Journal of Behavioral Decision Making, *J. Behav. Dec. Making* (2015)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/bdm.1889

Exploring the Generalization Process from Past Behavior to Predicting Future Behavior

LASANA HARRIS[1]*, VICTORIA K. LEE[1], ELIZABETH H. THOMPSON[1] and RACHEL KRANTON[2]
[1]*Psychology and Neuroscience, Duke University, Durham, NC, USA*
[2]*Economics, Duke University, Durham, NC, USA*

ABSTRACT

Substantial evidence in social psychology documents that traits predict behavior. Research in behavioral economics establishes prior behavioral information—the actual behavior of another person in the past—influences future decision making, suggestive of the role of traits in guiding future behavior, but agnostic to the specific psychological mechanism. Yet the entire generalization process from past behavior to predicting future behavior has not been fully explored. Additionally, previous paradigms do not adequately dissociate prediction from explanation, and provide participants with trait information, or rely on participants to generate the appropriate trait. Here, we combine literature and experimental approaches in social psychology and behavioral economics to explore the generalization process from prior behavior that guides future decisions. Across three studies utilizing consequential economic game paradigms and online questionnaires, an initial group of participants (employees) played a time estimation game and a charity donations game before a second group of participants (employers) viewed the behavior of the first group, then decided whether to invest in employees in a trust game and rock guessing game. Although participants infer trait warmth and competence from the behavioral information in the first two games, estimates of normative behavior predicted investment decisions on the warmth-relevant games better than trait inferences. These results dissociate generalizations guided by warmth and competence behavioral information, and question the extent to which traits always serve as heuristics to predict behavior. Copyright © 2015 John Wiley & Sons, Ltd.

Additional supporting information may be found in the online version of this article at the publisher's web-site

KEY WORDS person perception; social decision making; trait inferences; norms

People somehow predict what others will do after observing their prior behavior. Eighty years of social psychological research argues that trait inferences serve as heuristics—spontaneously generated mental shortcuts—to predict behavior, allowing people to generalize from prior behavior to predict future behavior (for early review, see Paunonen & Jackson, 1985; Pervin, 1985), often at the expense of base-rate normative information (Kahneman & Tversky, 1973). However, it may be time to reconsider the extent to which people prefer trait inferences to normative information when making predictions. Prediction and explanation are not different sides of the same coin, and although personality traits may be valuable explanatory devices, people may not actually use them to generalize from one behavior in service of predicting another. This may particularly be the case when the other person's behavior is consequential for the perceiver. Explanations require understanding abstract concepts, a feature that is unnecessary for prediction (Andrews, 2012). Instead, people may place more weight on a heuristic about the context to guide future interactions, as opposed to a more cognitively complex construct such as a personality trait inference. Stated differently, although personality traits are useful as folk explanatory conceptions, personality trait inferences may not be as useful as normative information when predicting behavior.

*Correspondence to: Pieter de la Court Building, 52 Wassennaarseweg, 2223AK, Leiden, Netherlands.
E-mail: l.t.harris@fsw.leidenuniv.nl

Traits provide adequate explanation for behavior because they locate the person as the causal agent, holding him or her responsible for initiating the behavior. Therefore, traits are very useful as abstract concepts that allow meaning making of social behavior. Certainly, on aggregate, traits describe or explain people's behavioral consistency over time (Beck, McCauley, Segal, & Hershey, 1988; Burke, Kraut, & Dworkin, 1984; Emmons & Diener, 1986; Funder & Colvin, 1991; Furr & Funder, 2004; Hettema & Hol, 1998; Koestner, Bernieri, & Zuckerman, 1989; Krahe, 1986; Leikas, Lonnqvist, & Verkasalo, 2012; Lippa & Mash, 1981; Magnusson & Ekehammar, 1978; Moskowitz, 1994; Welbourne, 2001). But traits only correlate with behavior across situations at $r = +.30$. Empirical results suggest methodological improvements could increase traits' predictive power if they are used to predict behavior in a specific-enough social context; if the social context from which the trait is inferred (previous context) and the social context that was being predicted (future context) are similar enough, then correlation coefficients rise above the modest mark of +.30, and traits become better predictors (Baird & Lucas, 2011; Hemmelgarn, James, Ladd, & Mitchell, 1995; Magnusson, 1976; Paunonen & Jackson, 1986; Van Mechelen, 2009).

However, traits may not be useful for generalizing from single or limited instances of behavior to predict another's behavior in a less related social context. Traits could only serve as useful tools for prediction if people always behaved according to their traits. However, social psychology repeatedly demonstrates the power of the social context to influence behavior. Therefore, when asked to make predictions about another person's behavior, people may also take into account normative information about the social context to guide their decisions.

Additionally, social neuroscience research is beginning to show dissociations in trait inference processing in the brain, suggesting that traits are not a homogeneous category. Most stable, enduring trait inferences activate a reliable pattern of brain activity, including medial prefrontal cortex, superior temporal sulcus, precuneus, and temporoparietal junction (see for review Amodio & Frith, 2006; Frith & Frith, 2001; Van Overwalle, 2009). These cortical brain regions also constitute nodes that contribute to other brain networks underlying cognitive processes such as economic valuation (Lee & Harris, 2013) and cognitive control (Botvinick & Cohen, 2014). However, inferring a person's trustworthiness depends on the amygdala (Engell, Haxby, & Todorov, 2007; Said, Baron, & Todorov, 2009; Todorov, 2008; Todorov, Baron, & Oosterhof, 2008), a subcortical brain structure implicated in emotional learning, social biases, and vigilance (Whalen & Phelps, 2009). This brain structure receives almost immediate input from primary sensory cortex, whereas the cortical structures receive their information via a different, relatively longer, pathway from primary sensory cortex, via many more synapses (Milner & Goodale, 2008). Supporting evidence comes from research demonstrating that subliminal presentation of untrustworthy faces drives the amygdala (Winston, Strange, O'Doherty, & Dolan, 2002). This suggests dissociation between different types of traits; some of which may require more complex cognitive processing, and others where such processing is unnecessary. Only traits such as trustworthiness are inferences that may serve as heuristics.

The argument for trait inferences as heuristics

Social psychology has gathered evidence that supports the hypothesis that trait inferences are heuristics (for review, see Booth-Butterfield, 1988; Furr, 2009). Within the thin-slice and spontaneous trait inference (STI) literatures, studies have demonstrated that trait inferences can be fairly automatic. For instance, STIs of trustworthiness occur within a couple hundred milliseconds of face perception (Willis & Todorov, 2006), suggesting that some trait inferences are fast, a horseman of automaticity (Bargh, 1994). Traits are also an efficient way of organizing information about a person (Fiske & Taylor, 2013), a second condition of automaticity.

Given descriptions of two behaviors of fictitious people, participants can identify the correct trait inference and predict the person's behavior in a similar novel situation (Newman, 1996). However, this study used behavior of fictitious people, suggesting that the "correct" prediction is just one that the researchers identified as correct because participants are not actually observing what the person did in the second situation. Additionally, this study also provided participants trait inferences instead of simply having participants generalize from previous behavior. If traits probe explanatory mechanisms and explanatory mechanisms are not used for prediction, then it is possible that asking people to make a trait inference triggers an explanatory mechanism, making this mechanism salient, and allowing it to masquerade as prediction. This presents a potential confound that may have led to support for the idea that traits predict behavior.

Another study in support of traits as predictive heuristics examined how people predict other people's behavior using STIs (McCarthy and Skowronski, 2011). Participants saw pictures of people paired with a one-sentence description of behavior before rating the likelihood of the pictured person performing three novel behaviors. Although this study did not provide participants with trait inferences (it demonstrates that traits could be inferred from picture–behavior pairings in a pilot study), it did not rule out perceived norms (frequency of the behaviors) as sources of predictive information. Specifically, participants could just as easily have inferred the likelihood that most people would perform each of the three behaviors and used this estimate of behavior to guide their responses. The current study extends these results by asking participants to estimate normative information.

The argument against trait inferences as heuristics

The oversight present in the studies described earlier and present in theorizing about trait inferences as heuristics occurs perhaps because investigators have long relied upon the domain of explanation by measuring self-reports or providing trait information to participants. Traits represent abstract concepts; the ability to understand such concepts develops later in life and should be a more controlled, system-two process, compared with a heuristic, a prototypical system-one automatic process. Maybe researchers have attempted to shoehorn trait inferences into a heuristic, largely ignoring the difference between prediction and explanation.

Additionally, evidence is weaker that traits seem to be inferred effortlessly; STIs depend on cognitive resources (Ferreira et al., 1986; Van Overwalle, van Duynslaeger, Coomans, & Timmermans, 2012; Wells, Skowronski, Crawford, Scherer, & Carlston, 2011). Indirect evidence from the stereotyping literature suggests that traits constitute individuated representations that occur later, after time and motivation (Fiske & Neuberg, 1990; Fiske, Neuberg, & Lin, 1999). There is also little evidence of trait inferences occurring without awareness because of the dependence on self-report measures. Therefore, with only two out of four automaticity conditions satisfied, traits may not be heuristics given that they require cognitive resources and do not have apparent physiological or other implicit, non-self-report, components.

Social psychological evidence suggests that cross-situational behavior consistency occurs because of language —traits serve as linguistic concepts that suggest the same behavior is applicable in a novel social context (Semin & Krahne, 1988). This also supports the idea that traits are not heuristics. More damning evidence comes from a number of studies that failed to replicate classic cross-situational behavior consistency effects (for instance, Chaplin & Goldberg, 1984; Eysenck & Wild, 1996), including cross-culturally (Church et al., 2008; Oishi, Diener, Napa Scollon, & Biswas-Diener, 2004), although at least one paper finds support for the folk intuition that traits serve as behavioral predictors cross-culturally (Norenzayan, Choi, & Nisbett,

DOI: 10.1002/bdm

2002). However, other studies have questioned whether the folk intuition exists (Heffler & Magnusson, 1979; Epstein & Teraspulsky, 1986).

Presumably, trait inferences facilitate generalizations that can guide future behavior. But people may be required to generalize from behavior to broad traits in order to predict behavior in different contexts. For instance, someone who engages in kind behavior in one context (helping an old lady cross the street) may be considered a trustworthy person (returns extra money received in error from the cashier). Kindness and trustworthiness, though both warmth dimension traits, are separable and do not necessarily co-existent in all people. As such, generalizations to the broad traits (like morally good or warm in the example earlier) may lead to inaccurate trait inferences, poor predictions, and suboptimal decisions once the context shifts from crossing the street to a monetary transaction with a cashier. Instead, participants may simply form a heuristic based on the person's previous behavior that serves as information to guide their behavior. Stated differently, participants may judge how deviant the person's behavior is from other people's behavior in the same social context (the norm) and use that information to predict behavior (Kelley, 1967). However, when asked to explain the person's behavior, they may then rely upon a trait inference.

