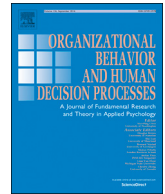




Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Free-riding and cost-bearing in discrimination

Xilin Li, Christopher K. Hsee*

The University of Chicago Booth School of Business, United States

A B S T R A C T

We study how the temporal positions in which a disadvantaged person (e.g., an unattractive-looking customer) and an advantaged person (e.g., an attractive-looking customer) encounter an actor (e.g., a vendor) influence the treatment they get from the actor (e.g., the prices the vendor offers). Three experiments, including a field experiment and a pre-registered experiment, incorporate three types of personal attributes (physical appearance, nationality, and gender) and find both a *free-riding effect* for the disadvantaged person and a *cost-bearing effect* for the advantaged person. Specifically, the disadvantaged person receives better treatment by following the advantaged person, and the advantaged person receives worse treatment by following the disadvantaged person. These effects occur only if the attribute that differentiates the disadvantaged and advantaged persons is perceived as unjustifiable, and they disappear if the attribute is perceived as justifiable, suggesting that these effects are due not to anchoring, but to the actor's need for justifiability. This research highlights the importance of choice architecture in discrimination and its prevention.

1. Introduction

An unattractive-looking customer and an attractive-looking customer are both strolling through a bazaar where prices are negotiable. Both customers are interested in buying a certain type of jacket from a vendor. Consider two alternative scenarios. In one, the unattractive-looking person approaches the vendor first and asks for the lowest price; afterward, the attractive-looking person approaches the vendor and asks the same question. In the other, the sequence of the two persons is reversed. Does sequence make a difference in the vendor's offer? If so, will the unattractive-looking person get a better (lower) price if he/she is the first or the second to encounter the vendor? What about the attractive-looking person?

More generally, suppose that two persons are similar except for one attribute (e.g., physical appearance), and this attribute is perceived as an unjustifiable ("should-not-consider") factor in the given situation. One person is relatively disadvantaged on the attribute (e.g., unattractive-looking), and one person is relatively advantaged (e.g., attractive-looking). The disadvantaged person and the advantaged person may encounter an actor (e.g., a vendor) in one of two sequences: the disadvantaged-person-first sequence, in which the disadvantaged person encounters the actor first and the advantaged person encounters the actor second, or the advantaged-person-first sequence, in which the advantaged person encounters the actor first and the disadvantaged person encounters the actor second. Does the sequence make a difference in the treatment of the two persons? If so, will the disadvantaged person get better treatment (e.g., a better price) in the disadvantaged-person-first sequence or in the advantaged-person-first sequence? What

about the advantaged person?

These questions are about choice architecture (Thaler & Sunstein, 2008), and in particular, they concern the temporal positions of the target persons. These questions are relevant not only to consumer price inquiries, but also to other domains in which multiple target persons are evaluated sequentially—for instance, job candidates may be interviewed sequentially, and faculty promotion cases may be voted on sequentially. Furthermore, these questions pertain to discrimination regarding not only physical appearance, but also other attributes, such as nationality, ethnicity, and gender. In this research, we use the word "discrimination" to mean differential treatment of individuals who are similar except for one personal attribute (such as physical appearance) that is generally perceived as an unjustifiable consideration in the given decision context.

To address the above questions regarding sequence, let us first consider what treatment each target person will receive if he/she is the first person to encounter the actor. Existing research indicates that discrimination is still widespread despite policies against it and education about it (e.g., Ayres, 1991; Bertrand & Mullainathan, 2004; Correll, Benard, & Paik, 2007; Eagly, Makhijani, & Klonsky, 1992; Langlois et al., 2000; Sellers & Shelton, 2003). For example, science faculty still favor male students over similarly competitive female students (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012), and employers still favor white job candidates over their African Americans counterparts (Bertrand & Mullainathan, 2004). In the context of the current research, the existing literature on discrimination suggests that a disadvantaged person who encounters an actor first will receive worse treatment than an advantaged person who encounters the

* Corresponding author at: Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, United States.

E-mail address: Chris.Hsee@ChicagoBooth.edu (C.K. Hsee).

<https://doi.org/10.1016/j.obhdp.2019.08.003>

Received 11 July 2018; Received in revised form 6 July 2019; Accepted 15 August 2019

0749-5978/© 2019 Elsevier Inc. All rights reserved.

same actor first. For example, an unattractive-looking customer who encounters a vendor first will receive a worse price than an attractive-looking customer who encounters the same vendor first.

Will the disadvantaged person still receive worse treatment than the advantaged person in the 2nd position? Perhaps not, research suggests that people want their decisions to be consistent and justifiable (Crandall & Eshleman, 2003; Hsee, 1995; Lerner & Tetlock, 1999; Schuman, Kalton, & Ludwig, 1983; Simonson & Tversky, 1992). For example, after a person has agreed to comply with a (small) request to support a certain cause, the person is more likely to comply with another (larger) request to support the same cause (Cialdini, 2001; Freedman & Fraser, 1966; Petrova, Cialdini, & Sills, 2007). This foot-in-the-door phenomenon reflects people's need for consistency, which, in turn, stems from their need for justification, as it would be unjustifiable to make inconsistent or contradictory decisions.

In the context of the present research, the need for justifiability means that once an actor has treated an advantaged person favorably, the actor will find it unjustifiable to treat a disadvantaged person differently, and therefore will also treat the disadvantaged person favorably. For example, once a vendor has offered an attractive-looking customer a good price, the vendor will find it unjustifiable to treat an unattractive-looking customer differently, and therefore will also offer the unattractive-looking customer a similarly good price. The same logic applies if we reverse the order: once an actor has treated a disadvantaged person unfavorably, the actor will also treat an advantaged person unfavorably. Together, this analysis predicts that the discrimination (the worse treatment of the disadvantaged person relative to the advantaged person) in the 1st position will disappear or even reverse in the 2nd position.

The position effect proposed above is consistent with previous findings that actors show less discrimination in joint evaluation than in single evaluation, because joint evaluation makes the unjustifiable attribute salient (Bohnet, Van Geen, & Bazerman, 2015; Li & Hsee, forthcoming; see also Bazerman, Tenbrunsel, & Wade-Benzoni, 1998; Hsee, 1996a; Hsee & Zhang, 2010; Moore, 1999). Although the actor in our research is never in a strict joint evaluation mode, the actor is relatively *more* in the joint evaluation mode when evaluating the second person than when evaluating the first person.

In summary, based on existing research on discrimination, we assume that a disadvantaged target person will receive worse treatment than an advantaged target person in the 1st position (i.e., when each target person is the first to encounter an actor). From this assumption and the literature on justifiability, we propose our main hypothesis:

Hypothesis 1 (position effect). The discrimination seen in the 1st position will disappear or reverse in the 2nd position (i.e., when each target person is the second to encounter the actor).

The general position effect proposed in Hypothesis 1 implies two more specific position effects, one about the disadvantaged person and one about the advantaged person. If the disadvantaged person receives worse treatment than the advantaged person as the first to encounter the actor, but similar treatment to the advantaged person as the second to encounter the actor, then the disadvantaged person will receive better treatment as the second person to encounter the actor than as the first person to encounter the actor. For example, an unattractive-looking customer will receive a better price by asking a vendor after an attractive-looking customer first asks the vendor than by approaching the vendor first. We refer to this phenomenon as the *free-riding effect*. The opposite applies to the advantaged person, who will be treated worse as the second person to encounter the actor than as the first person to encounter the actor. We refer to this phenomenon as the *cost-bearing effect*—the reverse of the *free-riding effect*.

Hypothesis 1A (free-riding effect). The disadvantaged person will receive better treatment in the 2nd position than in the 1st position.

