Fairness Perceptions in Demographic Targeting

Franklin Shaddy*

Elizabeth M.S. Friedman*

Olivier Toubia

Forthcoming, *Journal of Consumer Research*

Corresponding author: Franklin Shaddy

*The first and second authors contributed equally

Franklin Shaddy (franklin.shaddy@anderson.ucla.edu) is Assistant Professor of Marketing and Behavioral Decision Making at the UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, 90095. Elizabeth M.S. Friedman (e.friedman@columbia.edu) is Assistant Professor of Marketing at Columbia Business School, 645 West 130th Street, New York, NY 10027. Olivier Toubia (ot2107@gsb.columbia.edu) is the Glaubinger Professor of Business at Columbia Business School, 645 West 130th Street, New York, NY 10027. The authors thank UCLA Anderson, Columbia Business School, a Facebook Core Data Science (CDS) Research Gift, and the Morrison Center for Marketing and Data Analytics at UCLA Anderson for funding.

ABSTRACT

Deciding which customer segment(s) to serve, or targeting, is a cornerstone of marketing strategy. While companies commonly target based on demographic characteristics, recent cultural movements (e.g., #MeToo, Black Lives Matter) have heightened sensitivity to the differential treatment of certain demographic groups. Yet research to date has not examined whether demographic targeting *itself* seems fair or appropriate. Fourteen experiments (N = 9,399), 13 supplemental studies (N = 7,065), and two Facebook A/B Tests (N = 513,151) reveal that when consumers learn or infer that they or others have been targeted based on demographic characteristics, fairness perceptions and brand support suffer (relative to broad advertising). To explain why, we propose a conceptual model based on the extent to which consumers view demographic targeting as discriminatory (i.e., differential treatment based on attributes that are irrelevant and/or uncontrollable), and whether the discrimination is perceived as intentional (i.e., knowingly or willingly bringing about an avoidable outcome). Consequently, factors that (a) improve relevance (i.e., whether belonging to the targeted segment is diagnostic of preferences) or (b) increase controllability (i.e., whether consumers themselves determine their membership in the targeted segment) attenuate perceptions of unfairness, as do variables that (c) reduce perceived intentionality (e.g., firm size, industry norms).

Keywords: segmentation, targeting, fairness, discrimination, race, gender, marketing strategy

Word count: 200/200

Deciding which customer segment(s) to serve, or targeting, is a cornerstone of marketing strategy (Kotler and Keller 2011). Often, customer segments are defined by demographic characteristics—for example, by race (Aaker, Brumbaugh, and Grier 2000), gender (Winterich et al. 2015), age (Tepper 1994), socioeconomic status (SES; Shavitt, Jiang, and Cho 2016), or geography (Andreasen 1966).

While recent advances in digital marketing technologies have facilitated more sophisticated psychographic and behavioral approaches to targeting, they have also triggered stricter consumer privacy protections (Goldfarb and Tucker 2010; Rafieian and Yoganarasimhan 2021). For example, the European Union's General Data Protection Regulation (GDPR) prohibits behavioral tracking without explicit user consent. And a key feature of Apple's mobile operating system is its "Ask App Not to Track" functionality. As a result, firms have increasingly reverted back to demographic targeting (Moorman, Ryan, and Tavassoli 2022).

Yet cultural movements like Black Lives Matter and #MeToo have heightened sensitivity to the differential treatment of certain demographic groups, reflecting broader concerns about systemic bias and social justice (Grier et al. 2023; Nardini et al. 2021). Critically, differential treatment is an inescapable feature of targeting, regardless of its basis (demographic, behavioral, or psychographic). Some platforms, like Facebook, have responded by limiting certain forms of demographic targeting (Facebook 2021). Nevertheless, targeting based on increasingly sensitive attributes (e.g., race and gender) remains pervasive (Moshary, Tuchman, and Bhatia 2023).

In this research, we introduce a broadly applicable framework for understanding fairness perceptions in demographic targeting. We find that fairness perceptions suffer when consumers learn or infer that they or others have been targeted based on their demographic characteristics (relative to broad advertising, or mass marketing), decreasing their willingness to engage with, recommend, or purchase from the company.

To explain why, we propose a conceptual model that hinges on the extent to which consumers view demographic targeting as *discriminatory* (i.e., differential treatment based on attributes that are irrelevant and/or uncontrollable; Tomova Shakur and Phillips 2022), and whether the discrimination is perceived as intentional (i.e., knowingly or willingly bringing about an avoidable outcome). Consequently, factors that (a) improve relevance (i.e., whether belonging to the targeted segment is diagnostic of preferences) or (b) increase controllability (i.e., whether consumers themselves determine their membership in the targeted segment) attenuate perceptions of unfairness, as do variables that (c) reduce perceived intentionality (e.g., when the company is small and when demographic targeting is the norm).

Our theory contributes insights to at least three distinct literatures. First, while fairness research has largely focused on pricing (Xia et al. 2004), relatively less attention has been paid to targeting, an equally critical element of marketing strategy. Second, our findings build on past work exploring persuasion knowledge (Aaker, Brumbaugh, and Grier 2000; Campbell and Kirmani 2008; Friestad and Wright 1994), not only further exploring the intersection of fairness perceptions and marketplace metacognition (Bolton and Chen 2024; Wright 2002), but also illustrating a potential tension between the lay beliefs of consumers and the prevailing views of practitioners. For example, when we surveyed 164 full-time MBA students at a West Coast business school, these current and future managers did *not* view demographic targeting as unfair, relative to broad advertising (see Web Appendix Study WA1). Third, because demographic targeting often involves differential treatment of historically underrepresented groups (e.g., Black customers, women, lower SES consumers), this research answers recent calls to promote and explore diversity in marketing (Arsel, Crockett, and Scott 2022; Uduehi et al. 2025). In particular, we find that consumers respond more favorably to ads depicting diverse groups of people or ads placed in media outlets that tend to draw diverse target audiences. Our results suggest that diversity in advertising can thus create value for both consumers and firms alike (i.e., “win-win” outcomes; Chandy et al. 2021).

PAST RESEARCH

Fairness refers to the appropriateness, legitimacy, or justness of a process or outcome (Colquitt and Rodell 2015; Lupfer et al. 2000). And consumers’ perceptions of fairness can meaningfully constrain firms in the marketplace (Bhattacharjee, Dana and Baron 2017; Gal, Parker, and Li 2018; Kahneman, Knetsch, and Thaler 1986). For example, although raising prices in response to a demand shock might be profit-maximizing in theory, doing so can trigger accusations of unfairness in practice (Bolton, Warlop, and Alba 2003; Campbell 1999; Friedman and Toubia 2022). These beliefs harm profitability by increasing complaints (Huppertz, Arenson, and Evans 1978), causing dissatisfaction (Oliver and Swan 1989), and ultimately reducing purchase intentions (Bechwati and Morrin 2003).

Relatively less attention has been paid to the fairness of targeting, the practice of selecting which customer segment(s) to serve. These segments can be defined by everything from demographics (e.g., race, gender, age, socioeconomic status, geography) to behaviors (e.g., purchase patterns; Assael and Roscoe 1976) to psychographics (e.g., personality, lifestyle, identity; Wells 1975). Importantly, targeting allows firms to more efficiently allocate limited resources, increasing the likelihood of reaching an interested customer. Firms often favor targeting demographic characteristics—the focus of our theorizing—because these variables tend to be relatively observable and accessible.

Targeting is viewed negatively when promoted products could be harmful to vulnerable populations (e.g., advertising alcohol in poor communities; Smith and Cooper-Martin 1997), or when identity-based messages reinforce stereotypes about marginalized groups (e.g., BIC's pink pens "For Her"; Kim et al. 2023; Paul, Parker and Dommer 2020). Firms also price discriminate across segments (e.g., different prices for different groups), a practice that consumers regard as exploitative (Heyman and Mellers 2008; van der Sluis et al. 2023; Wang and Krishna 2012).

Advances in digital marketing technologies, meanwhile, have enabled more sophisticated forms of psychographic and behavioral targeting, but they have triggered new concerns about privacy (Acquisti, John, and Loewenstein 2013; Brough et al. 2022). For example, marketers can tailor advertisements to a specific user's search terms or browsing history (Summers, Smith, and Reczek 2016), and "retarget" advertisements from one website to the next (Lambrecht and Tucker 2013). However, trust and ultimately advertising effectiveness suffer when consumers believe personal information has been inappropriately shared "behind-the-scenes" (Aguirre et al. 2015; Kim, Barasz, and John 2019). As a result, many firms have reverted back to demographic targeting (Moorman et al. 2022).

In short, research to date has largely focused on how companies *execute* various targeting strategies, which can feel unfair and inappropriate when they facilitate price discrimination, promote harmful products, reinforce stereotypes, and violate privacy. But in our work, we offer a generalizable framework for answering a broader question: When is demographic targeting *itself* considered unfair or inappropriate? And what happens when consumers learn from or infer through promotion that they have been targeted based on their demographics?

These are critical questions, because although firms' demographic targeting strategies are not typically communicated to consumers directly, consumers often learn about them *indirectly*, such as through news coverage. For example, Facebook faced public backlash when journalists discovered features that allowed marketers to target advertisements to certain races and genders (Imana, Korolova, and Heidemann 2021; Isaac and Hsu 2021). Other times, consumers themselves want to know how firms make targeting decisions, leading some companies to disclose this information explicitly (e.g., “Why am I seeing this ad?” buttons on social media; Culnan 2000; Kim et al. 2019). As we show in our studies, when asked to consider the fairness of demographic targeting, consumers are sensitive to and pick up on subtle cues in ads that reveal which demographic groups the firm seems to be targeting (e.g., when only members of a certain race or gender are depicted).

Accordingly, our findings contribute new insights to the persuasion knowledge literature, which has similarly explored how consumers identify and respond to marketers’ persuasion attempts (Friestad and Wright 1994; Eisend and Tarrahi 2022; Isaac and Grayson 2017). For example, attitudes toward ads can depend on inferences about which segment the firm is seeking to persuade (i.e., target) in the first place, such as whether viewers believe they belong to the targeted segment and whether that segment is distinctive (i.e., minority vs. majority; Aaker et al. 2000). Our framework extends this work to examine implications for fairness and brand support (operating through appraisals of discrimination), thereby integrating theories of fairness and marketplace metacognition (Bolton and Chen; 2024; Wright 2002).

THEORETICAL FRAMEWORK

In this research, we propose a conceptual framework (Figure 1) to explain fairness perceptions in demographic targeting. We suggest that these beliefs depend on the extent to which consumers view demographic targeting as discriminatory (Fiske et al. 2002; Major, Quinton, and McCoy 2002) and whether the discrimination is perceived as intentional.

To define discrimination, we borrow from recent literature in organizational behavior highlighting two key factors: (a) relevance and (b) controllability (Tomova Shakur and Phillips 2022). Specifically, when assessing hiring decisions, people view the use of demographic characteristics as less discriminatory when they are either relevant (i.e., conveying credible diagnostic information; Song Hing et al. 2011) or controllable (i.e., people can change those attributes themselves; Weiner 2000).

We extend and expand upon this framework in the context of marketing strategy. First, we define *relevance* as whether belonging to the targeted demographic group is diagnostic of preferences (Shaddy and Shah 2018). For example, it could be easier to imagine how wants or needs might meaningfully differ between men and women, but not necessarily between Black and White customers (for certain products). Second, we define *controllability* as the extent to which consumers have agency over their own membership in the targeted group. For example, people generally exercise more control over where they live than over their race or gender. Third, we define *perceived intentionality* as the belief that an action or its consequences were knowingly or willingly brought about, when the outcome was otherwise avoidable. For example, consumers may be *less* likely to view discrimination resulting from demographic targeting as intentional when a firm is small and simply lacks the resources to advertise broadly.

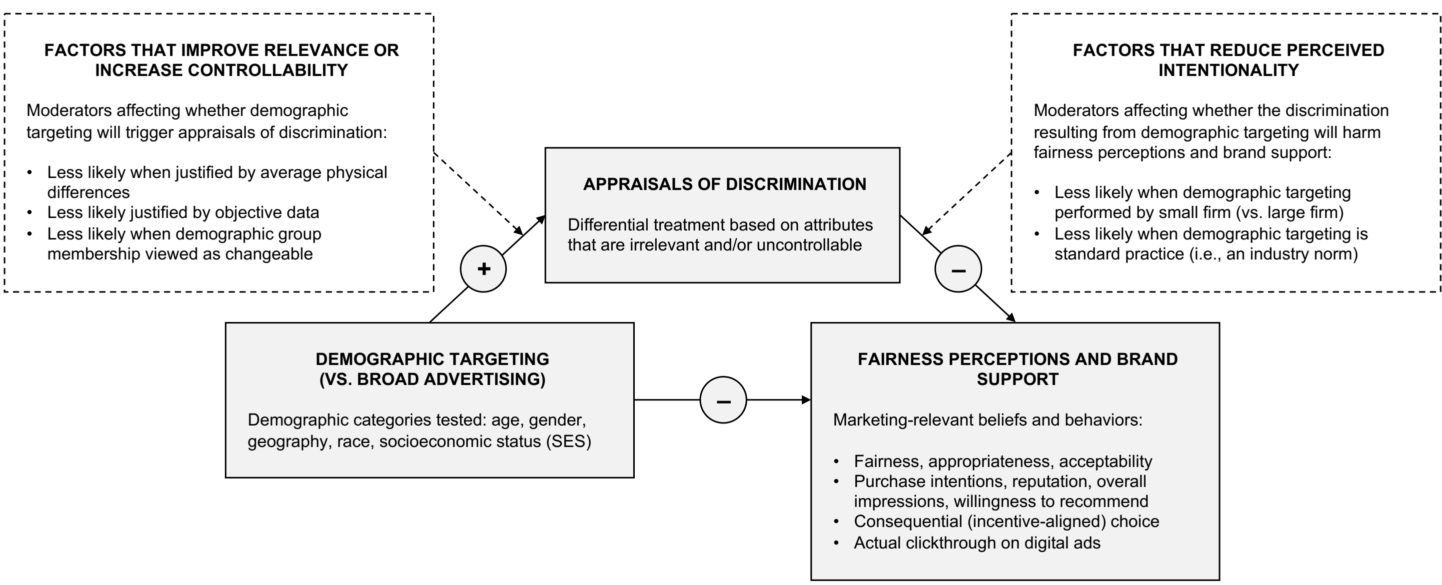
Our first set of hypotheses thus compares various forms of demographic targeting to broad advertising (i.e., not targeting at all), testing both fairness and brand support (e.g., purchase intentions, impressions of the firm, willingness to recommend, etc.), along with the underlying discrimination mechanism. Moreover, in line with recent research exploring fairness in consumer behavior (e.g., Evangelidis 2024; Friedman and Toubia 2022; Goenka and Bagchi 2024; Trupia and Shaddy 2025), we focus on *lay beliefs* about fairness (i.e., folk conceptions; Bhattacharjee and Dana 2024; Lupfer et al., 2000), defined as a general or global sense of whether an action feels appropriate, legitimate, and just (Tomova Shakur and Phillips 2022; Xia et al., 2004).

H_1: Targeting a particular demographic group, relative to broad advertising (a) is perceived as less fair and (b) reduces brand support.

H_2: Appraisals of discrimination (i.e., differential treatment based on attributes that are irrelevant and/or uncontrollable) mediate the effect of demographic targeting (vs. broad advertising) on perceptions of unfairness and brand support.

FIGURE 1

CONCEPTUAL MODEL



Theoretically Derived Moderators

A theoretical implication of our account is that demographic targeting should not always trigger negative reactions. Factors that (a) improve relevance, (b) increase controllability, or (c) reduce perceived intentionality should attenuate perceptions of unfairness. Our second set of hypotheses, therefore, not only corroborate our proposed mechanism, but also highlight the conditions under which demographic targeting will be viewed as fairer and more appropriate.

Relevance. First, when group membership actually signals preferences (Shaddy and Shah 2018), demographic targeting may prove useful in matching interested customers to products and services that uniquely suit their needs. For example, Band-Aid offers "Ourtone" bandages, which come in a variety of skin tone shades (D'Angelo, Dunn, and Valsesia 2024). Here, demographic targeting may feel less discriminatory, because consumers expect average physical differences at the population level (e.g., skin tone) across certain demographic groups to shape preferences (e.g., for different bandage colors). Other examples could include products designed for certain body sizes (e.g., average physical differences by gender) or services tailored to certain medical needs (e.g., average physical differences by age).

Second, firms can also explicitly justify their targeting decisions based on objective data. For instance, if a firm anticipated that a particular demographic group maintained stronger preferences for the promoted product, its managers might measure those preferences directly. Communicating those objective data to customers could confirm group membership is actually correlated with (i.e., relevant to) preferences (Shaddy and Shah 2021) and increase the credibility of such claims (Ford, Smith, and Swasy 1990), thereby assuaging concerns about discrimination.

H_{3a}: Factors that increase the perception that membership in a demographic category is relevant (i.e., diagnostic of preferences) attenuate the negative effect of demographic targeting (vs. broad advertising) on perceptions of fairness.

Controllability. When membership in the targeted demographic group is believed to be changeable, targeting that group may seem more justifiable. For example, people often disagree about the extent to which socioeconomic status (SES) is controllable (Bullock, Williams, and Limbert 2003; Davidai 2018; Dolifka, Christensen, and Shaddy 2024). We expect beliefs about whether people have agency to determine certain demographic attributes themselves (e.g., SES), therefore, to moderate concerns about discrimination.

H_{3b}: The negative effect of demographic targeting (vs. broad advertising) on perceptions of fairness attenuates when membership in a demographic category is viewed as more controllable.

Perceived Intentionality. When actions resulting in discrimination are believed to be less intentional, demographic targeting may seem more acceptable. In other words, a relevant consideration is whether the firm could or should have behaved differently. For example, a local mom-and-pop store might need to stretch every ad dollar as far as possible, limiting its ability to advertise broadly. On the other hand, a massive multinational conglomerate can presumably afford to advertise broadly, so its decision to engage in demographic targeting would seem more intentional—possibly driven by opportunism, exploitation, and profit-seeking (Bhattacharjee et al. 2017; Lu et al. 2020). Indeed, large firms feel more "corporate" (Reich and Hanson 2024), and are believed to possess more market power (Paharia, Avery, and Keinan 2014; Yang and Aggrawal 2019) and financial resources (Woolley, Kupor, and Liu 2023).

Additionally, consumers may not realize demographic targeting is standard practice. When a firm conforms to normative (i.e., common) behaviors, consumers may infer less intentionality (e.g., it is simply "following along"; Bellezza, Gino, and Keinan 2014; Li, Gordon, and Gelfand 2017). However, if demographic targeting were *not* standard practice, and a firm chose to do so anyway, that decision to override a norm could communicate greater intentionality. This reasoning is broadly consistent with work showing that practices initially deemed unfair can become more acceptable when those practices become normative, or more common (e.g., dynamic pricing for flights and hotels; Haws and Bearden 2006; Kimes 1994).

Consumers may thus reason that large firms and firms that violate norms could or should have behaved differently (but instead chose to engage in demographic targeting). Both factors suggest the resulting discrimination was an avoidable outcome that was knowingly or willingly brought about. Indeed, consumers naturally try to gauge firms' intentions (Newman, Gorlin, and Dhar 2004; Reich, Kupor, and Smith 2018). And past work has shown that *negative* side effects—such as the discrimination resulting from demographic targeting predicted by our framework—are perceived as more intentional than positive side effects (Knobe 2003; Papadopoulos and Hayes 2018; Uttich and Lombrozo 2010).

H_{3c}: Variables that reduce perceptions of intentionality attenuate the negative effect of demographic targeting (vs. broad advertising) on perceptions of fairness.

Empirical Roadmap and Theoretical Scope

We explore this account across 14 experiments (*N* = 9,399; Table 1), 13 supplemental studies (*N* = 7,065), and two Facebook A/B Tests (*N* = 513,151), in which we ask participants directly to assess the fairness of various demographic targeting strategies. In Studies 1A–3B, we test the basic effect of demographic targeting on fairness perceptions and brand support (H_1), using various naturalistic stimuli (e.g., social media disclosures, realistic ads, news coverage), naturalistic behaviors (e.g., purchase intentions, consequential choice), and naturalistic samples (e.g., participants from the targeted segment). In Studies 4A–7B, we systematically test each element of our conceptual model ($H_{2–3C}$) by describing various targeting strategies to participants and measuring perceptions of fairness. Finally, in Studies 8A–B, we report two large-scale

Facebook ad campaigns, observing the real behavior (i.e., clickthrough) of more than 500,000 users. This empirical approach allows us to paint a comprehensive theoretical picture describing both consumer responses to being targeted and consumer judgments of marketing strategy.

