

Own experience bias in evaluating the efforts of others*

Edward Cartwright [†] Myrna Wooders [‡]

This version: March 2020

Keywords: Bayesian updating, discrimination, attribution error, false consensus, own experience bias

JEL Codes: J71, D11, C91

Abstract

We develop a model with which to explore how an individual's own experience in the labor market may influence her assessment of efforts of other individuals. Specifically, we consider a two stage process in which individuals first learn, through experience, whether effort is rewarded and then subsequently have to estimate the effort of others. We derive a theoretical benchmark based on rational inference and then explore how own experience may lead to systematic bias from this benchmark. Our theoretical results suggest that those who are not rewarded for high effort will underestimate the effort of other individuals while those for whom effort is rewarded will (slightly) overestimate the effort of others. We empirically test and confirm this prediction in the lab.

*This research was partly funded by: Nuffield Foundation Social Sciences Small Research Grant SGS/36604, 'An experimental investigation of prejudice and economic discrimination'; NSF Grant 1526860; the Douglas May Funds for Economics at Vanderbilt University; New York University Abu Dhabi, and an Equity, Diversity and Inclusion Grant at Vanderbilt University.

[†]Department of Economics and Marketing, De Montfort University, Leicester, UK. edward.cartwright@dmu.ac.uk

[‡]Department of Economics, Vanderbilt University, myrna.wooders@vanderbilt.edu

1 Introduction

There is a growing recognition in academic research and in the popular press of the role that *unconscious bias* can play in labor market discrimination. In short, an individual may be unaware that her judgment of others is influenced by, say, gender or race. Unconscious bias differs fundamentally from more conventional explanations of discrimination, such as taste-based discrimination, where the individual simply prefers one group over another, or statistical discrimination, where the individual uses observable characteristics to make unbiased inferences about unobservable characteristics. Disentangling unconscious bias from other potential sources of discrimination is, however, difficult because individuals have unique experiences and, when assessing others, may simply be making the best assessments possible given their own experiences and the information available to them. In our work we seek to isolate the effect of unconscious bias by randomly exposing individuals to different experiences.

The starting point for our work is a recognition that different individuals will naturally face different opportunities and incentives because of their backgrounds. For instance, a white male from a rich family has a very different starting point in life to a black female from a poor neighbourhood. As individuals go through their lives they learn from personal experience the rewards for their effort (at school, the workplace etc.). This experience may shape the way they view others and influence the decisions they make (as manager, colleague etc.). To capture this, our model has two stages. An individual, whom we will call Alice, is randomly allocated a type that determines her returns from effort. Alice does not know own type. In the *learning stage* Alice engages in a series of decision problems where she exerts effort in search of bonuses. This gives Alice an opportunity to learn her type. In the subsequent *evaluation stage* Alice is presented with the bonuses received by another individual, whom we call Bob, and asked to estimate Bob's effort. Alice is rewarded for correctly estimating Bob's average effort. In interpretation we think of Alice as being a manager looking to appraise the past performance of Bob. Our main research question is whether the own experience of Alice biases in predictable ways her estimate of Bob's effort.

We consider a transparent, dichotomous setting in which Alice (similarly Bob) is one of two types. For one type, called Effort Matters (EM), effort increases the probability that Alice will receive a bonus. For the other type, called Luck Matters (LM), the probability of a bonus is independent of effort.

Alice is not told her type but has an opportunity in the learning stage to learn whether effort pays for her *personally*. We derive a theoretical benchmark based on Bayesian updating and rational behavior by Bob and Alice. We then consider what happens if Alice’s beliefs are biased away from this benchmark by her own experience. Our theoretical results lead to testable hypotheses: If Alice is type EM then she will be systematically biased towards over-estimating the effort of Bob but this bias will be *small*. By contrast, if she is type LM she will be systematically biased towards under-estimating the effort of Bob and this bias will be *large*. The predicted asymmetry between EM and LM types is novel and, as we shall discuss, distinguishes own-experience bias from related concepts such as the (false) consensus effect.

We test and support the predictions of the model through an experiment. The experiment, like the model, consists of two stages. In the learning stage subjects have the opportunity, over 20 periods, to earn bonuses, where the probability of a bonus depends on type and choice of the subject. In the evaluation stage of the experiment subjects are shown the sequence of bonuses received by 12 other subjects and asked to estimate the average effort choice made by each of these subjects. Our experimental results support our main hypotheses: subjects exogenously assigned type EM over-estimate the average effort of others and, on average, the bias is small; subjects assigned type LM under-estimate the average effort of others and the bias is relatively large and statistically significant.

Our results support the conclusion that those for whom luck matters are more biased in their judgment of others than those for whom effort matters. Consider, for example, an academic who believes, through her own experience, that paper acceptance depends on ‘luck’ or connections and positions obtained by chance. This person may not value the amount of effort that others put into getting papers published. Our results suggest you may not want such an academic as department chair. It is worth highlighting, however, that subjects appeared on average to be good at discerning the effort choices others had made. In particular, there was a strong correlation between estimates of effort and the number of bonuses an employee received. This suggests that subjects were relatively sophisticated in their judgments. Moreover, we do not rule out the possibility that there could be settings (different to those considered here) where those for whom luck matters are less biased. It is the potential for own-experience to cause asymmetric bias that motivates our work.

That judgment of others could be biased by own experience seems natural

and has significant indirect support in the psychology literature. Extensive evidence shows that people are biased in overweighting own experience because of, for example, hindsight bias (e.g. Christensen-Szalanski and Willham, 1991; Hoffrage, Hertwig and Gigerenzer, 2000), overconfidence (e.g. Dunning et al. 1990), the availability heuristic (e.g. Tversky and Kahneman, 1973, 1974), and anchoring (Furnham and Boo, 2011). Moreover, a growing literature shows evidence of projection bias, a bias towards assuming that others are similar to oneself (e.g. Ross, Greene and House, 1977; Mullen et al. 1985; Krueger and Clement, 1994; Breitmoser 2019). Our work introduces a new theoretical model of own experience bias and confirms that such biases can be developed in a laboratory setting.

To appreciate the wider contribution of our paper, consider the extensive literature, field and experimental, on discrimination in labor markets (e.g. Bertrand and Mullainathan, 2004; Anderson, Fryer and Holt, 2006; Mobius and Rosenblat, 2006; Carlsson and Rooth, 2007; Pager, Western and Bonikowski, 2009; Charness and Kuhn, 2011; Fang and Moro, 2011; Arceo-Gomez and Campos-Vazquez, 2014; Azmat and Petrongolo, 2014). Our work is distinguished from this literature in two respects - the type of bias we study and the way we study it. We explain each in turn:

Type of bias. We consider the extent to which individuals are biased by own experience, a question of fundamental importance in understanding labor markets. This type of bias that has been largely overlooked and motivates our research. The existing literature focuses on a manager differentiating *between* candidates based on *observable* characteristics; for instance, a manager may, consciously or unconsciously, expect the productivity of a male to be higher than that of a female.¹ In our work we look at whether a manager is systematically biased in evaluating *any and all* candidates she observes; for instance, a manager may overestimate the productivity of both male and female workers (Bordalo et al. 2019).