Hypotheses

Here, we explore the entire generalization process from past behavior to predicting future behavior across the two primary person perception traits, warmth and competence (Asch, 1946; Fiske, Cuddy, & Glick, 2007). We created a paradigm where participants make predictions about other people's behavior during social interactions. We employ a behavioral economics approach through the first two studies; we record actual behavior from Phase 1 participants "employees" during a trait warmth-relevant donation game, and a trait competence-relevant time estimation game, before giving a separate group of Phase 2 participants "employers" the opportunity to make trait inferences from this behavior to guide future decisions regarding employees in a trait warmth-relevant trust game and a trait competence-relevant rock estimation game. As such, employers could choose to invest with employees based on inferred warmth and competence. Unlike previous paradigms, the participants predict how the other person would behave by behaving themselves, not commenting on the person's traits, or about what that person may do in a novel situation not involving the participant.

We address alternate explanations for the pattern of results in Study 1 with Study 2; we balance the reward for employees across trust and rock estimation games, ruling out motivational differences. Secondly, we collect normative estimates of donation game behavior from employers to determine alternate cutoff criteria for "good" investment decisions in the trust game, and to test whether these normative estimates predicted participants' decisions in Phase 2 better than trait inferences. Additionally, participants would not be aware of the actual social norm present among employees during the game until they approached the last trial of the experiment. As a result, they may employ one of a number of normative estimates of behavior to guide their decisions, such as the perceived ideal amount to donate, the estimated amount the person believes should be donated, or the estimated amount on average that everyone will donate. This strategy allows us to determine which norms best match actual behavior. In Study 3, we employ a social psychological approach to explicitly check whether the behavioral information in our profiles leads to trait inferences about the employees. We then use these trait ratings to predict Phase 2 behavior in our first two studies.

STUDY 1

Method

Employees

Participants. Sixty-four individuals from Duke University and the surrounding community contributed to the player database. Four employees were removed from the sample, because of their wish not to have their picture taken (see succeeding text for significance of photographs), resulting in 60 employees in the database. The sample's mean age was 26.73 years ($SD = 11.34$), and 56.7% of the employees were female. All employees gave written consent prior to beginning the experiment, in accordance with university standards, and were paid \$10 for their participation.

Employees' tasks

Employees sequentially completed four games/tasks: a time estimation game and charity donation game, with the order counterbalanced across employees, and a trust game scenario task and rock estimation task, again with order counterbalanced across employees.

Time estimation. We used the time estimation game to record competence behavioral information from the employees. Employees viewed varying amounts of time (e.g., 7.2 seconds) and pressed a button to estimate the time interval. We scored accuracy as within 500 milliseconds of the designated time. After each of 60 trials, employees were given accuracy feedback for that trial and their cumulative accuracy score for the entire game.

Charity donation. We used the charity donation game to record warmth behavioral information from the employees. Employees viewed 60 charities from within 10 different categories, six charities per category. For each trial, employees viewed a short description of a charity's mission statement and indicated what percentage of their participation fee (\$10) they would like to donate, if any, to the charity. Percentages could range from 0% (i.e., \$0) to 100% (i.e., \$10). Employees were instructed that each decision should be independent of the other donation decisions. Employees were also informed that one trial would be randomly selected and realized at the end of the study for both the employee and charity.

DOI: 10.1002/bdm

Trust game scenario. We used the trust game scenario to record warmth behavioral information from the employees. The player was assigned the role of trustee. An anonymous investor had the choice to either keep \$10 or invest it with the employee, at which point the money would be tripled, becoming \$30. Employees were told that the investor had chosen to invest with them. Employees were asked to indicate whether they would keep all of the profit or return half to the investor. Employees only answered the question once.

Rock estimation game. We used the rock estimation game to record competence behavioral information from the employees. We asked employees to estimate the number of rocks inside the jar. The correct number of rocks in the jar was 256. Employees estimated in a free response style in order to avoid anchoring effects. Each player only made one estimate. We informed employees that they could receive money from us in the future based on their performance in these latter two games.

Employee procedure

The experimenter paid employees their \$10 participation fee upon arrival. Employees were informed that, at some point in the experiment, they would have the opportunity to donate a portion of that money to charity if they so choose. The experimenter then explained the first two games (time estimation and donation games) with the help of a visual aid. Upon completing the time estimation and donation games, employees completed the trust game scenario task and the rock estimation task.

Employers

Participants. Thirty-two individuals from Duke University and the surrounding community participated as employers. The mean age was 21.86 years (SD = 4.56), and 56.3% of the subjects were female. Two participants were excluded from analysis because of incomplete data, for a total sample size of 30. Across both studies, participants earned \$10 for participating in the study and could earn up to an additional \$30 based on their decisions in two economic games. All participants gave written consent prior to beginning the experiment, in accordance with university standards.

Employers' stimuli

Player profiles. We created a profile for each Phase 1 employee to be used as behavioral information by Phase 2 employers on each trial, including the following: a photograph of the employee; general demographic information, including field of study, handedness, and age; and behavioral information from their performance on the time estimation and donation games, including overall average accuracy and generosity scores, respectively, as well as round-by-round performance. This information appeared as percentages and bar graphs.

Maze task. In order to allow employers ownership of the money used in the economic games, they completed a series of mazes. For each maze completed in under 2 minutes, the experimenter awarded the employer \$5. The employers continued completing mazes until they had earned \$20 (every employer except one accomplished this with the minimum of four mazes). Therefore, each employer had earned \$10 each to play the trust and rock estimation games.

Trust game. Phase 2 employers played the standard trust game in the role of the investor. On each trial, employers viewed an employee profile, followed by a decision screen where they indicated whether or not they wanted to invest their \$10. We then asked employers to rate their confidence *in their decision* on a scale from 0 *not at all confident* to 100 *very confident*. Employers did not receive trial-by-trial feedback during the game in order to avoid learning effects. We informed employers that we would choose one trial at random and realize the decision for both the employer and employee at the end of the experiment.

Rock estimation game. We modeled the rock estimation game after the trust game. Employers decided to bet or not bet their previously earned \$10 on whether the employees would accurately estimate the number of rocks in a jar. On each trial, employers viewed the profile of an employee (the same profile shown in the trust game), followed by a decision screen that asked whether or not they would like to invest with that employee. Employers then rated their confidence in their decisions.

We matched the likelihood of payout for employers in both the rock estimation and trust games. That is, we chose a range of rock estimations (50 rocks over or 50 under the actual amount, in other words estimates between 206 and 306) to be considered correct in order to match the number of employees who decided to share and keep in the trust game. This resulted in 40 employees' rock estimates deemed correct and 20 employees' rock estimates deemed incorrect to match the 40 employees that decided to share the money and 20 employees that decided to keep the money in the trust game scenario.

Employers' procedure

Employers performed the maze task until they had earned \$20. The experimenter then explained the study in detail to the employers, including a description of the donation and time estimation games completed by the 60 employees from Phase 1, the information on the profiles, a description of the trust and rock estimation games to be completed in this experiment, and finally a description of the confidence ratings. The experimenter additionally informed the employers that one trial from each game would be randomly selected at the end of the study and realized for both parties. Then, employers completed a short quiz to ensure complete understanding of the rules of the four games and the information on the employee profiles. Employers were required to successfully complete the quiz (100% correct answers) before proceeding.

DOI: 10.1002/bdm

We counterbalanced the order in which the employers played the trust and rock estimation games. On each trial, an employee profile was randomly selected without replacement and displayed. Employers then indicated their decision to invest or not and rated their confidence in their decision. The trials in both games were visually identical (profile screen, decision screen, and confidence rating screen), with the only difference being an instruction slide introducing the game at the outset of the 60 trials (Supporting information Figure 2). Each screen was self-paced. Finally, after playing the trust and rock estimation games, employers completed the memory task to determine what information, if any, they could remember from the employee profiles. Employers were not informed of the memory test until immediately before taking it to ensure that employers did not consciously commit any information to memory.

After all of the tasks had been completed, the experimenter randomly selected one trial from each game, that is, realized one trial, paid the employers accordingly, and thanked them for their participation. At the end of the study, the experimenters mailed the employees the money earned in the randomly selected trials.

Data analysis strategy

Investment decisions

To determine whether employers were using the employees' behavioral information to guide their decisions, we dissociated employees with a high likelihood of returning the investments from those with a lower likelihood of returning investments. Furthermore, we used different approaches to determine cutoff points for warmth and competence.

Traditional economics would treat competence as an individual trait and would measure competence by the individual's ability to reach accurate conclusions. Measuring warmth, on the other hand, would be possibly subjective; we adopted a framework where employees' warmth is measured according to the distribution of generosity scores, capturing that individual generosity or warmth relative to a social norm. To dissociate high from low warmth employees, we used employees' average generosity scores to calculate a cutoff percentage one standard deviation above the mean. For the sample of employees, this calculation resulted 20%. Therefore, we defined a high warmth employee as an employee with an average generosity greater than or equal to 20% and a low warmth employee as an employee with an average generosity less than 20%.

To dissociate high from low competence employees, we used employees' average accuracy scores to calculate whether the expected value of investing with the employee was greater than the expected value of not investing with the employee. We used the following expected value equation:

$$10 = 15x + (1 - x)*0$$

where x is the average accuracy score of a particular employee. This expected value equation sums up each possible outcome multiplied by the probability of that outcome. Therefore, the right side of the equation represents the expected value of investing, where $15x$ is the probability of winning the gamble multiplied by the outcome of winning the gamble and $(1 - x)*0$ is the probability of losing the gamble multiplied by the outcome of losing the gamble. The value 10 represents the expected value of not investing because the probability of receiving \$10 when not investing is 1 given that participants keep the \$10 they already earned. Because $(1 - x)*0$ will always be equal to 0, the equation can be simplified to

$$10 = 15x$$

We defined a high competence employee as an employee whose average accuracy resulted in the statement $10 < 15x$ and a low competence employee as an employee whose average accuracy resulted in the statement $10 \geq 15x$.

We then dummy coded employer's decisions as either "good" or "bad," based on these criteria. We defined good decisions as trials in which (a) the employee was high warmth or high competence and the employer did invest or (b) the employee was low warmth or low competence and the employer did not invest. The remaining two trial types were defined as bad decisions. We dummy coded good and bad decisions as 1 and 0, respectively, allowing us to examine the amount of time employers used the trait inferences from the profiles to guide investment decisions.

We performed the analysis described earlier for both match and mismatch conditions created by our 2×2 factorial design. Match conditions are those in which behavioral information from a trait domain guides decisions in the game of that same trait domain, and tell us whether trait inferences lead to generalizations *within* person perception domains. Mismatch conditions are those in which behavioral information from a trait domain guides decisions in the game of the other trait domain, and tell us whether trait inferences lead to generalizations *across* person perception domains. We averaged the dummy-coded scores across the 60 trials within each of the four match and mismatch conditions, resulting in four scores for each employer indicating the amount of time they used behavioral information to guide their investment decisions. Numbers closer to 1 indicate that the employer made decisions consistent with the hypothesis that behavioral information from employees would guide investment decisions. We performed a 2 (*trait*) × 2 (*decision*) repeated measures analysis of variance (ANOVA) on confidence ratings to determine whether employers felt more or less confident when deciding to invest or not. We computed paired sample *t*-tests with a Bonferroni corrected $\alpha = .0125$ for all significant main effects and interactions. We subtracted employees' actual time estimation accuracy and donations from employers' estimates of accuracy to create indices of memory accuracy. We first tested employers' recall accuracy by performing one-sample *t*-tests against the value zero.