In several studies, we also test a second non-demographic (but managerially relevant) baseline condition: behavioral targeting. Like any other form of targeting, behavioral targeting necessarily involves differential treatment. Yet we expect behaviors to be viewed as relatively more controllable and relevant to preferences than demographic characteristics. It should therefore be viewed more favorably than demographic targeting (as we find in Study 4). In other words, differential treatment only undermines fairness and brand support when it is based on attributes that are irrelevant and/or uncontrollable, consistent with our theorizing.

Finally, our studies examine reactions to demographic targeting primarily through the lens of promotion (one of the four "Ps"). We either state (in scenarios) or imply (through subtle cues in stimuli) that a firm is simply *advertising* to a particular demographic group, rather than changing the product, its placement, or price—though we expect our framework to apply to other areas of marketing strategy, including the four "Ps" (see General Discussion).

TABLE 1

OVERVIEW OF STUDIES

Study	Hyp.	Contribution	Summary of findings
1A	H_1	Fairness	Fairness, appropriateness, and acceptability were lower when a social media disclosure revealed demographic targeting (vs. broad advertising)
1B	H_1	Marketing implications	Purchase intentions, overall impressions, and willingness to recommend were lower when a social media disclosure revealed demographic targeting (vs. broad advertising)
1C	H_1	Consequential choice	Consequential choice of a gift card was lower when a social media disclosure revealed demographic targeting (vs. broad advertising)
2A	H_1	Inferences from ads	Fairness, appropriateness, and acceptability were lower when an ad depicted a single race (implying demographic targeting) versus multiple races (implying broad advertising)
2B	H_1	Inferences from ads	Fairness, appropriateness, and acceptability were lower when ad depicted a single gender (implying demographic targeting) versus multiple genders (implying broad advertising)
3A	H_1	Black participants	Black participants rated fairness, appropriateness, and acceptability lower when a news article described advertisements placed on media consumed by a particular race (implying demographic targeting), as opposed to multiple races (implying broad advertising)
3B	H_1	Female participants	Female participants rated fairness, appropriateness, and acceptability lower when a news article described advertisements placed on media consumed by a particular gender (implying demographic targeting), as opposed to multiple genders (implying broad advertising)
4	$H_{1, 2}$	Mediation through discrimination	Demographic targeting (based on race, gender, age, SES, and geography) was rated as less fair (vs. broad advertising), and appraisals of discrimination (coded based on open-ended explanations) mediated differences in fairness perceptions
5A	H_{3a}	Moderation by relevance	Targeting based on race was viewed as fairer when justified by average physical differences, such as different skin tones (attenuation), than by differences in preferences (basic effect)
5B	H_{3a}	Moderation by relevance	Targeting based on gender was viewed as fairer when justified by average physical differences, such as different nutritional needs (attenuation) than by differences in preferences (basic effect)
5C	H_{3a}	Moderation by relevance	Targeting based on gender was viewed as fairer when justified by objective data (attenuation) than when not justified by objective data (basic effect)
6	H_{3b}	Moderation by controllability	Targeting based on SES was viewed as fairer when belonging to the targeted segment was believed to be more controllable
7A	H_{3c}	Moderation by intentionality	Targeting based on gender was viewed as fairer when performed by a small company (attenuation) than when performed by a large company (basic effect)
7B	H_{3c}	Moderation by intentionality	Targeting based on gender was viewed as fairer when it was standard practice (attenuation) than when it was not standard practice (basic effect)
8A	H_1	Facebook A/B test	Women were less likely to click on a Facebook ad when the ad copy clearly communicated demographic targeting (e.g., "for women") than when it did not
8B	H_{3a}	Facebook A/B test	Women were less likely to click on a Facebook ad when the ad copy clearly communicated demographic targeting (e.g., "for women") than when it did not, but this difference attenuated when justified by average physical differences (e.g., nutritional needs)

Note: Failures of preregistered instructional manipulation checks (IMCs; Oppenheimer, Meyvis, and Davidenko 2009) were excluded prior to analysis (in some cases resulting in minor discrepancies between reported and preregistered sample sizes). Data, materials, and statistical code for reproducing analyses are publicly available (https://osf.io/3vksh/?view_only=301c4048005848928872ff2ff659955b).

STUDIES 1A–C: WHY AM I SEEING THIS AD?

Studies 1A–C test the basic effect (H_1), highlighting a common way that consumers learn about demographic targeting: "Why am I seeing this ad?" disclosures, which Facebook has offered since 2014 (Kozlowska 2018). We manipulated whether a social media disclosure revealed demographic targeting or broad advertising and measured perceptions of fairness (Study 1A), marketing consequences (Studies 1B), and consequential choice (Study 1C).

Study 1A Method and Results

We recruited 576 MTurk workers (M_{AGE} = 44.43; 235 women, 331 men, 8 other, 2 undisclosed) for Study 1A (aspredicted.org/45m4-xv89.pdf), which employed a single-factor (targeting: demographic vs. broad) between-subjects design. Participants read: "You're scrolling Facebook and come across the following ad. You click on the "Why am I seeing this ad?" button." We then displayed a screenshot of a Facebook ad promoting "a new hard seltzer" (Figure 2). Participants read they had been targeted "due to your gender, age, and/or race" or "as a member of the general public." Below the screenshot, we presented three counterbalanced fairness measures: "How [fair/appropriate/acceptable] is this advertising strategy?" ("Not at all [fair/appropriate/acceptable]" = 1; "Very [fair/appropriate/acceptable]" = 9).

We averaged the three fairness measures (α = .97). This composite was lower in the demographic targeting condition (M = 5.67, 95% CI = [5.39, 5.94]) than in the broad advertising condition (M = 6.85, 95% CI = [6.59, 7.12], $t(574) = 6.11$, $p < .001$, $d = .49$; Figure 3).

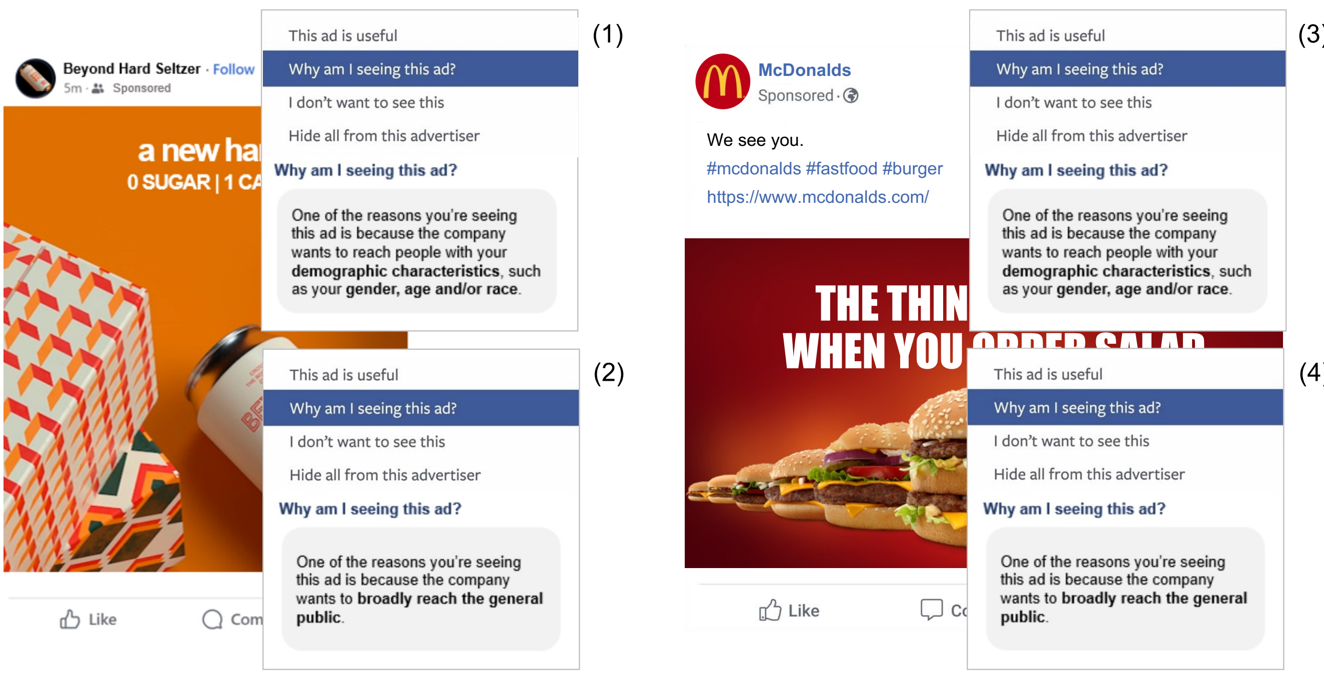
Study 1B Method and Results

We recruited 571 MTurk workers (M_{AGE} = 45.17; 288 women, 276 men, 7 other) for Study 1B, which was identical to Study 1A. However, instead of fairness, we measured three marketing consequences: “How does this advertising strategy affect your…” “willingness to purchase this product?” (“Definitely decreases” = 1; “Definitely increases” = 9); “impression of the company that makes this product?” (“Definitely worsens” = 1; “Definitely improves” = 9); “willingness to recommend this product?” (“Definitely reduces” = 1; “Definitely increases” = 9).

We averaged the three consequences measures ($\alpha = .90$). This composite was lower in the demographic targeting condition ($M = 4.18$, 95% CI = [3.98, 4.37]) than in the broad advertising condition ($M = 4.64$, 95% CI = [4.48, 4.81], $t(569) = 3.65$, $p < .001$, $d = .30$; Figure 3).

FIGURE 2

STUDIES 1A–B (LEFT PANEL): (1) DEMOGRAPHIC TARGETING AND (2) BROAD ADVERTISING. STUDY 1C (RIGHT PANEL): (3) DEMOGRAPHIC TARGETING AND (4) BROAD ADVERTISING



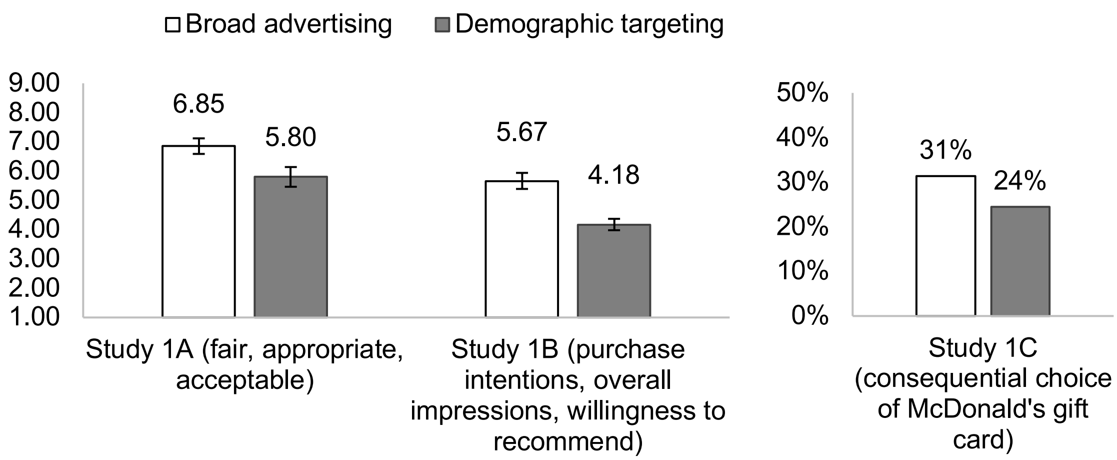
Study 1C Method and Results

We recruited 778 MTurk workers ($M_{AGE} = 45.26$; 412 women, 365 men, 1 other) for

Study 1C, which was identical to Study 1B, with two exceptions. First, we presented an ad for McDonald's (Figure 2). Second, after responding to the three marketing consequences measures, participants made a consequential, incentive-aligned choice (on a separate page): "At the conclusion of this study, we will randomly select one participant to receive a gift card of their choice. Which do you prefer?" ("$100 McDonald's gift card" vs. "$50 Amazon gift card").

FIGURE 3

STUDIES 1A–C: "WHY AM I SEEING THIS AD?" DISCLOSURES COMMUNICATING DEMOGRAPHIC TARGETING (VS. BROAD ADVERTISING) NEGATIVELY AFFECT FAIRNESS PERCEPTIONS, MARKETING CONSEQUENCES, AND CONSEQUENTIAL CHOICE (95% CONFIDENCE INTERVALS)



We averaged the three consequences measures ($\alpha = .94$). This composite was lower in the demographic targeting condition ($M = 4.17$, 95% CI = [4.00, 4.33]) than in the broad advertising condition ($M = 4.69$, 95% CI = [4.53, 4.85], $t(776) = 4.49$, $p < .001$, $d = .32$). Participants were also less likely to choose the McDonald's gift card over the Amazon gift card in the demographic condition (24.4%) than in the broad condition (31.3%; $\chi^2(1) = 4.47$, $p = .034$; Figure 3).

Studies 1A–C Discussion

Studies 1A–C tested a common way that consumers naturally learn about demographic targeting in the real world: social media disclosures. However, we made this information explicit to participants. A key question is whether consumers are sensitive to and pick up on subtle cues in ads that *implicitly* reveal which demographic groups the firm is targeting.

STUDIES 2A–B: INFERENCES FROM IMAGES

In Studies 2A–B, we manipulated only the image featured in otherwise identical ads. We expected participants to use the race or gender composition of those depicted to infer the underlying targeting strategy (Aaker et al. 2000), and that fairness perceptions would suffer when those images implied demographic targeting.

To confirm the images did not otherwise meaningfully differ across conditions, we conducted a pretest. Each participant viewed one image from Studies 2A–B and answered three questions: "How [attractive/appealing/interesting] is this picture?"). Average ratings in the broad advertising condition (M = 6.58, 95% CI = [6.05, 7.11]) did not differ from the race condition (M = 6.81, 95% CI = [6.28, 7.33]; $t(87)$ = .66, p = .510, d = .14) nor the gender condition (M = 6.36, 95% CI = [5.84, 6.88]; $t(88)$ = .51, p = .608, d = .11; see Web Appendix Study WA2).

Study 2A Method and Results

We recruited 397 MTurk workers (M_{AGE} = 41.51; 168 women, 224 men, 5 other; 308 White, 42 Black, 61 other)[1] for Study 2A, which employed a single-factor (targeting: race vs. broad) between-subjects design. All participants first read: "Please take a moment to consider the

---

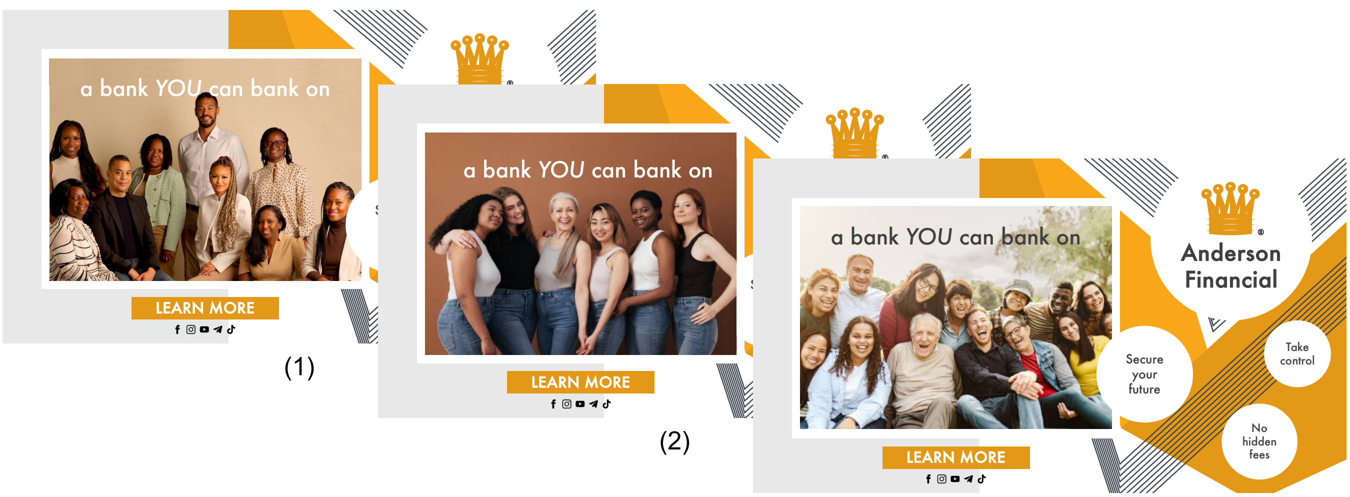
[1] We report participant race where available. Counts for race do not sum to 100% of N because participants could select multiple races. The "other" categories included: American Indian or Alaska Native, Asian, Hispanic or Latino, Middle Eastern or North African, Native Hawaiian or Pacific Islander, Other, and Prefer Not to Say.

below advertisement." All conditions featured the slogan "a bank YOU can bank on." The image in the race condition depicted people of a single race (Figure 4), while the image in the broad condition depicted people of multiple races. Beneath the ad, participants responded to the same fairness, appropriateness, and acceptability questions as in Study 1A (counterbalanced).

We averaged the three fairness measures ($\alpha = .96$). This composite was lower in the race targeting condition ($M = 6.49$, 95% CI = [6.18, 6.81]) that in the broad advertising condition ($M = 7.57$, 95% CI = [7.34, 7.79], $t(395) = 5.48$, $p < .001$, $d = .53$).

FIGURE 4

STUDIES 2A–B: (1) RACE TARGETING, (2) GENDER TARGETING, AND (3) BROAD ADVERTISING



(1) (2) (3)

Study 2B Method and Results

We recruited 401 MTurk workers ($M_{AGE} = 44.31$; 221 women, 174 men, 6 other; 327 White, 39 Black, 58 other) for Study 2B, which employed a single-factor (targeting: gender vs. broad) between-subjects design. Study 2B was identical to Study 2A, with one exception. We replaced the race condition with a gender condition that featured an image depicting people of a single gender (Figure 4).

We averaged the three fairness measures ($\alpha = .95$). This composite was lower in the gender targeting condition ($M = 6.28$, 95% CI = [5.99, 6.57]) than in the broad advertising condition ($M = 7.67$, 95% CI = [7.45, 7.89], $t(399) = 7.55$, $p < .001$, $d = .70$).

Studies 2A–B Discussion

Studies 2A–B reveal that consumers react negatively to *inferences* about demographic targeting, even when this information is not made explicit. In other words, consumers try to infer the firm's targeting strategy from subtle cues in the ads they encounter (Aaker et al. 2000). Notably, Studies 1A–2B employed diverse samples, suggesting consumers need not belong to a particular segment for unfairness perceptions to result from inferences about demographic targeting. Yet it is unclear whether the effect holds among *only* members of the targeted group—or if, alternatively, negative reactions are largely driven by outside observers.

STUDIES 3A–B: RECRUITING MEMBERS OF THE TARGETED SEGMENT

We recruited only members of the targeted segment as participants in Studies 3A–B (e.g., Black participants in Study 3A, women in Study 3B). We also tested a third, real-world source of information about demographic targeting: news coverage. We did this by adapting real news stories describing the actual targeting strategies of two well-known companies (e.g., Toyota in Study 3A, PepsiCo in Study 3B). We varied the advertising channels used in promotion, describing media consumed primarily by the targeted segment or broadly by the general public.

Study 3A Method

We recruited 490 prescreened Prolific users who identified as "Black/African American"

(M_{AGE} = 40.95; 316 women, 174 men; 2 White, 477 Black, 11 other)[2] for Study 3A (aspredicted.org/qyvr-yqqc.pdf), which employed a single-factor (targeting: race vs. broad) between-subjects design. All participants reviewed a fictional newspaper article that we adapted from a real *New York Times* story ("Different Ads, Different Ethnicities, Same Car"; Maheshwari 2017 [bit.ly/3CCY10Z]). The articles differed between conditions only with respect to the advertising channels Toyota would use to promote its "next-generation Camry" (Figure 5).

In the race condition, the article described television commercials airing during "Black-ish," "Scandal," and NBA games, as well as radio spots on local hip-hop and R&B stations. We mentioned "Scandal" and hip-hop music, in particular, because these media were specifically cited as examples of how Toyota designed the Camry's "Strut" campaign to appeal to Black consumers (Maheshwari 2017). In the broad condition, the article described television commercials airing during "Survivor," "The Price is Right," and the Olympics, as well as radio spots on local NPR and pop/rock stations.