Experimental approach. A crucial feature of our approach is that bias is created entirely within the lab through the random allocation of type and subsequent experience. By contrast, much of the work on discrimination looks at traits and characteristics bought from outside the lab, whether trust, cooperation, ability, race or gender (Castillo and Petrie, 2010; Reuben,

¹Her expectation may be consistent with optimal behavior given her own experience. The literature has looked to distinguish between statistical, taste-based, and implicit discrimination (e.g. Bertrand, Chugh and Mullainathan, 2005; Guryan and Charles, 2013; Ewens, Tomlin and Wang 2014).

Sapienza and Zingales, 2014; Bordalo et al., 2019). Our approach creates a ‘self-contained’ environment with which to control subject experience, allowing us to accurately measure bias.²

We proceed as follows. The model is described in Section 2. In Section 3 we provide our theoretical results and in Section 4 we report our experiment results. In Section 5 we conclude. Additional material is provided in an Appendix and Supplementary Material.

2 The model

We describe the learning and evaluation stage in turn.

2.1 Learning stage

We take as given a set $\{1, \dots, T\}$ of time periods and a set $E \subset [E_L, E_H]$, $0 < E_L < E_H$, of actions. We assume $E_L, E_H \in E$. We interpret an action as an effort level. In each period Alice must choose an effort level from E . Let $e_t \in E$ denote the effort level chosen in period t . Having chosen e_t , Alice learns whether she will receive a payoff bonus $B > 0$. Let $b_t \in \{0, 1\}$ record whether she received a bonus in period t , where $b_t = 1$ indicates she received a bonus and $b_t = 0$ otherwise. The payoff to Alice in period t is given by

$$u_t = Bb_t - e_t,$$

and consists of her bonus, if any, minus the cost of effort. The total payoff of Alice over the T periods is $u = \sum_{t \in T} u_t$.

Alice is one of two types: (1) effort matters (EM) or (2) only luck matters (LM). Each type is equally likely (and independent of the type of others). For a type EM individual the probability of receiving the bonus is linearly increasing in the amount of effort according to the formula

$$\pi^{EM}(e) = \alpha e - \beta$$

for all $e \in E$, where $\alpha > 0$. In order to guarantee $\pi^{EM}(e) \in (0, 1)$, we assume $\beta \in (\max\{0, \alpha E_H - 1\}, \alpha E_L)$. Thus, for a type EM individual, the higher her effort the higher the probability she will receive the bonus. For a type

²More ‘neutral, self-contained’ lab environments are typically reserved for studying statistical discrimination (e.g. Dickinson and Oaxaca, 2014).

LM individual the probability of receiving the bonus is independent of her effort and is given by

$$\pi^{LM}(e) = \gamma$$

for all e , where $\gamma \in (0, 1)$.

In our experiment we set $E_L = 17$, $E_H = 33$, $\alpha = 0.04$, $\beta = \gamma = 0.5$ and $B = 50$. Thus, if Alice is of type EM her probability of receiving a bonus is 0.82 if she chooses effort level 33 and 0.18 if she chooses effort level 17. This implies her expected payoff is increasing in effort. By contrast, if Alice is type LM her probability of receiving the bonus is 0.5 independent of effort and so expected payoff is decreasing in effort. *If* Alice knew her type it would be relatively simple for her to discern optimal effort. We explore in Section 3 what Alice should do given that she does not know her type.

2.2 Evaluation stage

In the evaluation stage we can think of Alice as a *manager* looking at the bonus record, the CV, of the *employee* Bob. By this point both Alice and Bob have completed the learning stage. The learning stage was identical for both Bob and Alice in terms of fundamental parameters. But crucially it is highly unlikely they had the same individual experience. As we shall see, *own experience* is crucial to our approach.

Alice observes whether Bob obtained bonuses in periods \bar{t} to T for some $\bar{t} > 1$. We refer to the sequence $\{b_{\bar{t}}, \dots, b_T\}$ as the CV of Bob. Note that Alice does not observe Bob's effort in any period nor his bonuses before period \bar{t} . We shall ask what Alice infers about the effort of Bob and whether this inference depends on her own experience. We assess Alice's evaluation of Bob based on her prediction of Bob's average effort. To formalize this let \bar{e} denote the average effort of Bob (in periods \bar{t} up to T) and let g denote Alice's estimate of Bob's effort (in periods \bar{t} up to T). We set the payoff of Alice according to a threshold measure of accuracy. Specifically, she receives payoff $Q_k > 0$ if $|\bar{e} - g| \leq k$, for an exogenous threshold $k > 0$, and receives 0 otherwise. Note that this incentivizes Alice to say the effort level that she thinks it was most likely Bob chose.³ In measuring accuracy we will also discuss prediction error $\bar{e} - g$ and absolute prediction error $|\bar{e} - g|$.

³One could imagine more complex quadratic scoring rules for rewarding accuracy. The threshold measure has the advantage of simplicity, particularly for experimental subjects. To illustrate, suppose Alice thinks there was a 40% chance Bob's average effort was 17 and a 60% chance it was 33. With our threshold rule she should choose 33 (irrespective

Let us highlight that there is no moral sense in which high effort is ‘better’ than low effort. Bob is simply responding to exogenously given incentives. Carrying this logic forward we should not think of Alice as wanting to reward employees who chose high effort in the past. Her task is to correctly assess Bob’s past effort and she is rewarded for doing so.

3 Theoretical analysis

In this section we outline some theoretical benchmarks around which to judge the behaviour and inferences of Alice. Recall that our objective is to see how Alice’s judgment of others (in the evaluation stage) may be biased her own-experience (in the learning stage). We, therefore, focus on two basic issues: (1) Alice’s experience in the learning stage and (2) how this experience may bias her estimate of Bob’s effort. We consider these two issues in turn.

3.1 Learning stage

Recall that Alice does not know her type. Let $p_0 \in (0, 1)$ denote the probability that she initially assigns to being type EM. A strategy determines an action in each period t for any plausible history $h_t = (e_1, b_1; \dots; e_t, b_t)$. Once we know history $h_T = (e_1, b_1; \dots; e_T, b_T)$ the total payoff $u = \sum_{t \in T} u_t$ is determined. With a slight abuse of notation let $u(h_T)$ denote total payoff given history $h_T \in H_T$. For any strategy s and initial belief p_0 it is possible to determine the probability distribution over the set of outcomes. Let $\Pr(h_T | s, p_0)$ denote the probability of outcome h_T given strategy s and initial belief p_0 . The objective of Alice in the learning stage is to choose a strategy s that maximizes expected total payoff,⁴

$$U(s) = \sum_{h_T \in H_T} \Pr(h_T | s, p_0) u(h_T). \tag{1}$$

Through experience Alice can update her beliefs on own type. Let p_t denote the probability she assigns to being type EM given history h_t . It is

of risk aversion). With a quadratic scoring rule it may be optimal to choose 33 or some intermediate number (e.g. 26) depending on the specifics of the rule and risk aversion.