Hypothesis 1B (cost-bearing effect). The advantaged person will

receive worse treatment in the 2nd position than in the 1st position.

Hypothesis 1, along with Hypothesis 1A and Hypothesis 1B, also yields the following conclusion: the advantaged-person-first sequence dominates the disadvantaged-person-first sequence. In other words, the advantaged-person-first sequence is better than the disadvantaged-person-first sequence for both target persons, because the advantaged-person-first sequence improves the treatment of the disadvantaged person (by positioning her after the advantaged person) without hurting the advantaged person (because the advantaged person is the first to encounter the actor), whereas the disadvantaged-person-first sequence hurts the treatment of the advantaged person (by positioning her after the disadvantaged person) without helping the disadvantaged person (because the disadvantaged person is the first to encounter the actor).

According to our theory, a key reason for the position effect proposed in Hypothesis 1 is that the actor finds it unjustifiable to treat the two persons inconsistently. That is, the actor in the advantaged-person-first sequence finds it unjustifiable to treat the disadvantaged person less favorably than he/she has just treated the advantaged person, and the actor in the disadvantaged-person-first sequence finds it unjustifiable to treat the advantaged person more favorably than he/she has just treated the disadvantaged person. This theory suggests that if the attribute that differentiates the target persons is a justifiable factor in the given decision context, then the actor will no longer feel obligated to treat the two persons consistently, and the differential treatment of the disadvantaged person relative to the advantaged person in the 1st position will persist in the 2nd position. In other words, the position effect postulated in Hypothesis 1 will disappear. This is our next hypothesis:

Hypothesis 2 (justifiability effect). The position effect postulated in Hypothesis 1 will disappear if the attribute that differentiates the target persons is justifiable.

Hypothesis 2 is important for two reasons. First, it directly tests the underlying mechanism of the proposed position effect. Second, it distinguishes our justifiability-based account for the position effect from anchoring as an alternative explanation (Epley & Gilovich, 2006; Jacobowitz & Kahneman, 1995; Tversky & Kahneman, 1974). According to the anchoring alternative explanation, the actor treats the second person similarly to the first person not because the actor finds it unjustifiable to treat the two persons differently, but because the actor anchors his/her treatment of the second person to his/her treatment of the first person. If that is the case, then manipulation of the justifiability of the focal attribute should not moderate the proposed position effect postulated in Hypothesis 1. If the manipulation does moderate the position effect, as Hypothesis 2 posits, then anchoring cannot be a viable alternative explanation.

Next, we report three studies that tested our hypotheses. Study 1—a field experiment—concerned appearance discrimination in a purchase context and tested Hypothesis 1 (our main hypothesis). Study 2 concerned nationality discrimination in a rental context, manipulated justifiability, and tested all hypotheses. Study 3—a pre-registered study—concerned gender discrimination in a disciplinary punishment context, and it also manipulated justifiability and tested all hypotheses. (In addition, we conducted another study concerning ethnicity discrimination in college admissions. Because that study was similar to Study 3, we report it in an appendix instead of in the main text; see Appendix A.)

2. Study 1

Study 1 was a field experiment that concerned physical appearance discrimination in price negotiations, as in the opening example. Evidence abounds that unattractive-looking individuals receive worse treatment than attractive-looking individuals (e.g., Cryder, Botti, & Simonian, 2017; Dion, Berscheid, & Walster, 1972; Hamermesh &

Biddle, 1994; Langlois, Ritter, Roggman, & Vaughn, 1991; Langlois et al., 2000; Mobius & Rosenblat, 2006; Slater et al., 1998; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Study 1 tested whether temporal positions influence appearance discrimination. The study consisted of two between-participants sequence conditions: disadvantaged-person-first and advantaged-person-first.

2.1. Method

The study took place at a large clothing market located on the east coast of China over a two-day period during the hours when most stores were open (10:30 am–12:00 pm; 1:00 pm–5:30 pm). Comprising over two thousand independent vendors, the market primarily sells leather products such as jackets, shoes, and purses. Prices for most products are negotiable.

This study involved two female confederates of similar ages who were blind to the purpose of the experiment. In a pretest ($N = 57$), one confederate was rated as significantly less attractive-looking than the other on a 7-point scale on which 1 meant “unattractive-looking” and 7 meant “attractive-looking” ($M = 1.91$, $SD = 0.81$, versus $M = 5.30$, $SD = 0.89$, $t(56) = -19.51$, $p < .001$). During the study, the two confederates wore similar outfits, including a black down jacket and a pair of blue jeans.

The study consisted of two between-participants sequence conditions: disadvantaged-person-first and advantaged-person-first. In the disadvantaged-person-first condition, the unattractive-looking confederate first walked into a store, looked around, picked up a woman’s leather jacket, went to the vendor, and used a scripted line to ask the vendor for the price: “Hi, how much is this leather jacket?” After the vendor offered an initial price, the confederate said, “Oh, that’s too expensive. What’s the lowest price you can give me?” After the vendor offered the final price, the confederate said, “I need to think about it,” and then stepped away from the vendor. While the unattractive-looking confederate was talking to the vendor, the attractive-looking confederate entered the store while talking on the phone. She always kept a distance from the unattractive-looking confederate and always talked on the phone, thus giving the vendor the impression that she could not hear the vendor’s conversation with the unattractive-looking confederate. After the unattractive-looking confederate walked away, the attractive-looking confederate slowly wandered toward the jacket, picked it up, and asked the vendor the same questions as the unattractive-looking confederate had asked. (Within a given store, the two confederates always picked up the same jacket; across different stores, the confederates tried to pick similar jackets.)

The procedure in the advantaged-person-first condition was the same as in the disadvantaged-person-first condition except that the order of the confederates was reversed. We alternated between the disadvantaged-person-first and the advantaged-person-first conditions every hour.

Over the two-day period, the two confederates approached 62 vendors. Of these approached vendors, 54 (87.1%) were female, and we found no meaningful gender differences.

2.2. Results and discussion

Fig. 1 presents the results. Two notes are in order. First, although the study consisted of only two between-participants sequence conditions (disadvantaged-person-first vs. advantaged-person-first), we present the results in a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) format for ease of exposition; see Appendix B for a detailed explanation of the data presentation format. Second, the study had two dependent variables—initial price and final price. Because these two prices were highly correlated ($r = 0.937$, $p < .001$), we based our analyses on the average of the two prices. (Analyzing the two prices separately produced essentially the same results.) When we say “prices” below, we mean the average of the two

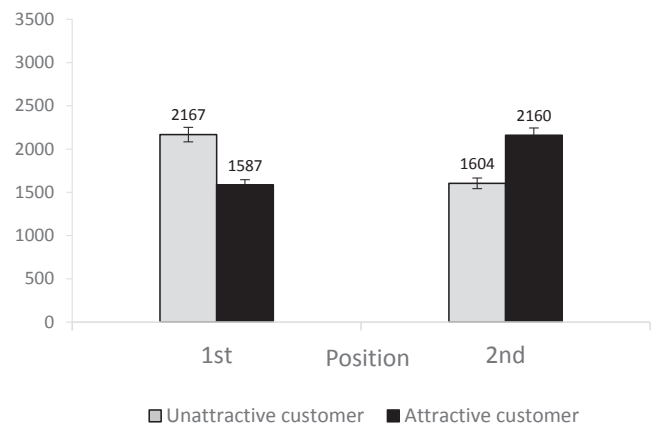


Fig. 1. Study 1 Results: Price in Chinese Yuan. (Taller columns mean higher prices, i.e., worse treatment; the error bars represent $\pm$ one SE). Note: The columns represent the averages of the initial and the final prices (in RMB) offered to the customers. The mean initial prices for the four columns from left to right are 2297.42 ($SD = 510.60$), 1779.35 ($SD = 403.83$), 1790.97 ($SD = 414.88$), and 2297.42 ($SD = 510.60$), respectively, and the mean final prices for the four columns from left to right are 2036.45 ($SD = 431.73$), 1394.52 ($SD = 274.83$), 1416.77 ($SD = 289.42$), and 2023.55 ($SD = 428.79$), respectively.

prices.