RESULTS

Employee behavioral information statistics

Mean accuracy scores from the time estimation game (which provided *competence behavioral information*) resulted in a

DOI: 10.1002/bdm

normal distribution (skew = −0.62 and kurtosis = −0.45), ranging from 0% to 90% (M = 56.08, SD = 24.78; Figure 1b). The mean generosity scores from the donation game (which provided *warmth behavioral information*) resulted in a positively skewed distribution (skew = 2.89 and kurtosis = 9.90) ranging from 0% to 72.67% (M = 6.94%, SD = 13.38%; Figure 1a). Employees' estimates of behavior from the rock estimation game (*competence-relevant game*) ranged from guesses of 60 to 550, with a mean estimate of 197.77 (SD = 93.63). They were normally distributed: skew = 1.41 and kurtosis = 3.57 (Figure 1d). Finally, 40 of the 60 employees (66.67%) indicated they would return the investment in the trust game scenario (*warmth-relevant game*; Figure 1c).

We found no correlation between employees' accuracy scores and a correct rock estimate, r (58) = .12, p = .356, or between employees' generosity scores and their decision to share the money in the trust game, r_{pbi} (58) = .19, p = .141. Therefore, warmth and competence-relevant games used to generate player behavioral information are *unrelated* to the warmth and competence-relevant games in which employers made decisions about said employees. Specifically, an employee's ability to estimate time and rocks is not related, neither is an employee's willingness to donate money to charity or return an investment in the trust game. Therefore, there is no added predictive value when employers use accuracy or generosity scores as behavioral information when making decisions. However, because of the poverty of other available information and the perceived relationship between the pairs of behaviors along warmth and competence trait domains, we hypothesize that employers will in fact use the behavioral information to guide their investment decisions.

Employer investment decisions

We performed a 2 (*behavioral information*: *warmth/competence*) × 2 (*game*: *trust/rock*) repeated measures ANOVA on employers' investment decisions. There was a significant main effect of *behavioral information*, F (1, 29) = 22.32, $p = 5.40 \times 10^{-5}$, partial $\eta^2 = 0.44$, $\Omega = 1.00$, such that employers used warmth information (M = 78.50%, SD = 14.05%) more than competence information (M = 66.39%, SD = 9.08%). This main effect was qualified by a significant *behavioral information* × *game* interaction, F (1, 29) = 21.66, $p = 6.60 \times 10^{-5}$, partial $\eta^2 = 0.43$, $\Omega = 0.99$. We followed up this interaction with paired sample t-tests with a Bonferroni corrected $\alpha = .0125$. When comparing *games* within *behavioral information*, we found that *warmth behavioral information* was used significantly more in the *trust game* than *rock estimation game*, t (29) = 3.57, p = .001, while *competence information* was used significantly more in the *rock estimation game* than *trust game*, t (29) = 4.79, $p = 4.53 \times 10^{-5}$. This suggests that employers generalized from the corresponding trait-relevant behavioral information, using it more frequently to guide their decisions within the relevant than irrelevant game. Comparing *behavioral information* within *game*, we found that in the *trust game*, *warmth behavioral information* was used significantly more than *competence*, t (29) = 11.16, $p = 5.17 \times 10^{-12}$, but in the *rock estimation game*, *competence behavioral information* was not used significantly more than *warmth*, t (29) = 0.33, p = .075 (Figure 2). This suggests that while employers generalized trait warmth from the relevant behavioral information to more frequently guide decisions in the warmth-relevant game, they generalized both trait warmth

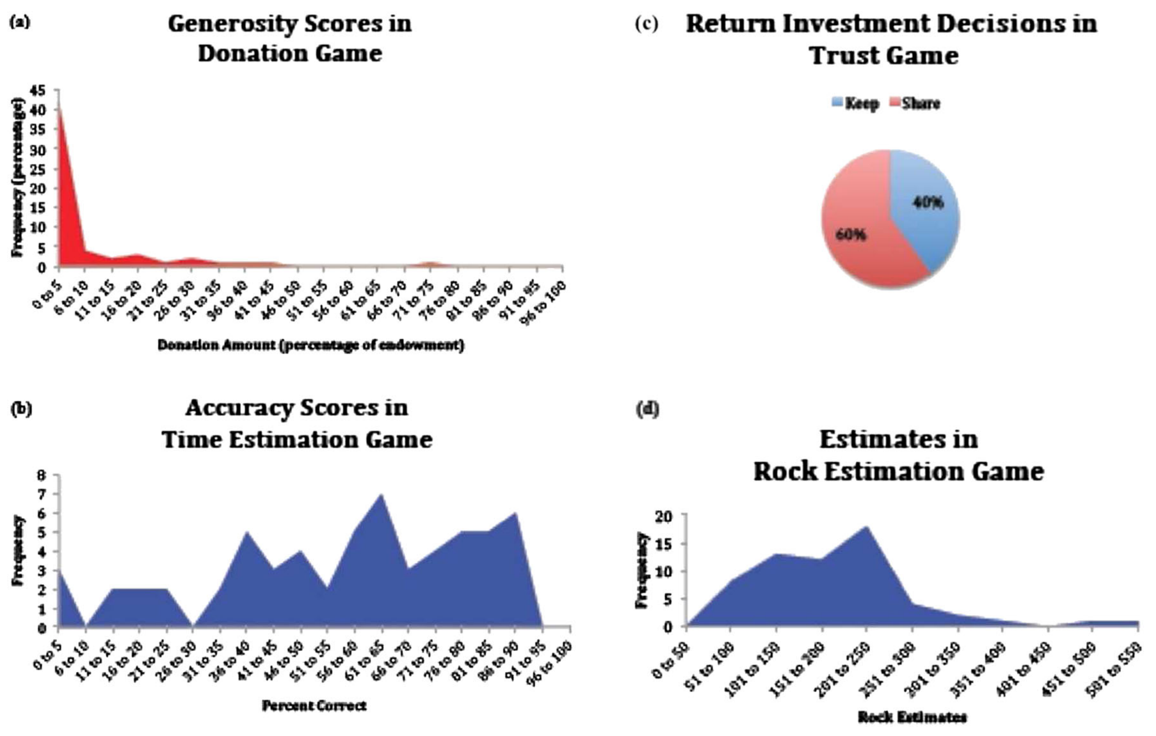

Figure 1. Player behavioral information. Average behavioral information from the sample of employees on the (a) donation game, (b) time estimation game, (c) trust game scenario, and (d) rock estimation game

DOI: 10.1002/bdm

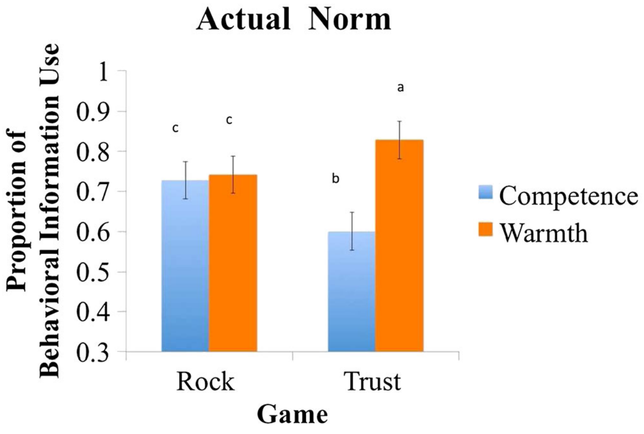

Figure 2. Investment decisions in Study 1. Bar graphs depict the proportion of behavioral information used in each game based on actual norms in Study 1 as cutoff criteria. Error bars reflect standard error of the mean. Different letters represent statistically significant differences between means, that is, different letters are significantly different from each other (e.g., the a's are the same, but different from b and c)

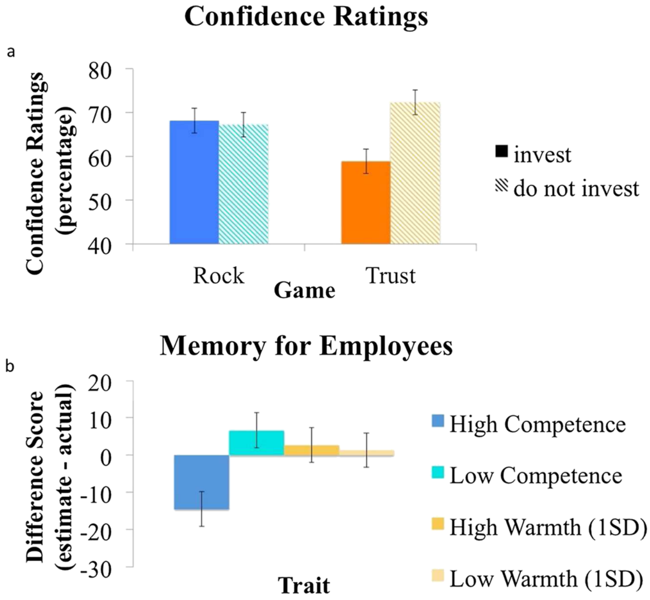

Figure 3. Confidence ratings and memory for employees. Bar graphs depict employers' reported (a) confidence in their investment decisions and employers' (b) memory recall of employees' performance in Study 1. Error bars reflect standard error of the mean

and competence from the relevant behavioral information to guide decisions in the competence-relevant game.

Employer confidence ratings

We performed a 2 (*game*: *rock*/*trust*) × 2 (*decision*: *invest*/*not invest*) repeated measures ANOVA on confidence ratings to determine whether employers felt more or less confident when deciding to invest or not in the warmth-related trust game and competence-related rock estimation game. We found a significant *game* × *decision* interaction, F (1, 25) = 12.16, p = .002, partial $\eta^2 = 0.33$, $\Omega = 0.92$. We then computed paired sample t-tests with a Bonferroni corrected α = .0125. We found a significant difference between confidence ratings in the trust game, t (26) = −3.36, p = .002, such that confidence was higher when employers decided *not to invest* (M = 72.09, SD = 17.66) than *invest with* an employee (M = 61.12, SD = 18.48). These results suggest that employers were least confident in their decisions to trust the employees. Neither of the main effects revealed significant differences.

We then tested whether these confidence ratings differed for employees high or low on each trait dimension. For *warmth*, we found a significant difference, t (26) = 5.25, $p = 1.3 \times 10^{-5}$, such that employers were more confident when making decisions regarding high warmth (M = 70.87, SD = 20.49) rather than low warmth employees (M = 55.89, SD = 18.94). We found a similar effect for *competence*, t (26) = 3.85, p = .001, where employers were also more confident making decisions for high competence (M = 70.31, SD = 16.90) rather than low competence employees (M = 57.86, SD = 19.76; Figure 3a). These findings suggest that employers perhaps intuitively dissociated "good" from "bad" employees along both trait dimensions.