⁴The set H_T need not be finite but, for any strategy s and initial belief p_0 , the set of outcomes that can occur with positive probability is finite.

optimal for Alice to update her beliefs over time using Bayes rule.⁵ Doing this, $p_t > p_{t-1}$ if and only if

$$b_t(\alpha e_t - \beta) + (1 - b_t)(1 + \beta - \alpha e_t) > b_t\gamma + (1 - b_t)(1 - \gamma).$$

This yields an important cut off effort level

$$e^* := \frac{\beta + \gamma}{\alpha}.$$

Alice optimally increases the probability she assigns to being type EM if and only if either (1) $e_t > e^*$ and $b_t = 1$ or (2) $e_t < e^*$ and $b_t = 0$. Thus, if Alice puts in high effort and is successful then, according to Bayes rule, she increases the probability she assigns to being of type EM. If she puts in low effort and is unsuccessful then she similarly increases the probability she assigns to being type EM. Otherwise she increases the probability she assigns to being type LM.

3.1.1 Optimal strategy in the learning stage

In maximizing expected total payoff Alice, in period t , has to trade off two potentially competing objectives: (1) maximize expected payoff in period t , and (2) maximize information about type to enable a more informed choice in periods $t+1, t+2, \dots, T$. In general, there may be a conflict between these two objectives.⁶ To make the problem more tractable (and our experiment design more transparent) we chose parameter values so that there is no conflict between the two objectives. The optimal strategy for Alice is then to simply maximize expected payoff each period. This is captured in our first result, proved in an appendix.

Proposition 1: Suppose that $\frac{\beta + \gamma}{\alpha} = \frac{E_H + E_L}{2}$ and $\gamma = \frac{1}{2}$. Then Alice can maximize expected total payoff in the learning stage by choosing, in any period $t \in \{1, \dots, T\}$, action E_H if $p_{t-1}\alpha B \geq 1$ and E_L if $p_{t-1}\alpha B < 1$.

⁵If $\Pr(b_t|\star)$ denotes the probability of b_t given Alice is type \star then

$$p_t = \frac{p_{t-1} \Pr(b_t|type EM)}{p_{t-1} \Pr(b_t|type EM) + (1 - p_{t-1}) \Pr(b_t|type LM)}.$$

Where $\Pr(b_t|type EM) = b_t(\alpha e_t - \beta) + (1 - b_t)(1 + \beta - \alpha e_t)$ and $\Pr(b_t|type LM) = b_t\gamma + (1 - b_t)(1 - \gamma)$.

⁶For instance, it could be that Alice should choose E_L to maximize her expected period 1 payoff but choose E_H in order to better learn her type and maximize expected payoff in future periods.

To illustrate Proposition 1 consider the parameter values we use in our experiment ($\beta = \gamma = 0.5, \alpha = 0.04, E_L = 17, E_H = 33$). The optimal strategy dictates Alice choose E_L if $p_{t-1} < 0.5$ and E_H if $p_{t-1} > 0.5$. If Alice uses the optimal strategy then expected effort relatively quickly converges on the ‘correct’ effort level for type (see Supplementary Material and Figure 1 to follow). That is, effort converges towards 33 if Alice is type EM and 17 if type LM.

3.1.2 Extremeness aversion

Abundant evidence suggests that many people are averse to choosing extreme options (see Neumann, Bckenholt and Sinha (2016) for a review). We, therefore, explore what happens if Alice deviates from the optimal strategy by choosing effort levels other than the extremes E_H and E_L . We begin by introducing the following definition. Alice’s strategy is said to be *belief responsive* if in any period $t \in \{1, \dots, T\}$, she chooses an action $e_t \geq e^*$ if $p_{t-1}\alpha B \geq 1$ and action $e_t \leq e^*$ if $p_{t-1}\alpha B < 1$. If a strategy is belief responsive then Alice is learning from past experience whether to choose ‘high’ or ‘low’ effort.

To illustrate consider a belief responsive strategy in which Alice chooses $E_H - y$ and $E_L + y$ in any period t , for some value $y \geq 0$. If $y = 0$ we have the optimal strategy. The larger is y then the higher the extent of extremeness aversion. Figure 1 plots the proportion of periods in which Alice chooses effort on the appropriate side of e^* for her type, i.e. $E_H - y$ if type EM and $E_L + y$ if type LM. We give the proportion over all periods and from period \bar{t} onwards. Provided y is not too large, which is a reasonable assumption, the proportion of choices on the appropriate side of e^* remains relatively high.⁷

3.1.3 Summary from learning stage

The optimal strategy in the learning stage requires Alice to update her beliefs using Bayes rule and to choose ‘extreme’ effort levels E_H and E_L . In reality, and in the lab, neither of these is likely to hold. We know that individuals

⁷Alice will see a significant loss in payoff from extremeness aversion. To illustrate consider $y = 3$. If Alice is type LM then choosing 20 costs her 3 more than choosing 17. Similarly, if of type EM the probability of getting the bonus drops from 0.82 to 0.7 if Alice chooses effort 30 rather than 33. This is a drop in expected bonus of 6 which is only partly compensated by spending 3 less on effort.

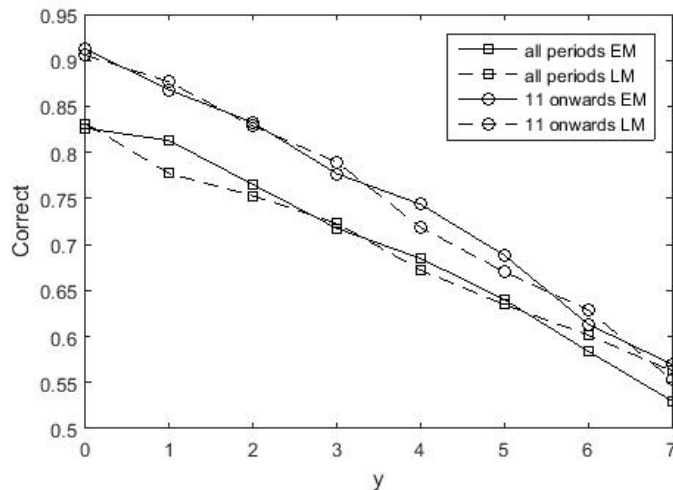


Figure 1: Proportion of choices that are ‘correctly’ above 25 for a type EM and below 25 for a type LM as a function of extremeness aversion (y).

deviate from Bayes updating and exhibit extremeness aversion. This, though, is not crucial to our model. The key feature driving our model is that Alice, by the time she reaches the evaluation stage, has a good idea, conscious or not, of her type and whether to choose high or low effort. Our analysis shows that if Alice behaves optimally then she quickly learns type. Extremeness aversion and failure to use Bayes updating will clearly slow learning, but we can still reasonably expect Alice to have a good idea of whether it is best for her to choose high or low effort by the time she reaches the evaluation stage.

