuristic Model of Persuasion*,” In Social influence: the Ontario symposium*, 5, 3-39.

Chaiken, Shelly, and Alice H. Eagly (1983), “Communication Modality as a Determinant of Persuasion: The Role of Communicator Salience,” *Journal of Personality and Social Psychology*, 45 (2), 241-56.

Chaiken, Shelly, and Durairaj Maheswaran (1994), “Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment,” *Journal of Personality and Social Psychology*, 66 (3), 461-73.

Chang, Ching Ching (2004) “Country of origin as a heuristic cue: The effects of message ambiguity and product involvement,” *Media Psychology*, 6 (2), 169-92.

Chen, Serena, Kimberly Duckworth, and Shelly Chaiken (1999), “Motivated Heuristic and Systematic Processing,” *Psychological Inquiry*, 10 (1), 44-9.

Dellaert, Benedict GC, and Gerald Häubl (2012), “Searching in Choice Mode: Consumer Decision Processes in Product Search with Recommendations,” *Journal of Marketing Research*, 49 (2), 277-88.

Dhar, Ravi (1997), “Consumer Preference for a No-Choice Option,” *Journal of Consumer Research*, 24 (2), 215-31.

Dhar, Ravi, and Itamar Simonson (2003), “The Effect of Forced Choice on Choice,” *Journal of Marketing Research*, 40 (2), 146-60.

Dhar, Ravi, and Stephen M. Nowlis (1999), "The Effect of Time Pressure on Consumer Choice Deferral," *Journal of Consumer Research*, 25 (4), 369-84.

Diemand-Yauman, Connor, Daniel M. Oppenheimer, and Erikka B. Vaughan (2011), "Fortune Favors The (): Effects Oo Disfluency on Educational Outcomes," *Cognition*, 118 (1), 111-15.

Frederick, Shane (2005), "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19 (4), 25-42.

Gunasti, Kunter, and William T. Ross Jr (2010), "How and When Alphanumeric Brand Names Affect Consumer Preferences," *Journal of Marketing Research*, 47 (6), 1177-92.

Hong, Sung-Tai, and Robert S. Wyer Jr. (1989), "Effects of Country-Of-Origin and Product-Attribute Information on Product Evaluation: An Information Processing Perspective," *Journal of Consumer Research*,16 (2), 175-87.

Huh, Young Eun, Joachim Vosgerau, and Carey K. Morewedge (2011), "Social Defaults: Observed Choices Become Choice Defaults," *Journal of Consumer Research*, 41 (3), 746-60.

Inbar, Yoel, Simona Botti, and Karlene Hanko (2011), "Decision Speed and Choice Regret: When Haste Feels Like Waste," *Journal of Experimental Social Psychology*, 47 (3), 533-40.

Inman, J. Jeffrey, Anil C. Peter, and Priya Raghubir (1997), "Framing the Deal: The Role of Restrictions in Accentuating Deal Value," *Journal of Consumer Research*, 24 (1), 68-79.

Iyengar, Sheena S. and Mark R. Lepper (2000), "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?," *Journal of Personality and Social Psychology*, 79 (6), 995–1006.

Janiszewski, Chris, and Tom Meyvis (2001), "Effects of Brand Logo Complexity, Repetition, and Spacing on Processing Fluency and Judgment," *Journal of Consumer Research*, 28 (1), 18-32.

Kahneman, Daniel, and Shane Frederick (2002), "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," *Heuristics and biases: The psychology of intuitive judgment*, 49, 1-30.

Labroo, Aparna A., Ravi Dhar, and Norbert Schwarz (2009), "Of Frog Wines and Frowning Watches: Semantic Priming, Perceptual Fluency, and Brand Evaluation," *Journal of Consumer Research*, 34 (6), 819-31.

Lee, Angela Y., and Aparna A. Labroo (2004), "The Effect of Conceptual and Perceptual Fluency on Brand Evaluation," *Journal of Marketing Research*, 41 (2), 151-65.