We first compared the results of the target persons in the 1st position and found significant discrimination: the unattractive-looking confederate received significantly worse prices when inquiring first than the attractive-looking confederate received when inquiring first, $t(60) = 5.64$, $p < .001$.

To test Hypothesis 1, we conducted a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA on prices, and we found the predicted interaction, $F(1, 60) = 30.32$, $p < .001$. The discrimination observed in the 1st position reversed in the 2nd position; the unattractive-looking confederate now received better prices than the attractive-looking confederate, $t(60) = 5.36$, $p < .001$.

To test Hypothesis 1A, we compared the prices received by the unattractive person in the 1st versus in the 2nd position, and we found the hypothesized free-riding effect: she received better prices when inquiring second than when inquiring first, $t(60) = 5.41$, $p < .001$. To test Hypothesis 1B, we compared the prices received by the attractive person in the 1st versus in the 2nd position, and we found the cost-bearing effect: she received worse prices when inquiring second than when inquiring first, $t(60) = 5.60$, $p < .001$.

This field experiment supported our hypotheses: When inquiring first, the unattractive-looking confederate received worse prices than the attractive-looking confederate; when inquiring second, the attractive-looking confederate received better prices than the attractive-looking confederate. In particular, the unattractive-looking confederate enjoyed a free-riding effect, while the attractive-looking confederate experienced a cost-bearing effect. Of note, although our findings may seem obvious in hindsight, they are not ex ante. We asked a group of customers shopping at the same clothing market whether they would go first or second if they were attractive-looking versus unattractive-looking, and found that most of them preferred the opposite sequence to what would be beneficial to them.

3. Study 2

Study 2 sought to extend the findings of Study 1 to another domain and another type of discrimination. While Study 1 tested Hypothesis 1 (position effect), Hypothesis 1A (free-riding effect), and Hypothesis 1B (cost-bearing effect), Study 2 also tested Hypothesis 2 (justifiability

Table 1
Stimuli in Study 2.

Justifiability condition	Target person
Low-justifiability	Disadvantaged = Saudi tenant making some noise Advantaged = Canadian tenant making some noise
High-justifiability	Disadvantaged = Australian tenant making a lot of noise Advantaged = Australian tenant making a little noise

effect) by including both a low-justifiability and a high-justifiability condition.

Participants assumed the role of a homeowner (the actor) and encountered and quoted rents to two prospective tenants (the target persons), one disadvantaged and one advantaged. We manipulated the order of the encounters. We also manipulated justifiability by varying the type of attribute that differentiated the tenants. In the low-justifiability condition, the tenants differed in nationality: either Saudi or Canadian. Based on existing research showing discrimination against Arabs (Ibish, 2003; Moradi & Hasan, 2004), we considered the Saudi tenant disadvantaged relative to the Canadian tenant. In the high-justifiability condition, the tenants differed in the noise they would make: either a lot (relatively disadvantaged) or a little (relatively advantaged). See Table 1 for a summary. We verified the justifiability manipulation through a pretest, as we report later.

3.1. Method

Study 2 included two between-participants justifiability conditions (low vs. high), and each justifiability condition included two sequence conditions (disadvantaged-person-first vs. advantaged-person-first). (In addition to the four conditions, we also ran a strict joint-evaluation condition using the same scenario. Because that condition was not designed to test any of the hypotheses in this research, we defer the description of that condition to the General Discussion section.)

Participants in the main study (not including the additional joint-evaluation condition) were 400 workers recruited from MTurk in the US (221 females, $M_{\text{age}} = 38.82$). (See Porter, Outlaw, Gale, & Cho, 2019 for pros and cons of using MTurkers as research participants.) In all the conditions, participants were asked to imagine the following:

Your house has a fully furnished basement, but you don't really need it and plan to rent it out. You are sensitive to noise and hope that the tenant is quiet at night. You posted a rental ad in a local newspaper, asking prospective tenants to check out the basement this Sunday and inquire about the rent. You are flexible about how much rent to charge per month. It could range from less than \$400 to more than \$1000. The more you are willing to rent the basement to someone, the less you would charge; conversely, the less you are willing to rent it to someone, the more you would charge. It is Sunday now. The first prospective tenant arrives at your place to see the basement.

In the low-justifiability/disadvantaged-person-first condition, participants continued to read the following about the first tenant (disadvantaged):

He is a middle-aged man. You have learned through a casual conversation that he is Saudi Arabian, born and raised in Riyadh. He is now in the US on a business visa. He sometimes talks to friends on the phone at night and makes some mild noise.

The participants then read, "He asks you how much the rent is. How much would you tell him?" They responded by choosing one of four options: (a) \$400 or less, (b) \$600, (c) \$800, or (d) \$1000 or more. (When analyzing data, we treated the first option as \$400 and the last option as \$1000.) While their answer was still on the screen, participants were told that soon after the first prospective tenant left, a second

prospective tenant (advantaged) arrived:

He is also a middle-aged man. You have learned through a casual conversation that he is Canadian, born and raised in Quebec City. He is now in the US on a business visa. He sometimes talks to friends on the phone at night and makes some mild noise.

The participants were asked the same question and chose among the same options as for the first prospective tenant.

The low-justifiability/advantaged-person-first condition was identical to the low-justifiability/disadvantaged-person-first condition except that the order of the two tenants was reversed.

In the high-justifiability/disadvantaged-person-first condition, participants read the same questionnaire as in the low-justifiability/disadvantaged-person-first condition, except that the first tenant (disadvantaged) was described as follows:

He is a middle-aged man. You have learned through a casual conversation that he is Australian, born and raised in Canberra. He is now in the US on a business visa. Three or four times a week, he would talk to friends on the phone at night and make noise.

The second tenant (advantaged) was described as follows:

He is a middle-aged man. You have learned through a casual conversation that he is Australian, born and raised in Canberra. He is now in the US on a business visa. Once or twice a month, he would talk to friends on the phone at night and make noise.

The high-justifiability/advantaged-person-first condition was identical to the high-justifiability/disadvantaged-person-first condition except that the order of the two tenants was reversed.

3.2. Results and discussion

3.2.1. Pretest

We assumed that when deciding how much to charge a prospective tenant, participants would perceive the nationality of the prospective tenant to be an unjustifiable consideration, but the noise level of the prospective tenant to be a justifiable consideration. To verify our assumptions, we conducted a pretest in which we presented a separate group of MTurkers ($N = 60$) with the same rental situation as in the main study. We asked participants whether they should consider the nationality of a prospective tenant and whether they should consider the noise level of a prospective tenant when deciding rent. Participants answered each question on a 6-point scale on which 1 meant "should not" and 6 meant "should." Supporting our assumptions, participants rated nationality as a should-not-consider factor ($M = 1.85$, $SD = 1.13$), $t(59) = 11.29$, $p < .001$, compared with the midpoint of the scale, and noise level as a should-consider factor ($M = 5.30$, $SD = 1.11$), $t(59) = 12.57$, $p < .001$, compared with the midpoint of the scale.

3.2.2. Main results

Fig. 2 displays the rent that participants would charge each of the two prospective tenants in each condition. Let us first focus on the low-justifiability condition, in which the two prospective tenants differed in nationality. We assumed that the Saudi tenant would receive worse treatment in the 1st position, and indeed, he received a significantly worse rent when appearing first than the Canadian tenant received when appearing first, $t(198) = 3.90$, $p < .001$.