3.2 Evaluation stage

Since Alice does not observe Bob’s bonuses between period one and period $\bar{t} - 1$ she cannot recreate Bob’s learning process. She can still, however, infer something about his type and his effort from his CV. In this section we derive a benchmark threshold based on Alice having an unbiased prior of the type of Bob and a correct estimate of Bob’s extremeness aversion. We then subsequently introduce own experience bias and explore the consequences of extremeness aversion.

3.2.1 Benchmark threshold

Suppose that if Bob is type EM then he chooses E_h in each period $t > \bar{t}$ where $e^* < E_h \leq E_H$. Also suppose that if he is type LM he chooses E_l in each period, where no assumptions are imposed on E_l . Then if Bob is of type EM the likelihood of obtaining the specific sequence of bonuses in his CV is⁸

$$\Pr(CV|EM) = \prod_{t=\bar{t}}^T (\alpha E_h - \beta)^{b_t} (1 + \beta - \alpha E_h)^{1-b_t} = (\alpha E_h - \beta)^X (1 + \beta - \alpha E_h)^{M-X}.$$

If he is of type LM, then irrespective of the effort he chooses, the likelihood of his CV is

$$\Pr(CV|LM) = \prod_{t=\bar{t}}^T \gamma^{b_t} (1 - \gamma)^{1-b_t} = \gamma^X (1 - \gamma)^{M-X}.$$

Note that the above expressions only depend on the number of bonuses, X , and not the full CV, $\{b_{\bar{t}}, \dots, b_T\}$. In the following we, therefore, focus on X .

Recall that Q_k is Alice's payoff if she correctly estimates Bob's effort within k . So, Alice's expected payoff if she estimates Bob's effort as $g = E_h$ is $Q_k \Pr(EM|X)$ and her payoff if she estimates $g = E_l$ is $Q_k \Pr(LM|X)$. So, if $\Pr(EM|X) > 0.5$ it is optimal to infer Bob is type EM and estimate $g = E_h$.⁹ If $\Pr(EM|X) < 0.5$ then it is optimal to infer Bob is type LM and estimate $g = E_l$. Suppose that Alice has prior belief p_A that Bob is type EM. Then we get $\Pr(EM|X) > 0.5$ if and only if $X > X^*(p_A)$ where¹⁰

$$X^*(p_A, E_h) := \frac{M (\ln(1 - \gamma) - \ln(1 + \beta - \alpha E_h)) + \ln(1 - p_A) - \ln(p_A)}{\ln(1 - \gamma) - \ln(1 + \beta - \alpha E_h) + \ln(\alpha E_h - \beta) - \ln \gamma}. \quad (2)$$

⁸Recall that Alice observes the sequence of bonuses in each period. If she only observed that the bonus was given X times, the likelihood would be slightly different because there are multiple ways of getting X bonuses.

⁹Given the margin of error k it is optimal to estimate within k of e_h .

¹⁰We use Bayes rule

$$\Pr(EM|X) = \frac{p_A \Pr(X|EM)}{p_A \Pr(X|EM) + (1 - p_A) \Pr(X|LM)} = \frac{1}{2}.$$

If Alice has an unbiased prior then $p_A = p_0 = 1/2$. From this we suggest the following rule for inferring type and choice.

Benchmark Threshold: If $X > X^*(p_0, E_h)$ Alice infers that Bob was of type EM and chose effort E_h in each period. If $X \leq X^*(p_0, E_h)$ Alice infers that Bob was of type LM and chose effort E_l in each period.

The benchmark threshold provides a relatively simple way to estimate Bob’s effort. For instance, for the parameter values of our experiment we obtain the benchmark threshold of $X^*(p_0, E_H) = 6.7$. So, an employee with 7 or more bonuses would be inferred to be a type EM who chose E_H in each period and an employee with 6 or less bonuses a type LM who chose E_L in each period. The error from using this threshold can be decomposed into two main portions: (1) Bob gets more than X^* bonuses despite being type LM and choosing low effort. In this case Alice will, on average, overestimate the effort of Bob. (2) Bob gets less than X^* bonuses despite being type EM and choosing high effort. In this case Alice will, on average, underestimate the effort of Bob.

The top row of Table 1 breaks down the resulting error for the parameter values in our experiment (and an assumption Bob uses the optimal strategy in the learning stage). With an unbiased prior of 0.5 the average prediction error, $\bar{e} - g$, if Bob gets more than X^* bonuses is -2.55 ; Alice tends to overestimate effort. The average prediction error if Bob gets less than X^* bonuses is 2.99 ; Alice tends to underestimate effort. Given that she is almost equally likely to overestimate as underestimate effort the overall predicted error is only 0.3 . The average absolute error, $|\bar{e} - g|$, is 2.81 and 76% of predictions are within 2 of actual effort. As well as being simple the benchmark threshold proves to be relatively accurate in minimizing Alice’s error.¹¹ Moreover we demonstrate in the supplementary material that this error is robust to Bob not knowing his type by period \bar{t} .

¹¹In particular, while it is possible to come up with more complex strategies to infer effort these alternative strategies hardly improve accuracy. With the parameter values we have in our experiment, the benchmark threshold gives us a 76% accuracy in predicting effort compared to a theoretical maximum of 78% accuracy. See the Supplementary Material for more information.

3.2.2 Own experience bias

The benchmark threshold is based on Alice having an unbiased prior belief $p_A = p_0$ on the probability that Bob is of type EM. In the introduction we gave a list of reasons why a person’s own experience may unconsciously influence her beliefs (availability heuristic, projection bias etc.). Here we assume that this effect of own experience manifests itself in prior beliefs on type. In particular, we consider the possibility Alice exhibits *own experience bias* and overweights the probability others are the same type as she is.¹² We capture this as deviations in p_A from p_0 influenced by beliefs, p_T , about herself.

In our context, own experience bias can be modelled as Alice using a threshold model but with a biased prior.

Own Experience Bias: If Alice is type EM she has a biased prior $p_A > p_0$ and if type LM she has biased prior $p_A < p_0$. In either case, if $X > X^*(p_A, E_h)$ Alice infers that Bob is of type EM and chose effort E_h in each period. If $X \leq X^*(p_A, E_l)$ Alice infers that Bob is of type LM and chose effort E_l in each period.

Own experience bias (as we have defined it) says that Alice behaves in accordance with the benchmark threshold except *she starts with a biased prior*. Her prior is more biased the further it deviates from p_0 . Our next result shows that own experience bias leads to fundamental changes in Alice’s estimate of the effort of Bob. To formalize this let $\bar{e}(b, p_A)$ denote Alice’s estimate of Bob’s effort using the benchmark threshold with prior p_A when Bob has CV of b . Note that estimates with the benchmark threshold, $\bar{e}(b, p_0)$, provide our key benchmark for comparison.

Proposition 2: Suppose Bob used a belief responsive strategy in the learning stage and Alice exhibits own experience bias. (1) If $p_A > p_0$ then Alice is biased towards over-estimating the effort of Bob, $\bar{e}(b, p_A) \geq \bar{e}(b, p_0)$. (2) If $p_A < p_0$ then Alice is biased towards under-estimating the effort of Bob, $\bar{e}(b, p_A) \leq \bar{e}(b, p_0)$.