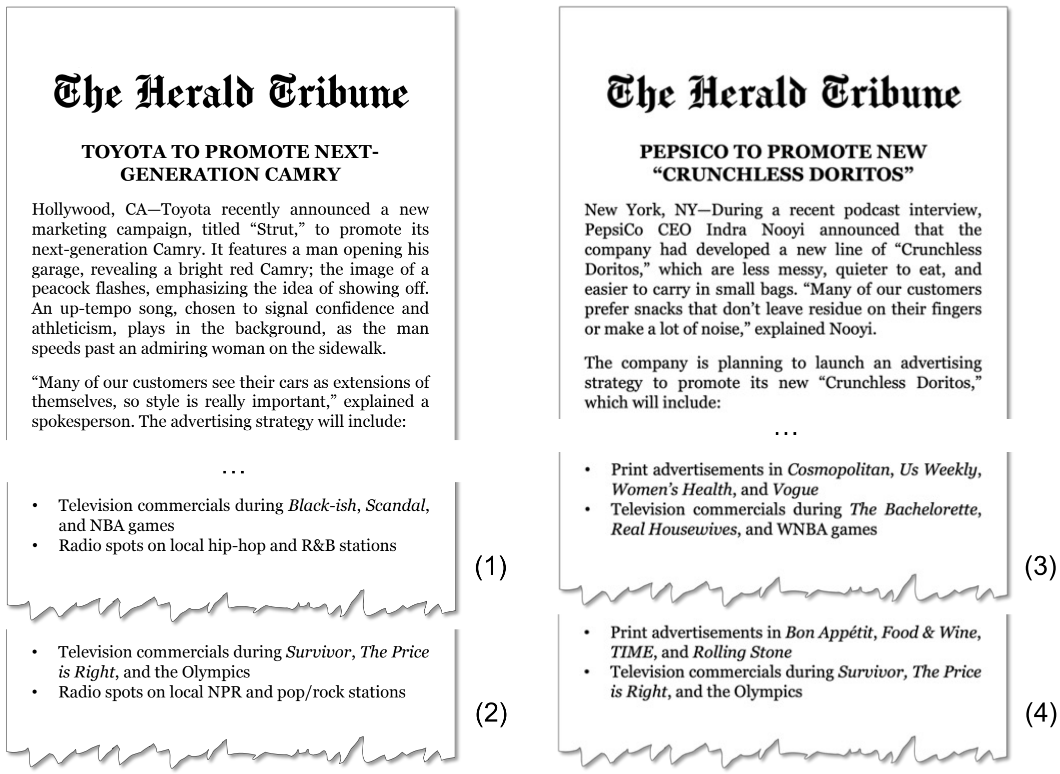
On the first page, we presented five counterbalanced questions below the article: "Do you believe the advertising strategy described above is targeted at the general public or targeted at a specific [race/gender/age group/socioeconomic status/geographic region]?" ("Definitely targeted at the general public" = 1; "Definitely targeted at a specific [race/gender/age group/socioeconomic status/geographic region]" = 7). The question about race served as the manipulation check, but we embedded it among the four other questions asking about gender, age group, SES, and geography to obscure our intentions and reduce the potential for demand effects. On the second page, we presented the same fairness, appropriateness, and acceptability measures as in Study 1A (counterbalanced) directly below the article.

[2] We limited eligibility to "Black/African American," but two participants self-reported only "White" (participants were free to select multiple races). Removing these observations does not qualitatively change the Study 3A results.

FIGURE 5

STUDY 3A (LEFT PANEL): (1) RACE AND (2) BROAD CONDITIONS.

STUDY 3B (RIGHT PANEL): (3) GENDER AND (4) BROAD CONDITIONS



The Herald Tribune

TOYOTA TO PROMOTE NEXT-GENERATION CAMRY

Hollywood, CA—Toyota recently announced a new marketing campaign, titled “Strut,” to promote its next-generation Camry. It features a man opening his garage, revealing a bright red Camry; the image of a peacock flashes, emphasizing the idea of showing off. An up-tempo song, chosen to signal confidence and athleticism, plays in the background, as the man speeds past an admiring woman on the sidewalk.

“Many of our customers see their cars as extensions of themselves, so style is really important,” explained a spokesperson. The advertising strategy will include:

…

- Television commercials during *Black-ish*, *Scandal*, and NBA games
- Radio spots on local hip-hop and R&B stations

(1)

- Television commercials during *Survivor*, *The Price is Right*, and the Olympics
- Radio spots on local NPR and pop/rock stations

(2)

The Herald Tribune

PEPSICO TO PROMOTE NEW “CRUNCHLESS DORITOS”

New York, NY—During a recent podcast interview, PepsiCo CEO Indra Nooyi announced that the company had developed a new line of “Crunchless Doritos,” which are less messy, quieter to eat, and easier to carry in small bags. “Many of our customers prefer snacks that don’t leave residue on their fingers or make a lot of noise,” explained Nooyi.

The company is planning to launch an advertising strategy to promote its new “Crunchless Doritos,” which will include:

…

- Print advertisements in *Cosmopolitan*, *Us Weekly*, *Women’s Health*, and *Vogue*
- Television commercials during *The Bachelorette*, *Real Housewives*, and WNBA games

(3)

- Print advertisements in *Bon Appétit*, *Food & Wine*, *TIME*, and *Rolling Stone*
- Television commercials during *Survivor, The Price is Right*, and the Olympics

(4)

Study 3A Results

Confirming the manipulation, participants believed the campaign was designed to target a specific race more in the race condition (M = 5.33, 95% CI = [5.07, 5.59]) than in the broad condition (M = 2.64, 95% CI = [2.42, 2.86]; $t(488) = 15.59$, $p < .001$, $d = 1.15$). This difference (e.g., inferences about race targeting) was the largest among all five manipulation checks (e.g., compared to gender, age, SES, and geography). The average of the three fairness measures (α = .97) was lower in the race targeting condition (M = 6.06, 95% CI = [5.80, 6.32]) than in the broad advertising condition (M = 7.11, 95% CI = [6.90, 7.33], $t(488) = 6.25$, $p < .001$, $d = .54$).

Study 3B Method

We recruited 495 prescreened Prolific users who identified as "Woman (including Trans Female/Trans Woman)" (M_{AGE} = 43.87; 479 women, 6 men, 10 other; 357 White, 87 Black, 92 other race)[3] for Study 3B (aspredicted.org/44dt-syr6.pdf), which employed a single-factor (targeting: gender vs. broad) between-subjects design. All participants reviewed a fictional newspaper article that we adapted from a real *New York Times* story ("Lady Doritos? Pepsi Wants a Do-Over"; LaForge 2018 [bit.ly/40YOdHX]). As in Study 3A, the articles differed between conditions only with respect to the advertising channels PepsiCo would use to promote its "Crunchless Doritos" (Figure 5).

In the gender condition, the article described print advertisements in publications like *Cosmopolitan* and *Women's Health*, as well as television commercials airing during "The Bachelorette" and WNBA games (i.e., media primarily consumed by women). In the broad condition, the article described print advertisements in publications like *TIME* and *Rolling Stone*, as well as television commercials airing during "Survivor" and the Olympics (i.e., media generally consumed by all genders; broad condition). The manipulation check (first page) and fairness measures (second page) were identical to Study 3A.

Study 3B Results

Confirming the manipulation, participants believed the campaign targeted a specific gender more in the gender condition (M = 5.00, 95% CI = [4.71, 5.29]) than in the broad condition (M = 2.10, 95% CI = [1.92, 2.28]; $t(493)$ = 16.82, p < .001, d = 1.21). This difference (e.g., inferences about gender targeting) was the largest among all five manipulation checks (e.g., compared to race, age, SES, and geography). We next averaged the three fairness measures (α = .98). This composite was lower in the gender condition (M = 6.33, 95% CI = [6.09, 6.58]) than in

[3] We limited eligibility to "Woman (including Trans Female/Trans Woman)," but six participants self-reported "male." Removing these observations does not qualitatively change the Study 3B results.

the broad condition (*M* = 7.39, 95% CI = [7.18, 7.60], $t(493) = 6.43$, $p < .001$, $d = .55$).

Studies 3A–B Discussion

Studies 1A–3B offered evidence for the basic effect (H_1), testing various naturalistic stimuli (e.g., social media disclosures, realistic ads, news coverage), measuring naturalistic behaviors (e.g., purchase intentions, consequential choice), and drawing from naturalistic samples (e.g., recruiting participants from the targeted segment). We next turn to our broader theoretical framework. To test each element of our model, we use experimental designs in which we describe firms' targeting strategies to participants and measure their fairness perceptions, offering process evidence through both mediation (Study 4) and moderation (Studies 5A–7B).

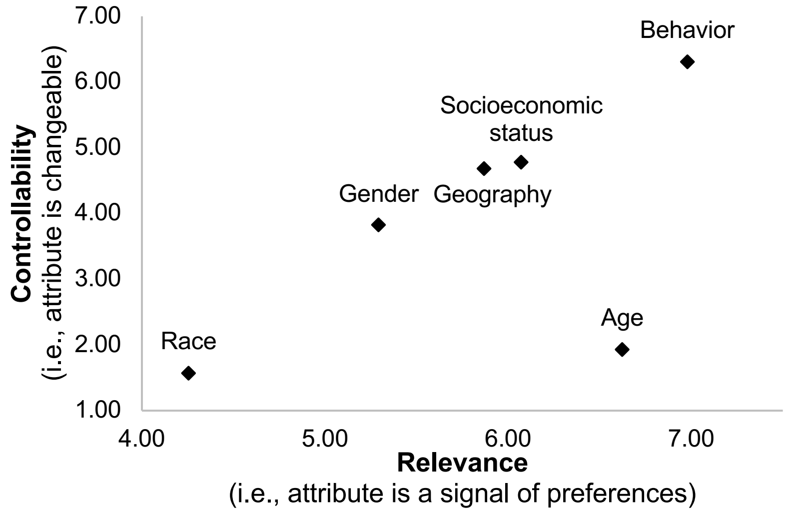
STUDY 4: MEDIATION BY BELIEFS ABOUT DISCRIMINATION

Study 4 not only tests for mediation through appraisals of discrimination (H_2), but also compares five common forms of demographic targeting (e.g., race, gender, age, SES, geography) against two baselines for comparison (e.g., broad advertising and behavioral targeting).

We define discrimination as differential treatment based on attributes that are irrelevant and/or uncontrollable. To confirm that beliefs about relevance and controllability vary across demographic attributes, we conducted a pretest. Participants were randomly assigned to one of six targeting bases and rated its relevance (e.g., "a signal of their preferences" and "relevant to most products and services") and controllability (e.g., "under their personal control" and "easily changeable"; adapted from Tomova Shakur and Phillips 2022; see Web Appendix Study WA3). Behavior was rated as the most controllable and relevant, race the least (Figure 6).

FIGURE 6

STUDY 4 PRETEST: RELEVANCE AND CONTROLLABILITY ACROSS VARIOUS TARGETING BASES



Study 4 Method

We recruited 1,388 Prolific users (M_{AGE} = 40.34; 702 women, 662 men, 24 other) for Study 4 (aspredicted.org/n6pq-dqbp.pdf), which employed a single-factor (targeting: race vs. gender vs. age vs. socioeconomic status vs. geography vs. behavioral vs. broad) between-subjects design.

All participants first read: "A company has developed a new product, which they believe will appeal to some customers more than others." The broad condition read: "They plan to advertise the product broadly to the general public." The targeting conditions read: "They plan to advertise the product to [members of a particular race/people of a particular gender/members of a particular age group/people of a particular socioeconomic status/people who live in a particular geographic region/people who have purchased similar products in the past], rather than broadly to the general public." We asked: "How fair is this plan?" ("Not at all fair" = 1; "Very fair" = 9).

On the next page, participants submitted open-ended explanations: "We are interested in understanding what you were thinking on the previous page." In the broad condition, we asked, "what specifically about advertising broadly to the general public makes it feel fair or unfair?" In the targeting conditions, we asked, "what specifically about advertising based on [race/gender/age/socioeconomic status/geography/similar past purchases] makes it feel fair or unfair?" We instructed all participants to "write down any thoughts or descriptions that came to mind" and to "list any phrase(s) or word(s)." They were free to submit up to five thought-listings in each of five open text fields.

Study 4 Results

A fairness ANOVA revealed a main effect of targeting ($F(6, 1{,}381) = 45.30$, $p < .001$). Each of the five demographic targeting conditions was rated as less fair than broad advertising; the behavioral targeting and broad advertising conditions did not differ (Table 2).

We preregistered two approaches for coding the open-ended explanations. For our primary preregistered analysis, we asked human coders to rate the extent to which each thought-listing invoked the concept of discrimination. To do this, we recruited a separate sample of 696 Prolific users ($M_{AGE} = 39.69$; 397 women, 291 men, 8 other), who read:

> "In a previous survey, we asked people to list their thoughts regarding whether a company's advertising strategy seemed fair or unfair. To explain their reasoning, they wrote down up to five descriptions, phrases, or words that came to mind. We are now interested in learning whether people brought up the concept of discrimination when listing their thoughts. We define discrimination as differential treatment based on attributes that are irrelevant and/or uncontrollable."

We told the raters that they would rate 20 thought listings, each on a separate page: "We'd like you to indicate, for each, whether the thought-listing invokes the concept of discrimination."

For every thought-listing from Study 4, we asked: "Does this thought-listing invoke the concept of discrimination?" ("Not at all" = 1; "A little bit" = 2; "Somewhat" = 3; "Very much so" = 4; "This thought-listing seems nonsensical or nonresponsive" = "N/A"). On each page, we also reminded participants of our definition of discrimination. To account for heterogeneity in each rater's use of the scale, we *z*-scored their discrimination ratings (i.e., for each rating, we subtracted the average of that rater's 20 ratings and divided by the corresponding standard deviation). Then, for each thought-listing, we averaged all the *z*-scored discrimination ratings. Each thought-listing garnered an average of 9.51 discrimination ratings.

For our secondary preregistered analysis, we asked generative AI to rate the extent to which each thought-listing from Study 4 invoked the concept of discrimination (see Web Appendix Study 4 Secondary Analysis Methodological Details).[4]

TABLE 2

STUDY 4: FAIRNESS AND DISCRIMINATION ACROSS VARIOUS TARGETING BASES

Targeting basis	Fairness		Targeting basis	Discrimination ratings: human coders		Discrimination ratings: GPT	
Race	3.71 [3.46, 3.96]		Race	+.621 [+.530, +.713]		4.54 [4.32, 4.76]	
SES	4.64 [4.39, 4.89]	A	SES	+.235 [+.142, +.329]	A	3.62 [3.39, 3.85]	A
Gender	4.92 [4.68, 5.15]	AB	Gender	+.186 [+.101, +.270]	A	3.62 [3.41, 3.82]	A
Geography	5.12 [4.90, 5.34]	B	Age	–.003 [–.088, +.082]	B	3.23 [3.03, 3.42]	B
Age	5.17 [4.97, 5.38]	B	Geography	–.099 [–.186, –.012]	B	3.03 [2.82, 3.24]	B
Behavioral	5.80 [5.65, 5.95]	C	Broad advertising	–.507 [–.564, –.449]	C	2.33 [2.14, 2.52]	C
Broad advertising	5.88 [5.69, 6.06]	C	Behavioral	–.562 [–.614, –.509]	C	2.26 [2.09, 2.42]	C

Note: Comparisons sharing a letter do not differ at p = .05. Brackets contain 95% CIs. Mediation results reflect bootstrapped CIs (10,000 resamples) for the indirect effect of targeting on fairness through discrimination. Conditions sorted from least to most fair, and most to least discriminatory

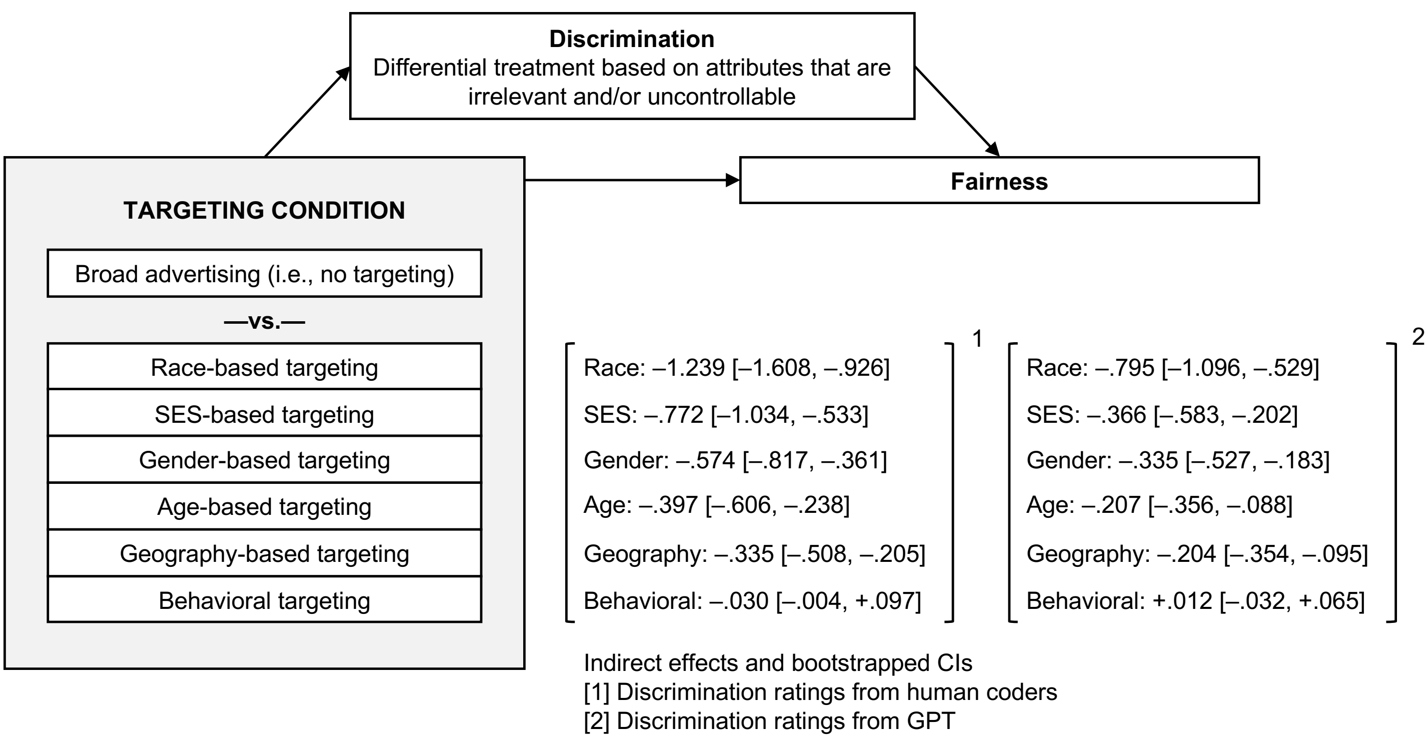
We submitted the human coders' discrimination ratings to an ANOVA, which revealed a main effect of targeting (F(6, 1,377) = 106.42, p < .001). The open-ended explanations invoked

[4] While we aimed to follow best practices for the use of large language models in behavioral research (e.g., preregistration, open materials, model specification, validation against human data; Abdurahman et al. 2025), there are well documented limitations of generative AI as a substitute for human judgment (e.g., Brucks and Toubia 2025; Bojić et al. 2025; Pavlovic 2024).

the concept of discrimination more in each of the five demographic targeting conditions than in the broad advertising condition; the behavioral targeting and broad advertising conditions did not differ (Table 2). We also submitted the AI-generated discrimination ratings to the same ANOVA, which revealed the same main effect of targeting ($F(6, 1{,}375) = 60.13$, $p < .001$; Table 2), along with same the same rank-ordering of demographic targeting conditions (i.e., from most- to least-discriminatory). The discrimination ratings supplied by our human coders and generative AI were highly correlated ($r = .707$, $p < .001$; see General Discussion for implications regarding the use of generative AI for coding unstructured text).

FIGURE 7

STUDY 4: MEDIATION MODEL



Finally, as preregistered, we tested whether the human coders' discrimination ratings mediated the differences in fairness between broad advertising and each of the six targeting bases.[5] Specifically, we fit six separate mediation models, with 10,000 bootstrapped resamples (i.e., one per targeting basis), comparing each to broad advertising. Every model included fairness as the dependent variable, discrimination rating as the mediator, and a dummy-coded independent variable (0 = broad advertising; 1 = targeting condition). Discrimination ratings mediated the effect for each of the five demographic targeting conditions, but not for the behavioral targeting condition (Figure 7). Though we did not preregister the same mediation analysis with the AI-generated discrimination ratings, it yielded similar overall results.

