

## Social Decoys: Leveraging Choice Architecture to Alter Social Preferences

## Abstract

Many of society's most significant social decisions are made over sets of individuals: for example, evaluating a collection of job candidates when making a hiring decision. Rational theories of choice dictate that decision makers' preferences between any two options should remain the same irrespective of the number or quality of other options. Yet people's preferences for each option in a choice set shift in predictable ways as function of the available alternatives. These violations are well documented in consumer behavior contexts: for example, the decoy effect, in which introducing a third inferior product changes consumers' preferences for two original products. The current experiments test the efficacy of *social decoys* and harness insights from computational models of decision-making to examine whether choice set construction can be used to change preferences in a hiring context. Across seven experiments ( $N = 6312$ ) we find that participants have systematically different preferences for the exact same candidate as a function of the other candidates in the choice set (Experiments 1a-1d, 2) and the salience of the candidate attributes under consideration (Experiments 2, 3a, 3b). Specifically, compromise and (often) asymmetric-dominance decoys increased relative preference for their yoked candidates when candidates were counter-stereotypical (e.g., high warmth/low competence male candidate). More importantly, we demonstrate for the first time that we can mimic the effect of a decoy in the absence of a third candidate by manipulating participants' exposure to candidates' attributes: balanced exposure to candidates' warmth and competence information significantly reduced bias between the two candidates.

Word count: 248

**Social Decoys: Leveraging Choice Architecture to Alter Social Preferences**

Psychologists often study how people evaluate and treat different ethnic and cultural groups (and their members) in isolation from one another, across a series of sequential judgments. In the real world, however, people commonly make judgments and decisions over *sets* of people (and groups, e.g., Biernat & Manis, 1994; Oakes, Haslam, & Turner, 1998; Judd & Park, 1993; Trope & Mackie, 1987; Wyer, Sadler, & Judd, 2002). This process of joint evaluation characterizes many consequential decision-making contexts including hiring, housing, and voting decisions.

Rational theories of choice dictate that decision makers' preferences between any two options should remain the same irrespective of the number or quality of other options: a property known as independence of irrelevant alternatives. And yet, humans, monkeys, birds, insects, even amoeboid organisms reliably violate this axiom (Bateson, Healy, & Hurly, 2003; Huber, Payne, & Puto, 1982; Hurly & Oseen, 1999; Latty & Beekman, 2011; Louie, Khaw, & Glimcher, 2013; Shafir, Waite, & Smith, 2002; Simonson, 1989). Specifically, decision-makers' preferences for each option in a choice set shift in predictable ways as a function of the available alternatives. These violations are well documented in consumer behavior contexts: for example, the decoy effect, in which introducing a third inferior product changes consumers' preferences for two original products. What if instead of purchasing, the decision was "who do we hire?" What if instead of price and warranty duration, the attributes of interest were gender and race, or two cardinal dimensions of social cognition: warmth and competence? These context-dependent rationality violations act like levers, systematically increasing or decreasing decision makers' preferences for specific options; however, research often fails to account for these effects in social decision-making. Furthermore, discrimination research has typically emphasized

perceivers' stereotypes and attitudes as the primary target of interventions. However, context dependent phenomena suggest that even if evaluators were entirely free from bias against any individual candidate, decoys would reliably change evaluators' judgments outside of their awareness, which could still lead to discrimination if the choice set construction process was biased. The current research shifts the target of inquiry from perceivers' minds to the decision-making context, and aims to harness insights from formal models of decision-making to examine whether choice set construction can be used to alter preferences in hiring.

### *Context Dependence in Decision-Making*

Context dependence in choice behavior is a widely documented and studied phenomenon (Tversky & Simonson, 1993). Altering the options under consideration (e.g., Huber et al., 1982; Simonson, 1989) or manipulating the salience of alternatives (e.g., Tversky & Kahneman, 1991) reliably changes individuals' preferences for each option within a choice set. Dynamic models of decision-making propose that these context effects emerge over time as choice attributes are sequentially sampled and options are compared (e.g., Roe, Busemeyer, & Townsend, 2001; Usher & McClelland, 2001).

Attraction, or *decoy effects* refer to the class of context effects whereby inclusion of an inferior third option increases preference for the option it most closely resembles (i.e., in a tradeoff scenario where each target is superior on a dimension the competitor lacks). In the compromise effect, an extreme third option shifts preference to the option that is now viewed as a compromise. The compromise effect has been widely documented in consumer choices (Pettibone & Wedell, 2000; Simonson, 1989) and perceptual decision-making tasks (Trueblood, Brown, Heathcote, & Busemeyer, 2013). In a second type of decoy effect—the asymmetric dominance effect (Huber et al., 1982)—a decoy superior to option A on only one attribute, but

inferior to option B on *both* attributes increases preferences for option B. Though the asymmetric dominance effect has also been documented in consumer choices (Doyle, O'Connor, Reynolds, & Bottomley, 1999; Huber et al., 1982; Simonson, 1989; Wedell, 1991) and perceptual judgments (Choplin & Hummel, 2005; Tsetsos, Chater, & Usher, 2012; Trueblood et al., 2013), its practical significance has recently come into question (Frederick, Lee, & Baskin, 2014).

Earlier investigations of the mechanisms underlying decoy effects emphasized the relative weights of attributes present in the options (e.g., Ariely & Wallsten, 1995) as well as changes in attributes' values (e.g., Wedell & Pettibone, 1996). Building on earlier models adopting a sequential accumulation framework (e.g., Roe et al., 2001; Usher & McClelland, 2004; see also attentional drift diffusion models, Krajbich & Rangel, 2011) the recent Associative Accumulation Model (AAM; Bhatia, 2013) integrates these explanations. Stated simply, the AAM assumes that attributes of an option are attended to at random and accumulated over time into preferences, but that attributes present in many options within a choice set will be relatively more accessible. Thus, adding a decoy to a choice set increases the accessibility of (and therefore sampling of) the attribute on which a decoy and its yoked target are high, increasing the likelihood their associated values are aggregated into preferences (i.e., increasing relative preference for the option that dominates the decoy). The utility of models like the AAM is that they can account for a variety of choice-set dependent effects. These models make explicit the computational processes that give rise to context-dependence in decision-making and highlight the dynamic nature of the decision-making process.

### *Context Dependence in Social Decisions*

Researchers have documented context dependence across numerous important social contexts, including but not limited to mate selection (Sedikides, Ariely, & Olsen, 1999), voting

behavior (O’Curry, & Pitts, 1995), and policy preferences (Herne, 1997). Most relevant is the work on context dependence in the domain of hiring.

One experiment manipulated whether the two superior candidates in a choice set of five candidates for an engineering position had relatively more education or less experience. Crucially, one superior candidate was given a male name and the other a female name. Even though participants saw education as more important than experience, they chose to hire the male candidate the majority of the time, even when he was less educated—demonstrating the pernicious effects of inconsistent weighting of choice attributes on decision-making (Norton, Vandello, & Darley, 2004). Consistent with the previous study, another study examined the quality (indexed by interviews) of MBA candidates in a choice set and found that interviewers’ ratings of a particular candidate’s aptitude varied according to the ratings made of candidates who had interviewed earlier in the day (Simonsohn & Gino, 2013). Only one study of which we are aware has directly examined decoy effects like those described in the previous section: researchers manipulated job candidate interview and promotability ratings and found evidence of asymmetric dominance and phantom decoys—decoys that are no longer available—in a hiring context (Highhouse, 1996).

What is the precise mechanism by which decoys exert their effects? A recent investigation examining the impact of joint versus separate evaluation in hypothetical hiring decisions on gender bias revealed that in a comparison between a female and male candidate, in a context that stereotypically favors men (i.e., a math task), participants relied more on candidate performance information and less on gender when they evaluated the targets jointly rather than separately (Bohnet, van Geen, & Bazerman, 2015). This effect is likely driven by the fact that people rely more on internal referents (e.g., stereotypes) in separate evaluation because there is a

lack of concrete comparison information (Kahneman & Miller, 1986). However, different combinations of attributes in a decoy can have markedly different effects (in some cases even suppressing preferences for the yoked target; Bhatia, 2013), therefore it is incumbent upon researchers to characterize the precise effects different decoys have on preferences and choice.

We draw on models of dynamic decision-making to shift from qualitative to quantitative predictions of the specific effects of multiple types of decoys and to better elucidate the cognitive mechanisms by which they have their effects.

#### *Stereotype Content Dimensions: Warmth and Competence*

In many cases, decision-makers are forced to trade off on multiple attributes. But which attributes matter most in social decision-making contexts? The Stereotype Content Model (SCM; Cuddy, Fiske, & Glick, 2007; Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002) organizes beliefs about social and cultural groups along two fundamental dimensions: perceived warmth and competence. Whether a social group is cooperative or competitive will determine if they apparently have intent to help or harm the culturally dominant group (or in-group), which guides people's perceptions of that social group's warmth. Likewise, whether a social group does or does not have a high status will determine if they apparently have capability to harm the in-group, which will guide people's perceptions of the social group's competence.

This  $2$  (low/high warmth)  $\times$   $2$  (low/high competence) mapping describes four broad stereotype categories and the emotional responses those categories elicit. Groups high on both warmth and competence (e.g., Americans, college students) elicit pride, whereas groups low on both warmth and competence (e.g., homeless people, drug addicts) elicit disgust. Groups falling in the mixed quadrants elicit ambivalent emotions; pity is elicited by people perceived as low in competence but high in warmth (e.g., elderly people, disabled people), whereas envy is reserved

for people perceived as high in competence but low in warmth (e.g., rich people, investment bankers). The ‘pity’ and ‘envy’ quadrants represent the canonical multi-attribute tradeoff decision because each one has a high degree of the attribute the other lacks. Here we use the labels ‘pity’ and ‘envy’ as shorthand for low-competence/high-warmth targets and high-competence/low-warmth targets, respectively. (At no point do we assess participants’ emotional responses to any of the targets.) Furthermore, because this is the first step in this line of research, we provide participants with warmth and competence information rather than allowing them to generate these trait attributes based on targets’ group membership. Nevertheless, we test our effects with both male and female candidates to investigate whether decoy effects depend on the (counter-) stereotypicality of the candidate.

#### *Overview of the Current Experiments and Hypotheses*

Across seven experiments, we manipulate choice architecture to change participants’ evaluations of low-competence/high-warmth (‘pity’) candidates and high-competence/low-warmth (‘envy’) candidates. Specifically, we examine (i) how people weigh the two fundamental dimensions of stereotype content—warmth and competence—to make decisions in a hiring context for male and female targets, and (ii) whether social decoys can increase preference for their yoked candidates. In Experiments 1a through 1d, we predict that asymmetric dominance and/or compromise social decoys will increase decision-makers’ preferences for their yoked candidates relative to when the decoys are absent. In Experiment 2, we replicate and extend the first experiments to explore whether decoy effects are sensitive to which attributes are emphasized by the context (i.e., manipulating whether the firm values competence or warmth more). In Experiments 3a and 3b, we mimic the effect of social decoys by manipulating participants’ sequential exposure to candidates’ attribute information. We predict that increasing

participants' exposure to candidates' ratings on the warmth attribute versus competence attribute will lead participants to weigh warmth more heavily, increasing participants' preference for the 'pity' candidate. Experiment materials, data, and analysis code for all seven experiments in the paper are available for download:

[https://osf.io/eg6w5/?view\\_only=cfb3ca2f44604197a528491add4bf14](https://osf.io/eg6w5/?view_only=cfb3ca2f44604197a528491add4bf14)

### **Experiment 1a: Compromise and Asymmetric Dominance Decoys in Hiring**

Experiment 1a investigated whether Pity Compromise and Pity Asymmetric Dominance decoys increase preferences for a high-warmth/low competence male candidate in a hiring context (and similarly whether Envy Compromise and Envy Asymmetric Dominance decoys increase preferences for a low-warmth/high-competence candidate). We also included willingness-to-pay as a dependent variable (WTP) and predicted that decoys would increase wages for their yoked candidates.

### Method

*Participants and exclusions.* To reduce the widespread sampling bias associated with including only Western, educated, industrialized, rich, and democratic participants (i.e., university undergraduates; Henrich, Heine, & Norenzayan, 2010), participants for all experiments were recruited from Amazon.com's Mechanical Turk. mTurk provides high quality data from a relatively diverse sample (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010). This approach allows us to collect data from a more representative cross-section of the U.S. along age, ethnic, and SES dimensions.

Across all experiments, participants provided informed consent; all procedures complied with the university's institutional review board's guidelines.

We aimed for a minimum of 150 participants per condition after exclusions to achieve 80% power to detect a small effect size. We recruited 775 participants from Amazon Mechanical Turk to participate in this experiment. Of these, 2 participants were excluded from all analyses for failing to report likelihood of hiring for all candidates in their assigned condition. This resulted in a final sample size of  $N = 773$  participants (350 female, 416 male<sup>1</sup>;  $M_{age} = 33.78$  years,  $SD = 10.84$ ).