To test Hypothesis 1, we performed a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA and found a significant interaction, $F(1, 198) = 4.11$, $p = .04$. The discrimination found in the 1st position disappeared in the 2nd position; the Saudi tenant now received a similar rent as the Canadian tenant, $p = .863$.

To test Hypothesis 1A, we compared the rents offered to the Saudi tenant in the 1st versus in the 2nd position, and we found the predicted

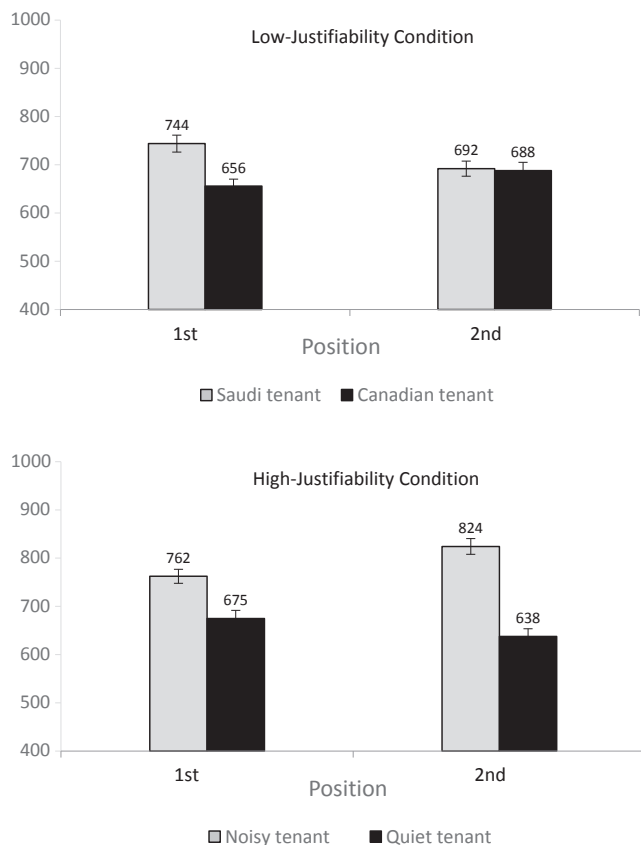


Fig. 2. Study 2 Results: Rent in US Dollars. (Taller columns mean higher rent quotes, i.e., worse treatment; the error bars represent $\pm$ one SE).

free-riding effect: he indeed received a better rent when appearing second than when appearing first, $t(198) = 2.21$, $p = .028$. To test Hypothesis 1B, we compared the rents offered to the Canadian tenant in the 1st versus in the 2nd position, and we found that he received a worse rent when appearing second than when appearing first, though the difference was not significant, $p = .152$.

We now turn to the high-justifiability condition, in which the two prospective tenants differed in noise level. To compare the results in the high-justifiability condition with the results in the low-justifiability condition, we ran a 2 (justifiability: low vs. high) $\times$ 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA and found a significant 3-way interaction, $F(1, 396) = 9.90$, $p = .002$. In support of Hypothesis 2, this 3-way interaction revealed that the 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) 2-way interaction found in the low-justifiability condition changed in the high-justifiability condition. Specifically, in the high-justifiability condition, a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA also yielded a significant 2-way interaction, $F(1, 198) = 5.88$, $p = .016$, but in the opposite direction from the 2-way interaction found in the low-justifiability condition. Thus, while the discrimination against the disadvantaged person (the Saudi tenant) in the low-justifiability condition was weaker in the 2nd position than in the 1st position, the “discrimination” against the disadvantaged person (the tenant making more noise) in the high-justifiability condition was stronger in the 2nd position than in the 1st position.

3.2.3. Discussion

Study 2 replicated the position effect on discrimination found in Study 1, this time using nationality as the differentiating attribute. In addition, Study 2 found that the effect disappeared when the attribute was replaced by noise level, a justifiable attribute. This latter finding

supported our justifiability-based account of the position effect and cast doubt on the anchoring account, because if the position effect were due to anchoring, it would have persisted regardless of whether the focal attribute was justifiable or unjustifiable.

4. Study 3

Study 3 was a conceptual replication of Study 2. Study 3 tapped a different context—disciplinary decisions—and concerned a different type of discriminatory factor—gender. Study 3 adopted a 2 (justifiability: low vs. high) $\times$ 2 (sequence: disadvantaged-person-first vs. advantaged-person-first) design. Unlike Study 2, which manipulated justifiability by using either a low-justifiability attribute (nationality) or a high-justifiability attribute (noise level), Study 3 used the same attribute (gender) and manipulated its justifiability by varying the presence or absence of a reason for discrimination (high justifiability vs. low justifiability, respectively).

Specifically, participants assumed the role of a member on a student disciplinary committee, which was tasked with evaluating two sexual harassment cases and recommending punishment for the wrongdoers. The two cases were similar except that the wrongdoer was a male student (and the victim a female student) in one case, and the wrongdoer was a female student (and the victim a male student) in the other case. Based on existing research suggesting more negative stereotypes against male wrongdoers than female wrongdoers in sexual harassment cases (Cummings & Armenta, 2002; Madera, Podratz, King, & Hebl, 2007), we considered the male wrongdoer to be the relatively disadvantaged person and the female wrongdoer the relatively advantaged person. We manipulated the order of the two cases and the justifiability of treating the two cases differently. We verified the effectiveness of our manipulation in a pretest (see below).

It is worth noting one other difference between Study 2 and Study 3: In Study 2, the rent quoted to the first prospective tenant remained on the screen while participants viewed and quoted a rent to the second prospective tenant. In Study 3, the judgment on the first case disappeared from the screen before participants viewed and made a judgment on the second case. This was a more subtle way to test our position effect. Because of this, we set a larger sample size for Study 3 than for Study 2.

Study 3 was pre-registered at <https://aspredicted.org/>, and the anonymized pre-registration document can be accessed through <http://aspredicted.org/blind.php?x=ga29qy>.

4.1. Method

Study 3 included four between-participants conditions, representing a 2 (sequence: disadvantaged-person-first vs. advantaged-person-first) $\times$ 2 (justifiability: low vs. high) design. Each condition involved a male wrongdoer (disadvantaged) and a female wrongdoer (advantaged).

Participants in the study were 603 workers recruited from MTurk in the US (297 females, $M_{\text{age}} = 38.39$). In all conditions, participants were asked to imagine that they were serving on an ad hoc college student disciplinary committee that was tasked with reviewing two cases.

In the low-justifiability/disadvantaged-person-first condition, participants first read a case involving a male wrongdoer and a female victim:

Case 1. A college senior named Jacob had a crush on his classmate Karen but knew she was not interested in him. During a party last week when Karen was drunk, Jacob hugged and kissed Karen without her consent.

The participants were asked, “What punishment would you give Jacob?” They answered by choosing one of five options: (a) warning, (b) suspension for one week, (c) suspension for one month, (d) suspension for one semester, or (e) expulsion. After that, participants

advanced to the next screen and read a second case, this time involving a female wrongdoer and a male victim:

Case 2. A college senior named Joanne had a crush on her classmate Kevin but knew he was not interested in her. During a party last week when Kevin was drunk, Joanne hugged and kissed Kevin without his consent.

The participants were asked the same question and chose among the same options as in the first case.

The low-justifiability/disadvantaged-person-first condition was the same as the low-justifiability/disadvantaged-person-first condition except that the order of the two cases was reversed.