In interpretation, Proposition 2 means that if Alice is type EM and suffers from own experience bias she will have a tendency to overestimate the effort

¹²This can be related to the law of small numbers, where a small sample is overweighted in importance (Rabin 2002). Connections with the false-consensus effect will be discussed more later.

of Bob. If she is type LM then she will have a tendency to underestimate the effort of Bob.

3.2.3 Asymmetric bias

To explore the consequences of bias let us return to Table 1. Here we first compare error with type EM bias ($p_A = 0.7, 0.9$) with a similar type LM bias ($p_A = 0.3, 0.1$). If we focus on overall error we observe a notable asymmetry in which type LM bias leads to a larger error than type EM bias. For instance, when $p_A = 0.9$ the error is -1.9 while when $p_A = 0.1$ it is 4.84 . To give a broader picture we consider what happens if prior beliefs of a type EM are uniformly distributed between p_0 and 1 (EM uniform) and those of a type LM between 0 and p_0 (LM uniform). Additionally, we consider what happens if prior beliefs have a triangular distribution with peak at 0.5 (EM peak and LM peak).¹³ Again, we see a notable asymmetry in which overall error is smaller with type EM bias than type LM bias.

To appreciate why we observe this asymmetry consider the relatively extreme priors of 0.1 and 0.9. With a prior of 0.1 Alice estimates that if Bob received 8 or less bonuses then he chose low effort. In reality, even if Bob is type EM and chose high effort there is a good chance he will get 8 or less bonuses. Alice, therefore, has a ‘one-sided’ tendency to underestimate the effort of Bob. This results in a relatively large overall error of 4.84 . With a prior of 0.9 Alice estimates that if Bob received 6 or more bonuses he chose high effort. This results in her overestimating the effort of those with 6 or more bonuses. That, however, is offset by the tendency to underestimate the effort of those with 5 or less bonuses. So, while Alice is less accurate in her predictions (compared to the benchmark threshold) the overall error is only -1.9 .

In summary, both EM and LM bias lead to a similar divergence from the benchmark. It is in the consequences of this shift in the benchmark we see the asymmetry. With LM bias, Alice uses a higher threshold which Bob (whatever his type) is less likely to reach. So, she has a one-sided tendency to underestimate Bob’s effort, resulting in a high error in both absolute and overall terms. By contrast, with EM bias, Alice uses a lower threshold which Bob may or may not reach (whatever his type). EM bias, therefore,

¹³So, for a type EM the distribution is proportional to $1 - p_A$ and for a type LM proportional to p_A .

Prior	X^*	$\Pr(X > X^*)$	Error		Overall	Absolute	$k = 2$ % correct
			$X > X^*$	$X < X^*$			
0.5	6.71	0.48	-2.55	2.99	0.30	2.81	76
0.7	6.18	0.48	-2.55	2.99	0.30	2.81	76
0.3	7.30	0.35	-1.08	4.40	2.49	3.21	74
0.9	5.29	0.62	-4.51	2.39	-1.9	3.68	71
0.1	8.19	0.20	-0.36	6.15	4.84	4.99	63
EM uniform		0.57	-3.70	2.68	-1.02	3.41	73
LM uniform		0.33	-1.14	4.65	2.80	3.67	71
EM peak		0.52	-3.08	2.86	-0.27	3.05	75
LM peak		0.38	-1.51	4.09	1.99	3.26	74

Table 1: For different priors the probability Bob gets more than the threshold number of bonuses, the prediction error if Bob has more bonuses than the threshold, less bonuses than the threshold, and overall error.

essentially adds noise without any strong tendency to under or overestimate effort. This results in a large absolute error but relatively small overall error.

The preceding argument allows us to predict when the asymmetry between EM and LM bias will be small (or potentially reversed). Specifically, the smaller the gap between E_l and E_h the more likely it is that Bob will, on average, choose effort around e^* (because the difference between γ and $\alpha E_h - \beta$ goes to 0 as E_h is close to e^*). Consequently, the LM bias from inferring Bob chose low effort will be larger if E_l is far from e^* (e.g., for our parameters, $E_l = 17$ and $E_h = 27$, with $e^* = 25$) and smaller if E_l is close to e^* (e.g. $E_l = 23$ and $E_h = 33$). In the supplementary material we provide results consistent with this prediction. In doing so we demonstrate that, for our model, a notable asymmetry between EM and LM exists unless E_l is very close to e^* .

The predicted asymmetry between EM and LM types allows us to more generally pinpoint the difference between our concept of experience bias and the false-consensus effect (or anchoring effect) that is familiar in the literature. The false-consensus effect maps own *action* to biased estimates about others actions (see, for example, Engelmann and Strobel 2012 for a definition). In our setting this would equate to Alice estimating that Bob chose a similar effort level to her. So, if Alice chose effort level E_h in every period then she would estimate a higher effort level for Bob than if she chose effort

level E_l in every period. Own experience bias, by contrast, maps own *type* to a biased belief about others types. So, if Alice is type EM she overestimates the probability Bob is type EM and if she is type LM she overestimates the probability Bob is type LM. This then *indirectly* influences her estimate of Bob’s effort.

Two practical aspects of the differences between the false consensus effect and own experience bias are: (1) There is no reason, a-priori, to expect any asymmetry in terms of false-consensus effect. Type EM’s should be as biased upwards as type LMs are downwards. As we have seen, however, own experience bias (which is symmetrical in terms of type) can lead to asymmetry in which one type is more biased than another in estimating effort. (2) The false-consensus effect has nothing to say on how the manager should interpret the CV of the employee. It simply says that Alice will estimate Bob chose ‘like her’. By contrast, own experience bias influences the way the manager interprets the CV.

3.2.4 Extremeness aversion and own-experience bias

In the above discussion we took it as given that Alice knows E_h .¹⁴ This allowed us to focus on the bias that comes from $p_A \neq p_0$. Here we consider what happens if Alice does not know E_h . Let E_h^A denote Alice’s prediction of E_h .¹⁵ We compare estimated effort if Alice uses her threshold $X^*(p_A, E_h^A)$ with that obtained using the benchmark threshold $X^*(p_0, E_h)$.

Proposition 3: If $\gamma = 0.5$ then the value of $X^*(p_A, E_h^A)$ is weakly increasing in E_h^A and weakly decreasing in p_A .

The larger is E_h^A relative to E_h the higher will be Alice’s threshold for estimating Bob is type EM. If this were the case it could partly compensate for type EM own experience bias, $p_A > p_0$. To illustrate the size of this effect Table 2 details the average error ($\bar{e} - g$) and absolute error ($|\bar{e} - g|$) if Alice sets $E_h^A = 33$ and Bob uses either the optimal strategy (this is the same as Table 1), a belief responsive strategy with effort 20 and 30 or 23 and 27. We can see that average error is higher with LM bias while lower with EM bias. Absolute error is similar with EM and LM bias. This reinforces the idea that LM bias leads to a systematic ‘one-sided’ tendency to underestimate

¹⁴For this discussion the value of E_l is not relevant.