Study 4 Discussion

Study 4 reveals that consumers view various forms of demographic targeting as less fair than broad advertising, and that beliefs about discrimination play a mediating role.[6] Moreover, consistent with our conceptualization, the degree to which discrimination ratings mediated differences in fairness perceptions varied across demographic categories. As revealed by the pretest, some demographic attributes feel more relevant and controllable than others. For example, race was rated as least controllable and least relevant (Figure 6). And targeting based on race was rated as least fair and most discriminatory. Accordingly, across the six targeting bases, we observed a negative correlation between relevance and controllability (averaged from the pretest) and discrimination (from Study 4; $r = -.87$, $p = .023$).

Additionally, neither fairness nor discrimination differed between behavioral targeting and broad advertising, further underscoring the role of relevance and controllability to our

[5] The modest correlation between fairness perceptions and the human coders' discrimination ratings ($r = -.53$) helps distinguish fairness and discrimination as unique and independent constructs. For example, Voorhees et al. (2015) suggest that correlations under $r = .60$ are sufficient for establishing discriminant validity.

[6] We also conducted a preregistered conceptual replication of this Study 4 mediation result with a different measure of discrimination (i.e., a seven-point scale, embedded among six other decoy measures) and observed qualitatively similar results (see Web Appendix WA4). The rank-ordering of fairness and discrimination ratings across targeting bases in Web Appendix WA4 were also nearly identical to Study 4.

conceptual definition of discrimination. Behavioral targeting, like any other form of targeting, necessarily involves differential treatment. However, as confirmed by the pretest, behaviors are viewed as more controllable and relevant than demographic characteristics (Figure 6). Thus, behavioral targeting is viewed more favorably than demographic targeting, because differential treatment only undermines fairness and brand support when it is based on attributes that are irrelevant and/or uncontrollable—that is, when it is viewed as discriminatory.

With evidence for mediation consistent with our account, we next turn to a process-by-moderation approach (Spencer, Zanna, and Fong 2005). Specifically, we manipulate factors that we hypothesize either improve relevance or increase controllability—and thus, according to our model, should attenuate perceptions of unfairness in demographic targeting.

STUDIES 5A–C: MODERATION BY RELEVANCE

Two factors that we expect to improve relevance (H_{3a}) are (a) justification based on average physical differences across demographic groups (e.g., skin tone or nutritional needs; Studies 5A–B), and (b) justification based on objective data (Study 5C).

To confirm these moderators shape beliefs about relevance as theorized, we conducted a pretest. Participants were randomly assigned to one of the justification conditions from Studies 5A–C. However, rather than rating fairness, they rated relevance and controllability (as in the Study 4 Pretest; see Web Appendix Study WA5). Relevance was higher in the justification by average physical differences conditions (5A: $M_{\text{AVG. PHYS. DIFF.}} = 7.31$ vs. $M_{\text{PREFERENCES}} = 6.42$; $F(1, 271) = 19.80$, $p < .001$, $d = .53$; 5B: $M_{\text{AVG. PHYS. DIFF.}} = 6.65$ vs. $M_{\text{PREFERENCES}} = 5.92$; $F(1, 359) = 15.67$, $p < .001$, $d = .41$) and the justification based on objective data condition (5C:

$M_{\text{OBJECTIVE DATA}} = 6.57$ vs. $M_{\text{NO OBJECTIVE DATA}} = 5.92$; $F(1, 358) = 11.68$, $p < .001$, $d = .36$). Controllability did not differ (ps > .262).

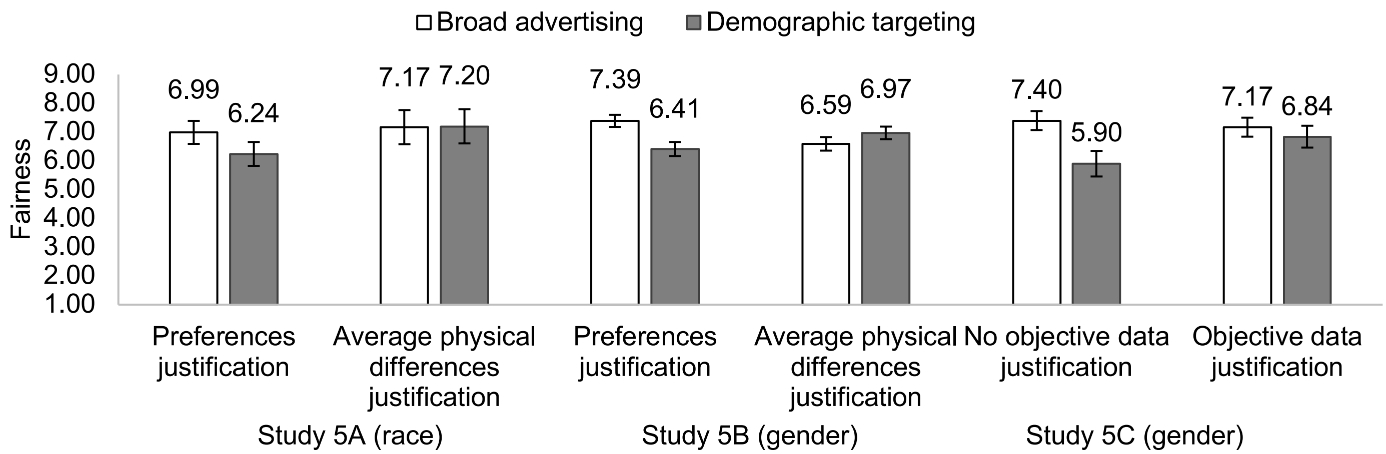
Study 5A Method and Results

The Study 5A context was inspired by recent events. Specifically, in 2020, the Band-Aid brand launched a diverse range of colors to "embrace the beauty of diverse skin," fulfilling a long-time unmet marketplace need (Alcorn 2020; D'Angelo et al. 2024). We hypothesized that justifying the decision to target Black consumers based on average physical differences in skin tone would be perceived as more relevant (see pretest) and thus fairer, because skin tone often varies by race (and skin tone should affect preferences for bandage colors). We recruited 399 MTurk workers ($M_{\text{AGE}} = 39.69$; 253 women, 141 men, 5 other; 329 White, 40 Black, 49 other) for Study 5A, which employed a 2 (targeting: race vs. broad) × (justification: average physical differences vs. preferences) between-subjects design.

Participants first read: "A company has developed a new line of Band-Aids that were designed to appeal to their Black customers." The preferences justification did not mention skin tone: "the Band-Aids are available in a variety of patterns and designs that were tested to appeal to the preferences of their Black customers." To heighten relevance, the average physical differences justification made an explicit connection between skin tone and race: "the Band-Aids are available in a variety of darker shades that were tested to match the skin tones of their Black customers." In the race condition, participants read: "Because they believe the products are best suited to their Black customers, they will advertise them to Black people directly." In the broad condition, participants read: "Even though they believe the products are best suited to their Black customers, they will advertise them broadly to the general public." Participants then rated fairness: "How fair is this advertising plan?" ("Not at all fair" = 1; "Very fair" = 9).

FIGURE 8

STUDIES 5A–B: DEMOGRAPHIC TARGETING IS VIEWED AS LESS UNFAIR (RELATIVE TO BROAD ADVERTISING) WHEN JUSTIFIED BY AVERAGE PHYSICAL DIFFERENCES AND OBJECTIVE DATA (95% CONFIDENCE INTERVALS)



A fairness ANOVA revealed a marginal main effect of targeting ($F(1, 395) = 3.57$, $p = .060$), qualified by a marginal interaction ($F(1, 395) = 3.06$, $p = .081$; Figure 8). There was a simple effect of targeting for the preferences justification ($F(1, 395) = 6.64$, $p = .010$), such that fairness perceptions were lower in the race targeting condition ($M = 6.24$, 95% CI = [5.82, 6.65]) than in the broad advertising condition ($M = 6.99$, 95% CI = [6.59, 7.39], $d = .36$). However, this simple effect of targeting disappeared for the average physical differences justification ($F(1, 395) = 0.01$, $p = .921$; $M_{\text{RACE}} = 7.20$, 95% CI = [6.78, 5.61] vs. $M_{\text{BROAD}} = 7.17$, 95% CI = [6.76, 5.57], $d = .01$). Within just the race targeting condition, the average physical differences justification (vs. preferences justification) improved fairness perceptions ($F(1, 395) = 10.10$, $p = .002$).

Study 5B Method and Results

Study 5B serves as a conceptual replication of Study 5A, but focuses instead on average physical differences across genders. We hypothesized that justifying the decision to target female consumers based on average physical differences in vitamin and nutrient needs would be perceived as more relevant (see pretest) and thus fairer, because different genders have different biological requirements for certain vitamins and nutrients. We recruited 1,047 MTurk workers (M_{AGE} = 36.88; 497 women, 499 men, 51 other; 743 White, 77 Black, 245 other) for Study 5B (https://aspredicted.org/NJG_5H1), which employed a 2 (targeting: gender vs. broad) × (justification: average physical differences vs. preferences) between-subjects design.

Participants first read that a company had "developed a new line of snacks" and was "developing a marketing plan." The preferences justification did not mention biological needs: "Initial testing showed that, due to the taste and texture profile of the snacks, the snacks are better suited to the preferences of their female customers." To heighten relevance, the average physical differences justification made an explicit connection between biological needs and gender: "Initial testing showed that, due to the vitamin and nutrient profile of the snacks, the snacks are better suited to the biological needs of their female customers." In the gender condition, participants read: "Because they believe the snacks are best suited to their female customers, they will advertise the snacks to women directly." In the broad condition, participants read: "Even though they believe the snacks are best suited to their female customers, they will advertise the snacks broadly to the general public." Participants then rated fairness: "How fair is this advertising plan?" ("Not at all fair" = 1; "Very fair" = 9).

A fairness ANOVA revealed a main effect of targeting (F(1, 1,043) = 6.65, p = .010), qualified by an interaction (F(1, 1,043) = 33.30, p < .001; Figure 8). There was a simple effect of targeting for the preferences justification (F(1, 1,043) = 34.86, p < .001, d = .54), such that fairness perceptions were lower in the gender targeting condition (M = 6.41, 95% CI = [6.16, 6.65]) than in the broad advertising condition (M = 7.39, 95% CI = [7.18, 7.60]). This simple effect of targeting reversed for the average physical differences justification ($F(1, 1{,}043) = 5.09$, $p = .024$; $M_{\text{GENDER}} = 6.97$, 95% CI = [6.75, 7.19] vs. $M_{\text{BROAD}} = 6.59$, 95% CI = [6.36, 6.83], $d = .22$). Within just the gender targeting condition, the average physical differences justification (vs. preference justification) improved fairness perceptions ($F(1, 1043) = 11.42$, $p < .001$).

Study 5C Method and Results

Study 5C complements Studies 5A–B by highlighting another way to justify demographic targeting. We hypothesized justification based on objective data would be perceived as more relevant (see pretest) and thus fairer, because objective data has been shown to increase the credibility of marketing claims (Ford et al. 1990). We recruited 401 MTurk workers ($M_{\text{AGE}} = 40.90$; 233 women, 163 men, 6 other; 328 White, 50 Black, 41 other) for Study 5C (aspredicted.org/W29_3B1), which employed a 2 (targeting: gender vs. broad) × (objective data justification: present vs. absent) between-subjects design.[7]

All participants first read: "A snack foods company has developed a new line of snacks. Initial testing showed that, due to the taste and texture profile of the snacks, the snacks are better suited to the preferences of their female customers." The objective data absent condition presented no other information. To heighten relevance, the objective data present condition read: "Specifically, testing shows that approximately 80% of women like the taste and texture, compared with only 30% of men." In the broad condition, participants next read: "Even though they believe the snacks are best suited to their female customers, they will advertise the snacks broadly to the general public." In the gender condition, participants next read: "Because they

[7] In the Web Appendix, we report three follow-up studies (Studies WA6–8), in which we systematically varied the interest level described by objective data. Consistent with both our theorizing and the results of Study 5C, demographic targeting was viewed as relatively fairer when there were large and discernable differences in reported preferences across segments. These results suggest consumers are sensitive to the strength of the evidence for relevance provided by objective data.

believe the snacks are best suited to their female customers, they will advertise the snacks to women directly." Finally, participants rated fairness: "How fair is this advertising plan?" ("Not at all fair" = 1; "Very fair" = 9).

A fairness ANOVA revealed a main effect of targeting ($F(1, 397) = 24.19$, $p < .001$), qualified by an interaction ($F(1, 397) = 9.37$, $p = .002$; Figure 8). There was a simple effect of targeting in the absence of objective data ($F(1, 397) = 31.75$, $p < .001$, $d = .77$), such that fairness perceptions were lower in the gender targeting condition ($M = 5.90$, 95% CI = [5.45, 6.34]) than in the broad advertising condition ($M = 7.40$, 95% CI = [7.07, 7.73]). However, this simple effect of targeting attenuated in the presence of objective data ($F(1, 397) = 1.73$, $p = .189$; $M_{\text{GENDER}} = 6.84$, 95% CI = [6.46, 7.22] vs. $M_{\text{BROAD}} = 7.17$, 95% CI = [6.84, 7.50], $d = .17$). Within just the gender targeting condition, the presence (vs. absence) of an objective data justification improved fairness perceptions ($F(1, 397) = 12.29$, $p < .001$).

Studies 5A–C Discussion

As noted in the introduction, targeting is useful for firms, allowing them to more efficiently allocate limited resources and increase the likelihood of reaching an interested customer. But Studies 1A–4 reveal that any potential positive effect of *demographic* targeting, in particular, may be offset by the negative effect of triggering appraisals of discrimination. Firms may therefore benefit from clearly *justifying* their targeting decisions above and beyond simply assuming or claiming different demographic groups maintain different preferences. Importantly, Studies 5A–C show how the credibility of these claims—and thus the extent to which demographic attributes will seem more diagnostic of preferences—can be by bolstered by justifications based on average physical differences (Studies 5A–B) and objective data (Study 5C). Consequently, demographic targeting is viewed more positively when it seems to help match consumers to relevant products and services. Study 6 next tests for moderation by the second factor in our theoretical framework: controllability.

STUDY 6: MODERATION BY CONTROLLABILITY

In Study 6, we leverage naturally occurring heterogeneity in the perceived controllability of a particular demographic attribute: socioeconomic status (SES). Past work has found that some people believe differences in SES are attributable to controllable factors like skill, merit, and effort; others point to uncontrollable forces like institutions, norms, and policies (Bullock et al. 2003; Davidai 2018; Dolifka et al. 2024). Our model implies, therefore, that individual differences in beliefs about the controllability of SES should moderate beliefs about whether SES-based targeting is fair (H_{3b}).

Study 6 Method

We recruited 975 MTurk workers (M_{AGE} = 38.49; 475 women, 483 men, 17 other) for Study 6 (aspredicted.org/H3C_9L9), which employed a single factor (targeting: SES vs. broad) between-subjects design. Participants first read: “A company has developed a new product, which they believe will appeal to some customers more than others.” Those in the SES condition next read: “They plan to advertise the product to people of a particular socioeconomic status, rather than broadly to the general public.” Those in the broad condition next read: “They plan to advertise the product broadly to the general public, rather than to a particular group.” On the same page, participants answered: “How fair is this plan?” (“Not at all fair” = 1; “Very fair” = 9). On the following page, participants answered: “How much control do people have over their socioeconomic status?” (“No control at all” = 1; “A lot of control” = 9). As control measures, participants also answered: “How would you characterize your own socioeconomic status?” (“Lower (working) class” =1; “Middle class” = 5; “Upper class” = 9) and “How would you characterize your political beliefs?” (“Very conservative” = 1; “Very liberal” = 9).

Study 6 Results

A linear regression of fairness on targeting (contrast-coded: –1 = broad advertising; 1 = SES targeting), perceived control over SES (mean-centered), and their interaction revealed a main effect of targeting ($B = –.88$, $SE = .06$, $t(971) = 14.25$, $p < .001$, $d = .82$), such that fairness perceptions were lower in the SES targeting condition ($M = 5.91$, 95% CI = [5.71, 6.12]) than in the broad advertising condition ($M = 7.70$, 95% CI = [7.56, 7.84]). However, this main effect was qualified by an interaction ($B = .18$, $SE = .04$, $t(971) = 4.92$, $p < .001$), such that perceived control over SES was more strongly associated with higher fairness in the SES targeting condition than in the broad advertising condition.

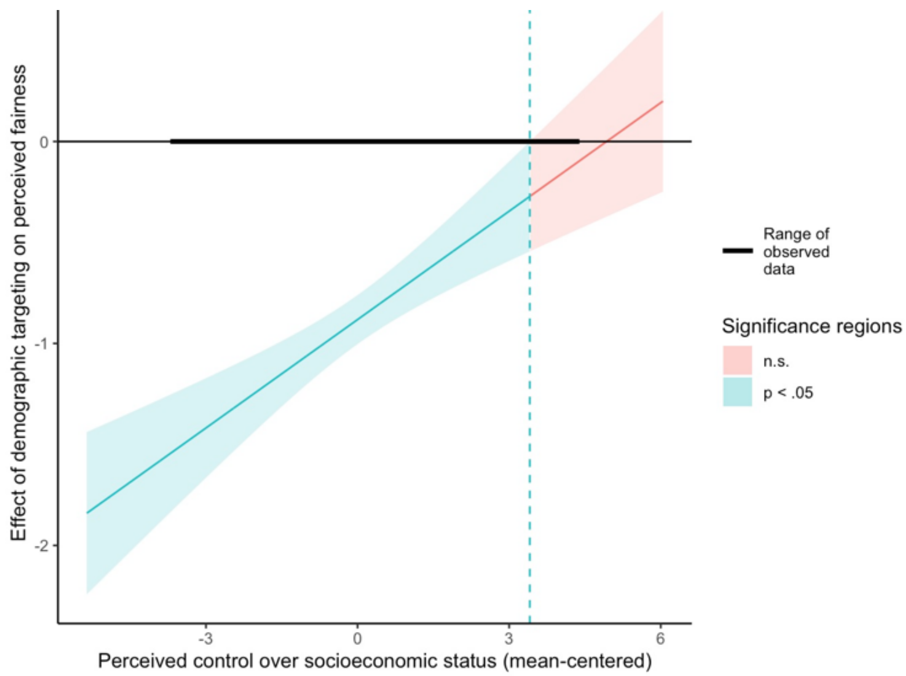
We used the Johnson–Neyman technique to identify the range(s) of perceived control over SES for which the simple effect of targeting was significant (Spiller et al. 2013). This analysis revealed a significant negative effect of targeting on fairness for perceived control over SES less than 3.41 (mean-centered; Figure 9). This interaction persisted ($B = .18$, $SE = .04$, $t(969) = 5.02$, $p < .001$) after controlling for self-reported SES and political beliefs.

Study 6 Discussion

Study 6 offers support for the controllability dimension of our framework. When consumers believe they have greater agency to determine whether they belong to the targeted demographic group, it feels more acceptable to target that group. Studies 7A–B test a third and final component of our model: perceived intentionality. Our account suggests that even when demographic targeting satisfies our definition of discrimination (based on relevance and controllability), there may be situations in which the discrimination is perceived as less intentional, and thus less unfair.

FIGURE 9

STUDY 6: DEMOGRAPHIC TARGETING IS VIEWED AS LESS UNFAIR WHEN PERCEIVED CONTROL OVER MEMBERSHIP IN THE DEMOGRAPHIC CATEGORY IS HIGHER (FLOODLIGHT ANALYSIS)



STUDIES 7A–B: MODERATION BY INTENTIONALITY

Two variables that we hypothesize (H_{3c}) reduce perceived intentionality—defined as knowingly or willingly bringing about an avoidable outcome—are (a) when the firm is small (i.e., constrained by limited resources), and (b) when demographic targeting is standard practice (i.e., an industry norm). For example, when a firm is large and has virtually unlimited resources (Woolley et al. 2023), it can presumably afford to advertise broadly. Its decision to instead employ demographic targeting could reinforce assumptions about opportunism, exploitation, and profit-seeking (Bhattacharjee et al. 2017; Lu et al. 2020), signaling greater intentionality. Similarly, if demographic targeting were *not* standard practice, and a firm chose to do so anyway, that decision to override a norm (and potentially trigger negative side effects; Knobe 2003) could also signal greater intentionality.