Lee, Kyoungmi, and Sharon Shavitt (2009), "Can Mcdonald's Food Ever Be Considered Healthful? Metacognitive Experiences Affect the Perceived Understanding of a Brand," *Journal of Marketing Research*, 46 (2), 222-33.

Maheswaran, Durairaj (1994) "Country of Origin as a Stereotype: Effects of Consumer Expertise and Attribute Strength on Product Evaluations," *Journal of Consumer Research*, 21 (2), 354-65.

Maheswaran, Durairaj, and Joan Meyers-Levy (1990), "The Influence of Message Framing and Issue Involvement," *Journal of Marketing Research*, 27 (3), 361-67.

Maheswaran, Durairaj, Diane M. Mackie, and Shelly Chaiken (1992), "Brand Name as a Heuristic Cue: The Effects of Task Importance and Expectancy Confirmation on Consumer Judgments," *Journal of Consumer Psychology*, 1 (4), 317-36.

Mariadassou, Shwetha, Christopher J. Bechler, and Jonathan Levav, (2023), "The Effect of Auditory and Visual Recommendations on Choice," *Psychological Science, 34* (1), 47-59.

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), e16-e30.

Mohsenin, Shahryar and Kurt Munz (2023), "Gender-Ambiguous Voices and Social Disfluency," working paper.

Mukherjee, Ashesh, and Seung Yun Lee (2016), "Scarcity Appeals in Advertising: The Moderating Role of Expectation of Scarcity," *Journal of Advertising*, 45 (2), 256-68.

Munz Kurt, Morwitz Vicki (2019), "Not-So Easy Listening: Roots and Repercussions of Auditory Choice Difficulty in Voice Commerce," New York: NYU Stern School of Business.

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

Payne, John W., James R. Bettman, and Eric J. Johnson (1988), "Adaptive Strategy Selection in Decision Making," J*ournal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (3), 534-52.

Petty, Richard E. and John T. Cacioppo (1986), "The Elaboration Likelihood Model of Persuasion," in *Advances in Experimental Social Psychology*, Vol. 19, ed. Leonard Berkowitz, New York: Academic Press, 123-205.

Pocheptsova, Anastasiya, Aparna A. Labroo, and Ravi Dhar (2010), “Making Products Feel Special: When Metacognitive Difficulty Enhances Evaluation,” *Journal of Marketing Research*, 47 (6), 1059-69.

Pocheptsova, Anastasiya, On Amir, Ravi Dhar, and Roy F. Baumeister (2009), “Deciding Without Resources: Resource Depletion and Choice in Context,” *Journal of Marketing Research*, 46 (3), 344-55.

Preacher, Kristopher J., and Andrew F. Hayes (2004), “SPSS and SAS Procedures for Estimating Indirect Effects in Simple Mediation Models,” *Behavior Research Methods, Instruments, and Computers*, 36, 717-31.

Ratneshwar, Srinivasan, and Shelly Chaiken (1991), “Comprehension’s Role in Persuasion: The Case of Its Moderating Effect on the Persuasive Impact of Source Cues,” *Journal of Consumer Research*, 18 (1), 52-62.

Schwarz, Norbert (2004), “Metacognitive Experiences in Consumer Judgment and Decision Making,” *Journal of Consumer Psychology*, 14 (4), 332-48.

Schwarz, Norbert, Madeline Jalbert, Tom Noah, and Lynn Zhang (2021), “Metacognitive Experiences as Information: Processing Fluency in Consumer Judgment and Decision Making,” *Consumer Psychology Review*, 4 (1), 4-25.

Sela, Aner, and Jonah Berger (2012), “How Attribute Quantity Influences Option Choice,” *Journal of Marketing Research*, 49 (6), 942-53.

Senecal, Sylvain, and Jacques Nantel (2004), “The Influence of Online Product Recommendations on Consumers’ Online Choices,” *Journal of Retailing*, 80 (2), 159-69.

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.

Singh, Anne-Marie, Nadine Marcus, and Paul Ayres (2012), "The Transient Information Effect: Investigating the Impact of Segmentation on Spoken and Written Text," *Applied Cognitive Psychology*, 26 (6), 848-53.