¹⁵We assume that this prediction is not influenced by the CV of Bob.

Bob	Alice	Error	Absolute
17 or 33	Benchmark	0.30	2.82
	EM bias	-1.30	3.60
	LM bias	3.01	3.87
20 or 30	Benchmark	2.73	6.15
	EM bias	0.62	6.27
	LM bias	5.45	6.90
23 or 27	Benchmark	4.75	7.69
	EM bias	2.34	7.68
	LM bias	6.74	7.83

Table 2: Accuracy of inferences with the benchmark threshold and own experience bias with a uniformly distributed prior

effort. In contrast, suppose that $E_h^A < E_h$. Then in this case it could partly compensate for type LM own experience bias, $p_A < p_0$.

3.2.5 Summary of evaluation stage

We have focused on Alice using a threshold strategy in which she infers that Bob chose high effort if he received more than a threshold level of bonuses. A threshold strategy is intuitive and simple and can be close to optimal if the threshold is set at the right level. We modelled own experience bias as Alice having a biased prior on the probability that Bob is type EM. This, in turn, influences the threshold Alice uses. In particular, we have suggested that if Alice is type LM then she may have a too high threshold because ‘in her experience’ bonuses were due to luck. Similarly, if she is type EM then she may have a too low threshold because in her experience effort was rewarded. This is despite incentives to make a correct estimate of Bob’s effort. The key prediction that comes out of our model is that Alice will, on average, be less biased in her estimates of Bob’s effort if she is type EM than if she is type LM.

4 Experiment

We performed an experiment that recreates the decision making environment described. The experiment consisted of two distinct parts, with the instruc-

tions for Part 2 not given until after Part 1 was completed. The experiment instructions are available in Supplementary Material. Here we detail the most salient features. Note that the language used in the experiment was deliberately neutral and so made no mention of effort or employment etc. In each session we arranged for a near equal split (within 1) of subjects into type EM and LM. A subject was told they had a 50-50 chance of being each type.¹⁶

In Part 1 of the experiment subjects were exposed to the decision making task detailed in Section 2, with parameters $E = \{17, 18, \dots, 33\}$, $\alpha = 0.04$, $\beta = \gamma = 0.5$ and $T = 20$. In each period subjects were able to see the complete history of their effort levels and bonuses. In Part 2 of the experiment each subject played the role of a manager estimating the effort level of employees. They were each shown the sequence of bonuses received by an employee in the last 10 periods, i.e. $\bar{t} = 11$. We shall continue to call this the CV of the employee. The subject was then asked to estimate the average effort level of the employee over these 10 periods. A subject saw a total of 12 employees CVs, one at a time. Let us emphasize that each employee corresponded to a subject randomly chosen (without replacement) from that particular experimental session.¹⁷

Subjects were incentivized in both Parts 1 and 2 of the experiment. In Part 1 the payoff accumulated over the 20 periods was converted into money. Subjects were given an endowment that more than covered any potential losses they could make during the experiment. In Part 2 a subject was paid £0.50 for every employee whose average effort was within two of their estimate of that effort (i.e. we used threshold accuracy $k = 2$). A total of 155 subjects took part in the experiment spread relatively evenly over 10 sessions. There were 77 subjects of type EM and 78 of type LM. Subjects were recruited from across the student population of the University of Kent. The experiment took place in a computer lab using z-Tree (Fischbacher 2007). A typical session lasted around 45 minutes with an average payment of around

¹⁶Subjects were told ‘Before the first round begins the computer will randomly decide whether you are of a type X or type Y. More on what this means later. You have an equal chance of being type X and type Y. You will not be told what type you are.’

¹⁷This means that different subjects were exposed to different CVs. There are disadvantages to this approach; for instance, a subject may be exposed to 12 similar CVs and so it becomes difficult for us to identify any bias. We believe, however, that the genuine randomness in our approach avoids any form of experimenter effect or bias that could come from selecting CVs in some way.

£10.50.

4.1 Experimental hypotheses

Based on our theoretical analysis we make the following hypotheses. The first hypothesis says that managers will be biased by own type and the second that the bias, as measured by average error, will be larger for those of type LM.

Hypothesis 1: The estimate of employee effort is, *ceteris paribus*, higher if the manager is of type EM than if the manager is of type LM.

Hypothesis 2: Managers of type LM have a larger average error in estimating employee effort than managers of type EM.

Hypotheses 1 and 2 are stated in terms of own type. Alternatively they could have been stated in terms of inferred type, where inferred type takes into account the actual experience of the subject in Part 1 of the experiment. In the following we focus on own type because this is exogenously assigned. In the Supplementary Material we show that our results are robust to using inferred type.

We would like to highlight that, whether we use own type or inferred type, the focus is on the manager unconsciously expecting others to face the same *incentive structure* as she does. This is different than expecting others to have, say, the same luck as she had in obtaining bonuses. Our approach, therefore, works from the idea that managers are not biased (controlling for type) by factors such as the number of bonuses they themselves received. We shall use this observation in the following to provide a further robustness check of our results.

4.2 Learning stage

Our focus in the experiment is on the evaluation stage and whether subjects are biased by own type. Given, however, that our predictions are based on the manager using an optimal or belief responsive strategy we briefly look at behaviour in the first part of the experiment. Figure 2 plots the distribution of choices in round 11 (where the CV initiates).¹⁸ You can see that around a

¹⁸The picture is similar for subsequent periods.

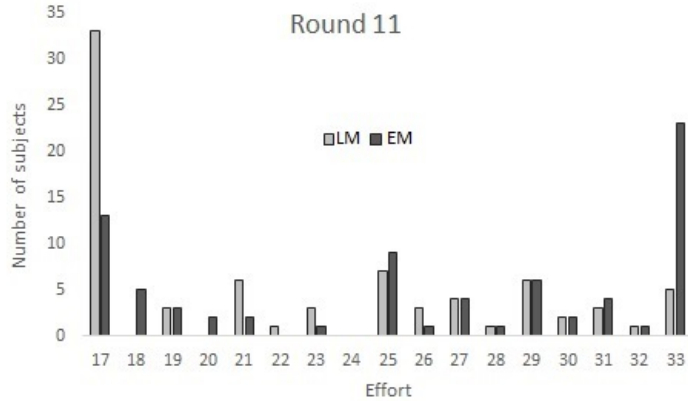


Figure 2: Distribution of choices in round 11 by type.

third of subjects chose appropriate effort levels for their type, with 17 and 33 the modal choices. Even so, there is evidence of extremeness aversion with a significant proportion of choices in the intermediate range.