In the high-justifiability/disadvantaged-person-first condition, participants read the same cases as in the low-justifiability/disadvantaged-person-first condition, but they also read the following additional information before answering the punishment question for each case:

According to scientific research, there are considerable gender differences in psychological responses to sexual harassment. Therefore, the chair of the disciplinary committee asks that for the same sexual misconduct, you give male offenders more severe punishment and female offenders less severe punishment.

(The statement above regarding gender differences in psychological reactions to sexual harassment was inspired by existing research on the topic, e.g., Rotundo, Nguyen, & Sackett, 2001; Russell & Trigg, 2004.)

The high-justifiability/disadvantaged-person-first condition was identical to the high-justifiability/disadvantaged-person-first condition except that the order of the two cases was reversed.

4.2. Results and discussion

4.2.1. Pretest

We assumed that the gender of the wrongdoer would be perceived as a should-not-consider factor in the low-justifiability condition and a should-consider factor in the high-justifiability condition. To verify the assumptions, we recruited a separate group of MTurkers ($N = 122$), randomly assigned them to read either the low-justifiability scenario or the high-justifiability scenario from the main study, and asked them to rate, on a 6-point scale, whether they should consider the gender of the wrongdoer in their punishment decision; 1 meant “should not” and 6 meant “should.” As expected, participants in the low-justifiability condition rated gender as a should-not-consider factor ($M = 2.25$, $SD = 1.61$), $t(60) = 6.09$, $p < .001$, compared with the midpoint of the scale, and participants in the high-justifiability condition rated gender as a should-consider factor ($M = 4.08$, $SD = 2.12$), $t(60) = 2.15$, $p = .036$, compared with the midpoint of the scale. (Of note, these are empirical results from the research and should not be interpreted as the authors’ view on whether male and female wrongdoers should be treated differently.)

4.2.2. Main results

We coded the five punishment options as 1 through 5, with greater numbers representing more severe punishment. Fig. 3 displays the level of punishment given to the two wrongdoers in the different conditions. We first focus on the low-justifiability condition. We assumed that the male wrongdoer would receive worse treatment in the 1st position, and indeed, the male wrongdoer received significantly harsher punishment when evaluated first than the female wrongdoer received when evaluated first, $t(298) = 4.16$, $p < .001$.

To test Hypothesis 1, we conducted a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA and found the predicted interaction, $F(1, 298) = 10.91$, $p = .001$. The differential treatment found in the 1st position reversed in the 2nd position: the female wrongdoer now received harsher punishment than the male wrongdoer, $t(298) = 2.25$, $p = .025$.

To test Hypothesis 1A, we compared the punishments for the male

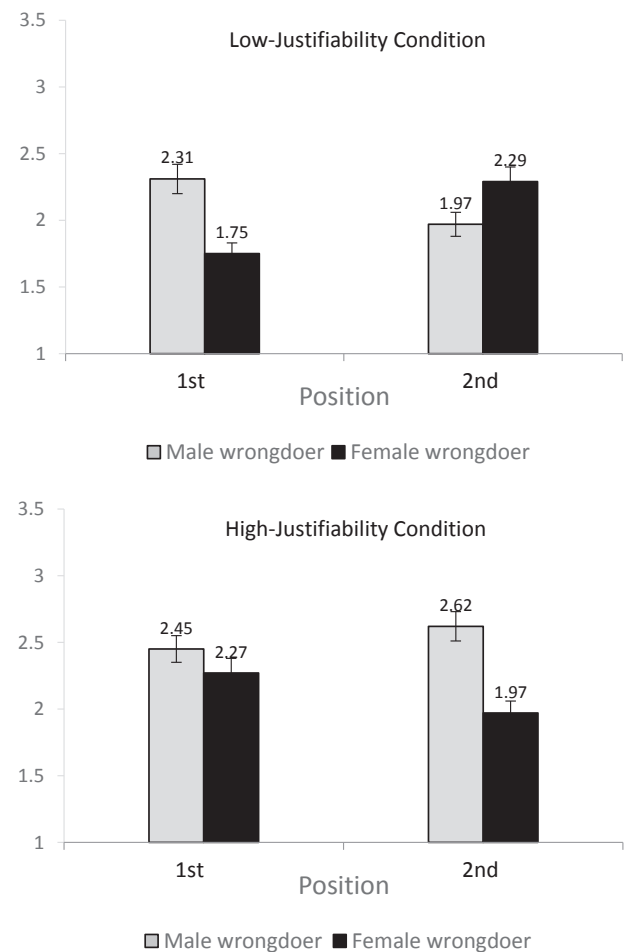


Fig. 3. Study 3 Results: Severity of Punishment. (Taller columns mean more severe punishment, i.e., worse treatment; the error bars represent $\pm$ one SE).

wrongdoer in the 1st versus in the 2nd position, and we found the hypothesized free-riding effect; he received milder punishment when evaluated second than when evaluated first, $t(298) = 2.43$, $p = .016$. To test Hypothesis 1B, we compared the punishments for the female wrongdoer in the 1st versus in the 2nd position, and we found the cost-bearing effect; she received harsher punishment when evaluated second than when evaluated first, $t(298) = 3.94$, $p < .001$.

To compare the results in the high-justifiability condition with the results in the low-justifiability condition, we ran a 2 (justifiability: low vs. high) $\times$ 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA and found a significant 3-way interaction, $F(1, 599) = 12.37$, $p < .001$. In support of Hypothesis 2, this 3-way interaction revealed that the 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) 2-way interaction found in the low-justifiability condition changed in the high-justifiability condition. Specifically, in the high-justifiability condition, a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) ANOVA found a marginally significant 2-way interaction, $F(1, 301) = 2.89$, $p = .090$. Thus, while discrimination against the male wrongdoer relative to the female wrongdoer in the 1st position was reversed in the 2nd position in the low-justifiability condition, discrimination was stronger in the 2nd position than in the 1st position in the high-justifiability condition. (It is curious to note that the difference in treatment between the male and the female wrongdoers in the 1st position in the high-justifiability condition was smaller than the difference in the low-justifiability condition. We do not know why and suspect it was a fluke.)

4.2.3. Discussion

Study 3 replicated the position effect found in Study 1 and Study 2, this time in the domain of disciplinary decisions. Specifically, Study 3 showed that a male wrongdoer in the 1st position received much harsher punishment than an analogous female wrongdoer in the 1st position, and that this differential treatment reversed in the 2nd position.

Unlike Study 2, which manipulated justifiability by using different attributes (nationality versus noise), Study 3 used the same attribute (gender) but varied the presence or absence of a reason for gender discrimination. Like Study 2, Study 3 demonstrated the moderating role of justifiability, thus further supporting our justifiability-based account for the position effect and further ruling out anchoring as an alternative account. The results of Study 3 also demonstrated that the position effect could happen even when the participants' judgment of the first target person was not displayed during their subsequent evaluation of the second target person.

5. General discussion

Whether we like it or not, individuals who differ in characteristics such as physical appearance and gender are often treated differently. The current research proposes and demonstrates that the worse treatment of a disadvantaged person who is the first to encounter an actor relative to an advantaged person who is the first to encounter the same actor disappears or even reverses when each person is the second to encounter the actor. In particular, the disadvantaged person receives better treatment by following the advantaged person—a free-riding effect—while the advantaged person receives worse treatment by following the disadvantaged person—a cost-bearing effect. However, this temporal position effect arises only if the attribute that differentiates the target persons is considered unjustifiable, and it disappears if the attribute is considered justifiable.