Sirota, Miroslav, Andriana Theodoropoulou, and Marie Juanchich (2021a), "Disfluent Fonts Do Not Help People to Solve Math and Non-Math Problems Regardless of Their Numeracy," *Thinking and Reasoning*, 27 (1), 142-59.

Sirota, Miroslav, Chris Dewberry, Marie Juanchich, Lenka Valuš, and Amanda C. Marshall (2021b), "Measuring Cognitive Reflection Without Maths: Development a Validation of the Verbal Cognitive Reflection Test," *Journal of Behavioral Decision Making*, 34 (3), 322-43.

Song, Camilla Eunyoung, and Aner Sela (2023), "Phone and Self: How Smartphone Use Increases Preference for Uniqueness," *Journal of Marketing Research*, 60 (3), 473-88.

Song, Hyunjin, and Norbert Schwarz (2008), "Fluency and the Detection of Misleading Questions: Low Processing Fluency Attenuates the Moses Illusion," *Social Cognition*, 26 (6), 791-99.

Suri, Rajneesh, and Kent B. Monroe (2003), "The Effects of Time Constraints on Consumers' Judgments of Prices and Products," *Journal of Consumer Research*, 30 (1), 92-104.

Thomas, Manoj, and Vicki G. Morwitz (2009), "The Ease-Of-Computation Effect: The Interplay of Metacognitive Experiences and Naive Theories in Judgments of Price Differences," *Journal of Marketing Research*, 46 (1), 81-91.

Thompson, Debora V., and Elise Chandon Ince (2013), “When Disfluency Signals Competence: The Effect of Processing Difficulty on Perceptions of Service Agents,” *Journal of Marketing Research*, 50 (2), 228-40.

Thompson, Valerie A., Jamie A. Prowse Turner, and Gordon Pennycook (2011), “Intuition, Reason, and Metacognition,” *Cognitive Psychology*, 63 (3), 107-40.

Thompson, Valerie A., Jamie A. Prowse Turner, Gordon Pennycook, Linden J. Ball, Hannah Brack, Yael Ophir, and Rakefet Ackerman (2013), “The Role of Answer Fluency and Perceptual Fluency as Metacognitive Cues for Initiating Analytic Thinking,” *Cognition*, 128 (2), 237-51.

Wadhwa, Monica, and Kuangjie Zhang (2015), “This Number Just Feels Right: The Impact of Roundedness of Price Numbers on Product Evaluations,” *Journal of Consumer Research*, 41(5), 1172-85.

Walter, Maik, Christian Hildebrand, Gerald Häubl, and Andreas Herrmann (2022), “Mixing It Up: Unsystematic Product Arrangements Promote the Choice of Unfamiliar Products,” *Journal of Marketing Research*, 57 (3), 509-26.

Wright, Peter (1975), “Consumer Choice Strategies: Simplifying vs. Optimizing,” *Journal of Marketing Research*, 12 (1), 60-7.

Yan, Dengfeng, and Rod Duclos (2013), “Making Sense of Numbers: Effects of Alphanumeric Brands on Consumer Inference,” *International Journal of Research in Marketing*, 30 (2), 179-84.

CHAPTER 3

BENEFITS OF DISFLUENCY: WHEN FEELING WRONG IS RIGHT

It has been long documented that, in comparison to fluency, disfluency leads to lower judgments and evaluations (see Alter and Oppenheimer 2009) or choice deferral (Novemsky et al. 2007). Our focus on disfluency is intended to demonstrate how framing information as either prevention-focused claims or promotion-focused claims affects purchase deferral. Regulatory focus theory suggests that consumer judgments, decisions, and behaviors are influenced by two regulatory orientations: promotion and prevention focus (e.g., Higgins et al. 2001; Wang and Lee 2006). Therefore, building on regulatory fit theory, we predict that, when experiencing feelings of disfluency, people experience less regulatory fit when exposed to promotion-focused claims rather than prevention-focused information. This reduction in regulatory fit leads to reduced purchase deferral due to the advantage of being wrong when faced with difficulty in processing information.