To look in more detail at whether subjects behaved optimally we calculated deviation from the effort level dictated by the optimal strategy. This is calculated consistent with Bayesian updating and the actual experience of the subject. Figure 3 plots total deviation (summed over periods 11 to 20). For comparison we provide the distribution one would expect with random choice. A total of 16 (out of 155) subjects choose the optimal strategy. A large proportion of subjects are ‘close’ to the optimal strategy in that the deviation from optimal is relatively small. We estimate (see the Supplementary Material for more information) that around a third to two thirds of subjects (depending on how we measure ‘closeness’) behaved consistent with a belief responsive strategy.

The crucial feature from our perspective is that subjects responses to their incentives led to a high correlation of effort levels, type and bonuses. Indeed, there was a clear threshold in the data consistent with our theoretical model. In particular it was optimal to infer a subject with less than 6 bonuses was type LM (for instance, 62% of subjects with 5 bonuses were type LM) and to infer that a subject with 6 or more bonuses was type EM (for instance, 70% of subjects with 6 bonuses were type EM).

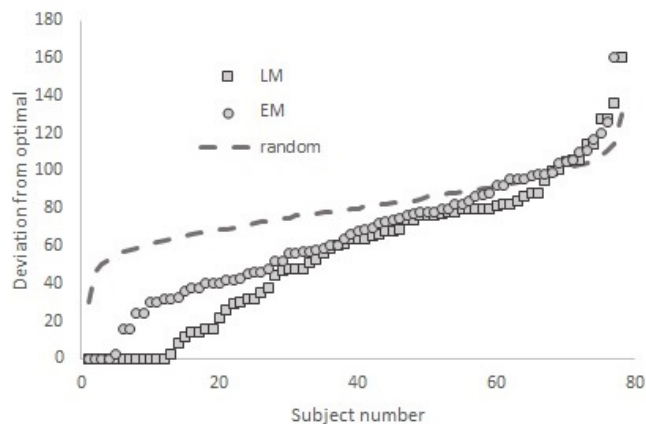


Figure 3: Distribution of absolute deviation from optimal strategy in Part 1 between periods 11 and 20.

4.3 Estimates are own-experience biased

Table 3 provides data for own effort (averaged over the last 10 periods), the estimated effort of others (averaged over the 12 employees seen) and four measures of estimate accuracy. We see the effort of Alice, her estimate of Bob’s effort and the accuracy of this estimate, distinguishing by the type of Alice. Recall that different managers were exposed to different employees. The numbers in Table 3 (and the analysis to follow) represent the actual error between estimate and effort for each manager, employee pair.

Consistent with the optimal strategy subjects of type EM choose significantly higher effort than those of type LM ($p < 0.001$ for own type and inferred type using a two-sided Mann-Whitney test of average effort with each subject as a unit of observation). Our main concern is subjects’ estimates of the efforts of others. Because own type is uncorrelated with the types of others these estimates should *not* depend on own type. Consistent, however, with Hypothesis 1 we see that the average estimate of type EM’s is higher than that of type LM’s ($p = 0.026, 0.023$ for own and inferred type respectively using two-sided Mann Whitney test). The average estimate of type EM managers was around 0.7 higher than that of type LM managers. This compares to a difference in own effort of around 3.5. Hence the size of bias is approximately one fifth of the difference in own effort.

	Own type	
	LM	EM
Own effort	23.3	26.8
Estimate of others effort	23.9	24.6
Prediction error	0.82	-0.14
Absolute error	4.8	4.3
Within 1 (%)	18	18
Within 2 (%)	35	36

Table 3: Own effort and the average estimated effort of others based on own type and treatment.

The numbers in table 3 are directly comparable to those in Table 1 derived from our theoretical model. Consistent with Hypothesis 1 we find that those of type LM underestimate effort (the prediction error term is positive) while those of type EM overestimate effort. This difference is highly significant ($p = 0.008$, 0.002 two-sided Mann Whitney). Consistent with Hypothesis 2 we also find that those of type EM are more accurate in estimating effort than those of type LM. Consistent with Table 1, the absolute prediction error ($p = 0.052$, 0.025 two-sided Mann Whitney) and threshold accuracy ($p > 0.5$) show a much less marked difference.

4.4 The CV of the employee

We complete the analysis by considering how estimated effort is influenced by the CV of the employee. Figure 4 depicts the relationship between the number of bonuses an employee received and the average estimated effort by managers. Own experience bias is captured by the gap between the EM and LM estimates. Overall we can see a strong positive relationship between the number of bonuses received by the employee and estimated effort. At the individual level we also find that subjects were responsive to the CV of the employee. For instance, the average difference between a subject’s maximum and minimum estimate was 11.25 (compared to a theoretical maximum of 16). Also we observe a positive relationship between employee bonuses and estimated effort for 117 of the 155 subjects. (See the Supplementary Material for additional details).

We exploit the panel nature of the data set and report in Table 4 the

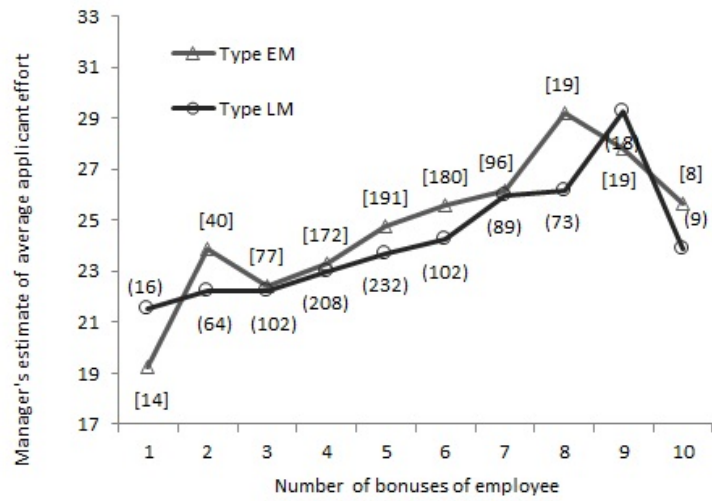


Figure 4: Manager's estimate of average effort of employee distinguishing between own type. The number of observations is given for type EM \square and LM \circ .

results of four random effects linear regressions with either estimated effort or estimate error as the dependent variable. For instance, we estimate $g_{in} = a_1 + a_2T_i + a_3X_n + \varepsilon_{in}$ where g_{in} is subject i 's estimate of n 's average effort, T_i is an indicator variable for whether i has type LM (the default is type EM), and X_n is the number of bonuses of subject n .

Consider first the results for estimated effort, which allow us to evaluate Hypothesis 1. We can see that estimated effort is significantly lower if the manager is type LM. The size of effect (-0.71 to -0.76) is similar to that resulting from the employee having one extra bonus on the CV (0.64).