To confirm that firm size and industry norms shape inferences about intentionality, we conducted a pretest. Participants first read: "Suppose a company engages in discriminatory behavior, such as targeting advertisements to some demographic groups and not others." We then described either the two conditions from Study 7A (size: large company vs. small company) or the two conditions from Study 7B (industry norm: baseline vs. standard practice). However, rather than rating fairness, participants rated intentionality (e.g., "…how intentional would their behavior seem?"; see Web Appendix Study WA9). Consistent with our theorizing, intentionality was rated as lower for small companies (M = 5.64, 95% CI = [5.17, 6.11) than for large companies (M = 7.62, 95% CI = [7.21, 8.03]), $t(88)$ = 7.38, p < .001, d = .78), and also when it was standard practice (M = 6.60, 95% CI = [6.19, 7.01]) than when it was not (M = 7.34, 95% CI = [6.98, 7.70]), $t(87)$ = 2.73, p = .008, d = .29).

Study 7A Method

Study 7A tests the company size moderator. We recruited 732 MTurk workers (M_{AGE} = 38.64; 373 women, 345 men, 14 other; 539 White, 82 Black, 179 other) for Study 7A (aspredicted.org/s4pd-bfx6.pdf), which employed a 2 (targeting: gender vs. broad) × 2 (size: large company vs. small company) between-subjects design.

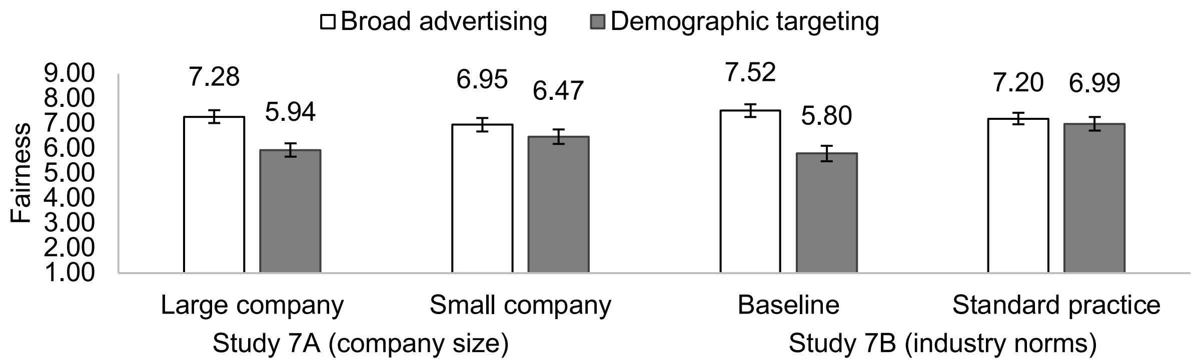
In the large company condition, participants read: "A large diversified multinational corporation has developed a new product, which they believe will appeal to their female customers. This multinational corporation has virtually unlimited resources, so it can spend its large advertising budget however it sees fit." In the small company condition, participants read: "A small mom-and-pop local business has developed a new product, which they believe will appeal to their female customers. This local business has very limited resources, so it must spend its modest advertising budget as carefully as possible." In the gender condition, participants read: "They plan to advertise the snacks broadly to the general public, rather than specifically to women." In the broad condition, participants read: "They plan to advertise the snacks specifically to women, rather than broadly to the general public." Below this text, we presented three counterbalanced measures capturing fairness: "How [fair/appropriate/acceptable] is this advertising strategy?" ("Not at all [fair/appropriate/acceptable]" = 1; "Very [fair/appropriate/acceptable]" = 9).

Study 7A Results

We averaged the three fairness measures ($\alpha = .93$). A fairness ANOVA revealed a main effect of targeting ($F(1, 728) = 42.03$, $p < .001$), qualified by an interaction ($F(1, 728) = 9.39$, $p = .002$; Figure 10). There was a simple effect of targeting for the large company ($F(1, 728) = 45.25$, $p < .001$), such that fairness perceptions were lower in the gender targeting condition ($M = 5.94$, 95% CI = [5.67, 6.21]) than in the broad advertising condition ($M = 7.28$, 95% CI = [7.02, 7.54], $d = .69$). However, this simple effect of targeting was significantly attenuated for the small company ($F(1, 728) = 5.88$, $p = .016$; $M_{\text{GENDER}} = 6.47$, 95% CI = [6.18, 6.76] vs. $M_{\text{BROAD}} = 6.95$, 95% CI = [6.68, 7.22], $d = .25$). Within just the gender targeting condition, fairness perceptions were higher for the small company than for the large company ($F(1, 728) = 7.50$, $p = .006$).

FIGURE 10

STUDIES 7A–B: DEMOGRAPHIC TARGETING IS VIEWED AS LESS UNFAIR WHEN PERFORMED BY A SMALL COMPANY AND WHEN IT IS STANDARD PRACTICE (95% CONFIDENCE INTERVALS)



Study 7B Method

Study 7B tests the industry norm moderator. We recruited 749 MTurk workers (M_{AGE} = 37.98; 387 women, 353 men, 9 other; 536 White, 97 Black, 169 race) for Study 7B (aspredicted.org/chk6-ncvk.pdf), which employed a 2 (targeting: gender vs. broad) × 2 (industry norm: baseline vs. standard practice) between-subjects design.

All participants read: "A company has developed a new snack, which they believe will appeal to their female customers." In the broad condition, we wrote: "They plan to advertise the snacks broadly to the general public, rather than specifically to women." In the gender condition, we wrote: "They plan to advertise the snacks specifically to women, rather than broadly to the general public." Those in the baseline condition read nothing else. Those in the standard practice condition additionally read: "Segmentation and targeting based on demographic characteristics (like gender) is standard practice in marketing. This is because it increases the likelihood that advertisements will be shown to the customers who are most interested in the product or service." The dependent variables were identical to Study 7A.

Study 7B Results

We averaged the three fairness measures ($\alpha = .97$). A fairness ANOVA revealed a main effect of targeting ($F(1, 745) = 47.57$, $p < .001$; Figure 10), qualified by an interaction ($F(1, 745) = 29.05$, $p < .001$). There was a simple effect of targeting in the baseline condition ($F(1, 745) = 77.17$, $p < .001$), such that fairness perceptions were lower in the gender targeting condition ($M = 5.80$, 95% CI = [5.49, 6.11]) than in the broad advertising condition ($M = 7.52$, 95% CI = [7.26, 7.77], $d = .77$). This simple effect of targeting was eliminated in the standard practice condition ($F(1, 745) = 1.11$, $p = .292$; $M_{\text{GENDER}} = 6.99$, 95% CI = [6.72, 7.26] vs. $M_{\text{BROAD}} = 7.20$, 95% CI = [6.97, 7.43], $d = .12$). Within just the gender targeting condition, fairness perceptions were higher in the standard practice condition than in the baseline condition ($F(1, 745) = 38.04$, $p < .001$).

Studies 7A–B Discussion

Studies 7A–B reveal that consumers are sensitive to the context in which demographic targeting takes place, characterizing a third factor in our model: perceived intentionality. Specifically, when demographic targeting is performed by small (vs. large) firms, and when demographic targeting is the norm (vs. not the norm), the resulting discrimination is viewed as less intentional and thus regarded as fairer (relative to broad advertising). Altogether, Studies 5A–7B not only explore three key factors that moderate fairness perceptions in demographic targeting—relevance, controllability, and perceived intentionality—but also underscore the causal role of beliefs about discrimination. In our final studies, we report the results of two large-scale Facebook campaigns, which offer real-world proof-of-concept for our account.

STUDIES 8A–B: FACEBOOK A/B TESTS

In Studies 8A–B, we report findings from two Facebook campaigns that measured actual clickthrough (i.e., real behavior). We used the A/B Test feature, which allows marketers to assess the performance of advertising campaigns conducted on evenly split and demographically comparable subsets of the Facebook user base. As in Studies 3A–B, we only served ads to members of the targeted segment (i.e., women).

To shed light on how often consumers think about demographic targeting, and to confirm inferences about demographic targeting, we conducted a pretest (see Web Appendix Study WA10). Participants viewed one of the two ads used in Study 8A (Figure 11). Those in the gender condition were more likely to (a) think about why they were seeing the ad (M_{GENDER} = 5.76 vs. M_{BROAD} = 5.14; $F(1, 346) = 5.45$, $p = .020$), (b) believe that they knew why they were seeing the ad (M_{GENDER} = 4.98 vs. M_{BROAD} = 3.67; $F(1, 346) = 24.46$, $p < .001$), and (c) anticipate that the ad was targeted to women (vs. men) (M_{GENDER} = 90.7% vs. M_{BROAD} = 42.8%; $\chi^2(2) = 89.22$, $p < .001$). Both relevance (M_{GENDER} = 4.93 vs. M_{BROAD} = 5.81; $F(1, 346) = 17.64$, $p < .001$) and controllability (M_{GENDER} = 3.71 vs. M_{BROAD} = 5.04; $F(1, 346) = 44.22$, $p < .001$) were also rated as lower in the gender condition.

Study 8A Method

Study 8A employed a single factor (targeting: gender vs. broad) between-subjects design. Our campaign was limited to U.S. women aged 21–44 (i.e., the primary age group for hard seltzer; Nielsen 2020) and optimized for traffic (i.e., clicks). It ran for seven days with a $100 daily budget, yielding 135,189 unique impressions.

FIGURE 11

STUDY 8A: BROAD ADVERTISING (LEFT) AND GENDER TARGETING (RIGHT)

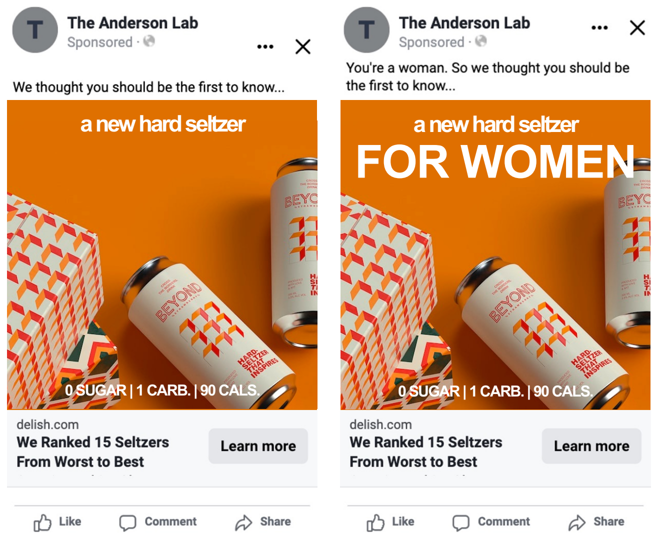
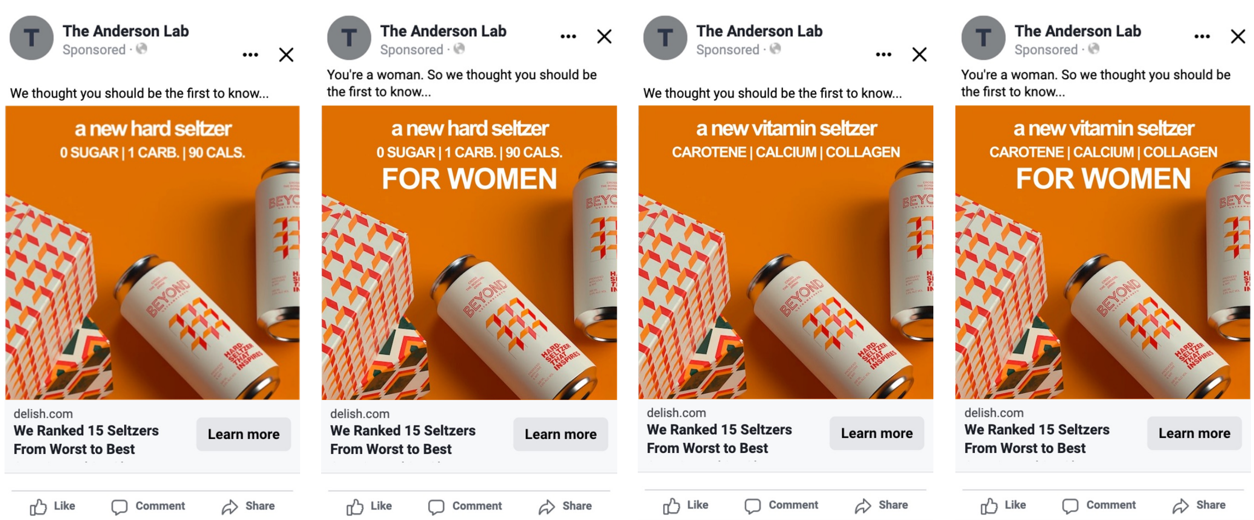


FIGURE 12

STUDY 8B: (1) HARD SELTZER BROAD ADVERTISING, (2) HARD SELTZER GENDER TARGETING, (3) VITAMIN SELTZER BROAD ADVERTISING, (4) VITAMIN SELTZER GENDER TARGETING



(1) (2) (3) (4)

In the broad advertising condition, we wrote: "We thought you should be the first to know." Beneath this text appeared an image depicting "a new hard seltzer" (Figure 11). In the gender targeting condition, we followed the basic format of a manipulation from previous research using the Facebook A/B Test (e.g., Tucker [2014] wrote: "As a fan of Beyoncé, you know that strong women matter"). We wrote: "You're a woman. So we thought you should be the first to know." The image text read: "a new hard seltzer. FOR WOMEN." We note that our decision to deliver advertisements only to women reflects a conservative test, since messaging so clearly targeting women (e.g., "You're a woman," "FOR WOMEN") should fare even worse if the audience also included men.

Study 8A Results

We divided the number of clicks in each condition by the audience exposed to each advertisement, to compute click-through rates. Clickthrough was higher in the broad advertising condition (1.39%, 95% CI = [1.29%, 1.48%]) than in the gender targeting condition (0.86%, 95% CI = [0.80%, 0.93%]; $\chi^2(1) = 84.59$, $p < .001$, $\varphi_c = .025$).[8]

Study 8B Method

Study 8B aimed to replicate both the results of Study 8A and the attenuating effect of the moderator featured in Study 5B (i.e., justification by average physical differences). Study 8B employed a 2 (targeting: gender vs. broad) × 2 (justification: average physical differences [vitamin seltzer] vs. preferences [hard seltzer]) between-subjects design, using the same campaign specifications as Study 8A. It ran for 10 days with a $200 daily budget, yielding 377,962 unique impressions. The hard seltzer condition was virtually identical to Study 8A.

However, for vitamin seltzer, we described "a new vitamin seltzer." Its formulation

[8] We held constant the daily advertising budget across all conditions, but because cost-per-click mechanically varies as a function of clickthrough rates, so too does audience size for each advertisement. In Study 8A, the broad condition comprised $N = 61{,}756$ (45.7%) and the gender targeting condition comprised $N = 73{,}433$ (54.3%). In Study 8B, the hard seltzer broad advertising condition comprised $N = 93{,}873$ (24.8%), the hard seltzer gender targeting condition comprised $N = 98{,}829$ (26.2%), the vitamin seltzer broad advertising condition comprised $N = 94{,}243$ (24.9%), and the vitamin seltzer gender targeting condition comprised $N = 91{,}017$ (24.1%).

included carotene, calcium, and collagen (Figure 12)—supplements commonly found in multivitamins designed for women, based on average physical differences in vitamin and nutrient needs across genders. A second pretest confirmed that these vitamins (e.g., carotene, calcium, and collagen) seemed more relevant to women than to men: “Do you believe these vitamins would be more relevant to men or to women?” (“Definitely men” = 1; “Equally relevant/no difference” = 5; “Definitely women” = 9; $M = 6.09$, 95% CI [5.86, 6.32]; comparison to scale midpoint: $t(142) = 9.46$, $p < .001$; see Web Appendix Study WA11).

Study 8B Results and Discussion

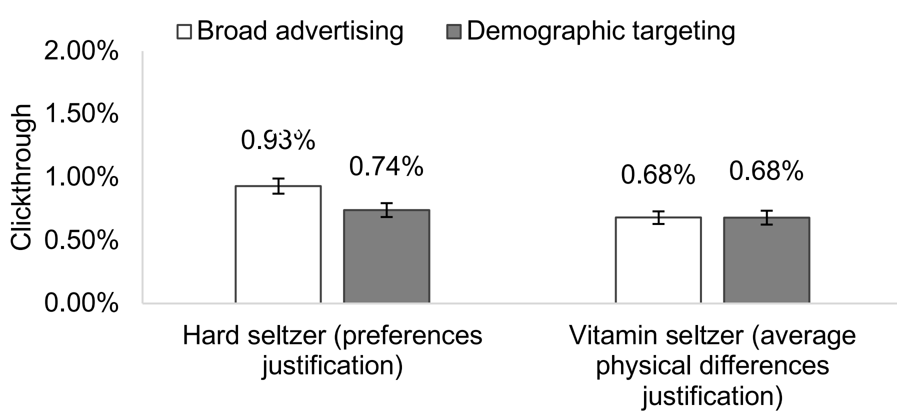
We fit a logistic regression of clickthrough on targeting (0 = broad advertising; 1 = targeting condition), product type ((0 = vitamin seltzer; 1 = hard seltzer), and their interaction. This regression revealed an interaction ($z = 3.00$, $p = .003$; Figure 13). Replicating Study 8A, for hard seltzer (preferences justification), clickthrough was higher in the broad advertising condition (0.93%, 95% CI = [0.87%, 0.99%]) than in the gender targeting condition (0.74%, 95% CI = [0.68%, 0.79%]; $\chi^2(1) = 21.51$, $p < .001$, $\varphi_c = .011$). However, for vitamin seltzer (average physical differences justification), there was no difference in clickthrough (gender targeting: 0.68%, 95% CI = [0.63%, 0.73%] vs. broad advertising: 0.68%, 95% CI = [0.62%, 0.73%]; $\chi^2(1) = .01$, $p = .909$, $\varphi_c = .000$).

Notably, Facebook’s A/B Test feature does not employ true randomization, particularly when ads contain very different imagery or text (Braun and Schwartz 2022; Eckles, Gordon, and Johnson 2018). However, Matz et al. (2018) argue that concerns about the internal validity of Facebook A/B Tests are mitigated when testing for an interaction, as in Study 8B, because any alternative interpretation would have to account for asymmetric effects of nonrandom assignment across conditions. The moderation observed in Study 8B implies the differences in clickthrough in both Study 8A and the hard seltzer condition of Study 8B were not simply attributable to differences in stimuli (e.g., if deleting "for women" made the ad more visually appealing).

Ultimately, we believe the strength of Studies 8A–B paradigm lies in its illustration of how this type of campaign would *actually* perform on Facebook, where advertising effectiveness would be jointly determined by the ad copy and Facebook's implementation of the campaign. To that end, these studies demonstrate that even when ads are targeted to those whom an algorithm expects to be most responsive (e.g., when the "for women" version is shown to users it predicts are likelier to click on ads mentioning women), gender-based appeals can backfire. Given that millions of advertisers use Facebook to demographically target billions of dollars of ads annually, Studies 8A–B serve as highly realistic proofs-of-concept attesting to the ecological validity and managerial relevance of our account.

FIGURE 13

STUDY 8B: GENDER TARGETING REDUCED CLICKTHROUGH (VS. BROAD ADVERTISING) FOR HARD SELTZER (PREFERENCES JUSTIFICATION), BUT NOT FOR VITAMIN SELTZER (AVERAGE PHYSICAL DIFFERENCES JUSTIFICATION) (95% CONFIDENCE INTERVALS)



GENERAL DISCUSSION

Fairness is a foundational topic in marketing, spanning decades and disciplines. But while this literature has largely focused on fairness perceptions in pricing (Xia et al. 2004), relatively less attention has been paid to other equally critical aspects of marketing strategy, such as segmentation and targeting. Importantly, while these strategies allow firms to increase the likelihood of reaching an interested customer, our results suggest that any potential positive effect of *demographic* targeting, in particular, may be counteracted by the negative effect on fairness perceptions triggered by appraisals of discrimination.

Across 14 experiments (N = 9,181), 13 supplemental studies (N = 7,065), and two Facebook A/B Tests (N = 513,151), we found that when consumers learned or inferred that they or others had been targeted based on demographic characteristics, fairness perceptions and brand support suffered, relative to both broad advertising and behavioral targeting. These differences in fairness perceptions were mediated by beliefs about discrimination and attenuated by factors that (a) improved relevance (e.g., justification by average physical differences or market research), (b) increased perceptions of controllability, and (c) reduced perceived intentionality (e.g., when the company was small, and when demographic targeting was the norm).