Theory development

According to regulatory fit theory (Higgins 2002), when individuals pursue a goal in a manner that aligns with their regulatory orientation (i.e., a promotion-focused person pursuing a goal in a promotion-oriented manner or a prevention-focused person pursuing a goal in a prevention-oriented manner), they experience a psychological sense of correctness. The alignment between regulatory orientation and the approach to goal pursuit, referred to as *regulatory fit* (Higgins 2002), generates a psychological advantage that can significantly impact subsequent judgments, decisions, and behaviors. This means that when there is a fit, individuals are more deeply engaged and feel a sense of correctness in their actions (Cesario, Grant, and Higgins 2004). As a result, subsequent evaluative reactions, whether positive or negative, can be intensified by this experience of fit (Avnet and Higgins 2006).

Wang and Lee (2006) demonstrated that individuals tend to place significant reliance on prevention-focused attributes that align with their regulatory focus. In other words, they are more likely to choose options with strong prevention claims. Levav, Kivetz, and Cho (2010) found that people exhibit a greater tendency to defer their choices in regulatory compatibility conditions when the product attribute (promotion vs. prevention) and the consumer's goal (promotion vs. prevention focus) are aligned, compared to regulatory incompatibility conditions. Novemsky, Dhar, Schwarz, and Simonson (2007) observed that individuals are more inclined to defer their decisions when faced with hard to process information. They argued that individuals attribute perceived disfluency to the decision's difficulty within a choice set, leading to increased choice deferral when the task appears difficult. This means that when information is easy to process, consumers are more inclined to make a purchase rather than deferring. However, when they experience subjective feelings of disfluency, this tendency shifts toward purchase deferral. Consequently, this research aims to apply regulatory fit theory to investigate how disfluency influences purchase deferral when product information is framed in terms of promotions and preventions. Our hypothesis is as follows:

H1. *People focus more on prevention strategies when it is hard to process information than when they feel at ease; choosing more often an option stronger on prevention featured claims.*

Building upon regulatory fit theory (Higgins, 2002) and the concept of "*feeling right*" (Cesario et al., 2004), our hypothesis posits that in situations characterized by information processing difficulty, individuals are more inclined to defer their choices when presented with a set of options emphasizing prevention-focused claims. This preference stems from their active selective attention to information that aligns with their regulatory focus, particularly a prevention focus. For instance, having faced difficulty in processing information, people might weigh a product featuring prevention claims strongly. Meaning that disfluency tunes their attention to this type of information, as they actively seek justifications for avoiding a purchase. In contrast, when it is easy to read information, consumers tend to place greater emphasis on promotion-focused claims, which align with their orientation toward fluency as they look for reasons to justify making a purchase.

As a result, encountering disfluency within a consumer choice context featuring information focused on prevention aspects leads consumers who face difficulty in processing information to experience a stronger regulatory fit (Higgins, 2002). This alignment creates a subjective sense of congruence between the message content and their approach to the decision-making task, fostering an increased likelihood of choice deferral. Conversely, individuals facing disfluency are less prone to deferring their choices, primarily due to the incongruity between their regulatory focus orientation, prompted by disfluency, and the messaging that predominantly highlights promotional claims (see Levav et al., 2010). Consequently, the sense of discord, often described as "feeling wrong" (i.e., indicating regulatory nonfit), diminishes the likelihood of choice deferral, particularly in situations where the information emphasizes promotion strategies rather than prevention claims under conditions of decision difficulty. Hence, we posit the following hypotheses:

H2a. *When it is hard to process information, consumers feel wrong (right) during a choice task when an option featuring promotion (prevention)-focused claims, which in turn decreases (increases) purchase deferral.*

H2b. *When it is easy to process information, consumers feel right (wrong) during a choice task when an option featuring promotion (prevention)-focused claims, which in turn increases (decreases) choice deferral.*

STUDY 1: DISFLUENCY AND REGULATORY FOCUS

This study aims to show the relationship between subjective feelings of difficulty and regulatory focus mindset. Since it has been shown that promotion-focused (preventions-focused) participants preferred an option with the strong promotion (prevention) features (Wang and Lee, 2006), we are running this study to demonstrate that if disfluency activates prevention foci mindset then they would show higher regulatory fit under the feelings of difficulty when a product framed with prevention benefits.