*Experimental design.* Participants were randomly assigned to one of five decoy conditions: Baseline (no decoy), Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, and Envy Compromise. In the Baseline condition, participants jointly evaluated two candidates: a Pity candidate and an Envy candidate. In the four Decoy conditions, participants jointly evaluate three candidates: the Pity and Envy candidates from the Baseline condition plus one additional decoy candidate.

*Materials.* In the joint evaluation phase of the experiment, participants were presented with identical silhouettes representing each candidate, along with gender information identifying each candidate as male, and global warmth and competence ratings for each candidate (see Figure 1a for example). Global Warmth was described as sincere, good-natured, warm, tolerant and Global Competence was described as competent, intelligent, confidence, competitive, independent (Cuddy et al., 2007). Because previous research indicates that the decoy effect can only be found when the attributes are quantified as numeric indices (Frederick et al., 2014), we visualized warmth ratings as number of hearts and competence ratings as numbers of stars.

We chose the relative amounts of warmth and competence for each candidate by defining the entire range of values of each dimension on a scale of 20 units. We wanted the Envy candidate to have warmth and competence values in the 33<sup>rd</sup> and 66<sup>th</sup> percentile, respectively

(and vice versa for the Pity candidate). For Asymmetric Dominance decoys, we took  $\frac{1}{4}$  of the difference between the Pity and Envy candidates and subtracted this value from the warmth and competence ratings of the yoked candidate. For Compromise decoys, we took  $\frac{3}{4}$  of the difference between the Pity and Envy candidates and added this value to the yoked candidate's dominant dimension and subtracted this value from the yoked candidate's inferior dimension (see Figure 1c for graphical representation of dimension values for each candidate and decoy).

*Procedure.* We asked participants to imagine that they had been put in charge of hiring a consultant to advise their boss on strategy. We told participants that the headhunting firm had sent over information about how each final candidate performed on two global personality assessments (warmth/competence) to help them make their decisions. After jointly evaluating the set of candidates for at least 10 seconds, participants reported likelihood of hiring (7-point scale; *very unlikely* to *very likely*) and WTP (slider scale; ranging from hourly rate of \$0 to \$200) for each candidate (in turn) in counterbalanced order such that participants reported all likelihood of hiring decisions first or all WTP decisions first. In the Baseline condition, participants reported likelihood of hiring and WTP of the Envy and Pity candidates in randomized order. In each of the decoy conditions, participants reported likelihood of hiring and WTP for the decoy first, followed by Envy and Pity candidates in randomized order. We fixed this sequence to anchor participants' ratings on the decoy.

*Analyses.* We conducted linear mixed effect models in R (version 3.4.2; R Core Team, 2017). We fit linear mixed models and post-hoc contrasts using *lme4* (Bates, Maechler, Bolker & Walker, 2015), *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017), *MuMIn* (Bartoń, 2017), and *emmeans* (Lenth, 2017) packages. Given our repeated measures design, all mixed models included participant as a random intercept. We fit two classes of models: (1) null models,

featuring participant as the sole predictor, and (2) hypothesis-driven models, which included additional manipulated factors and their interaction (e.g., candidate type, condition). We then used likelihood ratio tests (LRTs) to evaluate model fit by assessing whether including certain predictors significantly reduced residual variance. We estimated degrees of freedom using both the Kenward-Roger approximation method and Satterthwaite approximation method. Since results do not differ based on estimation method, we report all model-produced degrees of freedom using the Satterthwaite approximation method. While there is no agreed upon standardized effect size for within-subjects linear mixed models, we follow the approach taken in Rouder et al. (2012) in assuming that the  $\sigma$  in the classical Cohen's  $d$  formula refers to the standard deviation of the residuals.

## Results

*Likelihood of hiring.* We fit a linear mixed model to predict likelihood of hiring from the fixed effects of candidate type (Pity, Envy), decoy condition (Baseline, Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise), along with their interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $X^2(9) = 210.83, p < 0.001$ . We compared this hypothesis-driven model to a reduced model with no interaction term, and found that the interaction between candidate type and condition was significant,  $X^2(4) = 15.85, p = 0.003$ .

To explore the effect of response order (likelihood of hiring, WTP) on model fit, we compared the previous hypothesis-driven model to a model including response order as a fixed effect. We found that including response order significantly increased model fit,  $X^2(1) = 7.27, p = 0.007$ . Subsequently, we compared this model to a model including two two-way interactions, between candidate type and decoy condition and candidate type and response order. The

inclusion of the second interaction term between candidate type and response order significantly improved model fit,  $X^2(1) = 9.25, p = 0.002$ . We then compared this two two-way interaction model with a model including a three-way interaction between candidate type, decoy condition, and response order, and found no significant improvement in model fit,  $X^2(8) = 4.76, p = 0.783$ . Therefore, we report the results of the two, two-way interaction model, which includes the interaction between candidate type and decoy condition as well as the interaction between candidate type and response order,  $mR^2 = 0.137$ .

The significant interaction between candidate type and decoy condition,  $F(4,1546) = 4.07, p = 0.003$ , and the significant interaction between candidate type and response order,  $F(1, 1546) = 9.28, p = 0.002$ , qualified the main effects of candidate type,  $F(1,1546) = 205.29, p < 0.001$  and response order,  $F(1,1546) = 7.34, p = 0.007$ . To unpack the candidate type  $\times$  decoy condition interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of each decoy on the Pity candidate and Envy candidate, respectively. See Figure 2 for a summary of the results.

As we predicted, the Pity Asymmetric Dominance decoy increased likelihood of hiring the Pity candidate ( $M = 5.13, SE = 0.096$ ) in comparison to the Baseline condition (where there was no decoy;  $M = 4.73, SE = 0.095$ ), 95% CI [0.142,0.672],  $t(1546) = 3.01, p = 0.003$  (Holm-corrected  $p = 0.011$ ),  $d = 0.339$ . Similarly, the Pity Compromise decoy increased likelihood of hiring the Pity candidate ( $M = 5.07, SE = 0.097$ ) in comparison to the Baseline condition, 95% CI [0.081,0.613],  $t(1546) = 2.56, p = 0.011$  (Holm-corrected  $p = 0.032$ ),  $d = 0.289$ .<sup>2</sup>

In contrast, the Envy Asymmetric Dominance decoy did not increase likelihood of hiring the Envy candidate ( $M = 5.90, SE = 0.097$ ) in comparison to the Baseline condition (where there was no decoy;  $M = 5.67, SE = 0.095$ ), 95% CI [-0.041,0.491],  $t(1546) = 1.66, p = 0.098$  (Holm-

corrected  $p = 0.298$ ),  $d = 0.187$ . Similarly, the Envy Compromise decoy did not increase likelihood of hiring the Envy candidate ( $M = 5.91$ ,  $SE = 0.098$ ) in comparison to the Baseline condition, 95% CI [-0.024, 0.510],  $t(1546) = 1.79$ ,  $p = 0.074$  (Holm-corrected  $p = 0.298$ ),  $d = 0.202$ .<sup>3</sup>

To unpack the unpredicted candidate type and response order interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of response order on likelihood of hiring each candidate type. We found that reporting likelihood of hiring first significantly increased likelihood of hiring the Pity candidate ( $M = 5.08$ ,  $SE = 0.061$ ) as compared to reporting WTP first ( $M = 4.72$ ,  $SE = 0.061$ ), 95% CI [0.182, 0.521],  $t(1546) = 4.07$ ,  $p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = -0.293$ . In contrast, response order did not have a significant effect on likelihood of hiring the Envy candidate (hiring:  $M = 5.77$ ,  $SE = 0.061$ ; WTP:  $M = 5.79$ ,  $SE = 0.061$ , 95% CI [-0.190, 0.149],  $t(1546) = -0.24$ ,  $p = 0.812$  (Holm-corrected  $p = 0.812$ ),  $d = 0.017$ .

WTP. In contrast to our predictions, decoys had no effect on candidates relative to the Baseline condition,  $F(4,772.29) = 0.47$ ,  $p = 0.759$ . (see Supplementary Materials for WTP results). In hindsight, this dependent variable was ill-conceived because including a third candidate relative to a condition with only two candidates communicates a very different supply-demand ratio for a given position, driving down the value of every candidate in the set (which works in direct opposition to the decoy). As a result, we omit this dependent variable from subsequent experiments.

### Discussion

In general, participants were more likely to hire the Envy (high-competence/low-warmth) relative to the Pity (low-competence/high-warmth) candidate. In line with our predictions,

however, including Pity decoys—both Asymmetric Dominance and Compromise—in the choice set increased participants' self-reported likelihood of hiring that same Pity candidate in comparison to the Baseline condition where no decoys were present. In contrast to our predictions, neither of the Envy decoys increased preferences for the Envy candidate relative to the Baseline condition. We did not have a priori predictions about the effect of decoys on non-yoked candidates (i.e., Envy decoys on the Pity candidate or Pity decoys on the Envy candidate), but found no significant effect for decoys on any of their non-yoked candidates.

Finally, we did not predict an effect of response order, but found that when participants reported likelihood of hiring for all candidates first, participants reported higher likelihood of hiring the Pity candidate as compared to participants who reported WTP first. In order to ensure our results were not an artifact of this feature we replicated Experiment 1a excluding the WTP measure.

### **Experiment 1b: A Replication of Male Candidates and Decoys in Hiring**

Experiment 1b is a direct replication of Experiment 1a, excluding the WTP measure and changing the labels 'warmth' and 'competence' to 'attitude' and 'aptitude' to more closely approximate human resources language (e.g., Lim & Chan, 2001).

#### Method

*Participants and exclusions.* We aimed for a minimum of 150 participants per condition after exclusions to achieve 80% power to detect a small effect size. We recruited 801 participants from Amazon Mechanical Turk to participate in this experiment. Of these, one participant was excluded from all analyses for failing to report likelihood of hiring for all candidates in their assigned condition. This resulted in a final sample size of  $N = 800$  participants (469 female, 325 male<sup>4</sup>;  $M_{\text{age}} = 34.37$  years,  $SD = 11.43$ ).

*Experimental design.* Experimental design was identical to Experiment 1a.

*Materials.* Materials were identical to the ones used in Experiment 1a, with one exception; we changed the labels ‘warmth’ and ‘competence’ to ‘attitude’ and ‘aptitude.’

*Procedure.* The procedure was identical to Experiment 1a, with one exception; we did not ask participants to report WTP.

## Results

We fit a linear mixed model to predict likelihood of hiring from the fixed effects of candidate type (Pity, Envy) and decoy condition (Baseline, Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) along with their interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $X^2(9) = 103.98, p < 0.001$ . Furthermore, we compared the hypothesis-driven model with the interaction to a reduced model with no interaction term, and found that the interaction between candidate type and condition was significant,  $X^2(4) = 37.93, p < 0.001$ . Therefore, we report the results of the hypothesis-driven model including the interaction,  $mR^2 = 0.063$ .

The significant interaction between candidate type and decoy condition,  $F(4,1600) = 9.60, p < 0.001$ , replicated Experiment 1a and qualified the main effects of candidate type,  $F(1,1600) = 59.00, p < 0.001$  and decoy condition,  $F(4,1600) = 2.65, p = 0.032$ . To unpack the omnibus interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of each decoy on the Pity candidate and Envy candidate, respectively. See Figure 3 for a summary of the results.

As we predicted, the Pity Asymmetric Dominance decoy increased likelihood of hiring the Pity candidate ( $M = 5.53, SE = 0.094$ ) in comparison to the Baseline condition (where there

was no decoy;  $M = 5.24$ ,  $SE = 0.095$ ), 95% CI [0.025,0.550],  $t(1600) = 2.15$ ,  $p = 0.032$  (Holm-corrected  $p = 0.073$ ),  $d = 0.240$ . Similarly, the Pity Compromise decoy increased likelihood of hiring the Pity candidate ( $M = 5.60$ ,  $SE = 0.094$ ) in comparison to the Baseline condition, 95% CI [0.096,0.621],  $t(1600) = 2.68$ ,  $p = 0.008$  (Holm-corrected  $p = 0.030$ ),  $d = 0.299$ .<sup>5</sup>

In contrast, the Envy Asymmetric Dominance decoy did not increase likelihood of hiring the Envy candidate ( $M = 5.89$ ,  $SE = 0.094$ ) in comparison to the Baseline condition (where there was no decoy;  $M = 5.78$ ,  $SE = 0.095$ ), 95% CI [-0.158,0.367],  $t(1600) = 0.78$ ,  $p = 0.437$  (Holm-corrected  $p = 1$ ),  $d = 0.087$ . Similarly, the Envy Compromise decoy did not increase likelihood of hiring the Envy candidate ( $M = 5.79$ ,  $SE = 0.096$ ) in comparison to the Baseline condition, 95% CI [-0.260,0.270],  $t(1600) = 0.04$ ,  $p = 0.970$  (Holm-corrected  $p = 1$ ),  $d = 0.004$ .<sup>6</sup>

In two cases, the decoy had a significant effect on the non-yoked candidate. The Envy Compromise decoy decreased likelihood of hiring the Pity candidate ( $M = 4.94$ ,  $SE = 0.096$ ) in comparison to the Pity candidate's baseline, 95% CI [-0.569,-0.040],  $t(1600) = -2.26$ ,  $p = 0.024$  (Holm-corrected  $p = 0.073$ ),  $d = -0.254$ ; and the Pity Asymmetric Dominance decoy decreased likelihood of hiring the Envy candidate ( $M = 5.46$ ,  $SE = 0.094$ ) in comparison to the Envy candidate's baseline, 95% CI [-0.589,-0.062],  $t(1600) = -2.43$ ,  $p = 0.015$  (Holm-corrected  $p = 0.062$ ),  $d = -0.272$ .