Employer memory data

We subtracted employees' actual time estimation accuracy and donations from employers' estimates to create indices of memory accuracy. We first tested employers' recall accuracy by performing one-sample t-tests against the value zero. For competence, we found that employers' memory indices did not significantly differ for time estimation, t (29) = −0.78, p = .444, suggesting that they were accurate in recalling employees' competence information. We followed up this finding by testing whether employers displayed memory differences for employees who were categorized as high or low on the trait dimension, calculating memory accuracy indices separately for high and low competence employees. We found that employers showed a significant difference for highly competent employees, t (29) = −6.48, $p = 1.00 \times 10^{-6}$, such that they significantly underestimated their time estimation ability (M_{diff} = −14.52, SD = 13.11). Employers also inaccurately recalled low competence employees' time estimation ability, t (29) = 3.03, p = .005, such that they significantly overestimated (M_{diff} = 6.71, SD = 12.12). Together, these results suggest that although employers' accurately recalled average employee time estimation abilities, they did so by overestimating the ability of less skilled employees and underestimating the ability of more skilled employees.

We performed a similar analysis for warmth, where we found that employers' memory indices did not significantly differ for donations to charity, t (29) = 1.08, p = .288, suggesting that they were accurate in recalling employees' warmth information. We followed this up by separately testing memory recall for high and low warmth employees. Unlike competence, we found no significant difference for recall of donation behavior for either high, t (29) = 1.16, p = .256, or low warmth employees, t (29) = 0.82, p = .417, suggesting that employers accurately remembered warmth behavioral information (Figure 3b).

We next tested whether there was enhanced memory for warmth relative to competence information by performing a paired-samples t-test using the memory indices. We did not

DOI: 10.1002/bdm

find a significant difference, t (31) = 1.31, p = .200, suggesting that employers remembered warmth and competence information equally well.

CONCLUSION

Study 1 replicates previous behavioral economics studies showing that previous behavior affects future decisions using a novel paradigm that utilizes actual behavior (i.e., financially consequential decisions without the aid of deception) and highlights the primacy of warmth behavioral information. In particular, we demonstrated that competence behavioral information generalizes within person perception domain, and warmth behavioral information generalizes both within and across person perception domains. Thus, employers are using warmth information to help guide decision making even in situations in which warmth information should be irrelevant. This finding is consistent with the halo effect literature where traits and behaviors of one kind can carry over and influence traits and behaviors of another kind (Dion, Berscheid, & Walster, 1972; Landy & Sigall, 1974; Moore, Filippou, & Perrett, 2011). The fact that employers were significantly more confident when making good decisions in the warmth-relevant game (trust game) than in the competence-relevant game (rock estimation game) lends additional support; warmth information in particular seems both reliable and predictive. Finally, employers also better remembered warmth behavioral information than competence behavioral information, lending further support for the notion that warmth behavioral information is prioritized relative to competence behavioral information.

However, this study did not assess whether employers actually made trait inferences, so it cannot rule out the possibility that trait inferences are the mechanism that guides decision. Additionally, employees' payout structure for the rock estimation game was different from the trust game such that employers may infer a lack of effort from employees on the rock estimation game. Finally, the criteria used for characterizing high and low warmth employees relied on the mean and standard deviation of the labor market, which employees would not have an accurate estimation of until the end of the study. We addressed these methodological issues in Study 2.

STUDY 2

Study 2 is a replication of Study 1, with two changes to address potential flaws. First, we rewarded employees for accurate rock estimations, and second, we asked employers to indicate their estimation of normative donation behavior to determine alternate cutoff criteria for "good" investment decisions. This latter addition also allows us to test the extent to which perceived norms may have guided behavior in the warmth domain. This study used the same database of behavioral information from the employees in Study 1, but recruited a new sample of participants as employers.

Method

Participants

Thirty individuals from Duke University and the surrounding community participated in Study 2. The mean age was 22.37 years (SD = 6.36 years), and 66.7% of the subjects were female.

Stimuli

We used the same employee profiles from Study 1. At the end of Study 2, employers completed a post-study questionnaire that asked the following:

1. If you played the donation game, how much would you donate to charity from your money?
2. In general, how much do you think people should donate to charity from their money?
3. Before playing the game, what did you think the average donation to charity would be?

The first question assesses employers' personal norms for donating, while the second and third assess estimates of ideal norms and actual norms, respectively.

Procedure

The procedure was the same as Study 1, with one exception: employees were also rewarded for correct rock guesses if employers invested with them, and we realized that outcome.

Data analysis strategy

The analysis strategy was the same as Study 1. In addition, we also determined the cutoff criterion for warmth information individually for each employer based on his or her response to each question on the post-study questionnaire.

RESULTS

Employee norms

On average, employers reported that they would donate 16.07% (SD = 13.37%) of their earnings to charity (personal norms), that people should donate 12.59% (SD = 7.91%) of their earnings to charity (perceived ideal norms), and that before playing the game, they believed the average donation by an employee would be 10.72% (SD = 7.58%) of their earnings to charity (perceived actual norms). Only the reports of personal norms and perceived actual norms statistically significantly differed, t (29) = 2.63, p = .013, while personal norms and perceived ideal norms marginally differed, t (28) = 1.93, p = .064. These findings suggest that each of these reports could yield different results when used as estimates to distinguish high from low warmth employees (warmth behavior cutoff criteria).

Employer investment decisions

We performed a series of 2 (*behavioral information*) × 2 (*game*) repeated measures ANOVAs on the investment

DOI: 10.1002/bdm

behavior using the four different estimates of employers' normative behavior ascertained from the employee sample and the post-experiment questionnaire. We report each separately later.

Actual norms

We first replicated the analysis performed in Study 1 using the actual mean and standard deviation of the employees' donation behavior to generate cutoff criteria for warmth information (Figure 4a). This tested whether we could replicate Study 1. There was a significant main effect of *behavioral information*, F (1, 29) = 10.39, p = .003, partial $\eta^2 = 0.26$, $\Omega = 0.88$, such that employers used warmth more than competence behavioral information. This main effect was qualified by a significant *behavioral information × game* interaction, F (1, 29) = 27.47, $p = 1.3 \times 10^{-5}$, partial $\eta^2 = 0.47$, $\Omega = 1.00$. We computed paired sample *t*-tests with a Bonferroni corrected $\alpha = .0125$. When comparing *games* within *behavioral information*, we found that *warmth behavioral information* was used significantly more in the *trust* than *rock estimation game*, t (29) = 4.02, $p = 3.75 \times 10^{-4}$, while *competence information* was used significantly more in the *rock estimation game* than *trust game*, t (29) = 5.40, $p = 8.0 \times 10^{-6}$. This suggests that employers generalized from the corresponding trait-relevant behavioral information, using it more frequently to guide their decisions within the relevant than irrelevant game. Comparing *behavioral information* within *game*, we found that in the *trust game*, *warmth behavioral information* was used significantly more than *competence*, t (29) = 8.25, $p = 1.0 \times 10^{-7}$, but in the *rock estimation game*, *competence behavioral information* was not used significantly more than *warmth*, t (29) = 0.52, p = .606. There was no significant main effect for *game*, suggesting that employers did not differ in their use of behavioral information across both games. These findings replicate Study 1.

Personal norms

We next used the response to the question asking employers to report their own donation behavior to generate cutoff criteria (Figure 4b). There was a significant *behavioral information × game* interaction, F (1, 29) = 22.87, $p = 4.6 \times 10^{-5}$, partial $\eta^2 = 0.44$, $\Omega = 1.00$. We computed paired sample *t*-tests with a Bonferroni corrected $\alpha = .0125$. When comparing *games* within *behavioral information*, we found that *warmth behavioral information* was used significantly more in the *trust* than *rock estimation game*, t (29) = 3.24, p = .003, while *competence information* was used significantly more in the *rock estimation game* than *trust game*, t (29) = 5.40, $p = 8.0 \times 10^{-6}$. This suggests that employers generalized from the corresponding trait-relevant behavioral information, using it more frequently to guide their decisions within the relevant than irrelevant game. Comparing *behavioral information* within *game*, we found no significant differences. There was no significant main effect for *game or behavioral information*, suggesting that employers did not differ in their use of behavioral information across both games or in their use of warmth or competence behavioral information. These findings do not replicate the pattern of results in Study 1 or the actual norms as cutoff criteria,

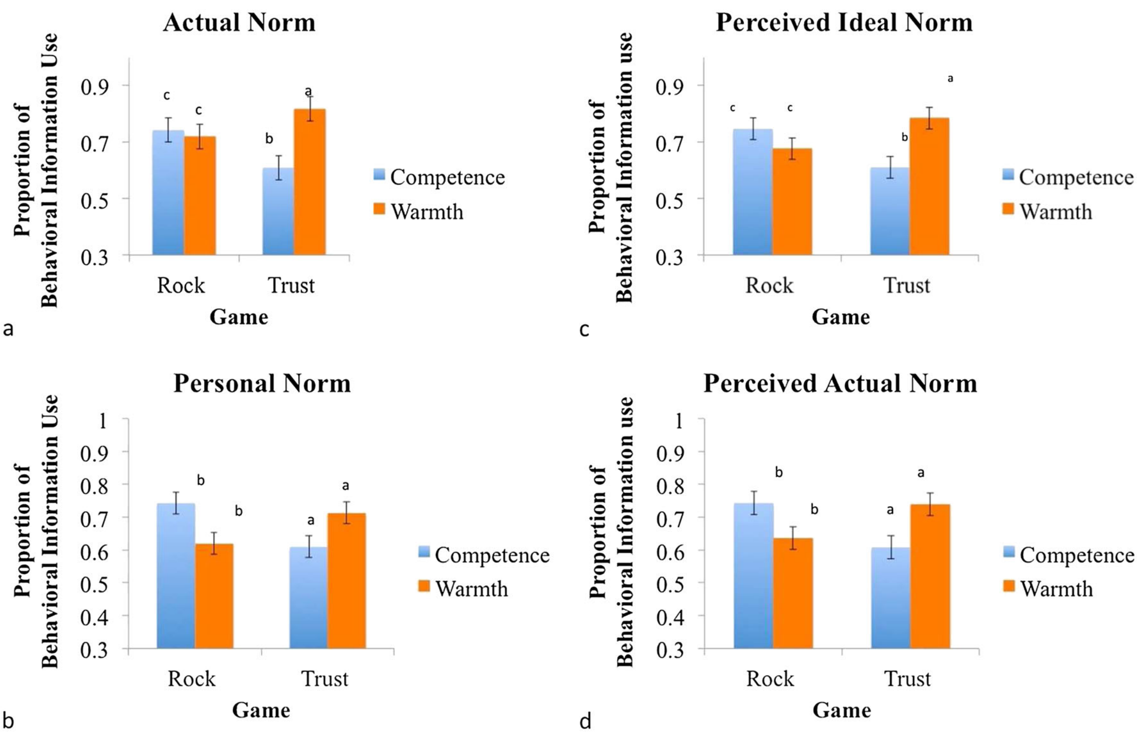

Figure 4. Investment decisions in Study 2. Bar graphs depict the proportion of behavioral information used in each game based on actual norms in (a) Study 2, and (b) personal norms, (c) perceived ideal norms, and (d) perceive actual norms as cutoff criteria. Error bars reflect standard error of the mean. Different letters represent statistically significant differences between means, that is, different letters are significantly different from each other (e.g., the a's are the same, but different from b and c)

DOI: 10.1002/bdm

suggesting that personal norms serve as less suitable substitutes for actual norms.