In addition to testing a broadly applicable framework for understanding fairness perceptions in demographic targeting, our studies highlighted both the various ways consumers themselves learn or infer they have been targeted (e.g., social media disclosures, ad images, ad copy, news coverage) and how consumers react when they do (e.g., reduced purchase intentions, consequential choice, clickthrough, etc.). For example, in the incentive-aligned Study 1C, participants were less likely to actually choose a McDonald's gift card (over a less valuable Amazon gift card) after learning that McDonald's engaged in demographic targeting.

Theoretical Implications

We view our account as contributing to three primary streams of research: fairness perceptions, persuasion knowledge, and diversity in marketing. First, our findings help expand a growing literature examining the perceived fairness of marketing tactics beyond pricing (e.g., versioning, planned obsolescence, “shrinkflation”; Gershoff, Kivetz and Keinan 2012; Evangelidis 2024; Kuppelwieser et al. 2019; Trupia and Shaddy 2024). And while recent work exploring fairness in promotion has studied differences in *content* (e.g., identity-based messaging backfires when it reinforces stereotypes; Kim et al., 2023), our account elucidates reactions to the strategic decision *itself* to target certain demographic groups.

Our work also connects to research on persuasion knowledge, which has yielded numerous insights relating to how consumers identify and cope with marketers’ persuasion attempts (Friestad and Wright 1994; Eisend and Tarrahi 2022; Isaac and Grayson 2017). We not only conceptually replicate the finding that consumers try to form a metacognitive assessment of the upstream strategic decision about targeting (Aaker et al. 2000), but also spotlight downstream implications for fairness perceptions and brand support.

Furthermore, we believe our research promotes diversity in marketing (Arsel et al. 2022; Ferraro, Haltman et al. 2025; Hemsley and Sands 2023; Park, Voss and Voss 2023; Uduehi et al. 2025). Our paradigms demonstrated that targeting strategies can be communicated visually through representations of diversity (or lack thereof) in the ad itself, or through the choice of media outlet(s) and the diversity of their associated audiences (Study 3A–B). In both cases, we found that consumers responded positively to diversity in advertising. Our findings may thus be interpreted as helping to make a business case for diversity, highlighting an area where it might not only be “good for the world,” but also “good for the firm” (Chandy et al. 2021).

Finally, we believe the methodological approach employed in Study 4 answers recent calls promoting the use of generative AI for unstructured text analysis (e.g., Arora, Chakraborty, and Nishimura 2025; Berger et al. 2022). For example, the correlation between the discrimination ratings supplied by our human coders and generative AI was strongly positive. Moreover, the rank-ordering of demographic targeting conditions (i.e., from most- to least-discriminatory), as well as the significance of the differences between them, were virtually identical. Yet the generative AI approach was significantly faster and more cost-effective.

Limitations, Managerial Implications, and Directions for Future Research

Several important limitations are worth acknowledging, which we believe characterize opportunities for future work. For instance, in the majority of our studies, we directly asked participants to consider the fairness of various targeting strategies. These paradigms mirror many real-world situations in which journalists, aggrieved customers, or regulators surface and publicize concerns about fairness—such as Facebook's potential violations of the Fair Housing Act and Equal Credit Opportunity Act (Imana et al. 2021; Isaac and Hu 2021). Follow-up research might therefore identify and explore the factors that spontaneously lead consumers to think about fairness in these contexts (e.g., building on Studies 8A–B)—not only with respect to targeting, but also in light of other marketing tactics, such as *place* (i.e., distribution). For example, when firms choose to open and operate retail stores in some neighborhoods and not others (e.g., resulting in food deserts; Walker, Keane and Burke 2010), consumers may infer unfair targeting practices and view these decisions as discriminatory (depending on relevance and controllability).

Moreover, while our theory should apply generally to all forms of demographic targeting, we tested five common bases (e.g., race, gender, age, SES, and geography). Future work could examine other demographics (e.g., ethnicity, sexual orientation, and religious affiliation) and other bases for segmentation and targeting. For instance, targeting psychographics (e.g., motherhood) could seem both more relevant and more controllable, and thus perceived as fairer, than targeting demographics (e.g., women). We also tested five operationalizations of our moderators (i.e., average physical differences, market research, changeability, firm size, and industry norms), and encourage exploration of others. For example, to reduce inferences about intentionality, businesses might consider crafting ads that convey sympathy. And opt-in policies could explicitly permit demographic targeting, potentially increasing controllability.

Additionally, while we found that consumers were more accepting of demographic targeting when relevance was high (i.e., when membership in the demographic group seemed sufficiently correlated with preferences), targeting decisions are often based on which segment seems most *persuadable*. Yet persuasiveness is conceptually distinct from relevance. Our findings thus underscore the need for managers to carefully navigate the efficiency-versus-fairness trade-off posed by demographic targeting—and, in particular, whether its benefits are outweighed by costs (e.g., to fairness perceptions). This is an important managerial question given the results of Study 1C, which demonstrated a negative effect of demographic targeting (vs. broad advertising) on the consequential choice of a gift card.

There may also be important differences across demographic groups in how unfair they perceive demographic targeting to be (relative to broad advertising), especially among historically underrepresented (Aaker et al. 2000) or marginalized groups (Uduehi, Saint Clair, and Crabbe 2024). Our studies were underpowered to conclusively test any moderating effects of sample demographics beyond participant gender (nor did we preregister such analyses). But in exploratory analyses of participant gender across Studies 1A–7B, we found no moderation by participant gender in 13 out of 14 studies.

Finally, a natural question is whether demographic targeting is even more aversive when it involves harmful, injurious, or damaging *products*. We tested common products considered beneficial (at best) or benign (at worst). We did this to isolate negative reactions to demographic targeting, in particular, as opposed to negative reactions to harmful products, in general. However, it is unclear how our framework might apply when products are unambiguously helpful (e.g., motivated by altruistic intentions) or, conversely, when they are clearly harmful.

To address this latter possibility, we report two preregistered supplemental studies in the Web Appendix (Studies WA12–13), which manipulate product harm (e.g., low-APR student loans vs. high-APR payday loans) and targeting (race vs. broad). For example, in Study WA12, we replicated the basic effect for low-APR (non-harmful) student loans (M_{RACE} = 3.28, 95% CI = [2.99, 3.56] vs. M_{BROAD} = 7.13, 95% CI = [6.92, 7.35], $F(1, 984) = 365.15$, $p < .001$, $d = 1.38$). However, for high-APR (harmful) payday loans, this difference attenuated (interaction: $F(1, 984) = 52.77$, $p < .001$; M_{RACE} = 2.27, 95% CI = [2.01, 4.39] vs. M_{BROAD} = 4.05, 95% CI = [3.70, 4.39], $F(1, 984) = 77.59$, $p < .001$, $d = .68$). This was because broad advertising was perceived as *much less* fair for high-APR (harmful) payday loans ($M_{\text{HIGH-APR}}$ = 4.05) than for low-APR (non-harmful) student loans ($M_{\text{LOW-APR}}$ = 7.13; $F(1, 984) = 233.03$, $p < .001$). Therefore, the relative fairness of demographic targeting to broad advertising actually attenuates for harmful products, because *any* promotion is deemed unfair (possibly due to a floor effect). Notably, demographic targeting was still less fair for high-APR (harmful) payday loans ($M_{\text{HIGH-APR}}$ = 2.27) than for low-APR (non-harmful) student loans ($M_{\text{LOW-APR}}$ = 3.28; $F(1, 984) = 25.12$, $p < .001$).

Conclusion

Understanding consumer reactions to demographic targeting is critical to marketing theory and practice, given the potential for these practices to be viewed as discriminatory. Our findings suggest that when consumers consider the fairness of targeting, they do not believe different groups should be treated differently, based on factors that are irrelevant and/or uncontrollable. Importantly, by probing the underlying psychology of these beliefs, our framework not only identifies their causes and consequences, but also offers concrete guidance for managers.

REFERENCES

Aaker, Jennifer L., Anne M. Brumbaugh, and Sonya A. Grier (2000). "Nontarget markets and viewer distinctiveness: The impact of target marketing on advertising attitudes." *Journal of Consumer Psychology*, 9 (3), 127–40.

Abdurahman, Suhaib, Alireza Salkhordeh Ziabari, Alexander K. Moore, Daniel M. Bartels, and Morteza Dehghani (2025). "A primer for evaluating large language models in social-science research." *Advances in Methods and Practices in Psychological Science*, 8 (2). DOI:10.1177/25152459251325174.

Acquisti, Alessandro, Leslie K. John, and George Loewenstein (2013). "What is privacy worth?" *The Journal of Legal Studies*, 42 (2), 249–74.

Aguirre, Elizabeth, Dominik Mahr, Dhruv Grewal, Ko De Ruyter, and Martin Wetzels (2015). "Unraveling the personalization paradox." *Journal of Retailing*, 91 (1) 34–49.

Alcorn, Chauncey (12 June 2020). "Band-Aid will make black and brown flesh-toned bandages." CNN Business. https://www.cnn.com/2020/06/12/business/black-band-aids/index.html

Andreasen, Alan R. (1966). "Geographic mobility and market segmentation." *Journal of Marketing Research*, 3 (4), 341–8.

Arora, Neeraj, Ishita Chakraborty, and Yohei Nishimura (2025). "AI–human hybrids for marketing research: Leveraging large language models (LLMs) as collaborators." *Journal of Marketing*, 89 (2), 43–70.

Arsel, Zeynep, David Crockett, and Maura L. Scott (2022). "Diversity, equity, and inclusion in the Journal of Consumer Research." *Journal of Consumer Research*, 48 (5), 920–33.

Assael, Henry, and A. Marvin Roscoe Jr (1976). "Approaches to market segmentation analysis."

Journal of Marketing, 40 (4), 67–76.

Bechwati, Nada Nasr, and Maureen Morrin (2003). “Outraged consumers: Getting even at the expense of getting a good deal.” *Journal of Consumer Psychology*, 13 (4), 440–53.

Bellezza, Silvia, Francesca Gino, and Anat Keinan (2014). “The red sneakers effect: Inferring status and competence from signals of nonconformity.” *Journal of Consumer Research*, 41 (1), 35–54.

Berger, Jonah, Wendy W. Moe, and David A. Schweidel (2023). “What holds attention? Linguistic drivers of engagement.” *Journal of Marketing*, 87 (5), 793–809.

Bhattacharjee, Amit, and Jason Dana (2024). “Lay economic reasoning: An integrative review and call to action.” *Consumer Psychology Review*, 7 (1), 3–39.

Bhattacharjee, Amit, Jason Dana, and Jonathan Baron (2017). “Anti-profit beliefs: How people neglect the societal benefits of profit.” *Journal of Personality and Social Psychology*, 113 (5), 671–96.

Bojić, Ljubiša, Olga Zagovora, Asta Zelenkauskaite, Vuk Vuković, Milan Čabarkapa, Selma Veseljević Jerković, and Ana Jovančević (2025). “Comparing large language models and human annotators in latent content analysis of sentiment, political leaning, emotional intensity and sarcasm.” *Scientific Reports*, 15 (1), 11477.

Bolton, Lisa E., and Haipeng Allan Chen (2024). “Behavioral pricing and price fairness.” In *New Directions in Behavioral Pricing*, pp. 51–78. World Scientific Publishing Co.

Bolton, Lisa E., Luk Warlop, and Joseph W. Alba (2003). “Consumer perceptions of price (un) fairness.” *Journal of Consumer Research*, 29 (4), 474–91.

Braun, Michael, and Eric M. Schwartz (2022). “The A/B test deception: Divergent delivery, response heterogeneity, and erroneous inferences in online advertising field

Experiments." SMU Cox School of Business Research Paper.

Brough, Aaron R., David A. Norton, Shannon L. Sciarappa, and Leslie K. John (2022). "The bulletproof glass effect." *Journal of Marketing Research*, 59 (4), 739–54.

Brucks, Melanie., and Olivier Toubia (2025). "Prompt architecture induces methodological artifacts in large language models." *PLoS One*, 20 (4), e0319159.

Bullock, Heather E., Wendy R. Williams, and Wendy M. Limbert (2003). "The impact of attributions and beliefs about inequality." *Journal of Poverty*, 7 (3), 35–56.

Campbell, Margaret C. (1999). ""Why did you do that?" The important role of inferred motive in perceptions of price fairness." *Journal of Product & Brand Management*, 8 (2), 145–53.

Campbell, Margaret C., and Amna Kirmani (2008). "I know what you're doing and why you're doing it." *Handbook of Consumer Psychology*, 549–74.

Chandy, Rajesh K., Gita Venkataramani Johar, Christine Moorman, and John H. Roberts (2021). "Better marketing for a better world," *Journal of Marketing*, 85 (3), 1–9.

Colquitt, Jason A., and Jessica B. Rodell (2015). "Measuring justice and fairness." *The Oxford Handbook of Justice in the Workplace*, 187–202.

Culnan, Mary J. (2000). "Protecting privacy online: Is self-regulation working?" *Journal of Public Policy & Marketing*, 19 (1), 20–6.

D'Angelo, Jennifer K., Lea Dunn, and Francesca Valsesia (2025). "Is this for me? Differential responses to skin tone inclusivity initiatives by underrepresented consumers and represented consumers." *Journal of Marketing*, 89 (2), 25–42.

Davidai, Shai (2018). "Why do Americans believe in economic mobility? Economic inequality, external attributions of wealth and poverty, and the belief in economic mobility." *Journal of Experimental Social Psychology*, 79, 138–48.

Dolifka, David, Katherine L. Christensen, and Franklin Shaddy (2024). “Highlighting opportunities (versus outcomes) increases support for economic redistribution.” *Social Psychological and Personality Science*. https://doi.org/10.1177/19485506241228469

Eckles, Dean, Brett R. Gordon, and Garrett A. Johnson (2018). “Field studies of psychologically targeted ads face threats to internal validity.” *Proceedings of the National Academy of Sciences*, 115 (23), E5254–5.

Eisend, Martin, and Farid Tarrahi (2022). “Persuasion knowledge in the marketplace: A meta-analysis.” *Journal of Consumer Psychology*, 32 (1), 3–22.

Evangelidis, Ioannis (2024). “Shrinkflation aversion: When and why product size decreases are seen as more unfair than equivalent price increases.” *Marketing Science*, 43 (2), 280–8.

Facebook (9 April 2021). “Removing certain ad targeting options and expanding our ad controls.” *Facebook*. https://www.facebook.com/business/news/removing-certain-ad-targeting-options-and-expanding-our-ad-controls

Ferraro, Carla, Alicia Hemsley, and Sean Sands (2023). “Embracing diversity, equity, and inclusion (DEI): Considerations and opportunities for brand managers.” *Business Horizons*, 66 (4), 463–79.

Fiske, Susan T., Amy J.C. Cuddy, Peter Glick, and Jun Xu (2002). “A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition.” *Journal of Personality and Social Psychology*, 82 (6), 878–902.

Ford, Gary T., Darlene B. Smith, and John L. Swasy (1990). “Consumer skepticism of advertising claims.” *Journal of Consumer Research*, 16 (4), 433–41.

Friedman, Elizabeth M.S., and Olivier Toubia (2022). “Pricing fairness in a pandemic: Navigating unintended changes to value or cost.” *Journal of the Association for*

Consumer Research, 7 (1), 89–97.

Friestad, Marian, and Peter Wright (1994). “The persuasion knowledge model: How people cope with persuasion attempts.” *Journal of Consumer Research*, 21 (1), 1–31.

Gal, David, Jeffrey Parker, and Tianyi Li (2018). “The weight of history: When historical corporate transgressions affect present-day consumer attitudes.” SSRN. https://ssrn.com/abstract=3182709

Gershoff, Andrew D., Ran Kivetz, and Anat Keinan (2012). “Consumer response to versioning: How brands’ production methods affect perceptions of unfairness.” *Journal of Consumer Research*, 39 (2), 382–98.

Goenka, Shreyans, and Rajesh Bagchi (2025). “Price partitioning of socio-moral surcharges.” *Journal of Consumer Research*, 51 (5), 873–95.

Goldfarb, Avi, and Catherine E. Tucker (2011). “Privacy regulation and online advertising.” *Management Science*, 57 (1), 57–71.

Grier, Sonya A., David Crockett, Guillaume D. Johnson, Kevin D. Thomas, and Tonya Williams Bradford (2024). “Race in consumer research: Past, present, and future.” *Journal of Consumer Research*, 51 (1), 56–65.

Haltman, Cory, Jianna Jin, Grant Donnelly, and Rebecca Walker Reczek (2025), “Equity or exclusion? How political ideology shapes response to the depiction of historically underrepresented identities in marketing communications,” Working Paper.

Haws, Kelly L., and William O. Bearden (2006). “Dynamic pricing and consumer fairness perceptions.” *Journal of Consumer Research*, 33 (3), 304–11.

Heyman, James E., and Barbara A. Mellers (2008). “Perceptions of fair pricing.” In Curtis P. Haugtvedt, Paul M. Herr, and Frank R. Kardes, eds. *Handbook of Consumer Psychology*.

Psychology Press, 683–98.

Huppertz, John, Sidney Arenson, and Richard Evans (1978). "An application of equity theory to buyer-seller exchange situations." *Journal of Marketing Research*, 15 (2), 250–60.

Imana, Basileal, Aleksandra Korolova, and John Heidemann (2021). "Auditing for discrimination in algorithms." *Proceedings of the Web Conference*, 3767–78.

Isaac, Mathew S., and Kent Grayson (2017). "Beyond skepticism: Can accessing persuasion knowledge bolster credibility?" *Journal of Consumer Research*, 43 (6), 895–912.

Isaac, Mike, and Tiffany Hsu (10 November 2021). "Meta plans to remove thousands of sensitive ad-targeting categories." *New York Times*. https://www.nytimes.com/2021/11/09/technology/meta-facebook-ad-targeting.html

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986). "Fairness as a constraint on profit seeking: Entitlements in the market." *American Economic Review*, 76 (4), 728–41.

Kim, Tami, Kate Barasz, and Leslie K. John (2019). "Why am I seeing this ad? The effect of ad transparency on ad effectiveness." *Journal of Consumer Research*, 45 (5), 906–32.

Kim, Tami, Kate Barasz, Michael I. Norton, and Leslie K. John (2023). "Calculators for women: When identity-based appeals alienate consumers." *Journal of the Association for Consumer Research*, 8 (1), 72–82.

Kimes, Sheryl E. (1994). "Perceived fairness of yield management." *Cornell Hotel and Restaurant Administration Quarterly*, 35 (1), 22–9.

Knobe, Joshua (2003). "Intentional action and side effects in ordinary language." *Analysis*, 63 (30), 190–4.

Kotler, Philip, and Keller L. Keller (2011). *A Framework for Marketing Management*. Toronto: Prentice Hall.

Kozlowska, Hanna (6 April 2018). "“Why am I seeing this ad” explanations on Facebook are incomplete and misleading." *Quartz*. https://qz.com/1245941/why-am-i-seeing-this-ad-explanations-on-facebook-are-incomplete-and-misleading-a-study-says

Kuppelwieser, Volker G., Phil Klaus, Aikaterini Manthiou, and Othman Boujena (2019). "Consumer responses to planned obsolescence." *Journal of Retailing and Consumer Services*, 47, 157–65.

LaForge, Patrick (6 February 2018). "‘Lady Doritos’? Pepsi wants a do-over." *New York Times*. https://www.nytimes.com/2018/02/06/business/lady-doritos-indra-nooyi.html

Lambrecht, Anja, and Catherine Tucker (2013). "When does retargeting work? Information specificity in online advertising." *Journal of Marketing Research*, 50 (5), 561–76.

Li, Ren, Sarah Gordon, and Michele J. Gelfand (2017). "Tightness–looseness: A new framework to understand consumer behavior." *Journal of Consumer Psychology*, 27 (3), 377–91.

Lu, Zhi, Lisa E. Bolton, Sharon Ng, and Haipeng Allan Chen (2020). "The price of power: How firm's market power affects perceived fairness of price increases." *Journal of Retailing*, 96 (2), 220–34.

Lupfer, Michael B., Kelly P. Weeks, Kelly A. Doan, and David A. Houston (2000). "Folk conceptions of fairness." *European Journal of Social Psychology*, 30 (3), 405–28.

Maheshwari, Sapna (12 October 2017). "‘Different Ads, Different Ethnicities, Same Car." *New York Times*. https://www.nytimes.com/interactive/2017/10/12/business/media/toyota-camry-ads-different-ethnicities.html

Major, Brenda, Wendy J. Quinton, and Shannon K. McCoy (2002). "Antecedents and consequences of attributions to discrimination: Theoretical and empirical advances." *Advances in Experimental Social Psychology*, 34, 251–330.