### Discussion

Replicating Experiment 1a, including Pity decoys—both Asymmetric Dominance and Compromise—in the choice set increased participants' self-reported likelihood of hiring that same Pity candidate in comparison to the Baseline condition where no decoys were present. Again, neither of the Envy decoys increased preferences for the Envy candidate relative to the Baseline condition.

We did not have a priori predictions about the effect of decoys on non-yoked candidates. Experiment 1b indicated that the Envy Compromise decoy suppressed preferences for the Pity candidate relative to Baseline, and the Pity Asymmetric Dominance decoy suppressed preferences for the Envy candidate relative to the Baseline condition. However, one could argue that these two decoys *were* effective, insofar as they suppressed preferences for the competing candidate.

The candidates evaluated in Experiments 1a and 1b were all male. However, we know that women who exhibit competence but not warmth in hiring contexts are often victims of backlash (Heilman, 2012; Heilman & Okimoto, 2007; Phelan, Moss-Racusin, & Rudman, 2008; Rudman & Fairchild, 2004; Rudman, Moss-Racusin, Phelan, & Nauts, 2012). Therefore, it is possible that an individual's priors on the stereotypes of the candidates under consideration could bias attribute weightings and subsequently determine which decoys will have an impact on evaluation. Specifically, when candidates are female, yoked decoys may be more effective for Envy candidates than Pity candidates. In contrast, if gender stereotypes play no role in biasing attribute weights, we predict that there would be no difference in the effectiveness of decoys between a female versus male candidate choice set. We test these competing hypotheses in Experiment 1c.

### **Experiment 1c: Female Candidates and Decoys in Hiring**

Experiment 1c is identical to Experiment 1b, changing the gender of all candidates and decoys to female instead of male, and including a more feminine silhouette graphic (Figure 1b).

#### Method

*Participants and exclusions.* We aimed for a minimum of 150 participants per condition after exclusions to achieve 80% power to detect a small effect size. We recruited 799 participants

from Amazon Mechanical Turk to participate in this experiment. We did not exclude any participants. This resulted in a final sample size of  $N = 799$  participants (489 female, 306 male<sup>7</sup>;  $M_{\text{age}} = 35.45$  years,  $SD = 10.99$ ).

*Experimental design.* Experimental design was identical to Experiment 1b.

*Materials.* Materials were identical to the ones used in Experiment 1b, with two exceptions; we changed the gender of all candidates and decoys to female and used a more feminine silhouette graphic.

*Procedure.* The procedure was identical to Experiment 1b.

### Results

We fit a linear mixed model to predict likelihood of hiring from the fixed effects of candidate type (Pity, Envy) and decoy condition (Baseline, Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) along with their interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $X^2(9) = 98.94, p < 0.001$ . Furthermore, we compared the hypothesis-driven model with the interaction to a reduced model with no interaction term, and found that the interaction between candidate type and condition was significant,  $X^2(4) = 23.55, p < 0.001$ . Therefore, we report the results of the hypothesis-driven model including the interaction,  $mR^2 = 0.060$ .

There was a significant interaction between candidate type and decoy condition,  $F(4,1598) = 5.93, p < 0.001$ , qualified by the main effect of candidate type,  $F(1,1598) = 75.51, p < 0.001$ . The main effect of decoy condition was not significant,  $F(4,1598) = 0.628, p = 0.642$ . To unpack the omnibus interaction, we conducted paired contrasts on the estimated marginal

means extracted from the model to examine the effect of each decoy on the Pity candidate and Envy candidate, respectively. See Figure 4 for a summary of the results.

In contrast to Experiments 1a and 1b, the Pity Asymmetric Dominance decoy ( $M = 5.52$ ,  $SE = 0.094$ ) did not increase likelihood of hiring the Pity candidate in comparison to the Baseline condition (where there was no decoy;  $M = 5.46$ ,  $SE = 0.093$ ), 95% CI [-0.200,0.319],  $t(1598) = 0.45$ ,  $p = 0.653$  (Holm-corrected  $p = 1$ ),  $d = 0.050$ . Similarly, the Pity Compromise decoy ( $M = 5.38$ ,  $SE = 0.094$ ) did not increase likelihood of hiring in comparison to the Pity candidate's baseline, 95% CI [-0.335,0.183],  $t(1598) = -0.58$ ,  $p = 0.565$  (Holm-corrected  $p = 1$ ),  $d = -0.064$ .

Similar to Experiments 1a and 1b, the Envy Asymmetric Dominance decoy ( $M = 5.98$ ,  $SE = 0.093$ ) did not increase likelihood of hiring the Envy candidate in comparison to the Baseline condition ( $M = 5.72$ ,  $SE = 0.093$ ), 95% CI [-0.004,0.513],  $t(1598) = 1.93$ ,  $p = 0.054$  (Holm-corrected  $p = 0.215$ ),  $d = 0.216$ . Likewise, the Envy Compromise decoy ( $M = 5.92$ ,  $SE = 0.093$ ) did not increase likelihood of hiring in comparison to the Envy candidate's baseline, 95% CI [-0.059,0.457],  $t(1598) = 1.51$ ,  $p = 0.131$  (Holm-corrected  $p = 0.393$ ),  $d = 0.168$ .<sup>8</sup>

In two cases, the decoy had a significant effect on the non-yoked candidate. Replicating Experiment 1b, the Envy Compromise decoy decreased likelihood of hiring the Pity candidate ( $M = 5.04$ ,  $SE = 0.093$ ) in comparison to the Pity candidate's baseline, 95% CI [-0.674,-0.158],  $t(1598) = -3.16$ ,  $p = 0.002$  (Holm-corrected  $p = 0.006$ ),  $d = -0.352$ . Similarly, the Envy Asymmetric Dominance decoy decreased likelihood of hiring the Pity candidate ( $M = 5.16$ ,  $SE = 0.093$ ) in comparison to the Pity candidate's baseline, 95% CI [-0.556,-0.039],  $t(1598) = -2.26$ ,  $p = 0.024$  (Holm-corrected  $p = 0.073$ ),  $d = -0.252$ .

## Discussion

Experiment 1c extended the results of Experiments 1a and 1b to female candidates. In general, participants were still more likely to hire the Envy candidate in comparison the Pity candidate, as found in Experiments 1a and 1b. Interestingly, the difference between the Baseline candidates was smaller (0.261) as compared to Experiments 1a (0.945) and 1b (0.544).

In line with the hypothesis that gender stereotypes result in different weighting of attributes, the presence of both an Envy Asymmetric Dominance decoy and Envy Compromise decoy significantly lowered likelihood of hiring the Pity candidate. In other words, when candidates were women and therefore stereotypically “warm”, the Envy candidates benefited from the yoked Envy decoys through lowering the likelihood of hiring the Pity candidates, thus widening the evaluation gap between the Baseline candidates.

Experiments 1b and 1c examined all-male and all-female candidate/decoy choice sets, but in actual hiring scenarios, people often have to evaluate mixed-gender choice sets. In Experiment 1d, we examine whether male and female decoys are equally effective in the presence of both counter-stereotypical female and male candidates.

#### **Experiment 1d: Mixed-Gender Candidates and Decoys in Hiring**

To our knowledge, only one previous study has examined the influence of categorical attributes on decoy effects (Ha, Park, & Ahn, 2009) in the context of consumer products such as vacation packages. Their results indicated that a category-level match between the target and asymmetrically-dominated decoy facilitated the editing out of the decoy early in the decision-making process, making the decoy less impactful on judgments of the yoked targets. For example, in a comparison between two vacation packages, one taking travelers to Italy and one to France, the introduction of a French asymmetrically-dominated decoy vacation package did not affect preferences for the yoked French package because the decoy was categorically similar

and eliminated early for being obviously inferior. In contrast, the compromise decoy remained equally effective in the vacation package scenario because the decoy was not totally dominated by the same-category target.

Because of these results and because Experiments 1a-1c indicated that the compromise decoy was more consistently effective, here we compare the influence of either a male or a female compromise decoy (relative to a baseline in which no decoy is present) for a counter-stereotypical woman (high-competence/low-warmth) competing with a low-competence/high-warmth male candidate. We chose this combination of candidates because it is the case in which gender bias might lead evaluators to favor the less competent candidate. Based on previous results in the consumer behavior literature we predict that the female compromise decoy will be effective for a female candidate competing against a male candidate. This is the first study of which we aware to also test whether a category-level match between a decoy and the *competing* option will be equally effective for the yoked target.

### Method

*Participants and exclusions.* We aimed for a minimum of 150 participants per condition after exclusions to achieve 80% power to detect a small effect size. We recruited 482 participants from Amazon Mechanical Turk to participate in this experiment. We did not exclude any participants. This resulted in a final sample size of  $N = 482$  participants (282 female, 200 male;  $M_{\text{age}} = 35.43$  years,  $SD = 11.21$ ).

*Experimental design.* Participants were randomly assigned to one of three decoy conditions: Baseline (no decoy), Envy Compromise Female, and Envy Compromise Male. In the Baseline condition, participants jointly evaluated two candidates: a male Pity candidate and a female Envy candidate. In the two Decoy conditions, participants jointly evaluate three

candidates: the Pity and Envy candidates from the Baseline condition plus one additional Envy decoy candidate (either male or female).

*Materials.* Materials were identical to those used in Experiments 1b and 1c, with two exceptions. Because we have choice sets consisting of both male and female candidates, we removed silhouette graphics, and instead used names to indicate candidates' gender. The Pity candidate was always named Hunter McGrath, and the Envy candidate was always named Laurie Andersen. In the Envy Compromise Female condition, the decoy was named Stephanie Nielsen. In the Envy Compromise Male condition, the decoy was named Seth Nielsen (we matched names on perceptions of whiteness and associated SES; Gaddis, 2017).

*Procedure.* The procedure was identical to Experiment 1b.

## Results

We fit a linear mixed model to predict likelihood of hiring from the fixed effects of candidate type (Pity, Envy) and decoy condition (Baseline, Envy Compromise Female, Envy Compromise Male) along with their interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $\chi^2(5) = 125.65, p < 0.001$ . Furthermore, we compared the hypothesis-driven model with the interaction to a reduced model with no interaction term, and found that the interaction between candidate type and condition was marginal,  $\chi^2(2) = 4.72, p = 0.095$ . That said, we planned to conduct contrasts a priori to replicate the analyses of the three preceding experiments. Therefore, we report the results of the hypothesis-driven model including the interaction,  $mR^2 = 0.122$  (see Tybout et al., 2001 for justification for examining planned contrasts in absence of significant interaction).

There was a marginal interaction between candidate type and decoy condition,  $F(2,964) = 2.36, p = 0.095$ , qualified by the main effect of candidate type,  $F(1,964) = 119.97, p < 0.001$  and

the main effect of decoy condition,  $F(2,964) = 4.83, p = 0.008$ . To unpack the omnibus interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of each decoy on the Pity candidate and Envy candidate, respectively. See Figure 5 for a summary of the results.

Similar to Experiments 1a-1c, the Envy Compromise Male decoy ( $M = 5.88, SE = 0.097$ ) did not increase likelihood of hiring the Envy candidate in comparison to the Baseline condition ( $M = 5.91, SE = 0.097$ ), 95% CI [-0.293,-0.243],  $t(964) = -0.18, p = 0.856$  (Holm-corrected  $p = 1$ ),  $d = -0.020$ . Likewise, the Envy Compromise Female decoy ( $M = 5.82, SE = 0.097$ ) did not increase likelihood of hiring the Envy candidate in comparison to the Envy candidate's baseline, 95% CI [-0.357,0.181],  $t(964) = -0.64, p = 0.520$  (Holm-corrected  $p = 1$ ),  $d = 0.072$ .