5.1. Open questions

The three studies reported in the main text, along with the additional study reported in [Appendix A](#), concern different types of discrimination (on the basis of appearance, nationality, gender, and ethnicity) and different types of decisions (pricing, rental, disciplinary, and college applications). We view this diversity as a strength of the present research because it demonstrates the generality of our hypotheses. However, this diversity also raises questions about whether the psychological processes in these studies are exactly the same. For example, in terms of the types of discrimination, preferences regarding physical appearances may be more hard-wired than preferences regarding nationality ([Cogsdill, Todorov, Spelke, & Banaji, 2014](#)). In terms of the types of decisions, disciplinary committee members may face a greater need to justify their decisions than a vendor setting a price for a jacket or a homeowner quoting a rent for a basement. While it is beyond the scope of this research to test these differences, it is important for future research to do so, because these differences could moderate the position effect proposed in this research.

A related question is whether the position effect proposed and found in this research is stronger for decisions made in public than for decisions made in private. We suspect so, because private decisions evoke the need for internal justification, only—to be self-consistent and avoid the feeling of guilt ([Amodio, Devine, & Harmon-Jones, 2007](#)). By contrast, public decisions additionally evoke the need for external justification—to impress others or to avoid criticisms or legal consequences ([Dowden & Robinson, 1993](#); [Tedeschi, 2013](#)).

Does the position effect still happen if the disadvantaged and the advantaged persons differ not only in a low-justifiability attribute (e.g., nationality) but also in high-justifiability attributes? For example, in the low-justifiability condition in Study 2 (involving different nationalities), would the position effect still occur if the Saudi tenant had an

annoying pet and the Canadian tenant had an unfavorable reference from his previous landlord, or vice versa? We suspect that in such cases, the position effect would be weaker, because the differences in other justifiable dimensions create additional leeway for the actor to favor the advantaged person over the disadvantaged person ([Hsee, 1996b](#); [Norton, Vandello, & Darley, 2004](#); [Snyder, Kleck, Strenta, & Mentzer, 1979](#)). [Hsee \(1996b\)](#) refers to the tendency to use differences in other dimensions to justify discrimination as “elastic justification,” and [Norton et al. \(2004\)](#) refer to this tendency as “casuistry.” Whether differences in other dimensions mitigate the position effect depends on the strength of the elastic justification and casuistry, and this is not theoretically central to the present research.

Does the position effect persist even if there is a long delay between the evaluations of the two target persons? We do not have empirical evidence to answer this question, but we suspect that both the free-riding effect and the cost-bearing effect are stronger when the second person immediately follows the first person than when there is a long delay between the appearances of the two persons, or when there are other persons in between. This speculation agrees with prior research on joint versus single evaluations and order effects (e.g., [Hsee, 1996a](#); [Moore, 1999](#)).

5.2. Implications for choice architecture

The present research highlights the importance of choice architecture, showing that the sequence in which a disadvantaged person and an advantaged person encounter an actor influences the treatment they receive. Anyone who ever influences evaluation sequences should be aware of this effect. Depending on the circumstances, the sequence may be determined by the target persons, the actor(s), or both. For example, suppose that two friends, one of whom is more attractive-looking than the other, are each interested in buying a Toyota Camry from a Toyota dealership. In this example, the target persons (the two friends) could decide between themselves who should negotiate with the dealer first. As another example, suppose that two assistant professors, who have similar credentials but different ethnicities, are coming up for promotion. Their cases will be discussed and voted on in the same senior faculty meeting. In this example, the actors (the senior faculty) usually decide whose case should go first, or perhaps the target persons (the junior faculty) and the actors (the senior faculty) may jointly weigh in. In both of these examples, our research suggests that the sequence of evaluations may influence the treatment of the target persons, affecting the prices offered to the car buyers in the first example and the votes cast in support of the assistant professors in the second example. Thus, all individuals involved in such cases should be aware of this position effect.

A key take-home message from this research is that the advantaged-person-first sequence is better than the disadvantaged-person-first sequence for both target persons, because the former sequence helps the disadvantaged person without hurting the advantaged person.

Furthermore, the advantaged-person-first sequence is potentially also better for both target persons than a strict joint-evaluation paradigm, in which the actor encounters the disadvantaged person and the advantaged person simultaneously. In order to avoid the perception of discrimination, the actor may treat the advantaged person worse in joint evaluation than alone, treat the disadvantaged person better in joint evaluation than alone, or a combination of the two. By contrast, the advantaged-person-first sequence does not hurt the treatment of the advantaged person, and it improves the treatment of the disadvantaged person, because by the time the actor encounters the disadvantaged person, the actor has already completed the treatment of the advantaged person.

As mentioned in the Method section of Study 2, we conducted a joint-evaluation condition ($N = 101$) in addition to the study's 2×2 conditions. That additional condition allowed us to compare the advantaged-person-first sequence with joint evaluation. The joint-

evaluation condition was the same as the advantaged-person-first condition (within the low-justifiability condition) of Study 2, except that the participants in the joint-evaluation condition were told that the two tenants arrived simultaneously and were shown the descriptions of the two tenants on the same page. As expected, both the disadvantaged (Saudi) tenant and the advantaged (Canadian) tenant received worse rent offers in the joint-evaluation condition ($M = 766.34$, $SD = 169.87$ and $M = 728.71$, $SD = 148.55$, respectively) than in the advantaged-person-first condition of the main study ($M = 692.00$, $SD = 156.79$, $t(199) = 3.22$, $p = .001$, and $M = 656.00$, $SD = 142.36$, $t(199) = 3.54$, $p < .001$, respectively), suggesting that the advantaged-person-first sequence is indeed better than joint evaluation for both target persons (see Hsee & Leclerc, 1998, for more information on when single evaluations lead to better evaluations than joint evaluations).

Much of the existing research on nudges and choice architecture rests on the idea that we can improve people's decisions by exploiting

their own biases and other psychological tendencies (Benartzi & Thaler, 2013; Benartzi et al., 2017; Goswami & Urminsky, 2016; Johnson & Goldstein, 2003; Milkman, Minson, & Volpp, 2013; Thaler & Sunstein, 2008). This idea also underlies the temporal position effect, especially the free-riding effect, that we investigated in this research. The free-riding effect reduces an actor's discrimination against a disadvantaged person by exploiting the actor's preference for the advantaged person and the actor's need for justifiability and consistency.

Acknowledgements

We thank Zhizhi Qi for collecting the data in Study 1. We thank Bowen Ruan for developing the data analysis to test our hypotheses. We thank Devin Pope for helpful discussions and suggestions on early versions of the paper, and we especially want to thank Katherine Milkman for her guidance throughout the review process.

Appendix A. Additional study on college admissions

This study tested our theory regarding yet another domain—college admissions. We compared the treatments of two academically-similar college applicants, one identified as Chinese and the other identified as Black. Both anecdotes and research suggest that Chinese applicants might face higher expectations than Black applicants (Norton et al., 2004; Reilly, 2018); it is in this sense that we considered the Chinese applicant more disadvantaged than the Black applicant. As in Study 3, we manipulated the justifiability of differential treatment by either giving or not giving a reason for it. (It is beyond the scope of this research to discuss whether it is legal or moral to treat applicants of different ethnic backgrounds differently. We simply used the opinions of pretest participants to define the justifiability of an attribute; see below for details.) We tested whether the sequence of evaluations affected the treatments of the two applicants.

A.1. Method

Like Study 3, this study included four between-participants conditions, which constituted a 2 (sequence: Chinese-first vs. Black-first) $\times$ 2 (justifiability: low vs. high) design. Participants were 396 persons recruited from MTurk (215 females, $M_{age} = 35.02$). All participants read the following:

Imagine the following: You work at the admissions office of a top university in the US. Your job is to review college applications as they come in, and indicate whether you would admit each applicant on a 7-point scale (1 = definitely NO; 7 = definitely YES).