Matz, Sandra C., Michal Kosinski, Gideon Nave, and Dadvid J. Stillwell (2018). "Facebook's optimization algorithms are highly unlikely to explain the effects of psychological targeting." *Proceedings of the National Academy of Sciences*, 115 (23), E5256–7.

Moorman, Christine, Megan Ryan, and Nader Tavassoli (2022), "Why marketers are returning to traditional advertising," *Harvard Business Review* (April) 29. https://hbr.org/2022/04/why-marketers-are-returning-to-traditional-advertising

Moshary, Sarah, Anna Tuchman, and Natasha Bhatia (2023), "Gender-based pricing in consumer packaged goods: A pink tax?" *Marketing Science*. doi.org/10.1287/mksc.2023.1452

Nardini, Gia, Tracy Rank-Christman, Melissa G. Bublitz, Samantha NN Cross, and Laura A. Peracchio (2021). "Together we rise: How social movements succeed." *Journal of Consumer Psychology*, 31 (1), 112–45.

Newman, George E., Margarita Gorlin, and Ravi Dhar (2024). "When going green backfires: How firm intentions shape the evaluation of socially beneficial product enhancements." *Journal of Consumer Research*, 41 (3), 823–39.

Oliver, Richard, and John E. Swan (1989). "Equity and disconfirmation perceptions as influences on merchant and product satisfaction." *Journal of Consumer Research*, 16 (3), 372–83.

Oppenheimer, Daniel M., Tom Meyvis, and Nicolas Davidenko (2009). "Instructional manipulation checks: Detecting satisficing to increase statistical power." *Journal of Experimental Social Psychology*, 45 (4), 867–72.

Paharia, Neeru, Jill Avery, and Anat Keinan (2014). "Positioning brands against large competitors to increase sales." *Journal of Marketing Research*, 51 (6), 647–56.

Pavlovic, Maja (2 April 2024). "Can large language models be used to label data?" *Medium*. medium.com/data-science/can-large-language-models-llms-label-data-2a8334e70fb8

Papadopoulos, Christopher, and B. K. Hayes (2018). “What matters when judging intentionality—moral content or normative status?” *Psychonomic Bulletin*, 25, 1170–7.

Park, Young Woong, Glenn B. Voss, and Zannie Giraud Voss (2023). “Advancing customer diversity, equity, and inclusion: Measurement, stakeholder influence, and the role of marketing.” *Journal of the Academy of Marketing Science*, 51 (1), 174–97.

Paul, Iman, Jeffrey R. Parker, and Sara Loughran Dommer (2020). “The influence of incidental tokenism on private evaluations of stereotype-typifying products.” *Social Psychology Quarterly*, 83 (1), 49–69.

Rafieian, Omid, and Hema Yoganarasimhan (2021). “Targeting and privacy in mobile advertising.” *Marketing Science*, 40 (2), 193–218.

Reich, Taly, Daniella M. Kupor, and Rosanna K. Smith (2018). “Made by mistake: When mistakes increase product preference.” *Journal of Consumer Research*, 44 (5), 1085–103.

Shaddy, Franklin, and Anuj K. Shah (2018). “Deciding who gets what, fairly.” *Journal of Consumer Research*, 45 (4), 833–48.

Shavitt, Sharon, Duo Jiang, and Hyewon Cho (2016), “Stratification and segmentation: Social class in consumer behavior.” *Journal of Consumer Psychology*, 26 (4), 583–93.

Smith, N. Craig, and Elizabeth Cooper-Martin (1997). “Ethics and target marketing: The role of product harm and consumer vulnerability.” *Journal of Marketing*, 61 (3), 1–20.

Son Hing, Leanne S., D. Ramona Bobocel, Mark P. Zanna, Donna M. Garcia, Stephanie S. Gee, and Katie Orazietti (2011), “The merit of meritocracy.” *Journal of Personality and Social Psychology*, 101 (3), 433–50.

Spencer, Steven J., Mark P. Zanna, and Geoffrey T. Fong (2005). “Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes." *Journal of Personality and Social Psychology*, 89 (6), 845–51.

Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch Jr, and Gary H. McClelland (2013). "Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression." *Journal of Marketing Research*, 50 (2), 277–88.

Summers, Christopher A., Robert W. Smith, and Rebecca Walker Reczek (2016). "Behaviorally targeted ads as implied social labels." *Journal of Consumer Research*, 43 (1), 156–78.

Tepper, Kelly (1994). "The role of labeling processes in elderly consumers' responses to age segmentation cues." *Journal of Consumer Research*, 20 (4), 503–19.

Tomova Shakur, Teodora K., and L. Taylor Phillips (2022). "What counts as discrimination? How principles of merit shape fairness of demographic decisions." *Journal of Personality and Social Psychology*, 123 (5), 957–82.

Trupia, Maria Giulia, and Franklin Shaddy (2024). ""No time to buy": Asking consumers to spend time to save money is perceived as fairer than asking them to spend money to save time." *Journal of Consumer Psychology*. https://doi.org/10.1002/jcpy.1444

Tucker, Catherine E. (2014). "Social networks, personalized advertising, and privacy controls." *Journal of Marketing Research*, 51 (5), 546–62.

Uduehi, Esther, Julian K. Saint Clair, Mitchell Hamilton, and Americus Reed (2025). "When diversity backfires: The asymmetric role of multicultural diversity on brand perception." *Journal of Consumer Research*. https://doi.org/10.1093/jcr/ucae068

Uduehi, Esther, Julian K. Saint Clair, and Rowena Crabbe (2024). "Intersectionality in marketing: A paradigm for understanding understudied consumers." *Journal of Marketing*. https://doi.org/10.1177/00222429241258493

Uttich, Kevin, and Tania Lombrozo (2010). "Norms inform mental state ascriptions: A rational explanation for the side-effect effect." *Cognition*, 116 (1), 87–100.

Van der Sluis, Helen, Adriana Samper, Andrea C. Morales, and Nathan D. Martin (2023), "Responses to gender-based price variation" Working paper.

Voorhees, Clay M., Michael K. Brady, Roger Calantone, and Edward Ramirez (2016). "Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies." *Journal of the Academy of Marketing Science*, 44, 119–34.

Walker, Renee E., Christopher R. Keane, and Jessica G. Burke (2010). "Disparities and access to healthy food in the United States." *Health & Place*, 16 (5), 876–84.

Wang, Yu, and Aradhna Krishna (2012). "Enticing for me but unfair to her: Can targeted pricing evoke socially conscious behavior?" *Journal of Consumer Psychology*, 22 (3), 433–42.

Weiner, Bernard (2000). "Attributional thoughts about consumer behavior." *Journal of Consumer Research*, 27 (3), 382–7.

Wells, William D. (1975). "Psychographics: A critical review." *Journal of Marketing Research*, 12 (2), 196–213.

Winterich, Karen Page, Robert E. Carter, Michael J. Barone, Ramkumar Janakiraman, and Ram Bezawada (2015). "Tis better to give than receive? How and when gender and residence-based segments predict choice of donation-versus discount-based promotions." *Journal of Consumer Psychology*, 25 (4), 622–34.

Woolley, Kaitlin, Daniella Kupor, and Peggy J. Liu (2023). "Does company size shape product quality inferences?" *Journal of Marketing Research*, 60 (3), 425–48.

Wright, Peter (2002). "Marketplace metacognition and social intelligence." *Journal of Consumer Research*, 28 (4), 677–82.

Xia, Lan, Kent B. Monroe, and Jennifer L. Cox (2004). "The price is unfair! A conceptual framework of price fairness perceptions." *Journal of Marketing*, 68 (4), 1–15.

Yang, Linyun W., and Pankaj Aggarwal (2019). "How company size affects consumer expectations and evaluations." *Journal of Consumer Research*, 45 (6), 1369–84.

Fairness Perceptions in Demographic Targeting

WEB APPENDIX

STUDY WA1: MBA STUDENTS

Study WA1 Method

We recruited 164 full-time MBA students at a West Coast business school (M_{age} = 28.00; 94 women, 70 men) for Study WA1, which employed a single-factor (targeting: demographic vs. broad) between-subjects design. All participants first read: "Suppose a company has developed a new product which they believe will appeal specifically to a particular demographic group." Depending on condition, participants read that the company would "advertise the product directly to people within that group" (demographic targeting) or "advertise the product broadly to the general public" (broad advertising). We then asked: "How fair is this advertising plan?" ("Not at all fair" = 1; "Very fair" = 9).

Study WA1 Results

Participants rated demographic targeting as fairer (M = 6.89, 95% CI = [6.48, 7.30]) than broad advertising (M = 6.45, 95% CI = [5.95, 6.96]), but this difference was not significant ($t(162) = 1.35$, $p = .180$, $d = .21$).

STUDY WA2: PRETEST FOR STUDIES 2A–B

Study WA2 Method

We recruited 135 Prolific workers (M_{age} = 35.07; 81 women, 53 men, 1 other) for Study WA2, which employed a single-factor (targeting: race vs. gender vs. broad) between-subjects design. All participants read: "Please take a look at the following picture and answer the questions below." We then presented an image depicting people of a single race (Study 2A), a single gender (Study 2B), or a diverse group (Figure WA1). Participants then responded to three questions (counterbalanced): "How attractive is this picture?" ("Very unattractive" = 1, "Very attractive" = 9), "How appealing is this picture?" (Very unappealing" = 1, "Very appealing" = 9), and "How interesting is this picture?" ("Very uninteresting" = 1, "Very interesting" = 9).

FIGURE WA1

STUDY WA2: (1) RACE TARGETING, (2) GENDER TARGETING, AND (3) BROAD ADVERTISING

(1) (2) (3)

Study WA2 Results

We averaged these measures to form a composite (α = .93). There was no difference between the race targeting condition and broad advertising condition (M_{RACE} = 6.81, 95% CI [6.28, 7.33] vs. M_{BROAD} = 6.58, 95% CI [6.05, 7.11], $t(87)$ = .66, p = .510, d = .14) and no difference between the gender targeting condition and broad advertising condition (M_{GENDER} = 6.36, 95% CI [5.84, 6.88] vs. M_{BROAD} = 6.58, 95% CI [6.05, 7.11, $t(88)$ = .51, p = .608, d = .11).

STUDY WA3: RELEVANCE AND CONTROLLABILITY OF SIX TARGETING BASES

Study WA3 Method

We recruited 545 MTurk workers (M_{age} = 34.78; 284 women, 258 men, 3 other) for Study WA3 which employed a single-factor (targeting basis: race vs. gender vs. age vs. socioeconomic status vs. geography vs. behavioral) between-subjects design.

We adapted four measures from Tomova Shakur and Phillips (2022) to capture relevance and controllability. For relevance, we asked: "To what extent is someone's [race/gender/age at any given moment in time/socioeconomic status/geographic location/purchasing behavior] relevant to most products and services?" and "To what extent is someone's [race/gender/age at any given moment in time/socioeconomic status/geographic location/purchasing behavior] a signal of their preferences (i.e., their wants and needs)?" For controllability, we asked: "To what extent is someone's [race/gender/age at any given moment in time/socioeconomic status/geographic location/purchasing behavior] under their personal control?" and "To what extent is someone's [race/gender/age at any given moment in time/socioeconomic status/geographic location/purchasing behavior] easily changeable?" Participants responded to each of the four questions, which were presented in random order, on nine-point scales ("Not at all" = 1; "Extremely" = 9).

Study WA3 Results and Discussion

We averaged the two controllability measures ($\alpha = .83$). A controllability ANOVA revealed a main effect of condition ($F(5, 542) = 106.50$, $p < .001$), with race rated as the least controllable ($M = 1.57$, 95% CI = [1.28, 1.86]) and behavior as the most controllable ($M = 6.31$, 95% CI = [6.05, 6.57]; Table WA1). We averaged the two relevance measures ($\alpha = .67$). A relevance ANOVA revealed a main effect of targeting basis ($F(5, 542) = 38.39$, $p < .001$), with race rated as the least relevant ($M = 4.26$, 95% CI = [3.86, 4.65]) and behavior as the most relevant ($M = 6.98$, 95% CI = [6.75, 7.21]).

TABLE WA1

STUDY WA1: RELEVANCE AND CONTROLLABILITY OF SIX TARGETING BASES

Targeting basis	Controllability		Targeting basis	Relevance
Race	1.57 [1.28, 1.86]	A	Race	4.26 [3.96, 4.65]
Age	1.93 [1.57, 2.28]	A	Gender	5.29 [4.92, 5.66]
Gender	3.82 [3.34, 4.30]		Geography	5.87 [5.59, 6.15]
Geography	4.68 [4.36, 5.00]	B	Socioeconomic status	6.07 [5.78, 6.36]
Socioeconomic status	4.78 [4.45, 5.11]	B	Age	6.62 [6.37, 6.88]
Behavioral	6.31 [6.05, 6.57]		Behavioral	6.98 [6.75, 7.21]

Note: Comparisons sharing a letter are not significantly different at $p = .05$. Brackets contain 95% CIs. Conditions are sorted from least to most controllability, and from least to most relevant

STUDY 4 SECONDARY ANALYSIS METHODOLOGICAL DETAILS

For our secondary preregistered analysis, we asked generative AI to rate the extent to which each thought-listing from Study 4 invoked the concept of discrimination. We queried GPT-4o-mini via the Application Programming Interface (API). We set the temperature to 0 and pinned the model to its 18 July 2024 version. We did this to minimize randomness (temperature) and limit variations in the underlying large language model (pinning). The prompt read: "The following are responses of people evaluating the fairness of different advertising strategies. To what extent does each response involve the concept of discrimination as part of the explanation? Discrimination is defined as 'differential treatment based on attributes that are irrelevant and/or uncontrollable.' Answer only with a number on a scale from 1 to 7 where 1 means 'not at all' and 7 means 'very much so.' Just give me a number from 1 to 7."

STUDY WA4: MEDIATION BY BELIEFS ABOUT DISCRIMINATION

Study WA4 Method

We recruited 1,390 Prolific users ($M_{\text{AGE}} = 40.34$; 702 women, 662 men, 24 other) for

Study 4 (aspredicted.org/2dx8-dshh.pdf), which employed a single-factor (targeting: race vs. gender vs. age vs. socioeconomic status vs. geography vs. behavioral vs. broad) between-subjects design.

All participants first read: "A company has developed a new product, which they believe will appeal to some customers more than others." The broad condition read: "They plan to advertise the product broadly to the general public." The targeting conditions read: "They plan to advertise the product to [members of a particular race/people of a particular gender/members of a particular age group/people of a particular socioeconomic status/people who live in a particular geographic region/people who have purchased similar products in the past], rather than broadly to the general public." We asked: "How fair is this plan?" ("Not at all fair" = 1; "Very fair" = 9).

On the next page, we asked participants to indicate the extent to which the strategy reflected seven potential descriptions (e.g., creative, discriminatory, efficient, impersonal, inefficient, rigid, and strategic). We randomized the order of the seven questions (i.e., one per descriptor): "Does advertising the product [based on race/based on gender/based on age/based on socioeconomic status/based on geography/based on similar past purchases/broadly to the general public] feel [creative/discriminatory/efficient/impersonal/inefficient/rigid/strategic]?" (for each: "Not at all" = 1; "Definitely" = 9). The descriptor of interest was "discriminatory," while the other six descriptors were decoys.

Study WA4 Results and Discussion

A fairness ANOVA revealed a main effect of targeting ($F(6, 1{,}383) = 64.02$, $p < .001$). Each of the five demographic targeting conditions was rated as less fair than broad advertising; the behavioral targeting and broad advertising conditions did not differ (Table WA1). A discrimination ANOVA revealed a main effect of targeting ($F(6, 1{,}383) = 83.80$, $p < .001$). Each of the five demographic targeting conditions was rated as more discriminatory than broad advertising; the behavioral targeting and broad advertising conditions did not differ (Table WA1).

TABLE WA2

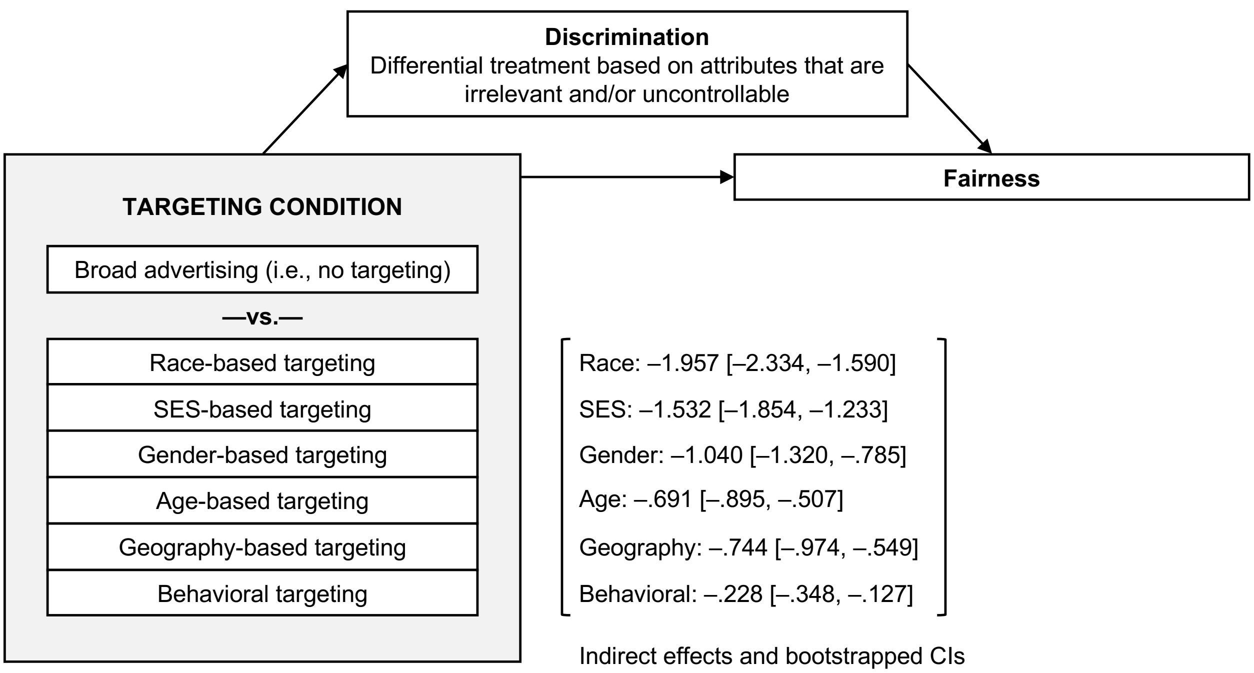
STUDY WA4: FAIRNESS AND DISCRIMINATION ACROSS VARIOUS TARGETING BASES

Targeting basis	Fairness		Targeting basis	Discrimination ratings	
Race	3.71 [3.47, 3.94]		Race	4.77 [4.51, 5.02]	
SES	4.57 [4.34, 4.80]	A	SES	4.22 [3.94, 4.49]	
Gender	4.72 [4.48, 4.97]	A	Gender	3.35 [3.07, 3.63]	A
Geography	5.03 [4.82, 5.24]		Geography	3.14 [2.87, 3.40]	AB
Age	5.39 [5.20, 5.58]		Age	3.00 [2.76, 3.24]	B
Behavioral	5.99 [5.86, 6.12]	B	Behavioral	2.14 [1.97, 2.31]	
Broad advertising	6.05 [5.89, 6.21]	B	Broad advertising	1.56 [1.41, 1.70]	

Note: Comparisons sharing a letter do not differ at p = .05. Brackets contain 95% CIs. Mediation results reflect bootstrapped CIs (10,000 resamples) for the indirect effect of targeting on fairness through discrimination. Conditions sorted from least to most fair, and most to least discriminatory

FIGURE WA2

STUDY WA4: MEDIATION MODEL



We next tested whether the discrimination ratings mediated differences in fairness between broad advertising and each of the six targeting bases. Specifically, we fit six separate mediation models, with 10,000 bootstrapped resamples (i.e., one per targeting basis), comparing each to broad advertising. Every model included fairness as the dependent variable, discrimination rating as the mediator, and a dummy-coded independent variable (0 = broad advertising; 1 = targeting condition). Discrimination ratings mediated the effect for each of the six targeting conditions, but to different degrees (Figure WA1).