However, the female decoy had a significant effect on the non-yoked candidate. Replicating Experiment 1b and 1c, the Envy Compromise Female decoy decreased likelihood of hiring the Pity candidate ( $M = 4.71, SE = 0.097$ ) in comparison to the Pity candidate's baseline ( $M = 5.19, SE = 0.097$ ), 95% CI [-0.749,-0.211],  $t(964) = -3.51, p < 0.001$  (Holm-corrected  $p = 0.001$ ),  $d = -0.392$ . In contrast, the Envy Compromise Male decoy ( $M = 5.11, SE = 0.097$ ) did not significantly change likelihood of hiring the Pity candidate in comparison to the Pity candidate's baseline, 95% CI [-0.355,0.181],  $t(964) = -0.64, p = 0.525$  (Holm-corrected  $p = 0.525$ ),  $d = -0.071$ .

## Discussion

Experiment 1d extended the results of Experiments 1b and 1c to mixed-gender candidate choice sets. First, the more competent female candidate received higher likelihood of hiring ratings than her competitor even at baseline. Second, the female decoy helped the female high-competence candidate by suppressing ratings of her male competitor (effectively replicating the

results of Experiment 1c which included only female candidates and decoys). In contrast, the *male* compromise decoy did not significantly impact ratings of either candidate. This pattern of results comports with previous findings that a category-level match—in our case gender—between a target (Envy Female candidate) and a decoy (Envy Compromise Female decoy) does not reduce the efficacy of the compromise decoy (it only reduces the efficacy of an asymmetric dominance decoy; Ha et al., 2009). However, no previous work in consumer behavior has examined the analog of our Envy Male Compromise condition: the case where there was a category-level mismatch between the target (Envy Female candidate) and compromise decoy (Envy Compromise Male decoy) and a category-level match between the competitor (Pity Male candidate) and compromise decoy. One possibility is that even though the Envy Male decoy was of the compromise variety, people ‘edited him out’ early in the evaluation process precisely because in one regard he was a mismatch with both targets. Specifically, participants were presented with two counter-stereotypic targets, and one stereotype-consistent decoy; perhaps this combination rendered him too distant to be comparable to either target, precluding him from exerting a decoy effect.

It is worth noting that ratings for the Envy candidate were close to ceiling across all of our initial experiments. These results suggest that participants’ priors when making these judgments include not only stereotypes of the candidates under consideration, but also the hiring context itself. In a hiring context, participants may place greater emphasis on aptitude as compared to attitude by default. If participants are indeed spontaneously weighing aptitude to a greater degree than attitude by default, Envy decoys—which theoretically exert their effects by increasing accessibility of the competence attribute—should have little to no effect as seen in Experiments 1a and 1b. In Experiment 2, we manipulate attribute emphasis to test this account.

### **Experiment 2: Manipulating Attribute Emphasis to Drive Decoy Effects in Hiring**

Experiments 1a and 1b indicated stronger effects of the Pity decoys on the male Pity candidate ratings relative to the Envy decoys on the male Envy candidate ratings; we observed the opposite pattern for female choice sets in Experiment 1c and mixed-gender choice sets in Experiment 1d, though in these latter cases the relative preference for the Envy candidate was driven by the Envy decoys' suppression of ratings of the Pity candidate. One possible explanation is that hiring contexts spontaneously make competence more salient for decision-makers, rendering the Envy decoy effects negligible relative to the context-driven weight on the competence attribute. (See also Keck & Tang, working paper, for different effects of decoys for male versus female targets in hiring contexts). Indeed, there is ample evidence to suggest that attentional bias (indexed by eye-tracking) in binary and trinary choice sets predicts choice bias (Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011). This work predicts that manipulations (including the context), which bias participants' attention, will also bias choices in favor of the most attended option when its value is positive.

In Experiment 2, we explicitly asked participants to weigh one dimension more heavily than the other while making decisions over a set of male candidates. If participants are spontaneously weighing aptitude more heavily than attitude, then explicitly asking them to weigh aptitude more heavily than attitude should reveal a pattern of results replicating Experiments 1a and 1b. In contrast, asking participants to weigh attitude more heavily than aptitude should reveal an effect of Envy decoys on the Envy candidate (but not Pity decoys on the Pity candidate).

#### Method

*Participants and exclusions.* We aimed for a minimum of 100 participants per condition after exclusions in order to have 80% power to detect a small to moderate effect size. We

recruited 2,363 participants from Amazon Mechanical Turk to participate in this experiment. Of these, 118 participants were excluded from all analyses for failing to report likelihood of hiring for all candidates in their assigned condition, 728 participants<sup>9</sup> were excluded from all analyses for failing the manipulation check (described below). This resulted in a final sample size of  $N = 1517$  (723 female, 790 male<sup>10</sup>;  $M_{\text{age}} = 33.38$  years,  $SD = 10.48$ ).

*Experimental design.* Participants were randomly assigned to one of ten conditions: 2 attribute emphasis (Aptitude, Attitude)  $\times$  5 decoy condition (Baseline, Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise).

*Materials.* Materials were identical to the ones used in Experiments 1a and 1b.

*Procedure.* The procedure was identical to Experiment 1b, with one exception. The initial prompt included an additional sentence telling participants that the company had explicitly asked them to weigh heavily the Global Attitude or Global Aptitude assessment. After jointly evaluating the set of candidates for at least 10 seconds and reporting likelihood of hiring for all candidates (identical to Experiment 1b), participants completed a manipulation check question asking them to identify which assessment attribute they were asked to weigh more heavily.

## Results

We fit a linear mixed model to predict hiring likelihood from the fixed effects of candidate type (Pity, Envy), attribute emphasis (Aptitude, Attitude), and decoy condition (Baseline, Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) along with the three-way interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $\chi^2(19) = 1016.4$ ,  $p < 0.001$ . Furthermore, we compared the hypothesis-driven model with the three-way interaction to a reduced model with no interaction terms, and found that the interaction between candidate type,

attribute emphasis, and decoy condition was significant,  $X^2(13) = 1002.5, p < 0.001$ . Therefore, we report the results of the hypothesis-driven model including the interaction,  $mR^2 = 0.285$ .

The significant three-way interaction between candidate type, attribute emphasis, and decoy condition,  $F(4,3034) = 3.17, p = 0.013$ , qualified the significant main effects of candidate type ( $F(1,3034) = 10.31, p = 0.001$ ) and attribute emphasis ( $F(1,3034) = 11.96, p < 0.001$ ) as well as the significant two-way interactions between candidate type and attribute emphasis ( $F(1,3034) = 1015.54, p < 0.001$ ), candidate type and decoy condition ( $F(4,3034) = 16.13, p < 0.001$ ), and attribute emphasis and decoy condition ( $F(4,3034) = 12.29, p < 0.001$ ). The main effect of decoy condition was not significant  $F(4,3034) = 0.12, p = 0.977$ . To unpack the omnibus three-way interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of each decoy on the Pity candidate and Envy candidate when aptitude was emphasized and when attitude was emphasized. See Figure 6a for a summary of the ‘aptitude-emphasized’ results and Figure 6b for the ‘attitude-emphasized’ results.

When aptitude was emphasized, we replicated the results of Experiments 1a and 1b: the Pity Asymmetric Dominance decoy increased likelihood of hiring the Pity candidate ( $M = 5.05, SE = 0.100$ ) in comparison to the Aptitude Baseline condition (where there was no decoy;  $M = 4.62, SE = 0.098$ ), 95% CI [0.150, 0.701],  $t(3034) = 3.03, p = 0.003$  (Holm-corrected  $p = 0.008$ ),  $d = 0.343$ . Similarly, the Pity Compromise decoy increased likelihood of hiring the Pity candidate ( $M = 5.19, SE = 0.099$ ) in comparison to the Aptitude Baseline condition, 95% CI [0.297, 0.845],  $t(3034) = 4.08, p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = 0.460$ .<sup>11</sup> As in Experiment 1b, we also observed a suppressing effect of the Envy Compromise decoy ( $M = 4.22,$

$SE = 0.096$ ) on likelihood of hiring the Pity candidate, 95% CI [-0.669,-0.129],  $t(3034) = -2.90$ ,  $p = 0.004$  (Holm-corrected  $p = 0.008$ ),  $d = -0.321$ .

Also replicating Experiments 1a and 1b, the Envy Asymmetric Dominance decoy did not increase likelihood of hiring the Envy candidate ( $M = 6.15$ ,  $SE = 0.098$ ) in comparison to the Aptitude Baseline condition (where there was no decoy;  $M = 6.11$ ,  $SE = 0.099$ ), 95% CI [-0.230,0.317],  $t(3034) = 0.31$ ,  $p = 0.756$  (Holm-corrected  $p = 0.756$ ),  $d = 0.035$ . Similarly, the Envy Compromise decoy did not increase likelihood of hiring the Envy candidate ( $M = 5.91$ ,  $SE = 0.096$ ) in comparison to the Aptitude Baseline condition, 95% CI [-0.467,0.073],  $t(3034) = -1.43$ ,  $p = 0.152$  (Holm-corrected  $p = 0.455$ ),  $d = -0.159$ .<sup>12</sup>

However, when attitude was emphasized, we saw the opposite, predicted pattern: the Envy Compromise decoy increased likelihood of hiring the Envy candidate ( $M = 5.01$ ,  $SE = 0.113$ ) in comparison to the Attitude Baseline condition (where there was no decoy;  $M = 4.38$ ,  $SE = 0.103$ ), 95% CI [0.325,0.924],  $t(3034) = 4.09$ ,  $p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = 0.504$ . Contrary to our prediction, the Envy Asymmetric Dominance decoy did not significantly change likelihood of hiring the Envy candidate ( $M = 4.39$ ,  $SE = 0.104$ ) as compared to the Attitude Baseline condition, 95% CI [-0.281,0.294],  $t(3034) = 0.04$ ,  $p = 0.965$  (Holm-corrected  $p = 0.965$ ),  $d = 0.005$ .<sup>13</sup> We also observed a suppressing effect of the Pity Asymmetric Dominance decoy on likelihood of hiring the Envy candidate ( $M = 4.05$ ,  $SE = 0.093$ ) as compared to the Attitude Baseline, 95% CI [-0.610,-0.066],  $t(3034) = -2.44$ ,  $p = 0.015$  (Holm-corrected  $p = 0.044$ ),  $d = -0.273$ .

Finally, and in line with our predictions, once we emphasized attitude, the Pity Asymmetric Dominance decoy did not increase likelihood of hiring the Pity candidate ( $M = 6.15$ ,  $SE = 0.093$ ) in comparison to the Attitude Baseline condition (where there was no decoy;  $M =$

$5.92$ ,  $SE = 0.103$ ), 95% CI [-0.044,0.500],  $t(3034) = 1.64$ ,  $p = 0.100$  (Holm-corrected  $p = 0.402$ ),  $d = 0.184$ . Similarly, the Pity Compromise decoy did not increase likelihood of hiring the Pity candidate ( $M = 5.87$ ,  $SE = 0.106$ ) in comparison to the Attitude Baseline condition, 95% CI [-0.344,0.234],  $t(3034) = -0.37$ ,  $p = 0.708$  (Holm-corrected  $p = 1$ )  $d = -0.044$ .<sup>14</sup>

### Discussion

Experiment 2 replicated and extended the results of Experiment 1b. When we told participants to emphasize candidates' aptitude, both Pity decoys increased participants' self-reported likelihood of hiring the Pity candidate in comparison to the Aptitude Baseline condition where no decoys were present. In contrast (and in line with our predictions), when we told participants to emphasize candidates' attitude, the Envy Compromise decoy increased likelihood of hiring the Envy candidate (and we saw no effect of Pity decoys on the Pity candidate). These results support our proposal that manipulations which bias participants' attention, will also bias choices in favor of the most attended option when its value is positive.

### **Experiment 3a: Manipulating Attribute Exposure in Hiring**

Ideally, bias-reduction interventions do not require fabricating or selectively including social decoys in choice sets. The aim of this experiment is to mimic the effect of social decoys (in the absence of actual decoys) by manipulating participants' exposure to candidates' attribute information.

According to the Associative Accumulation Model (AAM; Bhatia, 2013) decoy effects arise because attributes that are abundantly present in multiple options within a choice set will be more accessible relative to attributes that dominate in fewer options. Therefore, these abundant attributes will be sampled more frequently across iterations of sequential sampling, biasing attribute accumulation and resulting preferences. Both the AAM and the attentional Drift

Diffusion Model (Krajbich & Rangel, 2011) indicate that especially salient attributes (and the options in which those positive attributes are dominant) will be favored at choice. Thus, by directing participants' attention to (and therefore weighting of) specific attributes of the candidates, we can leverage their attention to bias their resulting preferences (Dawes & Corrigan, 1974).