In the low-justifiability/Chinese-first condition, participants saw the first applicant, as follows:

Name	Kevin X. Zhang
Gender	Male
Age	18
Ethnicity	Chinese
SAT math	650 (80th percentile)
SAT verbal	550 (60th percentile)
High school GPA	3.2 out of 4.0
Extracurricular activities	Unknown
Interests	Math and music

The participants were asked to decide whether they would admit the applicant on the aforementioned 7-point scale. While their answer was still on the screen, participants saw the profile of the second applicant, as follows:

Name	Darius P. Jefferson
Gender	Male
Age	18
Ethnicity	Black
SAT math	650 (80th percentile)
SAT verbal	550 (60th percentile)
High school GPA	3.2 out of 4.0
Extracurricular activities	Unknown
Interests	Math and music

The participants were asked the same decision question as for the first applicant.

The low-justifiability/Black-first condition was the same as the low-justifiability/Chinese-first condition except that the order of the two applicants was reversed.

The high-justifiability conditions were the same as their respective low-justifiability conditions except that before viewing the applicants, participants read the following additional statement, which stated a reason for discrimination:

Currently, a very high percentage of the students at your university are ethnically Chinese. To balance the racial composition of the student body, the university explicitly says that your office should consider the ethnicities of the applicants and use different (more stringent) standards for admitting Chinese applicants.

A.2. Results and discussion

A.2.1. Pretest

We assumed that ethnicity would be perceived as a should-not-consider factor in the control (low-justifiability) condition, in which we gave no reason to discriminate against Chinese applicants, but would be perceived as a should-consider factor in the high-justifiability condition, in which we stated a reason to discriminate against Chinese applicants. To verify our assumptions, we recruited a separate group of MTurkers ($N = 113$), randomly assigned them to read either the low-justifiability scenario or the high-justifiability scenario in the main study, and asked them to rate, on a 6-point scale, whether they should consider the ethnicity of an applicant; 1 meant “should not” and 6 meant “should.” As predicted, participants rated ethnicity as a should-not-consider factor in the low-justifiability condition ($M = 2.40$, $SD = 1.60$), $t(56) = 5.17$, $p < .001$, compared with the midpoint of the scale, and a should-consider factor in the high-justifiability condition ($M = 4.11$, $SD = 1.69$), $t(55) = 2.69$, $p = .010$, compared with the midpoint of the scale.

A.2.2. Main results

We first focus on the low-justifiability condition. We compared the ratings for the two applicants in the 1st position and found significant discrimination; the Chinese applicant received significantly lower ratings ($M = 3.98$, $SD = 1.43$) than the Black applicant ($M = 4.70$, $SD = 1.43$), $t(197) = 3.55$, $p < .001$.

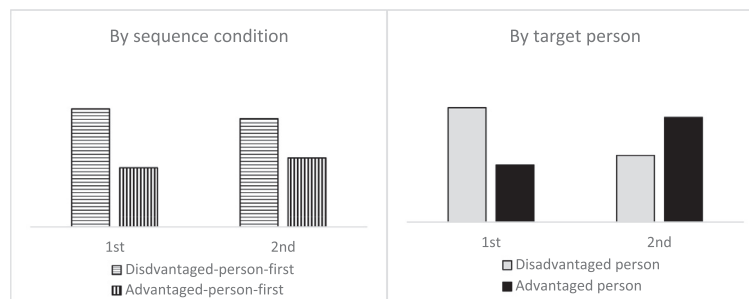
To test Hypothesis 1, we conducted a 2 (target person: Chinese vs. Black) $\times$ 2 (position: Chinese-first vs. Black-first) ANOVA and found the predicted interaction, $F(1, 197) = 9.57$, $p = .002$. The discrimination found in the 1st position reversed in the 2nd position; the Chinese applicant now received *higher* ratings ($M = 4.57$, $SD = 1.42$) than the Black applicant ($M = 4.07$, $SD = 1.37$), $t(197) = 2.52$, $p = .013$.

To test Hypothesis 1A, we compared the ratings for the Chinese applicant in the 1st versus in the 2nd position, and we found the hypothesized free-riding effect: he received better ratings when evaluated second than when evaluated first, $t(197) = 2.92$, $p = .004$. To test Hypothesis 1B, we compared the ratings for the Black applicant in the 1st versus in the 2nd position, and we found the predicted cost-bearing effect; the Black applicant received worse ratings when evaluated second than when evaluated first, $t(197) = 3.17$, $p = .002$.

We now turn to the high-justifiability condition. To compare the results in the high-justifiability condition with the results in the low-justifiability condition, we ran a 2 (justifiability: low vs. high) $\times$ 2 (target person: Chinese vs. Black) $\times$ 2 (position: Chinese-first vs. Black-first) ANOVA and found a significant 3-way interaction, $F(1, 392) = 6.02$, $p = .01$. In support of Hypothesis 2, this 3-way interaction meant that the 2-way interaction found in the low-justifiability condition changed in the high-justifiability condition. Specifically, in the high-justifiability condition, a 2 (target person: Chinese vs. Black) $\times$ 2 (position: Chinese-first vs. Black-first) ANOVA did not find a significant 2-way interaction, $F(1, 195) = 0.10$, $p = .748$. While the Chinese applicant in the low-justifiability condition was treated worse than the Black applicant in the 1st position but treated better than the Black applicant in the 2nd position, the Chinese applicant in the high-justifiability condition was treated worse than the Black applicant in both the 1st position ($M = 3.97$, $SD = 1.55$ vs. $M = 4.86$, $SD = 1.48$), $t(195) = 4.14$, $p < .001$, and the 2nd position ($M = 3.67$, $SD = 1.73$ vs. $M = 4.69$, $SD = 1.52$), $t(195) = 4.37$, $p < .001$. This study replicated Study 2 and Study 3 in another domain and supported all our hypotheses.

Appendix B. Note on two alternative formats for presenting the results

We can present our results in two alternative but equivalent formats, one based on our manipulated sequence conditions (disadvantaged-person-first vs. advantaged-person-first) and one based on the target persons (disadvantaged vs. advantaged). The following two figures are stylized illustrations of the two presentation formats. (Taller columns in the figures indicate *worse* treatments.)



The by-sequence-condition format is consistent with the designs of our studies, which are 2 (sequence: disadvantaged-person-first vs. advantaged-person-first) $\times$ 2 (position: 1st vs. 2nd). However, this format is not intuitive, because the horizontal-striped columns and the vertical-striped columns represent different target persons in the 1st versus in the 2nd positions. By contrast, the by-target-person format is more intuitive; the light columns always represent the disadvantaged person and the dark columns always represent the advantaged person.

Our main hypothesis (Hypothesis 1) appears to make different predictions in the two formats. In the by-sequence-condition format (the left panel), Hypothesis 1 predicts a sequence main effect, meaning that the overall treatment of the disadvantaged person and the advantaged person is better in the advantaged-person-first sequence than in the disadvantaged-person-first sequence. In the by-target-person format (the right panel), the same hypothesis predicts a 2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) 2-way interaction, meaning that the differential treatment in the 1st position disappears or reverses in the 2nd position.

Despite this apparent difference, these two formats are equivalent, and the sequence main effect in the by-sequence-condition format is equivalent to the 2-way interaction in the by-target-person format. We find the by-target-person format (the right panel) more intuitive, and therefore, we adopt that format in our paper.

The above discussion applies only to situations in which the attribute that differentiates the two target persons is of low justifiability. According to our theory (Hypothesis 2), Hypothesis 1 will no longer hold if the attribute is of high justifiability. In the by-sequence-condition format,

Hypothesis 2 means that the sequence main effect in the low-justifiability condition will diminish in the high-justifiability condition; equivalently, in the by-target-person format, Hypothesis 2 means that the 2-way interaction in the low-justifiability condition will diminish in the high-justifiability condition.