STUDY WA5: STUDIES 5A–C PRETEST

Study WA5 Method

We recruited 541 Prolific workers (M_{age} = 37.60; 252 women, 284 men, 5 other; 399 White, 33 Black, 131 other) for Study WA5. All participants evaluated three out of six potential scenarios, in randomized order (each scenario was designed to test a potential moderator for stimuli selection), and were reassigned conditions within each scenario. The two scenarios of interest were the band-aid scenario, resembling Study 5A, and the snack scenario, resembling Studies 5B–C.

In the band-aid scenario, participants were assigned to one of two conditions (justification: average physical differences vs. preferences). All participants read: “A company has developed a new line of Band-Aids that were designed to appeal to their Black customers.” Participants in the preferences justification condition read: “In particular, the Band-Aids are available in a variety of patterns and designs that were tested to appeal to the preferences of their Black customers.” Participants in the average physical differences justification condition read: “In particular, the Band-Aids are available in a variety of darker shades that were tested to match the skin tones of their Black customers.” All participants then read: “The company is developing a marketing plan to promote the new Band-Aids. Because they believe the products are best suited to their Black customers, they will advertise them to Black people directly."

In the snack scenario, participants were assigned to one of three conditions (justification: preferences vs. average physical differences vs. market research). The study was designed so the preferences justification condition would separately serve as the comparison group for both the average physical differences justification condition (Study 5B) and the objective data justification condition (Study 5C).

Participants in the preferences and objective data justification conditions first read: "A snack foods company has developed a new line of snacks. Initial testing showed that, due to the taste and texture profile of the snacks, the snacks are better suited to the preferences of their female customers." Participants in the objective data justification condition additionally read: "Specifically, testing shows that approximately 80% of women like the taste and texture, compared with only 30% of men." Participants in the average physical differences justification condition read: "A snack foods company has developed a new line of snacks. Initial testing showed that, due to the vitamin and nutrient profile of the snacks, the snacks are better suited to the biological needs of their female customers." All participants then read: "The company is developing a marketing plan to promote the snacks. Because they believe the snacks are best suited to their female customers, they will advertise the snacks to women directly."

In all scenarios, participants responded to four questions in counterbalanced order: "The company is advertising to customers based on factors that are under their personal control;" "The company is advertising to customers based on factors that are easily changeable;" "The company is advertising to customers based on factors that are relevant to this product;" and "The company is advertising to customers based on factors that are a signal of their preferences (i.e., their wants and needs)" ("Strongly disagree" = 1; "Strongly agree" = 9).

Study WA5 Results and Discussion

For each scenario, we averaged the two control measures (Band-aid: α = .81; snack: α = .76) and the two relevance measures (Band-aid: α = .60; snack: α = .63). Both justification by average physical differences and justification by objective data improved perceptions of relevance, but did not affect perceptions of controllability (Table WA3).

TABLE WA3

STUDIES 5A–C PRETEST: JUSTIFICATIONS BASED ON AVERAGE PHYSICAL DIFFERENCES AND OBJECTIVE DATA IMPROVE PERCEPTIONS OF RELEVANCE

Band-aid scenario (5A)	Justification by average physical differences	Justification by differences in preferences	Test statistic	Effect size
Relevance	7.31	6.42	$F(1, 271) = 19.80$, $p < .001$	$d = .53$
Controllability	4.37	4.46	$F(1, 271) = 0.06$, $p = .807$	$d = .03$
Snack scenario (5B)	Justification by average physical differences	Justification by differences in preferences	Test statistic	Effect size
Relevance	6.65	5.92	$F(1, 359) = 15.67$, $p < .001$	$d = .41$
Controllability	4.25	4.53	$F(1, 359) = 1.26$, $p = .262$	$d = .12$
Snack scenario (5C)	Justification based on objective data	No justification based on objective data	Test statistic	Effect size
Relevance	6.57	5.92	$F(1, 358) = 11.68$, $p < .001$	$d = .36$
Controllability	4.56	4.53	$F(1, 358) = 0.02$, $p = .888$	$d = .02$

STUDIES WA6–8: SENSITIVITY TO OBJECTIVE DATA

Studies WA6–8 further probe the role of justification by market research by capturing sensitivity to absolute and relative differences in interest across the targeted and non-targeted segments.

Studies WA6–8 Method

We recruited 747 MTurk workers (M_{age} = 37.91; 367 women, 369 men, 11 other) for Study WA6 (https://aspredicted.org/NDV_B8H), which held constant the interest of the non-targeted segment at 15% and varied the interest of the targeted segment in a four-cell between-subjects design (interest level: 80% vs. 60% vs. 40% vs. 20%). We recruited 736 MTurk workers (M_{age} = 37.91; 442 women, 288 men, 6 other) for Study WA7 (https://aspredicted.org/W3D_XVQ), which held constant the interest of the targeted segment at 85% and varied the interest of the non-targeted segment in a four-cell between-subjects design (interest level: 80% vs. 60% vs. 40% vs. 20%). We recruited 373 MTurk workers (M_{age} = 36.28; 206 women, 163 men, 4 other) for Study WA8 (https://aspredicted.org/86H_FPC), which employed a 2 (target interest: high-85% vs. medium-55%) × (non-target interest: medium-45% vs. low-15%) between-subjects design.

In all studies, participants first read: "A snack foods company has developed a new line of snacks. Initial testing showed that, due to the taste and texture profile of the snacks, the snacks are better suited to the preferences of their female customers. Specifically, testing shows that approximately [XX]% of women like the taste and texture, compared with only [XX]% of men." The exact percentages presented to participants varied according to the conditions described in each paradigm above. Participants then answered: "How fair is this advertising plan?" ("Not at all fair" = 1; "Very fair" = 9).

Studies WA6–8 Results and Discussion

For Study WA6, a fairness ANOVA revealed a significant effect of interest level ($F(3, 743) = 7.29$, $p < .001$), such that larger disparities in interest between the targeted and non-targeted segments improved fairness perceptions. For Study WA7, a fairness ANOVA revealed a significant effect of interest level ($F(3, 732) = 11.83$, $p < .001$), again such that larger disparities in interest between the targeted and non-targeted segments improved fairness perceptions. For Study WA8, a fairness ANOVA revealed main effects of target interest ($F(1, 369) = 16.37$, $p <$ .001) and non-target interest ($F(1, 369) = 4.34$, $p = .038$), conceptually replicating Studies WA6–7, but no interaction ($F(1, 369) = 2.26$, $p = .133$).

Notably, across Studies WA6–8, the decision to target was regarded as relatively fairer when there were large and discernable differences in reported preferences between the targeted and non-targeted segments (Table WA5). However, when such differences were small—in all cases where reported preferences differed by 5% or 10%, for example—the targeting decision was seen as less fair.

TABLE WA4

STUDIES WA6–8: LARGE AND DISCERNABLE DIFFERENCES IN REPORTED PREFERENCES BETWEEN THE TARGET AND NON-TARGETED SEGMENTS IMPROVED FAIRNESS PERCEPTIONS

Study WA6			
Targeted segment interest	Non-targeted segment interest	Fairness	
80%	15%	6.28 [5.98, 6.58]	A
60%	15%	6.35 [6.05, 6.65]	A
40%	15%	6.03 [5.73, 6.32]	A
20%	15%	5.44 [5.14, 5.74]	B

Study WA7			
Targeted segment interest	Non-targeted segment interest	Fairness	
85%	20%	6.10 [5.79, 6.42]	A
85%	40%	5.96 [5.64, 6.27]	A
85%	60%	5.70 [5.38, 6.02]	A
85%	80%	4.86 [4.54, 5.18]	B

Study WA8			
Targeted segment interest	Non-targeted segment interest	Fairness	
85%	45%	6.41 [6.01, 6.81]	A
85%	15%	6.53 [6.14, 6.91]	A
55%	45%	5.31 [4.92, 5.70]	B
55%	15%	6.02 [5.63, 6.41]	A

Note: Comparisons sharing a letter do not differ at $p = .05$. Brackets contain 95% CIs

STUDY WA9: STUDIES 7A–B PRETEST

Study WA9 Method

We recruited 177 Prolific workers (M_{age} = 38.95; 85 women, 91 men) for Study WA9. Participants were assigned to either respond to questions about company size or industry norms. All participants first read a description of discrimination: "Discrimination is defined as differential treatment based on attributes that are irrelevant and/or uncontrollable. We then wrote: "Suppose a company engages in discriminatory behavior, such as targeting advertisements to some demographic groups and not others." Participants assigned to respond to the company size questions then answered two questions in counterbalanced order: "If this company were large, with virtually unlimited resources at their disposal, how intentional would their behavior seem?" and "If this company were small, with very limited resources at their disposal, how intentional would their behavior seem?" ("Definitely not intentional" = 1; "Definitely intentional" = 9). Participants assigned to respond to the industry norm questions instead answered: "If this were standard in the industry, to what extent would you consider the company's behavior intentional?" and "If this were NOT standard in the industry, to what extent would you consider the company's behavior intentional?" on the same nine-point scale.

Study WA9 Results

Perceived intentionality was lower for small companies (M = 5.64, 95% CI = [5.17, 6.11) than for large companies (M = 7.62, 95% CI = [7.21, 8.03]), $t(88) = 7.38$, $p < .001$, $d = .78$), and when it was standard practice (M = 6.60, 95% CI = [6.19, 7.01]) than when it was not (M = 7.34, 95% CI = [6.98, 7.70]), $t(87) = 2.73$, $p = .008$, $d = .29$).

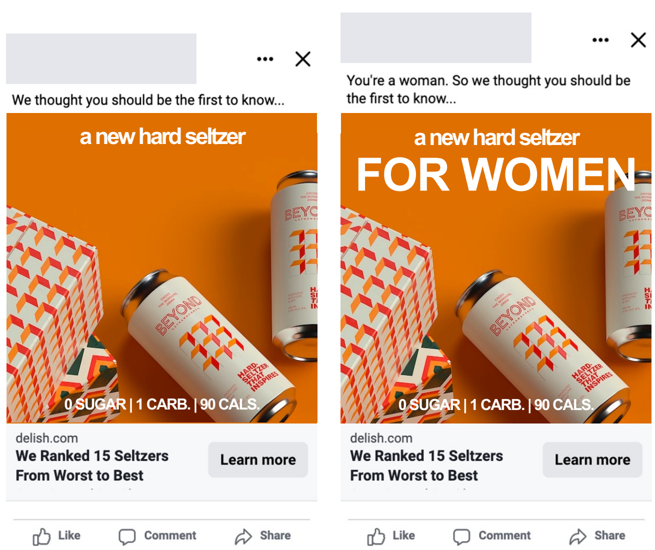
STUDY WA10: PRETEST FOR STUDY 8A

Study WA10 Method

We recruited 348 MTurk workers (M_{age} = 40.51; 348 women) for Study WA11, which employed a single-factor (targeting: gender vs. broad) between-subjects design. All participants first read: "Suppose you're browsing social media and come across the following ad. Please review the ad and answer the following questions." We then displayed one of the two advertisements used in Study 8A (Figure WA1) and asked three questions, each on a separate page: "If you were actually shown this ad on social media, would you think about why you were seeing this ad?" ("Definitely not" = 1; "Definitely" = 9); "If you were actually shown this ad on social media, would it be clear to you why you were seeing this ad?" ("Definitely not" = 1; "Definitely" = 9); and "Who do you think is seeing this ad?" ("Mostly men," "Mostly women," or "Both men and women").

FIGURE WA3

STUDY WA10: BROAD ADVERTISING (LEFT) AND GENDER TARGETING (RIGHT)



On the final page, participants rated relevance and controllability, indicating their agreement with the following four counterbalanced statements: "This company is advertising to customers based on factors that are [under their personal control/easily changeable/relevant to this product/a signal of their preferences]" ("Definitely disagree" = 1; "Definitely agree" = 9).

Study WA10 Results

Relative to those in the broad targeting condition, participants in the gender targeting condition were more likely to think about why they were seeing the ad (M_{GENDER} = 5.76, 95% CI [5.40, 6.12] vs. M_{BROAD} = 5.14, 95% CI [4.76, 5.52]; $F(1, 346) = 5.45$, $p = .020$), more likely to believe that they knew why they were seeing it (M_{GENDER} = 4.98, 95% CI [4.62, 5.34] vs. M_{BROAD} = 3.67, 95% CI [3.30, 4.05]; $F(1, 346) = 24.46$, $p < .001$), and more likely to anticipate that the ad was targeted to women (vs. men; M_{GENDER} = 90.7% vs. M_{BROAD} = 42.8%; $\chi^2(2) = 89.22$, $p < .001$). We next averaged the two measures of relevance ($\alpha = .75$) and the two measures of controllability ($\alpha = .68$). Both relevance (M_{GENDER} = 4.93, 95% CI [4.64, 5.21] vs. M_{BROAD} = 5.81, 95% CI [5.51, 6.11]; $F(1, 346) = 17.64$, $p < .001$) and controllability (M_{GENDER} = 3.71, 95% CI [3.44, 3.99] vs. M_{BROAD} = 5.04, 95% CI [4.76, 5.33]; $F(1, 346) = 44.22$, $p < .001$) were lower in the gender condition.

STUDY WA11: STUDY 8B PRETEST

Study WA11 Method

We recruited 143 Prolific workers (M_{age} = 38.01; 84 women, 57 men) for Study WA11. All participants read: "Suppose a seltzer contains additional vitamins in the form of carotene, calcium, and collagen. Do you believe these vitamins would be more relevant to men or to women?" ("Definitely men" = 1, "Equally relevant/no difference" = 5, "Definitely women" = 9).

Study WA11 Results

Participants viewed the vitamins as more relevant to women than men ($M = 6.09$, 95% CI [5.86, 6.32]; comparison to scale midpoint: $t(142) = 9.46$, $p < .001$).

STUDY WA12: PRODUCT HARM (STUDENT VS. PAYDAY LOANS)

Study WA12 Method

We recruited 988 MTurk workers ($M_{age} = 45.38$; 437 women, 544 men, 5 other, 2 undisclosed) for Study WA12 (https://aspredicted.org/5x9j-yndz.pdf), which employed a 2 (targeting: race vs. broad) × (loan type: student vs. payday) between-subjects design.

Participants read that a financial institution was offering either "student loans" that "must be repaid over 5–15 years and carry an annualized interest rate of 5%" or "payday loans" that "must be repaid over 2–4 weeks and carry an annualized interest rate of 500%." In the broad condition, participants read: "They plan to advertise the [student/payday] loans broadly to the general public, rather than to a particular group." In the race condition, participants read: "They plan to advertise the [student/payday] loans to members of a particular race, rather than broadly to the general public." We presented the three counterbalanced measures capturing fairness: "How [fair/appropriate/acceptable] is this advertising strategy?" ("Not at all [fair/appropriate/acceptable]" = 1; "Very [fair/appropriate/acceptable]" = 9). On the final page, as a manipulation check, we asked: "How harmful are the types of loans described on the previous page?" ("Not at all harmful" = 1; "Very harmful" = 9).

Study WA12 Results and Discussion

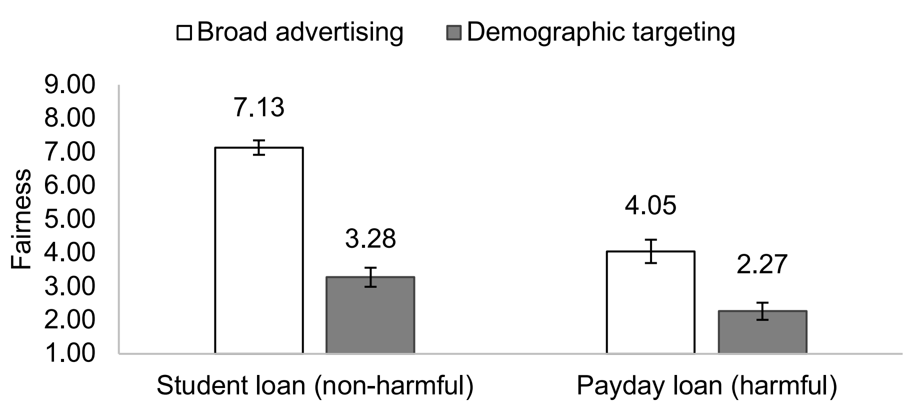
Confirming the effect of the manipulation, participants believed the payday loan was more harmful ($M = 8.11$, 95% CI = [7.97, 8.24]) than the student loan ($M = 4.97$, 95% CI =

[4.75, 5.18]; $t(986) = 24.43$, $p < .001$, $d = 2.51$).

We averaged the three fairness measures ($\alpha = .98$) to form a composite. A fairness ANOVA revealed a main effect of targeting ($F(1, 984) = 389.40$, $p < .001$; Figure WA4), a main effect of loan type ($F(1, 984) = 205.80$, $p < .001$), and an interaction ($F(1, 984) = 52.77$, $p < .001$). Decomposition revealed a simple effect of targeting in the student loan condition ($F(1, 984) = 365.15$, $p < .001$), such that fairness perceptions were lower in the race condition ($M = 3.28$, 95% CI = [2.99, 3.56]) than in the broad condition ($M = 7.13$, 95% CI = [6.92, 7.35], $d = 1.38$). However, this simple effect of targeting was attenuated in the payday loan condition ($F(1, 984) = 77.59$, $p < .001$), such that fairness perceptions were lower in the race condition ($M = 2.27$, 95% CI = [2.01, 4.39]) and the broad condition ($M = 4.05$, 95% CI = [3.70, 4.39], $d = .68$). Within just the broad condition, fairness perceptions were higher for the student loan than the payday loan ($F(1, 984) = 233.03$, $p < .001$). Within just the race condition, fairness perceptions were higher for the student loan than the payday loan ($F(1, 984) = 25.12$, $p < .001$).

FIGURE WA4

STUDY WA12: MODERATION BY PRODUCT HARM (95% CONFIDENCE INTERVALS)



STUDY WA13: PRODUCT HARM (HIGH VS. LOW APR LOANS)

Study WA13 Method

We recruited 778 MTurk workers (M_{age} = 43.84; 320 women, 428 men, 7 other, 23 undisclosed) for Study WA13 (https://aspredicted.org/6gbb-5sk6.pdf), which employed a 2 (targeting: race vs. broad) × (APR: low vs. high) between-subjects design.

Participants read: "A bank is offering a new line of personal loans, which can be used to fund anything from debt consolidation to home improvement projects to major expenses like weddings and vacations." In the low APR condition, we explained "the loans carry an annual percentage rate (APR) of 7.5%." In the high APR condition, we explained "the loans carry an annual percentage rate (APR) of 38.5%." In the broad condition, participants read: "They plan to advertise the loans broadly to the general public, rather than to a particular group." In the race condition, participants read: "They plan to advertise the loans to members of a particular race, rather than broadly to the general public." We presented the three counterbalanced measures capturing fairness: "How [fair/appropriate/acceptable] is this advertising strategy?" ("Not at all [fair/appropriate/acceptable]" = 1; "Very [fair/appropriate/acceptable]" = 9). On the final page, as a manipulation check, we asked: "How harmful are the types of loans described on the previous page?" ("Not at all harmful" = 1; "Very harmful" = 9).

Study WA13 Results and Discussion

Confirming the effect of the manipulation, participants believed the high APR loan was more harmful (M = 7.25, 95% CI = [7.04, 7.46]) than the low APR loan (M = 4.70, 95% CI = [4.44, 4.95]; $t(776) = 15.21$, $p < .001$, $d = .96$).

We averaged the three fairness measures ($\alpha = .99$) to form a composite. A fairness ANOVA revealed a main effect of targeting ($F(1, 774) = 611.77$, $p < .001$; Figure WA3), a main effect of APR ($F(1, 774) = 130.27$, $p < .001$), and an interaction ($F(1, 774) = 79.07$, $p < .001$). Decomposition revealed a simple effect of targeting in the low APR condition ($F(1, 774) = 573.13$, $p < .001$), such that fairness perceptions were lower in the race condition ($M = 2.62$, 95% CI = [2.33, 2.90]) than in the broad condition ($M = 7.75$, 95% CI = [7.52, 7.97], $d = 1.63$). However, this simple effect of targeting was attenuated in the high APR condition ($F(1, 774) = 123.80$, $p < .001$), such that fairness perceptions were lower in the race condition ($M = 2.23$, 95% CI = [1.98, 2.48]) and the broad condition ($M = 4.65$, 95% CI = [4.23, 5.07], $d = .91$). Within just the broad condition, fairness perceptions were higher for the low APR than the high APR loan ($F(1, 774) = 204.34$, $p < .001$). Within just the race condition, fairness perceptions were marginally higher for the low APR loan than the high APR loan ($F(1, 774) = 3.21$, $p = .074$).

FIGURE WA5

STUDY WA13: MODERATION BY PRODUCT HARM (95% CONFIDENCE INTERVALS)