In Experiment 3a, we manipulated the distribution of participants' exposure to candidates' warmth and competence scores. Specifically, we increased participants' exposure to candidates' warmth (relative to competence) ratings, their competence (relative to their warmth) ratings, or provided equivalent exposure to both attributes to examine whether increased exposure (and therefore increased salience) to warmth information increased preferences for the Pity candidate/decreased preferences for the Envy candidate. The hypotheses and exclusion criteria for this experiment are pre-registered on OSF:

[https://osf.io/frj5q/?view\\_only=14e54a53d22f4420a0e38bb1cfcdc9dd](https://osf.io/frj5q/?view_only=14e54a53d22f4420a0e38bb1cfcdc9dd)

### Method

*Participants and exclusions.* We aimed for a minimum of 320 participants per condition after exclusions in order to have 80% power to detect a small effect size. We recruited 1173 participants from Amazon Mechanical Turk to participate in this experiment. Of these, 151 participants were excluded for answering more than 10% (3 trials) of the exposure trials incorrectly. A further 49 participants were excluded for not completing the experiment. Finally, 2 participants were excluded not reporting likelihood of hiring for both candidates. This resulted in a final sample size of  $N = 971$  participants (506 female, 462 male<sup>15</sup>;  $M_{age} = 34.98$  years,  $SD = 11.02$ ).

*Experimental design.* Because so many participants failed the manipulation check in Experiment 2—in part due to the similarity of the words ‘Aptitude’ and ‘Attitude’—we used ‘Warmth’ and ‘Competence’ as the attribute labels for the third experiment. Participants were randomly assigned to one of three conditions: Equal exposure (50% warmth/50% competence trials), Warmth-biased exposure (80% warmth/20% competence trials), and Competence-biased exposure (20% warmth/80% competence trials; see Procedure for details).

*Materials.* Materials were identical to the ones we used in Experiments 1a, 1b, and 2 with two exceptions: (1) we replaced the personality assessment labels of ‘Aptitude’ and ‘Attitude’ with ‘Competence’ and ‘Warmth’, respectively; (2) on each trial we displayed *either* both candidates’ warmth *or* both candidates’ competence scores (see Figure 7).

*Procedure.* As in the previous experiments, we asked participants to imagine that they had been put in charge of hiring a human resources consultant to advise their boss on strategy. We then presented participants with 30 trials featuring the same two candidate silhouettes (Pity: high-warmth/low-competence; Envy: low-warmth/high-competence) and their respective warmth *or* competence ratings on each trial. We asked participants to report on each trial which candidate was warmer/more competent to make sure they were encoding each candidate’s scores. The position of the candidates (top or bottom) and the response buttons changed on each trial. Finally, we randomized the order of warmth and competence exposure trials across participants. After each participant completed 30 exposure trials, they reported how likely they were to hire each candidate in randomized order using a 7-point scale (*very unlikely* to *very likely*).

## Results

We fit a linear mixed model to predict hiring likelihood from the fixed effects of candidate type (Pity, Envy) and exposure condition (Equal exposure, Warmth-biased exposure, Competence-biased exposure), along with their interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $X^2(5) = 133.01, p < 0.001$ . Furthermore, we compared the hypothesis-driven model with the interaction to a reduced model with no interaction term, and found that the interaction between candidate type and condition was significant,  $X^2(2) = 11.89, p = 0.003$ . Therefore, we report the results of the hypothesis-driven model including the interaction,  $mR^2 = 0.066$ .

The significant interaction between candidate type and exposure condition,  $F(2,971) = 5.98, p = 0.003$ , qualified the main effect of candidate type,  $F(1,971) = 128.17, p < 0.001$ . The main effect of exposure condition was not significant,  $F(2,971) = 0.05, p = 0.951$ . To unpack the omnibus interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of each exposure condition on the Pity candidate and Envy candidate, respectively. See Figure 8 for a summary of the results.

Participants were least likely to hire the Pity candidate in the Competence-biased condition ( $M = 4.80, SE = 0.070$ ), but were significantly more likely to hire the same candidate in the Equal exposure condition ( $M = 5.04, SE = 0.070$ ), 95% CI [-0.442,0.053],  $t(1941.14) = -2.50, p = 0.013$  (Holm-corrected  $p = 0.038$ ),  $d = -0.199$ , and the Warmth-biased exposure condition ( $M = 4.98, SE = 0.070$ ); though the difference between Warmth-biased vs. Competence-biased was marginal: 95% CI [-0.383,0.005],  $t(1941.14) = -1.91, p = 0.057$  (Holm-corrected  $p = 0.113$ ),  $d = -0.152$ . In contrast to our prediction, we did not find a significant increase in likelihood of hiring the Pity candidate in the Warmth-biased exposure condition in

comparison to the Equal exposure condition, 95% CI [-0.253,0.135],  $t(1941.14) = -0.59$ ,  $p = 0.553$  (Holm-corrected  $p = 0.55$ ),  $d = -0.047$ .

For the Envy candidate, we find a similar but opposite pattern. Participants were more likely to hire the Envy candidate in the Competence-biased exposure condition ( $M = 5.71$ ,  $SE = 0.070$ ) as compared to the Equal exposure condition ( $M = 5.50$ ,  $SE = 0.07$ ), 95% CI [0.010,0.399],  $t(1941.14) = 2.06$ ,  $p = 0.040$  (Holm-corrected  $p = 0.119$ ),  $d = 0.164$ , and the Warmth-biased exposure condition ( $M = 5.53$ ,  $SE = 0.070$ ), 95% CI [-0.018,0.371],  $t(1941.14) = 1.78$ ,  $p = 0.075$  (Holm-corrected  $p = 0.149$ ; though this second comparison is marginal),  $d = 0.142$ . We find no significant difference in likelihood of hiring the Envy candidate in the Warmth-biased exposure condition and the Equal exposure condition, 95% CI [-0.167,0.222],  $t(1941.14) = 0.28$ ,  $p = 0.780$  (Holm-corrected  $p = 0.780$ ),  $d = 0.022$ .

### Discussion

In Experiment 3a, we demonstrated that increased exposure to specific attribute information across two candidates can mimic the decoy effect we observed in Experiments 1a, 1b, and 2. Based on the results of the previous experiments, we originally predicted that we would observe an increase in likelihood of hiring the Pity candidate in the Warmth-biased exposure condition compared to the Equal exposure condition. Though the data did not support that specific comparison, the results indicated that as attribute exposure shifted from competence-biased to equal exposure, the gap between the Envy and Pity candidate decreased. These results suggest that the Equal exposure condition may already function like an intervention, rather than a baseline (and that the Competence-biased exposure condition is a better approximation of the Baseline condition of Experiments 1a and 1b, and the aptitude-

emphasized baseline condition in Experiment 2). Thus, these results indicate that differential exposure to specific attributes can shift hiring preferences, even in the absence of a decoy.

There are, however, two major limitations to this and the preceding experiments. The first is a lack of ecologically validity in the cues we used (i.e., hearts and stars) to provide participants with warmth and competence information. In the real world, managers typically read resumes to make inferences about prospective employees' latent attitude and aptitude. The second concern is that these results may be unique to naïve participants and would not generalize to individuals who have experience making hiring evaluations and decisions. Experiment 3b replicates 3a and addresses these limitations.

### **Experiment 3b: Manipulating Attribute Exposure via Resume-based Cues**

In Experiment 3b, we replicate, and improve the design of Experiment 3a in two ways. First, instead of using heart and star ratings to signify warmth and competence, we provide participants with information ostensibly gleaned from applicants' resumes. This approach allows participants to build latent representations of the candidates' relative global warmth and competence. Resumes are often the first communication between an applicant and potential employer, and are used to screen applicants before deciding to reject or interview them (Cole, Feild, Giles & Harris, 2009). Previous research examining recruiters inferences from resume content suggest that these inferences are predictive of hireability perceptions (Burns, Christiansen, Morris, Periard, & Coaster, 2014; Cole, Feild, Giles, & Harris, 2004; Cole, Feild, & Stafford, 2005; Cole et al., 2009).

Second, we include a question at the end of the experiment to assess whether each participant has had any experience evaluating resumes and then examine whether our condition-based results interact with participants' status as experienced or inexperienced. This approach has

been taken by previous researchers (e.g., Burns et al., 2014) who found that there was no difference in the resume evaluations of mTurkers with versus without past hiring experience.

### Method

*Participants and exclusions.* We aimed for a minimum of 320 participants per condition after exclusions to have 80% power to detect a small effect size. We recruited 1351 participants from Amazon Mechanical Turk to participate in this experiment. Of these, 366 participants were excluded for answering more than 10% (3 trials) of the exposure trials incorrectly. A further 14 participants were excluded for not completing the experiment. Finally, 1 participant was excluded for not reporting likelihood of hiring for both candidates. This resulted in a final sample size of  $N = 970$  participants (501 female, 461 male<sup>16</sup>;  $M_{age} = 36.67$  years,  $SD = 11.21$ ). Additionally, 827 participants (85.3%) reported being currently employed or employed in the last six months, and 465 participants (47.9%) reported experience evaluating resumes as part of their current or past job.

*Experimental design.* Experimental design was identical to Experiment 3a.

*Materials.* Using HR resume examples we found online, and based on previous research on resume evaluations (Burns et al., 2014; Cole et al., 2004; Cole et al., 2005; Cole et al., 2009), we created 6 unique competence cues (e.g., Candidate A: Managers across different departments express 98% satisfaction with organization of job searches. vs. Candidate B: Managers across different departments express 83% satisfaction with organization of job searches.) and 6 unique warmth cues (e.g., Candidate A: Nominated for mentorship award. vs. Candidate B: Won a mentorship award.) for each candidate that were easily comparable.

*Procedure.* Procedure was identical to Experiment 3a, with a few exceptions. We did not change the position of the response buttons on each trial, because the information each

participant needed to encode for the exposure trials was no longer identical. In the Equal-exposure condition, for both attribute exposure trials, participants saw two repetitions of the 6 exposure items in randomized order, along with three randomly chosen items for each attribute, totaling 15 exposure trials for each attribute. In the biased exposure conditions (Warmth-biased, Competence-biased), participants saw 6 exposure trials for one attribute, and 24 exposure trials for the other. The 24 exposure trials consisted of the 6 exposure items repeated 4 times, while the 6 exposure trials consisted of the 6 exposure items. This allowed us to keep the variety and number of cues consistent across all three conditions. The order of the warmth and competence exposure trials was still randomized across participants.

## Results

We fit a linear mixed model to predict hiring likelihood from the fixed effects of candidate type (Pity, Envy) and exposure condition (Equal exposure, Warmth-biased exposure, Competence-biased exposure), along with their interaction, including participant as a random effect. This hypothesis-driven model provided a better fit than the null model,  $X^2(5) = 610.97, p < 0.001$ . Furthermore, we compared the hypothesis-driven model with the interaction to a reduced model with no interaction term, and found that the interaction between candidate type and condition was significant,  $X^2(2) = 609.84, p < 0.001$ . We also compared the hypothesis-driven model to a model with a three-way interaction between candidate type, exposure condition, and experience evaluating resumes, and similarly found no significant improvement in model fit,  $X^2(6) = 2.56, p = 0.862$ . Additionally, we compared the hypothesis-driven model to a model with a three-way interaction between candidate type, exposure condition, and employment status, and found no significant improvement in model fit,  $X^2(6) = 2.72, p = 0.843$ . Therefore, we

report the results of the hypothesis-driven model including the two-way interaction between candidate type and exposure condition,  $mR^2 = 0.270$ .

There was a significant interaction between candidate type and exposure condition,  $F(2,1940) = 358.29, p < 0.001$ . The main effects of candidate type,  $F(1,1940) = 0.07, p = 0.790$ , and exposure condition,  $F(2,1940) = 0.73, p = 0.482$ , were not significant. To unpack the omnibus interaction, we conducted paired contrasts on the estimated marginal means extracted from the model to examine the effect of each exposure condition on the Pity candidate and Envy candidate, respectively. See Figure 9 for a summary of the results.

Participants were significantly less likely to hire the Pity candidate in the Competence-biased condition ( $M = 4.50, SE = 0.063$ ), relative to the Equal exposure condition ( $M = 5.43, SE = 0.063$ ), 95% CI [-1.105, -0.756],  $t(1940) = -10.45, p < 0.001$  (Holm-corrected  $p < .001$ ),  $d = -0.822$ , and the Warmth-biased exposure condition ( $M = 6.22, SE = 0.063$ ), 95% CI [-1.888, -1.539],  $t(1940) = -19.25, p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = -1.513$ . Participants were also significantly more likely to hire the Pity candidate in the Warmth-biased exposure condition in comparison to the Equal exposure condition, 95% CI [-0.958, -0.609],  $t(1940) = -8.79, p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = -0.692$ .