The following table summarizes the predictions made by our hypotheses between these two ostensibly different, but actually equivalent, presentation formats:

	By-sequence-condition format	By-target-person format
Hypothesis 1 (position effect)	Sequence main effect, i.e., the overall treatments of the target persons are better in the advantaged-person-first sequence than in the disadvantaged-person-first sequence	2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) 2-way interaction, i.e., the discrimination seen in the 1st position disappears or reverses in the 2nd position
Hypothesis 1A (free-riding effect)	The disadvantaged person is treated better in the 2nd position than in the 1st position	The disadvantaged person is treated better in the 2nd position than in the 1st position
Hypothesis 1B (cost-bearing effect)	The advantaged person is treated worse in the 2nd position than in the 1st position	The advantaged person is treated worse in the 2nd position than in the 1st position
Hypothesis 2 (justifiability effect)	2 (sequence: disadvantaged-person-first vs. advantaged-person-first) $\times$ 2 (justifiability: low vs. high) 2-way interaction, i.e., the effect in H1 holds only in the low-justifiability condition and disappears in the high-justifiability condition	2 (target person: disadvantaged vs. advantaged) $\times$ 2 (position: 1st vs. 2nd) $\times$ 2 (justifiability: low vs. high) 3-way interaction, i.e., the effect in H1 holds only in the low-justifiability condition and disappears in the high-justifiability condition

Appendix C. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.obhdp.2019.08.003>.

References

- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, 18(6), 524–530.
- Ayres, I. (1991). Fair driving: Gender and race discrimination in retail car negotiations. *Harvard Law Review*, 817–872.
- Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review*, 23(2), 225–241.
- Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., ... Galing, S. (2017). Should governments invest more in nudging? *Psychological Science*, 28(8), 1041–1055.
- Benartzi, S., & Thaler, R. H. (2013). Behavioral economics and the retirement savings crisis. *Science*, 339(6124), 1152–1153.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4), 991–1013.
- Bohnet, I., Van Geen, A., & Bazerman, M. (2015). When performance trumps gender bias: Joint vs. separate evaluation. *Management Science*, 62(5), 1225–1234.
- Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon.
- Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science*, 25(5), 1132–1139.
- Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, 112(5), 1297–1338.
- Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, 129(3), 414–446.
- Cryder, C., Botti, S., & Simonyan, Y. (2017). The charity beauty premium: Satisfying donors' "Want" versus "Should" desires. *Journal of Marketing Research*, 54(4), 605–618.
- Cummings, K. M., & Armenta, M. (2002). Penalties for peer sexual harassment in an academic context: The influence of harasser gender, participant gender, severity of harassment, and the presence of bystanders. *Sex Roles*, 47(5–6), 273–280.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24(3), 285–290.
- Dowden, S., & Robinson, J. P. (1993). Age and cohort differences in American racial attitudes: The generational replacement hypothesis revisited. In P. M. Sniderman, P. E. Tetlock, & E. G. Carmines (Eds.), *Prejudice, politics, and the American Dilemma*. Eagly, A. H., Makhijani, M. G., & Klonksy, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111(1), 3–22.
- Epley, N., & Gilovich, T. (2006). The anchoring and adjustment heuristic: Why adjustments are insufficient. *Psychological Science*, 17, 311–318.
- Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, 4(2), 195–202.
- Goswami, I., & Urmitsky, O. (2016). When should the ask be a nudge? The effect of default amounts on charitable donations. *Journal of Marketing Research*, 53(5), 829–846.
- Hameremesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *The American Economic Review*, 84(5), 1174–1194.
- Hsee, C. K. (1995). Elastic justification: How tempting but task-irrelevant factors influence decisions. *Organizational Behavior and Human Decision Process*, 62, 330–337.
- Hsee, C. K. (1996a). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3), 247–257.
- Hsee, C. K. (1996b). Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes*, 66(1), 122–129.
- Hsee, C. K., & Leclerc, F. (1998). Will products look more attractive when presented separately or together? *Journal of Consumer Research*, 25(2), 175–186.
- Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science*, 5(4), 343–355.
- Ibish, I. (2003). *Report on hate crimes and discrimination against Arab Americans: The post-September 11 backlash, September 11, 2001–October 11, 2002*. Washington, DC: American-Arab Anti-Discrimination Committee.
- Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21, 1161–1167.
- Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science*, 302, 1338–1339.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126(3), 390–423.
- Langlois, J. H., Ritter, J. M., Roggman, L. A., & Vaughn, L. S. (1991). Facial diversity and infant preferences for attractive faces. *Developmental Psychology*, 27(1), 79–84.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255–275.
- Li, X. & Hsee, C. K. (forthcoming). Beyond preference reversal: Distinguishing justifiability from evaluability in joint versus single evaluations. *Organizational Behavior and Human Decision Processes*.
- Madera, J. M., Podratz, K. E., King, E. B., & Hebl, M. R. (2007). Schematic responses to sexual harassment complainants: The influence of gender and physical attractiveness. *Sex Roles*, 56(3–4), 223–230.
- Milkman, K. L., Minson, J. A., & Volpp, K. G. (2013). Holding the Hunger Games hostage at the gym: An evaluation of temptation bundling. *Management Science*, 60(2), 283–299.
- Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, 96(1), 222–235.
- Moore, D. A. (1999). Order effects in preference judgments: Evidence for context dependence in the generation of preferences. *Organizational Behavior and Human Decision Processes*, 78(2), 146–165.
- Moradi, B., & Hasan, N. T. (2004). Arab American persons' reported experiences of discrimination and mental health: The mediating role of personal control. *Journal of Counseling Psychology*, 51(4), 418–428.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41), 16474–16479.
- Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, 87(6), 817–831.
- Petrova, P. K., Cialdini, R. B., & Sills, S. J. (2007). Consistency-based compliance across cultures. *Journal of Experimental Social Psychology*, 43(1), 104–111.
- Porter, C. O., Outlaw, R., Gale, J. P., & Cho, T. S. (2019). The use of online panel data in management research: A review and recommendations. *Journal of Management*, 45(1), 319–344.
- Reilly, Katie. (2018, October 16). A lawsuit by Asian-American students against Harvard could end affirmative action as we know it. Retrieved from <http://time.com/5425147/harvard-affirmative-action-trial-asian-american-students/>.

- Rotundo, M., Nguyen, D. H., & Sackett, P. R. (2001). A meta-analytic review of gender differences in perceptions of sexual harassment. *Journal of Applied Psychology*, 86(5), 914–922.
- Russell, B. L., & Trigg, K. Y. (2004). Tolerance of sexual harassment: An examination of gender differences, ambivalent sexism, social dominance, and gender roles. *Sex Roles*, 50(7–8), 565–573.
- Schuman, H., Kalton, G., & Ludwig, J. (1983). Context and contiguity in survey questionnaires. *Public Opinion Quarterly*, 47(1), 112–115.
- Sellers, R. M., & Shelton, J. N. (2003). The role of racial identity in perceived racial discrimination. *Journal of Personality and Social Psychology*, 84(5), 1079–1092.
- Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, 29(3), 281–295.
- Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, 21(2), 345–354.
- Snyder, M. L., Kleck, R. E., Strenta, A., & Mentzer, S. J. (1979). Avoidance of the handicapped: An attributional ambiguity analysis. *Journal of Personality and Social Psychology*, 37(12), 2297–2306.
- Tedeschi, J. T. (2013). *Impression management theory and social psychological research*. Academic Press.
- Thaler, Richard H., & Sunstein, Cass R. (2008). *Nudge: Improving decisions on health, wealth, and happiness*. New Haven, CT: Yale University Press.
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, 66, 519–545.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.