Perceived ideal norms

We next used the response to the question asking employers to report their perceived ideal norm to generate cutoff criteria (Figure 4c). There was a significant *behavioral information* × *game* interaction, F (1, 28) = 27.12, $p = 1.6 \times 10^{-5}$, partial $\eta^2 = 0.49$, $\Omega = 1.00$. We computed paired sample t-tests with a Bonferroni corrected $\alpha = .0125$. When comparing *games* within *behavioral information*, we found that *warmth behavioral information* was used significantly more in the *trust* than *rock estimation game*, t (28) = 4.07, $p = 3.47 \times 10^{-4}$, while *competence information* was used significantly more in the *rock estimation game* than *trust game*, t (28) = 5.37, $p = 1.0 \times 10^{-5}$. This suggests that employers generalized from the corresponding trait-relevant behavioral information, using it more frequently to guide their decisions within the relevant than irrelevant game. Comparing *behavioral information* within *game*, we found that in the *trust game*, *warmth behavioral information* was used significantly more than *competence*, t (28) = 6.89, $p = 1.0 \times 10^{-7}$, but in the *rock estimation game*, *competence behavioral information* was not used significantly more than *warmth*, t (28) = 1.69, $p = .119$. There was no significant main effect for *game*, suggesting that employers did not differ in their use of behavioral information across both games. These results replicate the findings for actual norms as well as the result in Study 1, suggesting that perceived ideal norms serve as suitable substitutes for actual norms.

Perceived actual norms

Finally, we used the response to the question asking employers to report their perceived actual norms of the employees to generate cutoff criteria (Figure 4d). There was a significant *behavioral information* × *game* interaction, F (1, 28) = 23.65, $p = 4.0 \times 10^{-5}$, partial $\eta^2 = 0.46$, $\Omega = 1.00$. We computed paired sample t-tests with a Bonferroni corrected $\alpha = .0125$. When comparing *games* within *behavioral information*, we found that *warmth behavioral information* was used significantly more in the *trust* than *rock estimation game*, t (28) = 3.61, $p = .001$, while *competence information* was used significantly more in the *rock estimation game* than *trust game*, t (28) = 5.30, $p = 1.2 \times 10^{-5}$. This suggests that employers generalized from the corresponding trait-relevant behavioral information, using it more frequently to guide their decisions within the relevant than irrelevant game. Comparing *behavioral information* within *game*, we found no significant differences. There was no significant main effect for *game or behavioral information*, suggesting that employers did not differ in their use of behavioral information across both games or in their use of warmth or competence behavioral information. These findings do not perfectly replicate the pattern of results in Study 1 or using actual norms as cutoff criteria, suggesting that perceived actual norms serve as less suitable substitutes for actual norms.

Employer confidence ratings

We performed a 2 (*game*) × 2 (*decision*) repeated measures ANOVA on confidence ratings to determine whether employers felt more or less confident when deciding to invest or not in the trust and rock estimation games. We found a significant *decision* main effect, F (1, 26) = 6.23, $p = .019$, partial $\eta^2 = 0.19$, $\Omega = 0.67$, such that employers felt more confident when not investing ($M = 73.13$, $SD = 11.12$) rather than investing ($M = 66.09$, $SD = 17.21$). We computed paired sample t-tests with a Bonferroni corrected $\alpha = .0125$. This main effect was driven by a significant difference for trust game decisions, t (28) = −2.91, $p = .007$, such that confidence was higher when employers decided to not trust an employee ($M = 75.88$, $SD = 13.03$) than when they trusted an employee ($M = 66.67$, $SD = 17.76$; Figure 5a). These results suggest that employers were least confident in their decisions to invest in the employees. Neither the game main effect nor the interaction revealed significant differences. These results perfectly replicate Study 1.

We then tested whether these confidence ratings differed for employees rated as high or low on each trait dimension. For competence, we found a significant difference, t (27) = 5.79, $p = 4.0 \times 10^{-6}$, where employers were also more confident making decisions for high ($M = 75.20$, $SD = 10.10$) rather than low competence employees ($M = 64.24$, $SD = 14.58$). For warmth, we tested for differences in confidence ratings using each of the four cutoff criteria for a "good" decision.

Actual norms

We found a significant difference, t (27) = 2.43, $p = .022$, such that employers were more confident when making decisions regarding high ($M = 76.91$, $SD = 11.91$) than low warmth employees ($M = 63.30$, $SD = 28.59$).

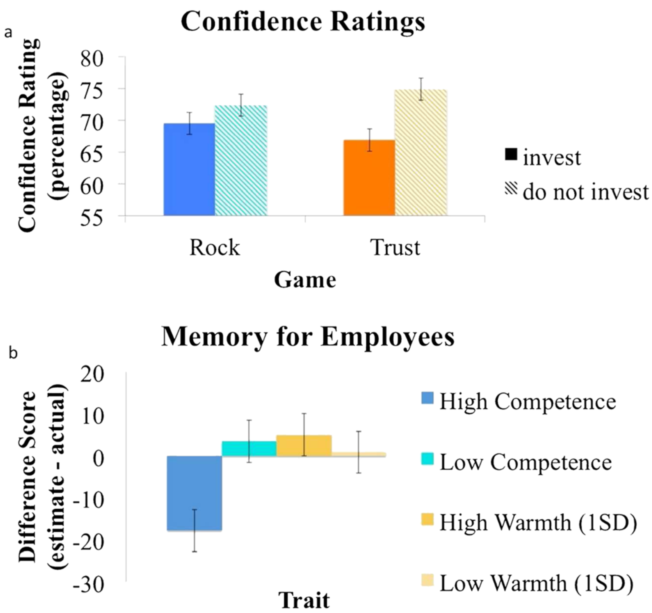

Figure 5. Confidence ratings and memory for employees. Bar graphs depict employers' reported (a) confidence in their investment decisions and employers' (b) memory recall of employees' performance in Study 2. Error bars reflect standard error of the mean

DOI: 10.1002/bdm

Behavioral intentions
We found a significant difference, $t\ (27) = 4.40$, $p = 1.53 \times 10^{-4}$, such that employers were more confident when making decisions regarding high ($M = 76.45$, $SD = 13.64$) than low warmth employees ($M = 60.37$, $SD = 17.08$).

Perceived ideal norms
We found a significant difference, $t\ (27) = 4.77$, $p = 5.7 \times 10^{-5}$, such that employers were more confident when making decisions regarding high ($M = 77.33$, $SD = 12.62$) than low warmth employees ($M = 63.17$, $SD = 16.60$).

Perceived actual norms
We found a significant difference, $t\ (26) = 4.56$, $p = 1.08 \times 10^{-4}$, such that employers were more confident when making decisions regarding high ($M = 75.80$, $SD = 13.31$) than low warmth employees ($M = 58.99$, $SD = 18.14$). Together, these results suggest that again employers were intuitively aware of “good” and “bad” employees regardless of the cutoff criteria, replicating the findings of Study 1.

Employer memory data
As in Study 1, we subtracted employees’ actual time estimation accuracy and donations from employers’ estimates to create indices of memory accuracy before performing one-sample *t*-tests against the value zero. For competence, we found that employers’ memory indices did significantly differ for time estimation, $t\ (29) = -2.09$, $p = .045$, $M_{diff} = -4.76$, $SD = 12.44$, such that they underestimated employees’ time estimation ability. We followed up this finding by testing whether employers displayed memory differences for employees who were categorized as high or low on the trait dimension, calculating memory accuracy indices separately for high and low competence employees. We found that employers showed a significant difference for highly competent employees, $t\ (29) = -6.50$, $p = 1.00 \times 10^{-7}$, such that they significantly underestimated their time estimation ability ($M_{diff} = -17.79$, $SD = 14.99$). However, employers accurately recalled low competence employees’ time estimation ability, $t\ (29) = 1.55$, $p = .131$ (Figure 5b). Together, these results suggest that employers’ accurately recalled low competence employees’ time estimation abilities, underestimating the ability of more skilled employees. These findings partially replicate the results of Study 1, demonstrating that employers’ memory for highly skilled employees underestimated their competence.

We performed a similar analysis for warmth, where we found that employers’ memory indices did not significantly differ for donations to charity, $t\ (29) = 0.98$, $p = .327$. We followed this up by separately testing memory recall for high and low warmth employees. We found no significant difference for recall of donation behavior for both high, $t\ (29) = 0.98$, $p = .336$ and low warmth employees, $t\ (29) = 0.85$, $p = .405$, suggesting that employers accurately remembered warmth behavioral information. These findings replicate the results of Study 1, demonstrating that employers accurately remembered warmth behavioral information for employees.

We next tested whether there was enhanced memory for warmth or competence information by performing a paired-samples *t*-test using the memory indices. There was no significant difference, $t\ (31) = 1.83$, $p = .078$, replicating Study 1.

CONCLUSION

There was no significant change in investment behavior in the rock estimation game after instituting the equivalent payoff paradigm. Additionally, we found little difference between our results when using the mean generosity score cutoff criterion for ‘good’ decisions employed in Study 1, and using the employers’ estimates of donation norm. This was particularly true when we used estimates of ideal norms, and less true when we used estimates of personal norms and perceived actual norms. This suggests that norms may be used to predict warmth-relevant behavior. Therefore, Study 2 replicated our results from Study 1 using norm estimates as determinants of high and low warmth employees. Study 2 also adds reliability to the findings in Study 1 while ruling out perceived player motivational differences between the competence and warmth games.

STUDY 3

The first two studies demonstrate that people use behavioral information to predict behavior, guiding consequential decisions. However, we have not yet tested whether trait inferences predict behavior. We intentionally avoided asking for self-reports of traits or measuring trait inferences directly during either of the first two studies because we did not want to engage an explanatory system. Therefore, we collected data in Study 3 to determine (1) whether it is possible to infer trait warmth and competence from the behavioral information in employee profiles and (2) whether trait warmth inferences better predict behavior than trait inferences in Studies 1 and 2.

Method
Participants
Three hundred seventeen individuals recruited from Amazon’s Mechanical Turk completed Study 3 online. The sample consisted of the following age breakdown: 34.7% were 18 to 29 years old, 27.1% were 30 to 39 years old, 14.8% were 40 to 49 years old, 15.8% were 50 to 59 years old, and 7.6% were over 60 years old; 51.1% of the subjects were female.

Stimuli
We replaced the employees’ photographs with computer-generated neutral faces (Todorov et al., 2013) to control for the impact of identity on any trait inferences or heuristic

DOI: 10.1002/bdm

processes. All other information in the 60 employee profiles remained unchanged from the first two studies.

Procedure

The procedure for employees was the same as the first two studies, except that it took place online. Participants still viewed employee profiles, but instead of making investment decisions, they rated 10 profiles on 10 warmth-related traits (helpful, sincere, trustworthy, moral, sociable, caring, unfriendly [reverse scored], insensitive [reverse scored], generous, and warm) and 10 competence-related traits (intelligent, skillful, creative, efficacy, capable, lazy-reverse scored, disorganized [reverse scored], punctual, precise, and competent). Therefore, 50–55 participants rated each profile.

Data analysis strategy

We first perform a cluster analysis with varimax rotation on the traits to reduce the number of traits to the warmth and competence dimensions. We then performed reliability analyses on each dimension before averaging across the ratings to create overall warmth and competence composites. We then compute correlation coefficients to test the association between trait inference and behavior.