For the Envy candidate, we find the opposite pattern. Participants were most likely to hire the Envy candidate in the Competence-biased exposure condition ( $M = 6.21, SE = 0.063$ ) as compared to the Equal exposure condition ( $M = 5.43, SE = 0.063$ ), 95% CI [0.605, 0.954],  $t(1940) = 8.76, p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = 0.689$ , and the Warmth-biased exposure condition ( $M = 4.55, SE = 0.063$ ), 95% CI [1.481, 1.830],  $t(1940) = 18.60, p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = 1.462$ . Similarly, they were less likely to hire the Envy

candidate in the Warmth-biased exposure condition relative to the Equal exposure condition, 95% CI [0.701,1.051],  $t(1940) = 9.84, p < 0.001$  (Holm-corrected  $p < 0.001$ ),  $d = 0.774$ .

### Discussion

In Experiment 3b, we demonstrated that increased exposure to specific attribute information across two candidates mimics the decoy effect we observed in Experiments 1a, 1b, and 2. In line with our hypothesis (pre-registered for Experiment 3a), we found that participants were most likely to hire the high-competence candidate in the Competence-biased condition, most likely to hire the high-warmth candidate in the Warmth-biased condition, and indifferent between the two candidates in the Equal exposure condition.

### General Discussion

Across seven experiments, we demonstrated that perceivers can have systematically different preferences for the exact same candidate as a function of the other candidates in the choice set (Experiments 1a-1d and 2) and the salience of the candidate attributes under consideration (Experiments 2, 3a and 3b). Across all the experiments, participants preferred the Envy (high-competence/low-warmth) to the Pity (high-warmth/low-competence) candidate (except in the attitude-emphasis condition in Experiment 2). However, in Experiments 1a and 1b, the presence of Pity Compromise and Pity Asymmetric Dominance decoys significantly increased participants' likelihood of hiring the Pity candidate (Envy decoys had no effect on the Envy candidate). In Experiment 1c with all female candidates, the presence of the Envy Compromise and Envy Asymmetric Dominance decoys significantly decreased participants' likelihood of hiring the Pity candidate, increasing the relative preference for the female Envy candidates. In Experiment 1d, the presence of a female Envy Compromise decoy increased the relative preference for a female Envy candidate relative to a male competitor (however a male

decoy had no effect on ratings of either candidate). In Experiment 2, we replicated our results from Experiments 1a and 1b when participants weighed aptitude more heavily, but observed the opposite, predicted pattern when they weighed attitude more heavily: now the presence of the Envy Compromise decoy increased likelihood of hiring the male Envy candidate and the Pity decoys had no effect on the male Pity candidate. In Experiment 3a, we manipulated participants' exposure to the candidates' attributes to mimic the effect of a decoy: when we balanced participants' exposure to warmth and competence information, their preference for the Envy candidate decreased and preference for the Pity candidate increased. We did not, however, observe the predicted preference inversion in the warmth-biased condition, in which we expected to see a greater preference for the Pity than the Envy candidate. Experiment 3b replicated 3a using resume-based cues instead of hearts and stars. When participants' exposure to warmth and competence information was balanced, they were equally likely to hire the Pity and Envy candidates. Furthermore, in this case we did see the predicted preference inversion in the warmth-biased exposure condition. More importantly, this result was not moderated by participants' past experience with resume evaluation suggesting that experienced evaluators are equally susceptible to these manipulations.

One possible reason for the divergent results in the warmth-biased exposure condition across Experiments 3a and 3b is that participants may have construed the warmth items in Experiment 3b as indices of interpersonal competence (whereas the hearts stood in for global warmth, irrelevant to past professional successes). Recall, however, that on each screen of 3b we asked participants “which candidate is more warm?” Furthermore, participants preferred the Pity candidate in the warmth-biased condition of 3b despite that candidate being judged as less competent on the competence-judgment screens. In other words, even if participants were

reappraising the warmth information as interpersonal competence, that candidate was still less competent on the “bottom-line competence” items relative to the competitor. In light of the results that the Envy candidate was preferred at baseline across all but one of our experiments, it is heartening that our more ecologically valid warmth-exposure manipulation inverted participants’ preferences entirely.

In Experiments 1b, 1c, 1d and 2 we also observed that two decoys had a suppressive effect on the unintended (i.e., non-yoked) candidate. Specifically, the Envy Compromise decoy suppressed preferences for the Pity candidate relative to Baseline, and the Pity Asymmetric Dominance decoy suppressed preferences for the Envy candidate relative to the Baseline condition. Thus it appears that decoys can reduce the gap in preferences between the Envy and Pity candidates, not only by increasing preferences for the yoked candidate, but by decreasing preferences for the competitor. For example, in Figure 3, the Pity Asymmetric Dominance decoy significantly increased preferences for the Pity candidate, and decreased preferences for the Envy candidate (relative to their respective baselines).

One possible interpretation of the results of Experiments 1a, 1b, and 2 is that participants interpreted the decoy as a signal of the firm’s priorities in the absence of any effect on attribute weighting (e.g., if I see two female, but only one male candidate in a pool, it may mean the firm would prefer to hire a woman; Johnson, Hekman, Chan, 2016). However, Experiments 3a and 3b addressed this alternative explanation because it directly manipulated participants’ attention to attributes in the absence of a third candidate. Furthermore, it appears that the Equal exposure condition already functioned as an effective preference intervention (i.e., reduced the gap between the Pity and Envy candidates). That is, in Experiments 3a Equal exposure was just as effective as Warmth-biased exposure, and could not have implicitly communicated that the firm

preferred a high-warmth candidate because the attributes were equally represented. These results also dovetail nicely with the results of Experiments 1a, 1b, and 2 in that the Competence-biased exposure condition most closely resembled the baseline conditions of the earlier experiments, providing further evidence that participants spontaneously weigh competence more in hiring decisions.

The results of Experiments 1c and 1d, including female candidates, documented a reversal of the effects we observed in Experiments 1a and 1b. In Experiment 1c the female Pity decoys had no effect on the candidate ratings relative to baseline, whereas the Envy decoys increased the relative preference for the female Envy candidate as compared to the female Pity candidate. This result replicated in Experiment 1d even when the high-competence/low-warmth female candidate was compared to a low-competence/high-warmth male competitor (but only when the decoy was female). These results suggest that people's priors, not only on the decision-making context (i.e., a hiring scenario), but also on the stereotypes of the targets under consideration can bias their attribute weightings, which then determine which decoys will have an impact on evaluation. Said another way, decoys were most effective for counter-stereotypical male and female candidates—precisely the targets that are most likely subject to discrimination in the real world.

Because we observed weaker effects of asymmetric dominance decoys, we ran a third replication of Experiments 1a and 1b, instead using range-asymmetric dominance decoys (see supplementary materials). Previous research on consumer behavior (Huber, Payne, & Puto 1982; Wedell 1991) and perceptual decision-making (Trueblood, Brown, Heathcote, & Busemeyer 2013), finds that range decoys (R-AD) produce the strongest effects, followed by range-frequency decoys (RF-AD, which we used across all our experiments), and then frequency

decoys (F-AD). We found that Pity R-AD decoy significantly decreased likelihood of hiring the Envy candidate as compared to Baseline and the Envy R-AD decoy significantly decreased likelihood of hiring the Pity candidate as compared to Baseline. Thus R-AD decoys were effective insofar as they suppressed ratings of the competing candidate. For example, the Pity R-AD decoy eliminated the gap between the Pity and Envy candidates. As Huber (1982) notes, it will be easier to detect dominance if one must consider only one dimension.

We have known for decades that one reason decision-makers make suboptimal choices is because they apply attribute weights inconsistently across options in a choice set (Dawes & Corrigan, 1974). For example, one may recruit a male prospective graduate student with less research experience instead of a female graduate student with more research experience because even though research experience is important, his higher GRE scores suddenly seem more diagnostic than his research experience (Norton et al., 2004). Our results indicate that structuring decision-makers' exposure to attribute information to decrease the error associated with inconsistent attribute weighting could help people make more equitable decisions (so long as they determined how they wanted to weigh different attributes consistently across candidates ahead of time).

The current findings are the first demonstration that manipulating weights on the cardinal dimensions of social cognition—either via decoys or attribute exposure—can affect hiring preferences among naïve participants. However, these experiments represent only the first step in a much broader program of research and serve more as a proof of principle. Future studies should explore different attributes and richer sources of attribute information (e.g., images of different social groups; facial masculinity; racial phenotypicality) and move beyond hypothetical scenarios with non-experts to field experiments in HR departments with participants who have

extensive experience hiring employees. Furthermore, there are many other consequential domains in which these effects could be tested, including but not limited to housing, education, and health. Going forward, incorporating these decision-making models into the research on social bias will significantly advance our understanding of how context gives rise to discrimination. These models make specific predictions about (i) the mechanisms by which social decoys influence individuals' decisions and (ii) the temporal dynamics underlying the decision process. Integrating insights from these models into the study of social-decision making allows for greater predictive precision and will stimulate innovative strategies for reducing bias.

Taken together, these experiments highlight a novel approach to increasing opportunity and access to marginalized social groups. Whereas previous bias reduction strategies have prioritized changing perceivers' stereotypes and implicit prejudices (which are resistant to long-term change, Lai et al., 2014; and may backfire, Vorauer & Sasaki, 2009), our approach seeks to debias the decision-making process itself. Our hope is that this and subsequent work will complement prejudice-reduction strategies to bring about greater social equity.

## References

Ariely, D., & Wallsten, T. S. (1995). Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect. *Organizational Behavior and Human Decision Processes*, 63(3), 223-232.

Bartoń, K. (2017). MuMIn: Multi-Model Inference. R package version 1.40.0. <https://CRAN.R-project.org/package=MuMIn>.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Bateson, M., Healy, S. D., & Hurly, T. A. (2003). Context-dependent foraging decisions in rufous hummingbirds. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1521), 1271-1276.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), 351-368.

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, 120(3), 522-543.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5-20.

Bohnet, I., Van Geen, A., & Bazerman, M. (2015). When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Management Science*, 62(5), 1225-1234.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science*, 6(1), 3-5.

Burns, G. N., Christiansen, N. D., Morris, M. B., Periard, D. A., & Coaster, J. A. (2014). Effects of applicant personality on resume evaluations. *Journal of Business and Psychology*, 29(4), 573-591.

Brown, B. K., & Campion, M. A. (1994). Biodata phenomenology: Recruiters' perceptions and use of biographical information in resume screening. *Journal of Applied Psychology*, 79, 897–908.

Choplin, J. M., & Hummel, J. E. (2005). Comparison-induced decoy effects. *Memory & Cognition*, 33(2), 332-343.

Cole, M. S., Feild, H. S., Giles, W. F., & Harris, S. G. (2004). Job type and recruiters' inferences of applicant personality drawn from resume biodata: Their relationships with hiring recommendations. *International Journal of Selection and Assessment*, 12, 363–367.

Cole, M. S., Feild, H. S., Giles, W. F., & Harris, S. G. (2009). Recruiters' inferences of applicant personality based on resume screening: Do paper people have a personality? *Journal of Business and Psychology*, 24, 5–18.

Cole, M. S., Feild, H. S., & Stafford, J. O. (2005). Validity of resume reviewers' inferences concerning applicant personality based on resume evaluation. *International Journal of Selection and Assessment*, 13, 321–324.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631-648.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81(2), 95-106.

Doyle, J. R., O'Connor, D. J., Reynolds, G. M., & Bottomley, P. A. (1999). The robustness of the asymmetrically dominated effect: buying frames, phantom alternatives, and in-store purchases. *Psychology & Marketing*, 16(3), 225-243.

Edelman, B. G., Luca, M., Svirsky, D. Racial discrimination in the sharing economy: Evidence from a field experiment. Forthcoming in *American Economic Journal: Applied Economics*.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77-83.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (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.

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. *Journal of Marketing Research*, 51(4), 487-507.

Gaddis, S. M. (2017). How Black are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies. *Sociological Science*, 4, 469-489.

Ha, Y. W., Park, S., & Ahn, H. K. (2009). The influence of categorical attributes on choice context effects. *Journal of Consumer Research*, 36(3), 463-477.

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113-135.

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communal deficit. *Journal of Applied Psychology*, 92, 81-92.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466(7302), 29-29.

Herne, K. (1997). Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy*, 13(3), 575-589.

Highhouse, S. (1996). Context-dependent selection: The effects of decoy and phantom job candidates. *Organizational Behavior and Human Decision Processes*, 65(1), 68-76.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, 113(16), 4296-4301.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1), 90-98.

Hurly, T. A., & Oseen, M. D. (1999). Context-dependent, risk-sensitive foraging preferences in wild rufous hummingbirds. *Animal Behaviour*, 58(1), 59-66.

Johnson, S. K., Hekman, D. R., & Chan, E. T. (2016, April 26). If there's only one woman in your candidate pool, there's statistically no chance she'll be hired. *Harvard Business Review*.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100, 109-128.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136-153.