Method

Three hundred US and Canadian participants will be recruited through Amazon Mechanical Turk to take part in this study in exchange for payment. Participants will be randomly assigned into two between conditions of a 2 (Disfluency: hard-to-read font vs. easy-to-read font; between) by 2 (Products' regulatory focus: promotion- vs. prevention- focused benefits; within) mixed design. All participants who complete the experiment will be included in the analysis.

The stimulus for this experiment is toothpaste. To manipulate subjective feelings of disfluency, the options are presented to half of the participants with the descriptions presented in an easy-to-read standard font (Arial font), while the others see the same descriptions in a hard-to-read font (Amarix SC, cursive font). Participants are asked to imagine that they needed to buy new toothpaste. Adopting from Wang and Lee (2006), in this study, we will activate either promotion or prevention foci mindset by presenting participants with description of two toothpaste. While one of the toothpaste has stronger promotion but weaker prevention product benefits, to activate prevention-focus, the other toothpaste features with stronger prevention but weaker promotion product claims. We'll ask participants to make a choice between two toothpaste by asking them, "If you were going to buy one, which toothpaste would you like to purchase?" (*Toothpaste A* vs. *Toothpaste B*). We'll counterbalance the order of the two toothpastes across all participants. Next, the Regulatory Focus Measurement (RFQ; Higgins et al. 2001) will be used to measure participants' regulatory focus, operationalized as the subjective measurement of promotion focus versus prevention focus. At the end, to check whether our manipulation of disfluency works, participants will be asked to rate "How difficult it was to read the products information?" (1 = Not at all; 7 = Very much).

We expect that people in hard-to-read font condition, choose more often the toothpaste with prevention foci claims as their goal is to reject more than people in easy-to-read font condition who are looking for a reason to justify their purchase leading to selecting more frequently the toothpaste with promotion foci attributes.

STUDY 2: SUBJECTIVE FEELINGS OF WRONGNESS AS A MEDIATOR

In Study 1, we are looking for evidence to show that people focus more on product prevention claims when it comes to subjective difficulty. Following the result of the first study, we are looking for the underlying mechanisms of this effect in a choice context. While we expect to see that people in easy-to-read font condition, defer less when options featuring with promotion attributes due to feeling rightness, we predict that when it is hard to process information, people feel right more when they are evaluating options featured with prevention attributes which in turn would decrease not choosing the available options compared to a choice settings consists of the same options featuring with promotion attributes.

Method

Three hundred US and Canadian participants will be recruited through Amazon Mechanical Turk to take part in this study in exchange for payment. Participants will be randomly assigned into four conditions of a 2 (Disfluency: hard-to-read font vs. easy-to-read font) by 2 (Products' regulatory focus: prevention- vs. promotion-focused benefits) between-participants design. All participants who complete the experiment will be included in the analysis.

In this study, we will use multivitamins as stimuli. To manipulate subjective feelings of disfluency, the options are presented to half of the participants with the descriptions presented with a background photo (disfluent condition), while the others see the same descriptions without any backgrounds (fluent condition). Following Study 1, we will activate either regulatory focus or promotion focus by providing participants with options featuring strong promotion or prevention benefits. Participants will be randomly assigned to one of the regulatory focus conditions. Having read product information, participants will be asked to make a decision whether want to buy a multivitamin or not. Then, to measure the underlying mechanism, perceived feeling right index (Cesario and Higgins 2008), participants will be asked to indicate how they feel right or wrong while they are evaluating options using 7-point scales. A feeling-right index is created by subtracting feeling-wrong ratings from feeling-right ratings; higher numbers indicate feeling more right. At the end, to check whether our manipulation of disfluency works, participants will be asked to rate "How difficult it was to read the products information?" (1 = Not at all; 7 = Very much).