Finally, we use the average trait rating across the raters as each profile's cutoff criteria to re-analyze the data collected in Studies 1 and 2. This strategy informs us whether actual ratings of trait warmth and competence based on the employees' behavioral information (independent of the employee's identity) guides participants' real consequential decisions. This allows us to test whether trait inferences serve as heuristics without relying on self-report from participants during the task.

RESULTS

Principle components factor analysis

We first completed a principle components factor analysis with varimax rotation to reduce the number of trait items. Because of our a priori hypothesis, we set criteria for convergence at two factors, which explained 96.49% of the variance. The screen plot confirmed a two-factor solution. A first warmth factor (listed in order of decreasing factor loadings) included items generous, helpful, caring, insensitive (reverse coded), warm, unfriendly, moral, sociable, sincere, and trustworthy, with eigenvalue that ranges from .968 to .832. A second competence factor emerged, consisting of items (listed in order of decreasing factor loadings) precise, skillful, punctual, capable, intelligent, efficacy, competent, disorganized (reverse coded), lazy (reverse coded), and creative, with eigenvalue that ranges from .982 to .741. Interestingly, generous and precise both loaded the highest on warmth and competence factors, respectively, suggesting that they best captured the specific warmth and competence traits inferred from the donation and time estimation behaviors.

Reliability estimates for scale warmth and competence

We then averaged the traits along the two factors, creating composites. A reliability analysis of the warmth composite resulted a Cronbach's alpha = .992, while the competence composited resulted a Cronbach's alpha = .987, suggesting that the items in the composite formed a reliable scale. The warmth and competence scales were significantly positively correlated, r (58) = .575, $p = 2.0 \times 10^{-6}$.

Trait–behavior correlations

We next ran correlations between the behavioral information in each profile and the inferred traits warmth and competence to test whether trait inferences corresponded to actual behavior. We found significant correlations between trait warmth inferences and donation behavior (r (58) = .839, $p = 1.0 \times 10^{-8}$), while trait competence correlated with both donation (r (58) = .329, $p = .010$) and time estimation behavior (r (58) = .847, $p = 1.0 \times 10^{-8}$). This suggests that behavioral information from both games underlie their respective trait inferences, although competence is also inferred from generous behavior.

Traits predicting behavior

Finally, we wondered whether trait inferences served as heuristics in Studies 1 and 2. To test this, we relied on the pattern of results in the first two studies to derive a number of axioms that traits should satisfy when examined with similar statistical analyses. These axioms concern the pattern of main effect and interactions that trait inferences should result if they serve as heuristics:

1. A significant main effect of *trait inference*, such that warmth trait inferences are more frequent than competence trait inferences.
2. A significant *trait inference* × *game* interaction, such that
   a. no difference exists between the occurrence of warmth and competence trait inferences in the rock estimation game,
   b. warmth trait inferences occur more frequently than competence trait inferences in the trust game,
   c. warmth trait inferences occur more frequently in the trust game than rock estimation game, and
   d. competence trait inferences occur more frequently in the rock estimation game than trust game.
3. No significant main effect of *game*, suggesting no difference between trait inferences used in the rock estimation game compared with trust game.

We employed a similar analysis as in Studies 1 and 2, characterizing good and bad decisions this time using the mean trait inferences from the profiles as cutoff criteria for high and low warmth and competence. We calculated mean trait ratings for all behavioral information; profiles with average trait ratings that fell at or above the mean were considered high on the trait dimension, while those that fell below the mean were considered low. We coded a good investment decision if employers invested in the rock or trust game and the employee was rated as high on the relevant traits or if the

DOI: 10.1002/bdm

employer did not invested in the employee in either game and the employee was rated as low on the relevant traits. We coded bad investment decisions when trait and behavior mismatches occurred.

We thus ran a series of 2 (*trait inference*) × 2 (*game*) repeated measures ANOVAs collapsed across employers' behavior in Studies 1 and 2, with *study* as a between-subjects covariate (the *study* main effect was not significant and did not interact with any other main effect or interaction across all our analyses; as such, it is no longer discussed, and results are collapsed across Studies 1 and 2). If employers were using trait inferences as heuristics to predict behavior, then these ANOVAs should replicate the pattern of results observed in the first two studies, satisfying all three axioms.

We ran this analysis first using broad trait inferences of warmth and competence resulting from the averages of the 10 warmth-related traits and the 10 competence-related traits. This allows us to test how a scale measure of warmth and competence fairs in our test of axioms. Specifically, this test informs us whether a general or broad sense of warmth or competence is estimated. We then test subsets of the specific trait attributions that make up the broad trait warmth and competence inferences. This second test assesses the same question as the first but allows dissociation; specifically, testing scale warmth and competence does not rule out whether a specific warmth or competence sub-trait (e.g., precision or generosity) is driving our findings. Testing how endorsement of the specific words "warm" and "competent" replicates this initial test without the influence of the nine other specific sub-traits. Furthermore, we perform two additional tests to examine the effects of specific warmth and competence sub-traits: generous and punctual, and trustworthy and precise. The first pair of traits tests the specific trait inference that may occur from the behavioral information contained in the employee profiles—generosity may be inferred from the charity donations game, and punctuality may be inferred from the time estimation game. The second pair of traits tests the specific trait inference appropriate for the decision facing the employer in the investment games—trustworthiness is relevant for the trust game, and precision is relevant for the rock estimation game. We employ a Bonferroni correction for multiple comparison across the four ANOVAs, resulting $\alpha = .0125$ for significance.

Scale warmth and competence trait inferences

We replicated the analysis performed in Studies 1 and 2 using the average warmth and competence ratings, collapsed across all warmth and competence traits, to generate cutoff criteria for high and low warmth and competence. There was a significant main effect of *game*, F (1, 58) = 147.69, $p < 1.0 \times 10^{-6}$, partial $\eta^2 = 0.72$, $\Omega = 1.00$, such that employers used trait inferences more in the rock estimation game than trust game (Figure 6a). This main effect was qualified by a significant *trait inference* × *game* interaction, F (1, 58) = 43.55, $p < 1.0 \times 10^{-6}$, partial $\eta^2 = 0.43$, $\Omega = 1.00$. There was no main effect of *trait inference*, F (1, 58) = 0.18, $p = .674$.

We computed follow-up paired sample *t*-tests with a Bonferroni corrected $\alpha = 3.125 \times 10^{-3}$. When comparing *games* within *trait inferences*, we found that *warmth trait inferences* (t (59) = 5.00, $p = 5.0 \times 10^{-6}$) and *competence trait*

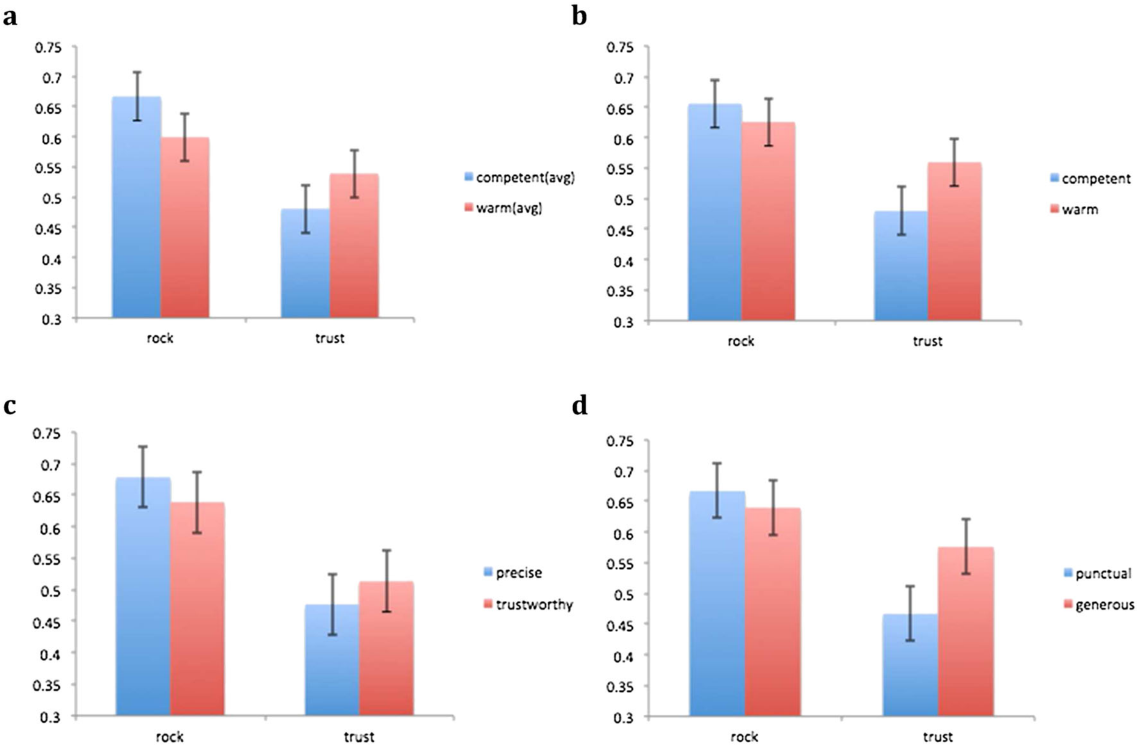

Figure 6. Investment decisions using trait inferences to predict behavior. Bar graphs depict the proportion of trait inferences used in each game based on trait inferences about (a) warmth and competence (scale), (b) warmth and competence, (c) trustworthy and precise, and (d) generous and punctual. Error bars reflect standard error of the mean

DOI: 10.1002/bdm

inferences (t (59) = 12.02, $p < 1.0 \times 10^{-6}$) were used significantly more in the *rock estimation* game than *trust* game. This suggests that both inferred traits more frequently guided decisions within the rock estimation game. Comparing *trait inferences* within *game*, we found that in the *trust game*, *warmth trait inferences* were used significantly more than *competence*, t (59) = 7.06, $p < 1.0 \times 10^{-6}$, and in the *rock estimation game*, *competence trait inferences* were used significantly more than *warmth*, t (59) = −3.56, p = .001. These results satisfy axioms 2, 2b, and 2d.

Specific warmth and competence trait inferences

We conducted the ANOVA for specific warmth and competence trait attribution (ratings of the terms "warm" and "competent") to generate cutoff criteria for high and low warmth and competence. There was a significant main effect of *game*, F (1, 58) = 139.37, $p < 1.0 \times 10^{-6}$, partial $\eta^2 = 0.76$, $\Omega = 1.00$, such that employers used trait inferences more in the rock estimation game than trust game (Figure 6b). There was also a marginally significant main effect of *trait inference*, F (1, 58) = 4.99, p = .029, partial $\eta^2 = 0.08$, $\Omega = 0.59$, such that employers used warmth more than competence trait inferences to guide behavior. These main effects were qualified by a significant *trait inference* × *game* interaction, F (1, 58) = 42.58, $p < 1.0 \times 10^{-6}$, partial $\eta^2 = 0.42$, $\Omega = 1.00$.