Keck, S., & Tang, W. (2015). Choice Sets and Gender Bias in Hiring Decisions: The Role of Decoys. *SSRN*.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292-1298.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33), 13852-13857.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26.

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L., Joy-Gaba, J. A., ... & Frazier, R. S. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4), 1765-1785.

Latty, T., & Beekman, M. (2011). Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1703), 307-312.

Lenth, R. (2017). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 0.9.1. <https://CRAN.R-project.org/package=emmeans>.

Lim, G. S., & Chan, C. (2001). Ethical values of executive search consultants. *Journal of Business Ethics*, 29(3), 213-226.

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, 110(15), 6139-6144.

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.

O'Curry, Y. P. S., & Pitts, R. (1995). The attraction effect and political choice in two elections. *Journal of Consumer Psychology*, 4(1), 85-101.

Oakes, P., Haslam, S. A., & Turner, J. C. (1998). The role of prototypicality in group influence and cohesion: Contextual variation in the graded structure of social categories. In S. Worchel, J. F. Morales, D. Paez, & J.-C. Deschamps (Eds.), *Social identity: International perspectives* (pp. 75–92). London, England: Sage.

Pager, D., Shepherd, H. (2008). The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets. *Annual Review of Sociology*. 38, 181-208.

Paolacci, G., Chandler, J., & Ipeirotis, P.G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5, 411–419.

Pettibone, J. C., & Wedell, D. H. (2000). Examining models of nondominated decoy effects across judgment and choice. *Organizational Behavior and Human Decision Processes*, 81(2), 300-328.

Phelan, J. E., Moss-Racusin, C. A., & Rudman, L. A. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly*, 32(4), 406-413.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108(2), 370-392.

Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, 87, 157-176.

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48, 165-179.

Sedikides, C., Ariely, D., & Olsen, N. (1999). Contextual and procedural determinants of partner selection: Of asymmetric dominance and prominence. *Social Cognition*, 17(2), 118-139.

Shafir, S., Waite, T. A., & Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, 51(2), 180-187.

Simonsohn, U. & Gino, F. (2013). Daily horizons: evidence of narrow bracketing in judgment from 10 years of MBA admissions interviews. *Psychological Science*, 24(2), 219-224.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16(2), 158-174.

Trope, Y., & Mackie, D. M. (1987). Sensitivity to alternatives in social hypothesis-testing. *Journal of Experimental Social Psychology*, 23(6), 445-459.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers context effects are fundamental to decision making. *Psychological Science*, 24(6), 901-908.

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences*, 109(24), 9659-9664.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039-1061.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, 39(10), 1179-1189.

Tybout, A., Sternthal, B., Keppel, G., Verducci, J., Meyers-Levy, J., Barnes, J., ... & Steenkamp, J. B. (2001). Analysis of variance. *Journal of Consumer Psychology*, 10(1-2), 5-35.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, 108(3), 550-592.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, 111(3), 759-769.

Vorauer, J. D., Gagnon, A., & Sasaki, S. J. (2009). Salient intergroup ideology and intergroup interaction. *Psychological Science*, 20(7), 838-845.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4), 767-778.

Wedell, D. H., & Pettibone, J. C. (1996). Using judgments to understand decoy effects in choice. *Organizational Behavior and Human Decision Processes*, 67(3), 326-344.

Wyer, N. A., Sadler, M. S., & Judd, C. M. (2002). Contrast effects in stereotype formation and change: The role of comparative context. *Journal of Experimental Social Psychology*, 38(5), 443-458.

## Footnotes

<sup>1</sup> Some participants chose not to report their gender ( $N = 3$ ), chose to identify with ‘other’ ( $N = 3$ ), or did not choose any of the given options ( $N = 1$ ).

<sup>2</sup> The Envy Asymmetric Dominance decoy did not significantly change likelihood of hiring the Pity candidate ( $M = 4.70$ ,  $SE = 0.097$ ) as compared to Baseline, 95% CI [-0.287,0.245],  $t(1546) = -0.15$ ,  $p = 0.878$  (Holm-corrected  $p = 0.878$ ),  $d = -0.017$ . Similarly, the Envy Compromise decoy did not significantly change likelihood of hiring the Pity candidate ( $M = 4.86$ ,  $SE = 0.098$ ) as compared to Baseline, 95% CI [-0.130,0.403],  $t(1546) = 1.00$ ,  $p = 0.316$  (Holm-corrected  $p = 0.632$ ),  $d = 0.114$ .

<sup>3</sup> The Pity Asymmetric Dominance decoy did not significantly change likelihood of hiring the Envy candidate ( $M = 5.64$ ,  $SE = 0.096$ ) as compared to the Baseline, 95% CI [-0.291,0.239],  $t(1546) = -0.19$ ,  $p = 0.848$  (Holm-corrected  $p = 1$ ),  $d = -0.022$ . Similarly, the Pity Compromise decoy did not significantly change likelihood of hiring the Envy candidate ( $M = 5.75$ ,  $SE = 0.097$ ) as compared to the Baseline, 95% CI [-0.185,0.347],  $t(1546) = 0.60$ ,  $p = 0.552$  (Holm-corrected  $p = 1$ ),  $d = 0.067$ .

<sup>4</sup> Some participants chose not to report their gender ( $N = 2$ ) or chose to identify with ‘other’ ( $N = 4$ ).

<sup>5</sup> The Envy Asymmetric Dominance decoy did not significantly change likelihood of hiring the Pity candidate ( $M = 5.06$ ,  $SE = 0.094$ ) as compared to Baseline, 95% CI [-0.447,0.078],  $t(1600) = -1.38$ ,  $p = 0.167$  (Holm-corrected  $p = 0.167$ ),  $d = -0.155$ .

<sup>6</sup> The Pity Compromise decoy did not significantly change likelihood of hiring the Envy candidate ( $M = 5.73$ ,  $SE = 0.094$ ) as compared to the Baseline, 95% CI [-0.313,0.212],  $t(1600) = -0.38$ ,  $p = 0.707$  (Holm-corrected  $p = 1$ ),  $d = -0.042$ .

<sup>7</sup> Some participants chose not to report their gender ( $N = 2$ ) or chose to identify with ‘other’ ( $N = 2$ ).

<sup>8</sup> Similar to Experiment 1a, the Pity Asymmetric Dominance decoy ( $M = 5.68, SE = 0.094$ ) did not increase likelihood of hiring the Envy candidate in comparison to the Baseline condition, 95% CI [-0.303,0.216],  $t(1598) = -0.33, p = 0.743$  (Holm-corrected  $p = 0.743$ ),  $d = -0.037$ . Similar to Experiments 1a and 1b, the Pity Compromise decoy ( $M = 5.84, SE = 0.094$ ) did not increase likelihood of hiring in comparison to the Envy candidate’s baseline, 95% CI [-0.137,0.381],  $t(1598) = 0.93, p = 0.354$  (Holm-corrected  $p = 0.709$ ),  $d = 0.104$ .

<sup>9</sup> We suspect we had higher numbers of failed manipulation checks because Attitude’ and ‘Aptitude’ look incredibly similar. Number of participants excluded by condition: Aptitude Baseline: 35, Aptitude Envy AD: 29, Aptitude Envy C: 30, Aptitude Pity AD: 31, Aptitude Pity C: 29, Attitude Baseline: 91, Attitude Envy AD: 124, Attitude Envy C: 113, Attitude Pity AD: 117, Attitude Pity C: 129.

<sup>10</sup> Some participants chose not to report their gender ( $N = 1$ ) or chose to identify with ‘other’ ( $N = 3$ ).

<sup>11</sup> The Envy Asymmetric Dominance decoy did not significantly change likelihood of hiring the Pity candidate ( $M = 4.44, SE = 0.098$ ) in comparison to the Aptitude Baseline condition, 95% CI [-0.453,0.093],  $t(3034) = -1.29, p = 0.197$  (Holm-corrected  $p = 0.197$ ),  $d = -0.145$ .

<sup>12</sup> The Pity Asymmetric Dominance decoy did not significantly change likelihood of hiring the Envy candidate ( $M = 5.87, SE = 0.100$ ) as compared to the Aptitude Baseline, 95% CI [-0.513,0.038],  $t(3034) = -1.69, p = 0.091$  (Holm-corrected  $p = 0.364$ ),  $d = 0.035$ . Similarly, the Pity Compromise decoy did not significantly change likelihood of hiring the Envy candidate ( $M$

$= 5.97$ ,  $SE = 0.099$ ) as compared to the Aptitude Baseline, 95% CI [-0.414,0.135],  $t(3034) = -1.00$ ,  $p = 0.319$  (Holm-corrected  $p = 0.637$ ),  $d = -0.112$ .

<sup>13</sup> The Pity Compromise decoy did not significantly change likelihood of hiring the Envy candidate ( $M = 4.18$ ,  $SE = 0.106$ ) as compared to the Attitude Baseline, 95% CI [-0.491,0.086],  $t(3034) = -1.37$ ,  $p = 0.169$  (Holm-corrected  $p = 0.339$ ),  $d = -0.163$ .

<sup>14</sup> The Envy Asymmetric Dominance decoy did not significantly change likelihood of hiring the Pity candidate ( $M = 6.07$ ,  $SE = 0.104$ ) as compared to the Attitude Baseline, 95% CI [-0.141,0.433],  $t(3034) = 1$ ,  $p = 0.318$  (Holm-corrected  $p = 0.954$ ),  $d = 0.118$ . Similarly, the Envy Compromise decoy did not significantly change likelihood of hiring the Pity candidate ( $M = 5.93$ ,  $SE = 0.113$ ) as compared to the Attitude Baseline, 95% CI [-0.291,0.308],  $t(3034) = 0.06$ ,  $p = 0.955$  (Holm-corrected  $p = 1$ ),  $d = 0.007$ .

<sup>15</sup> Some participants chose to report their gender as ‘not otherwise specified’ ( $N = 3$ ).

<sup>16</sup> Some participants chose not to report their gender ( $N = 4$ ) or chose to identify with ‘other’ ( $N = 4$ ).

Figure 1a. Example of joint evaluation screen in Pity Compromise condition. Note that the traits were labeled 'Global Warmth' and 'Global Competence' in Experiment 1a, and 'Global Attitude' and 'Global Aptitude' in Experiment 1b and 1c. In Experiment 1c, the gender of all candidates was female, and a more feminine silhouette (Figure 1b) was used.

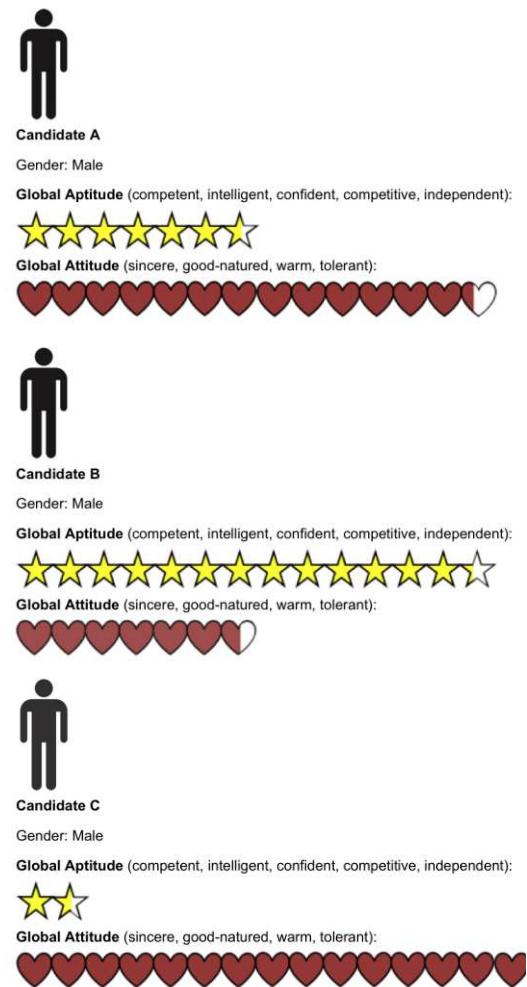


Figure 1b. Feminine silhouette used in Experiment 1c.



Figure 1c. Graphical representation of dimension values for each candidate and decoy.