We expect that people in disfluent condition defer less when evaluating multivitamins featuring with strong promotion claims due to feeling wrongness, while people in easy-to-read font condition would show the same pattern of result but due to the feeling of rightness. Thus, we expect that feeling of wrongness mediates the effect of subjective feelings of difficulty on purchase decision.

STUDY 3: ATTRIBUTION TO THE SOURCE OF FEELING RIGHT AS A MODERATOR

We plan to run this study to test for the hypothesized misattribution process or transferring of feeling wrongness from regulatory unfit to evaluation in the choice making context. That is, participants' attention either is or isn't drawn to the true source of their feeling right experience prior to their receiving the products information (e.g., Cesario et al., 2004). If the feeling wrong experience from regulatory unfit transfers to the choice context, as we hypothesize, then drawing participants' attention to the source of this unfit experience before they receive the message should reduce the choice deferral more which might lead to making the effect stronger compared with not drawing attention to the source of feeling wrongness. Therefore, a disfluency by attention condition interaction on participants' choice making is predicted. The standard condition in which participants' attention is not drawn to the source of the feeling wrong experience should show the usual effect on deferral, replicating the result of Study 2.

Method

Six hundred US and Canadian participants will be recruited through Amazon Mechanical Turk to take part in this study in exchange for payment. Participants will be randomly assigned into eight conditions of disfluency (background image vs. no background image) by products' regulatory focus (promotion vs. prevention-focused benefits) by attention (control vs. attention to source of feeling wrong) mixed design. All participants who complete the experiment will be included in the analysis.

Participants will be told that they are going to read some information about two face moisturizing creams. While in promotion focus attributes; both of them are featured with promotion attributes; in the prevention focus attributes they are featured with prevention attributes. First, participants will be randomly assigned in one of two attention conditions. In the standard condition, which served as a replication of the previous two studies, no instructions are given to direct attention to the source of the feeling wrong experience. In the attention to feeling wrong condition, participants' attention was directed to the source of their feeling wrong experience with the following instructions: "Sometimes thinking about avoiding negative outcomes can make you "feel better" about your goal pursuit." On the following scale, participants will indicate how much you "feel wrong" about your goal pursuit. Participants then indicated their response on a 7-point scale anchored at not at all and extremely.

REFERENCES

Alter, Adam L., and Daniel M. Oppenheimer (2009), "Uniting the tribes of fluency to form a metacognitive nation," *Personality and social psychology review,* 13 (3), 219-35.

Avnet, Tamar, and E. Tory Higgins (2006), "How regulatory fit affects value in consumer choices and opinions," *Journal of Marketing research,* 43 (1), 1-10.

Cesario, Joseph, and E. Tory Higgins (2008), "Making message recipients "feel right" how nonverbal cues can increase persuasion," *Psychological science,* 19 (5), 415-20.

Cesario, Joseph, Heidi Grant, and E. Tory Higgins (2004), "Regulatory fit and persuasion: Transfer from "feeling right."," *Journal of personality and social psychology,* 86 (3), 388-404.

Higgins, E. Tory (2002), "How self-regulation creates distinct values: The case of promotion and prevention decision making," *Journal of Consumer Psychology,* 12 (3), 177-91.

Higgins, E. Tory, Ronald S. Friedman, Robert E. Harlow, Lorraine Chen Idson, Ozlem N. Ayduk, and Amy Taylor (2001), "Achievement orientations from subjective histories of success: Promotion pride versus prevention pride," *European Journal of Social Psychology,* 31 (1), 3-23.

Levav, Jonathan, Ran Kivetz, and Cecile K. Cho (2010), "Motivational compatibility and choice conflict," *Journal of consumer research,* 37 (3), 429-42.

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

Wang, Jing, and Angela Y. Lee (2006). The role of regulatory focus in preference construction. *Journal of Marketing research,* 43(1), 28-38.