We computed follow-up paired sample t-tests with a Bonferroni corrected $\alpha = 3.125 \times 10^{-3}$. When comparing *games* within *trait inferences*, we found that *warmth trait inferences* (t (59) = 5.45, $p = 1.0 \times 10^{-6}$) and *competence trait inferences* (t (59) = 12.34, $p < 1.0 \times 10^{-6}$) were used significantly more in the *rock estimation* game than *trust* game. This suggests that both inferred traits more frequently guided decisions within rock estimation game. Comparing *trait inferences* within *game*, we found that in the *trust game*, *warmth trait inferences* were used significantly more than *competence*, t (59) = 8.82, $p < 1.0 \times 10^{-6}$, but in the *rock estimation game*, *competence trait inferences* were not used significantly more than *warmth*, t (59) = −1.72, p = .091. These findings satisfy axioms 2, 2b, and 2d.

Trustworthy and precise trait inferences

We next replicated the analysis performed in Studies 1 and 2 using trustworthy and precise trait attribution (ratings of the terms "trustworthy" and "precise") to generate cutoff criteria for high and low trustworthiness and precision. There was a significant main effect of *game*, F (1, 58) = 152.87, $p < 1.0 \times 10^{-6}$, partial $\eta^2 = 0.73$, $\Omega = 1.00$, such that employers used trait inferences more in the rock estimation game than trust game (Figure 6c). This main effect was qualified by a significant *trait inference* × *game* interaction, F (1, 58) = 22.07, $p = 1.7 \times 10^{-5}$, partial $\eta^2 = 0.28$, $\Omega = 1.00$. There was no main effect of *trait inference*, F (1, 58) = 0.03, p = .873.

We computed paired sample t-tests with a Bonferroni corrected $\alpha = 3.125 \times 10^{-3}$. When comparing *games* within *trait inferences*, we found that *trustworthiness trait inferences* (t (59) = 9.15, $p = 1.0 \times 10^{-6}$) and *precise trait inferences* (t (59) = 11.77, $p < 1.0 \times 10^{-6}$) were used significantly more in the *rock estimation* game than *trust* game. This suggests that both inferred traits more frequently guided decisions within the rock estimation game. Comparing *trait inferences* within *game*, we found that in the *trust game*, *trustworthiness trait inferences* were used significantly more than *precision*, t (59) = 5.80, $p < 1.0 \times 10^{-6}$, but in the *rock estimation game*, *precision trait inferences* were not used significantly more than *trustworthy*, t (59) = −2.55, p = .013. These findings satisfy axioms 2, 2b, and 2d.

Generous and punctual trait inferences

Next, we replicated the analysis performed in Studies 1 and 2 using generous and punctual trait attribution (ratings of the terms "trustworthy" and "precise") to generate cutoff criteria for high and low trustworthiness and precision. There was a significant main effect of *game*, F (1, 58) = 153.15, $p < 1.0 \times 10^{-6}$, partial $\eta^2 = 0.73$, $\Omega = 1.00$, such that employers used trait inferences more in the rock estimation game than trust game (Figure 6d). There was also a main effect of *trait inference*, F (1, 58) = 8.46, p = .005, partial $\eta^2 = 0.13$, $\Omega = 0.82$, such that generous trait inferences guided decisions more than punctual trait inferences. Both main effects were qualified by a significant *trait inference* × *game* interaction, F (1, 58) = 22.07, $p = 1.7 \times 10^{-5}$, partial $\eta^2 = 0.28$, $\Omega = 1.00$.

We computed follow-up paired sample t-tests with a Bonferroni corrected $\alpha = 3.125 \times 10^{-3}$. When comparing *games* within *trait inferences*, we found that *generous trait inferences* (t (59) = 4.67, $p = 1.8 \times 10^{-5}$) and *punctual trait inferences* (t (59) = 11.86, $p < 1.0 \times 10^{-6}$) were used significantly more in the *rock estimation* game than *trust* game. This suggests that both inferred traits more frequently guided decisions within the rock estimation game. Comparing *trait inferences* within *game*, we found that in the *trust game*, *generous trait inferences* were used significantly more than *punctual*, t (59) = 5.80, $p < 1.0 \times 10^{-6}$, but in the *rock estimation game*, *punctual trait inferences* were not used significantly more than *generous*, t (59) = −1.20, p = .235. These results satisfy axioms 1, 2, 2b, and 2d.

CONCLUSION

Study 3 demonstrated that people do infer traits from the behavioral information in the profiles. However, these trait inferences were informed by both warmth and competence behavioral information. It further demonstrated that these trait inferences only partially replicate the pattern of behavior observed when using behavioral information or norm estimates to predict warmth-related behavior. All trait inferences, whether to broad traits warmth and competence or specifically to traits that more narrowly described the context, failed to satisfy all three axioms. In particular, they all failed to satisfy axioms 2a, 2c, and 3. Only inferences to the specific traits (generous and punctual) from the initial behavioral information satisfied axiom 1, suggesting that these

DOI: 10.1002/bdm

traits most likely served as heuristics. However, the data suggest that none truly did.

GENERAL DISCUSSION

This research combines literature in social psychology and behavioral economics to demonstrate that people generalize to broad primary person perception dimensions trait warmth and competence from single instances of behavior, but those trait inferences do not guide future warmth-relevant decisions better than estimates of normative information. People often infer traits from single acts of behavior. They then generalize to broad trait domains to predict behavior. For instance, an observer witnesses an act of generosity, which leads to a trait inference that the person is generous. That observer may generalize across the warmth domain, thinking that the person may also be trustworthy, guiding the observer's decision to trust that person. Here, we demonstrate that such generalizations across broad trait domains warmth and competence exist, yet fail to predict warmth-relevant behavior as well as normative estimates of behavior.

We find that people depend on estimated norms relative to the social context to predict warmth-relevant behavior. When making decisions in a social context, people often have little or no information about how a specific person might behave. They do, however, have much more information available about the behavior of other people in that social context more generally. Because people are social agents, we often try to predict what others will do, relying on heuristics to guide those decisions. While previous studies have identified trait inferences as generalizations that guide these predictions, our results show that although participants are capable of making broad trait inferences, they are in fact more likely using previous behavior and estimated norms to generate heuristics when predicting warmth-relevant behavior.

All of the behaviors in the first two studies were the result of real people's consequential decisions. Interestingly, broad trait inferences occur across unrelated tasks—the ability to estimate time is not related to estimating rocks, and the amount of money donated to charity is not related to the likelihood of sharing money in the trust game—suggesting that broad trait inferences are not actually predictive of people's actual behavior in our experimental paradigm. Psychology has long ignored the question of accuracy of trait inferences (Zaki & Ochsner, 2011). By using consequential behavior in our task design, we are able to demonstrate that broad inferences are inaccurate and participants avoid using such inferences in predicting behavior: a reasonable occurrence.

Additionally, the dissociation of the two person perception domains suggests that generalizations do not occur uniformly for all types of social information. Trait warmth behavior affects investment decisions equally in both the warmth and competence domains, while competence behavior affects decisions more in its own domain than the warmth domain. Participants were also significantly more confident when making decisions in the warmth domain, and more accurately remembered warmth behavior, further suggesting a dissociation of processing between warmth and competence. Our findings also indicate that warmth behavioral information is used significantly more overall than competence behavioral information, and this holds true even across person perception domains. This is consistent with recent brain-imaging evidence (Lee & Harris, 2014).

It should be noted that we are not contesting the general premise that traits predict behavior: Indeed, we the authors hold this point of view because of the overwhelming data supporting this assertion. What the current data contest is that traits predict behavior even when explanatory mechanisms are not engaged. Indeed, we did not measure explanation in our samples along with prediction, so we cannot rule out the presence of explanation in our paradigm; future studies will address this question. However, our paradigm is one of the few where participants make consequential predictions and are not asked to report trait inferences, controlling for explanatory mechanisms. People make spontaneous trait attributions when observing behavior when they are asked to report on the specific attribution; our paradigm does not ask participants to make such a report and finds that norms do a better job.

Additionally, our interest is not in addressing whether people ignore base-rate information. In such studies, participants are often provided with this information, not asked to generate it themselves. Here, we simply use normative information as a subjective assessment of the social context by each individual employer, not as an objective truth regarding behavior. Therefore, our use of normative information deviates from the use of the term "base-rates" commonly operationalized in the literature; the norm may not be based on the past behavior, but on estimates of behavior based on an interpretation of the social context (this is a donation context, so people should donate a lot, but probably will donate a little). Employers may then evaluate the observed employee behavior in light of this standard, which may serve as a cutoff. Therefore, each prediction about future behavior may rely on where the employee's decision falls relative to the cutoff. Hence, the normative information we describe, provided not by the experimenter but self-generated by the employers, serves to create these arbitrary cutoffs, which are used to predict behavior. Interestingly, there is still a trait component to this process perhaps, where employers view an employee who comes in below the cutoff as less generous for instance, relative to other people, and therefore may be less trustworthy relative to other people. In fact, our use of the term normative information can be considered an individual difference measure because we are examining behavioral deviation from these estimated normative means for the behavior. However, it is difficult to test this assertion using the current paradigm. Future students can determine whether estimated norms are used differently for prediction than base-rate information framed as true or accurate information about the social context.

Our results beg the following questions: Why does warmth information hold a position of primacy, and why are estimates of ideal norms closer to actual norms than personal norms or estimates of actual norms? Also, we referred to the competence games as the investment and gambling games. There is often a stigma associated with "gambling."

DOI: 10.1002/bdm

It is therefore possible that participants might have been biased against betting in the gambling game. People also infer traits about others from photographs and faces (Todorov, Pakrashi, & Oosterhof, 2009; van't Wout & Sanfey, 2008). Although we used photographs of faces of employees in our initial two studies, we did not assess the extent to which employees' faces or perceived social identity influenced trait inferences that guided behavior; future research could address the extent to which trait generalization from these images interact with the prior behavior of the person to guide decisions. Finally, we did not test whether norm estimates predict competence-related behavior. Future research is needed to rule out alternate hypotheses presented by these potential limitations.

Another limitation surrounds the trait ratings in Study 3; the mean trait ratings determine the employee's standing on the trait. This fairly crude method of testing "inferred traits" suggests that the traits will not always match the decision maker's own trait inference, and thus, this may not provide a very sensitive test. Consequently, it may not be highly surprising or informative to find that trait inferences (based on means from another sample) are not all that predictive of decisions. However, this is a consequence of the paradigm and the attempts to make trait inferences less salient. Future research can require participants to also generate trait inferences of the past behavior, although at a time other than when participants make predictions.

In conclusion, these studies suggest that a general assumption about trait inferences as heuristics that guide behavior may be inaccurate. Given low cross-situational behavior consistency, people may be intuitively aware of the problems associated with using traits as predictive tools and rely instead on normative information and previous behaviors to generate heuristics that may guide future behavior. This study by no means provides definitive evidence that norms guide future behavior exclusively and traits play no role. Perhaps traits also interact with norms to guide behavior. Nonetheless, this research does highlight the difference between explanation and prediction and suggests there is value in conducting research that investigates prediction exclusively (without relying on self-report and tapping into explanatory mechanisms). Moreover, this research utilizes two experimental approaches across the same basic paradigm, shedding new light on an old phenomenon.