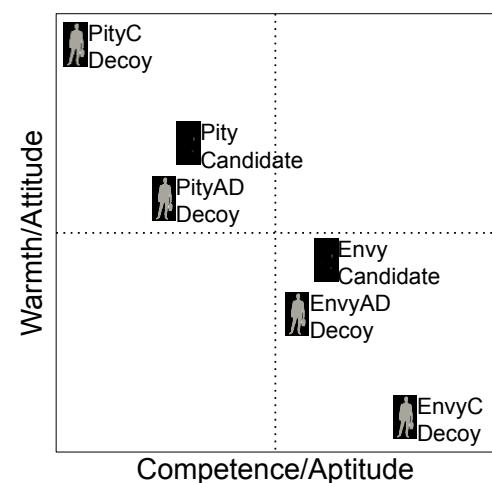


Figure 2. Estimated marginal means from Experiment 1a. The Baseline condition is represented by the dotted lines and the Decoy conditions (Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) are on the x-axis. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs. \*\*  $p < 0.01$ , \*  $p < 0.05$

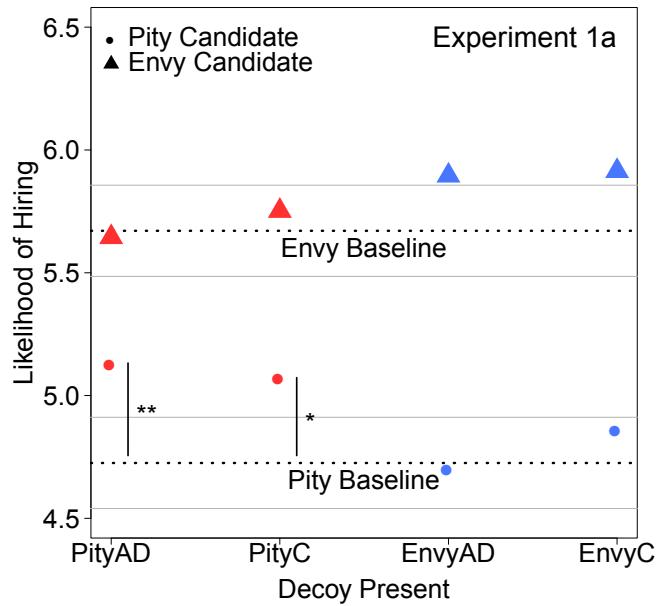


Figure 3. Estimated marginal means from Experiment 1b. The Baseline condition is represented by the dotted lines and the Decoy conditions (Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) are on the x-axis. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs. \*\*  $p < 0.01$ , \*  $p < 0.05$

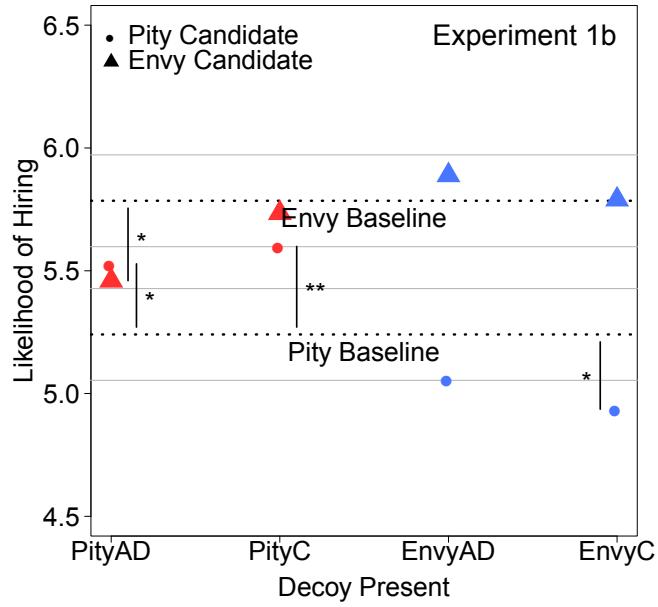


Figure 4. Estimated marginal means from Experiment 1c. The Baseline condition is represented by the dotted lines and the Decoy conditions (Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) are on the x-axis. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs. \*\*  $p < 0.01$ , \*  $p < 0.05$

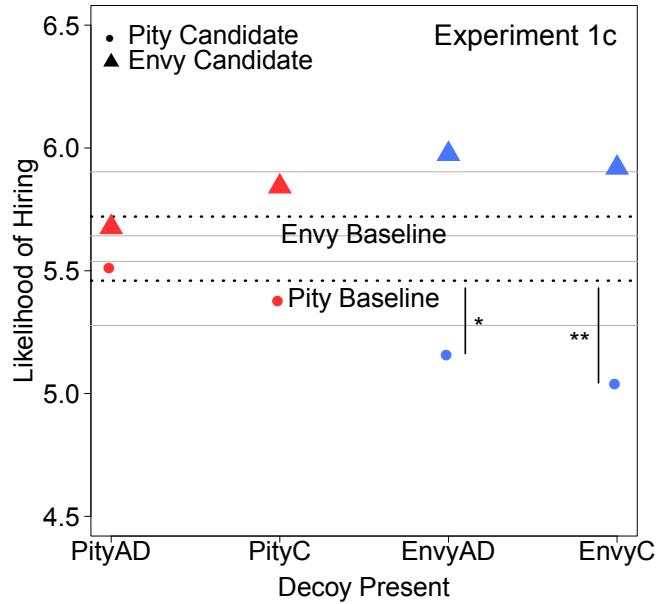


Figure 5. Estimated marginal means from Experiment 1d. The Baseline condition is represented by the dotted lines and the Decoy conditions (Envy Compromise Female, Envy Compromise Male) are on the x-axis. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs.

\*\*\*  $p < 0.001$

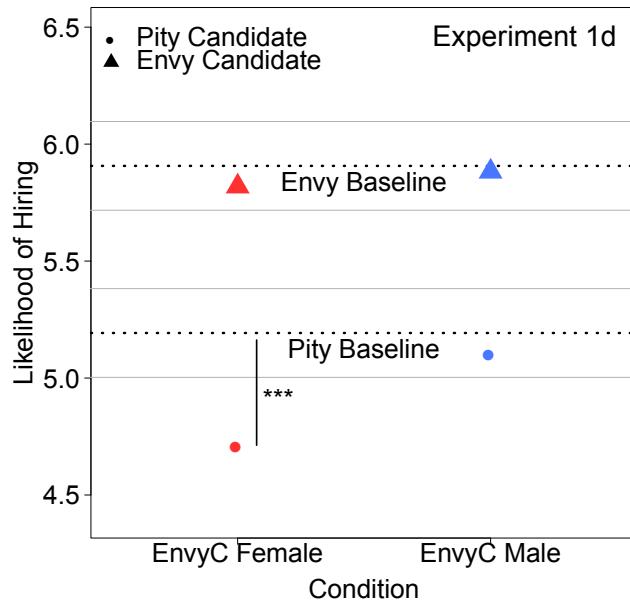


Figure 6. (a) Estimated marginal means from Experiment 2, Aptitude Emphasis. (b) Estimated marginal means from Experiment 2, Attitude Emphasis. The Baseline condition is represented by the dotted lines and the Decoy conditions (Pity Asymmetric Dominance, Pity Compromise, Envy Asymmetric Dominance, Envy Compromise) are on the x-axis. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

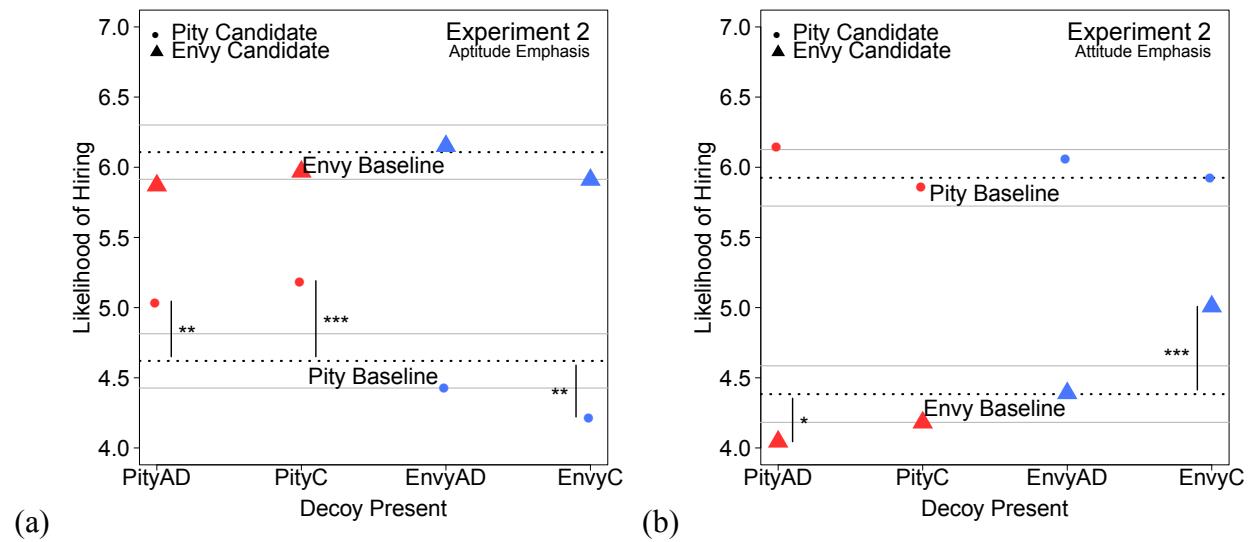


Figure 7. Examples of warmth (left) and competence (right) exposure trials from Experiment 3a.

Participants saw different proportions of each of these exposure trials, where the location of

Candidate A and B as well as the choice button labels were randomized.

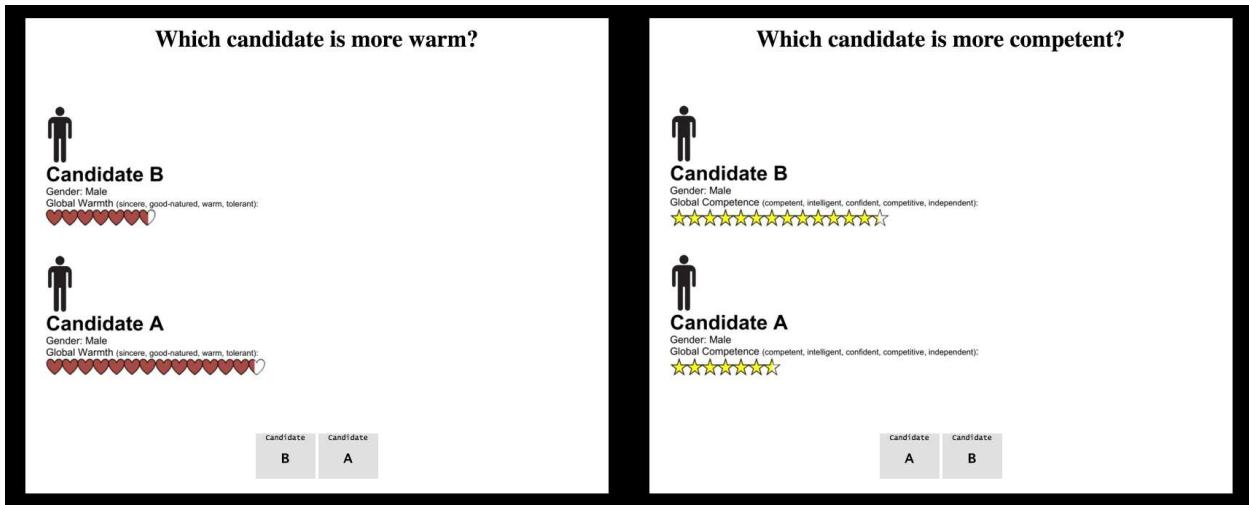


Figure 8. Estimated marginal means from Experiment 3a. The x-axis denotes the exposure condition. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs. \*  $p < 0.05$ , +  $p < 0.1$

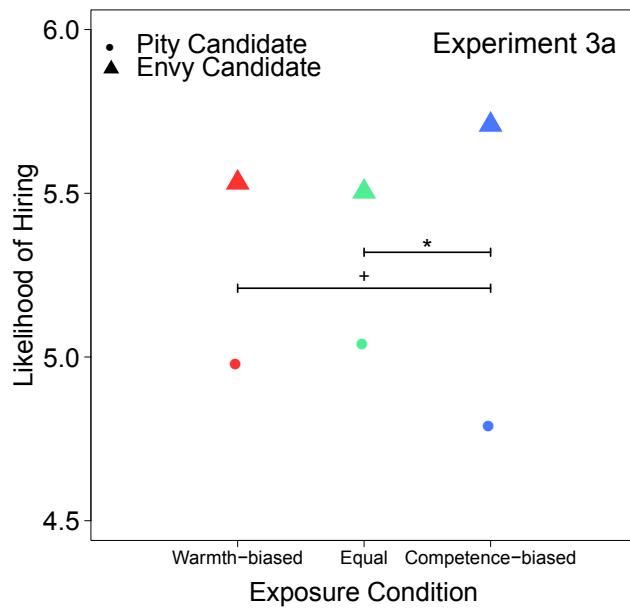


Figure 9. Estimated marginal means from Experiment 3b. The x-axis denotes the exposure condition. The y-axis shows the likelihood of hiring. Error bars denote 95% CIs. \*\*\*  $p < 0.001$