ACKNOWLEDGEMENTS

Start-up funds and a research award from Duke University to L. T. Harris and E. H. Thompson, respectively, funded this research. This work comprised the senior honor thesis for E. H. Thompson. These data were presented at the Society for Neuroeconomics Meeting in Lausanne Switzerland, September 2013. We would like to thank Naomi Ellermers and Yael Coen for their comments on earlier versions of this manuscript. L. T. Harris is now at the Department of Social and Organisational Psychology, Leiden University.

REFERENCES

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4): 268–277.

Andrews, K. (2012). *Do Apes Read Minds? Toward a new folk psychology*. MIT Press, Cambridge: MA.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal & Social Psychology*, *41*, 258–290.

Baird, B. M., & Lucas, R. E. (2011). "… And how about now?": Effects of item redundancy on contextualized self-reports of personality. *Journal of Personality*, *79*(5), 1081–1112.

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues.

Beck, L., McCauley, C., Segal, M., & Hershey, L. (1988). Individual differences in prototypicality judgments about trait categories. *Journal of Personality and Social Psychology*, *55*(2), 286.

Booth-Butterfield, S. (1988). Inhibition and student recall of instructional messages. *Communication Education*, *37*(4), 312–324.

Botvinick, M. M., & Cohen, J. D. (2014). The computational and neural basis of cognitive control: Charted territory and new frontiers. *Cognitive Science*, *38*, 1249–1285.

Burke, P. A., Kraut, R. E., & Dworkin, R. H. (1984). Traits, consistency, and self- schemata: what do our methods measure? *Journal of Personality and Social Psychology*, *47*(3), 568.

Chaplin, W. F., & Goldberg, L. R. (1984). A failure to replicate the Bem and Allen study of individual differences in cross-situational consistency. *Journal of Personality and Social Psychology*, *47*(5), 1074.

Church, A. T., Anderson-Harumi, C. A., del Prado, A. M., Curtis, G. J., Tanaka-Matsumi, J., Valdez Medina, J. L., & Katigbak, M. S. (2008). Culture, cross-role consistency, and adjustment: Testing trait and cultural psychology perspectives. *Journal Of Personality And Social Psychology*, *95*(3), 739–755.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*, 285–290.

Emmons, R. A., & Diener, E. (1986). A goal-affect analysis of everyday situational choices. *Journal of Research in Personality*, *20*(3), 309–326.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in human amygdala. *Journal of Cognitive Neuroscience*, *19*, 1508–1519.

Epstein, S., & Teraspulsky, L. (1986). Perception of cross-situational consistency. *Journal of Personality and Social Psychology*, *50*(6), 1152.

Eysenck, M. W., & Wild, J. (1996). Trait anxiety, social desirability and cross-situational variability. *Personality and Individual Differences*, *21*(2), 231–234.

Ferreira, M. B., Garcia-Marques, L., Hamilton, D., Ramos, T., Uleman, J. S., & Jerónimo, R. (2012). On the relation between spontaneous trait inferences and intentional inferences: An inference monitoring hypothesis. *Journal of Experimental Social Psychology*, *48*(1), 1–12.

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

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken, Y. Trope, S. Chaiken, Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231–254). New York, NY, US: Guilford Press.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1–74.

Fiske, S. T., & Taylor, S. E. (2013). *Social Cognition: From brains to culture*. Sage, London: UK.

DOI: 10.1002/bdm

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*(5), 773.

Furr, R. M., & Funder, D. C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centered, and person-centered approaches. *Journal of Research in Personality*, *38*(5), 421–447.

Furr, R. M. (2009). Personality psychology as a truly behavioural science. *European Journal of Personality*, *23*(5), 369–401.

Frith, U., & Frith, C. (2001). The biological basis of social interaction. *Current Directions In Psychological Science*, *10*(5), 151–155.

Heffler, B., & Magnusson, D. (1979). Generality of behavioural data: IV. Cross-situational invariance of objectively measured behaviour. *Perceptual and Motor Skills*, *48*(2), 471–477.

Hemmelgarn, P. L., James, P. L., Ladd, P. L., & Mitchell, D. K. (1995). Testing for cross-situational-consistency: A confirmatory factor analytic approach. *Journal of Management*, *21*(1), 121–139.

Hettema, J., & Hol, D. P. (1998). Primary control and the consistency of interpersonal behaviour across different situations. *European journal of personality*, *12*(4), 231–247.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237.

Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska Symposium On Motivation*, *15*, 192–238.

Koestner, R., Bernieri, F., & Zuckerman, M. (1989). Trait-specific versus person-specific moderators of cross-situational consistency. *Journal of Personality*, *57*(1), 1–16.

Krahé, B. (1986). Similar perceptions, similar reactions: An idiographic approach to cross-situational coherence. *Journal of Research in Personality*, *20*(3), 349–361.

Landy, D., & Sigall, H. (1974). Beauty is talent: Task evaluation as a function of the performer's physical attractiveness. *Journal of Personality and Social Psychology*, *29*, 299–304.

Lee, V. K., & Harris, L. T. (2013). How social cognition can inform social decision-making. *Frontiers in Neuroscience*, *1*, 1–13.

Lee, V. K., & Harris, L. T. (2014). Sticking with the nice guy: Trait warmth information impairs learning and modulates person perception brain network activity. *Cognitive and Behavioral Neuroscience*, *14*(4), 1420–1437.

Leikas, S., Lönnqvist, J. E., & Verkasalo, M. (2012). Persons, situations, and behaviors: Consistency and variability of different behaviors in four interpersonal situations. *Journal of Personality and Social Psychology*, *103*(6), 1007.

Lippa, R., & Mash, M. (1981). The effects of self-monitoring and self-reported consistency on the consistency of personality judgments made by strangers and intimates. *Journal of Research in Personality*, *15*(2), 172–181.

Magnusson, D. (1976). The person and the situation in an international model of behavior. *Scandinavian Journal of Psychology*, *17*(1), 253–271.

Magnusson, D., & Ekehammar, B. (1978). Similar situations—Similar behaviors?: A study of the intraindividual congruence between situation perception and situation reactions. *Journal of Research in Personality*, *12*(1), 41–48.

McCarthy, R. J., & Skowronski, J. J. (2011). What will Phil do next? Spontaneously inferred traits influence predictions of behavior. *Journal Of Experimental Social Psychology*, *47*(2), 321–332.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems reviewed. *Personality And Neuropsychologia*, *46*(3), 774–785.

Moore, F. R., Filippou, D., & Perrett, D. (2011). Intelligence and attractiveness in the face: Beyond the attractiveness halo effect. *Journal of Evolutionary Psychology*, *9*, 205–217.

Moskowitz, D. S. (1994). Cross-situational generality and the interpersonal circumplex. *Journal Of Personality And Social Psychology*, *66*(5), 921–933.

Newman, L. S. (1996). Trait impressions as heuristics for predicting future behavior. *Personality And Social Psychology Bulletin*, *22*(4), 395–411.

Norenzayan, A., Choi, I, & Nisbett, R. E. (2002). Cultural similarities and differences in social inference: Evidence from behavioral predictions and lay theories of behavior. *Personality and Social Psychology Bulletin*, *28*(1), 109–120.

Oishi, S., Diener, E., Napa Scollon, C., & Biswas-Diener, R. (2004). Cross-situational consistency of affective experiences across cultures. *Journal of Personality and Social Psychology*, *86*(3), 460.

Paunonen, S. V. & Jackson, D. N. (1985). Idiographic measurement strategies for personality and prediction: Some unredeemed promissory notes. *Psychological Review*, *92*(4), 486.

Paunonen, S. V. & Jackson, D. N. (1986). Nomothetic and idiothetic measurement in personality. *Journal of Personality*, *54*(2), 447–459.

Pervin, L. A. (1985). Personality: Current controversies, issues, and directions. *Annual review of psychology*, *36*(1), 83–114.

Said, C. P., Baron, S., & Todorov, A. (2009). Nonlinear amygdala response to face trustworthiness: Contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience*, *21*, 519–528.

Semin, G. R., & Krahé, B. (1988). Explaining perceived cross-situational consistency: Intuitive psychometrics or semantic mediation? *European Journal of Personality*, *2*(4), 239–252.

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. In A. Kingstone & M. Miller (Eds.), *The Year in Cognitive Neuroscience 2008*, Annals of the New York Academy of Sciences, 1124, 208–224.

Todorov, A., Baron, S., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social, Cognitive, & Affective Neuroscience*, *3*, 119–127.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*, 813–833.

Todorov, A., Dotsch, R., Porter, J., Oosterhof, N., & Falvello, V. (2013). Validation of Data-Driven Computational Models of Social Perception of Faces. *Emotion*, *13*, 724–738.

van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*, 796–803.

Van Mechelen, I. (2009). A royal road to understanding the mechanisms underlying person-in-context behavior. *Journal of Research In Personality*, *43*(2), 179–186.

Van Overwalle, F., Van Duynslaeger, M., Coomans, D., & Timmermans, B. (2012). Spontaneous goal inferences are often inferred faster than spontaneous trait inferences. *Journal of Experimental Social Psychology*, *48*(1), 13–18.

Welbourne, J. (2001). Changes in impression complexity over time and across situations. *Personality And Social Psychology Bulletin*, *27*(9), 1071–1085.

Wells, B. M., Skowronski, J. J., Crawford, M. T., Scherer, C. R., & Carlston, D. E. (2011). Inference making and linking both require thinking: Spontaneous trait inference and spontaneous trait transference both rely on working memory capacity. *Journal of Experimental Social Psychology*, *47*(6), 1116–1126.

Willis, J., & Todorov, A. (2006). First impressions making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature neuroscience*, *5*(3), 277–283.

Zaki, J., & Ochsner, K. (2011). Reintegrating accuracy into the study of social cognition. *Psychological Inquiry*, *22*, 159–182.

DOI: 10.1002/bdm

Authors' biographies:

Lasana T. Harris is a senior lecturer at University College London in Experimental Psychology and a guest lecturer in Social and Organisational Psychology at Leiden University. He received his PhD from Princeton University, where he studied both social psychology and cognitive neuroscience. His research interests explore the boundary conditions of social cognition (dis)engagement.

Rachel E. Kranton is James B. Duke Professor of Economics at Duke University. She earned her PhD in Economics at the University of California, Berkeley. She is a fellow of the Econometric Society and has held fellowships at the Russell Foundation and the Institute for Advanced Study. Her research brings social interactions, networks, and identity into formal economic models of decision making and strategic interaction.

Victoria K. Lee is a PhD candidate in the Social Psychology program at Duke University. She received her BA from Rutgers University, where she majored in psychology. Her research interests focus on examining how social cognition and person perception affect decision making in economic contexts.

Elizabeth Thompson earned her bachelors degree from Duke University. She is currently a strategist at a branding and design consultancy. Her research interests include social neuroscience and neuroeconomics.

Authors' addresses:

Lasana T. Harris, Psychology and Neuroscience, Duke University, Durham, NC, USA.

Victoria K. Lee, Psychology and Neuroscience, Duke University, Durham, NC, USA.

Elizabeth Thompson, Psychology and Neuroscience, Duke University, Durham, NC, USA.

Rachel E. Kranton, Economics, Duke University, Durham, NC, USA

DOI: 10.1002/bdm