Consider next the results for error estimate, which allow us to evaluate Hypothesis 2. Some care is needed here to take account of the negative coefficient for the bonuses of the employee. The vast majority of employees received between 4 and 7 bonuses. The predicted error of a type EM manager falls, therefore, in the range $1.33 - 4 \times 0.29 = 0.17$ to $1.33 - 7 \times 0.29 = -0.7$. We cannot reject the null hypothesis that the predicted error is zero when the number of bonuses is 4 ($p = 0.57$, F-test), 5 ($p = 0.60$) and 6 ($p = 0.12$) but is for 7 bonuses ($p = 0.017$). In interpretation this suggests type EM managers are generally unbiased but do overestimate the effort of employees with the best CVs. The predicted error of a type LM manager falls in the range 1.19 to 0.32. Here we can reject the null of zero error when the worker has 4 ($p < 0.001$), 5 ($p = 0.002$) and 6 bonuses ($p = 0.055$) but not 7 ($p = 0.44$). This suggests that type LM managers systematically underestimate the effort of employees. These results are further evidence in support of Hypothesis 2.

To provide an additional test of Hypothesis 2 we evaluate the null hypothesis that the predicted error is smaller for type LMs than EMs. We can reject this null if the worker has 4 bonuses ($p = 0.006$, F-test) and 5 bonuses ($p = 0.035$) but not 6 or 7 bonuses ($p > 0.4$). Given that 62% of the CVs observed by employers had 5 or less bonuses this evidence is again consistent with Hypothesis 2. In particular, over 12 CVs a type LM employer is predicted to have a larger average error than a type EM employer on over 98% of occasions (Monte Carlo simulations randomly drawing CVs).

To put things in perspective, compare an employee with 4 bonuses to one with 7 bonuses. Average estimates of effort for the employee with 4 bonuses are 23.0 and 23.3 for managers of type LM and EM respectively. For the employee with 7 bonuses the numbers are 26.0 and 26.2. The number of bonuses clearly seem to matter more than own type. We would argue, however, that this is not an excuse to ignore own experience bias. Indeed, that subjects took account of the worker's CV is evidence they understood

	Estimated effort		Error in estimate	
Constant	21.4*** (0.37)	22.0*** (0.85)	1.33*** (0.50)	0.50 (0.92)
Type LM	-0.71** (0.35)	-0.76** (0.36)	0.95** (0.37)	1.02*** (0.38)
Bonuses of employee	0.64*** (0.06)	0.64*** (0.06)	-0.29*** (0.08)	-0.29*** (0.07)
Bonuses of manager	—	-0.06 (0.07)	—	-0.08 (0.7)
Number of observations	1860	1860	1860	1860

Table 4: Results of random effects linear regressions with estimated employee effort as the dependent variable. Standard errors in brackets; ***, ** and * significant at 1%, 5% and 10% level respectively.

the decision task.

5 Conclusion

Depending on the characteristics of individuals and the prevailing norms within society, the relationship between effort and success in the workplace is likely to differ from one individual to another. To make our points we have considered a stark dichotomy between those for whom effort matters and those for whom it does not. More generally, we can think of each individual as having a personal payoff function (relating effort and expected payoff) that represents their earning potential. If there is discrimination or prejudice in labor markets then this will be reflected in the personal payoff function. For example, individuals whose parents are well educated may expect a high payoff with moderate effort and a low rate of increase of earnings as a function of effort. These are like our type LMs in that they do not need to make any special effort to succeed. By contrast, those individuals whose parents are not well educated may expect a low payoff with little effort and increasing earnings as a function of effort. These are like the type EMs in that success is highly dependent on effort.

Our contribution is to show that an individual’s own experience (based on their personal payoff function) can systematically influence their estimate

of the effort of others. In particular, we show that type LMs are likely to underestimate the effort of type EMs. Continuing our example, individuals whose parents are well educated may underestimate the effort of individuals from less advantaged backgrounds. To equate type, as we are now doing, with readily observable characteristics like parental education, gender or race, may seem at odds with our assumption that type is unknown. Note, however, that it is the returns from effort, captured in the personal payoff function, that are assumed unknown. So, a person may simply not appreciate the different constraints that others face.

In our model and experiment managers are merely asked to estimate the effort of employees, not make hiring decisions. Even so, it seems safe to conjecture that own experience bias could lead to inefficiency in labor markets. To develop this point let us reiterate the difference between our approach and that taken in the literature on labour market discrimination. That literature addresses questions of the form: does an individual treat, say, men differently than women. Our approach, by contrast, leads to questions of the form: might men, because of their different personal experiences, make systematically different judgments compared to women. Our results suggest that individuals may indeed be biased by their own experience. We conjecture that this will likely influence hiring decisions. For instance, an all male committee may look at the same set of CVs differently to an all female committee.

We would strongly argue that the current literature neglects the role of own experience and so this should be a priority for future work. This claim is largely consistent with the approach advocated by Akerlof and Shiller (2010, Chapter 13), who argue that to understand poverty, and discrimination, we need to take account of the different ‘stories’ and ‘views of the world’ that blacks have compared to whites. They write that, ‘[t]he role of stories, of us versus them, of the search for self-respect, and of fairness in the lives of the poor is absent from standard economic analysis of poverty’ (page 163). We would add that it is also absent from standard economic analysis of discrimination and prejudice.

Appendix

Proof of Proposition 1

The objective of the individual is to choose an action in each period to maximize her total payoff $u = \sum_{t \in T} u_t$. To address this problem we need a preliminary result. We first consider the individual in period $t \geq 1$, having observed her history $h_{t-1} = (e_1, b_1; \dots; e_{t-1}, b_{t-1})$, trying to maximize her payoff in period $g \geq t$. Thus, the individual is choosing an action in each period t, \dots, g so as to maximize payoff u_g .

Suppose that $t = g$ and so the individual simply wants to maximize her payoff in the current period. She has beliefs p_{t-1} . If the individual chooses effort level e_t then her expected payoff in period t is given by

$$\begin{aligned} E[u_t(e_t)] &= p_{t-1}(\alpha e_t - \beta)B + (1 - p_{t-1})\gamma B - e_t \\ &= e_t B(p_{t-1}\alpha - 1) + B(\gamma - p_{t-1}(\beta + \gamma)). \end{aligned}$$

Hence, the individual should choose E_H if $p_{t-1}\alpha B > 1$ and choose E_L if $p_{t-1}\alpha B < 1$.

Suppose next that $t < g$ so the individual is trying to maximize her payoff in some future period. Note the objective in choosing an action in periods $t, \dots, g - 1$ is to obtain as accurate a belief as possible, p_{g-1} , before choosing effort in period g . We proceed in two stages. First we show that there is a symmetry in effort choice and its effect on prior probabilities, and then we show that information is maximized by choosing either effort level E_L or E_H .

Symmetry: Take as given beliefs p_{t-1} . Let $E_M = (E_H + E_L)/2 = e^*$. We will compare effort level $e_t > E_M$ with effort level $e'_t = E_M - (e_t - E_M) = 2E_M - e_t$. We want to show that the probability distribution over beliefs p_t is the same whether the individual chooses e_t or e'_t . In other words, both e_t or e'_t provide the same information. This implies that the individual's expected payoff in period g is the same if she chooses effort level e_t or e'_t in period $t < g$. (Of course, this does not imply her payoff in period t is not effected by her effort in period t .)

Suppose the individual chooses effort level $e_t > E_M$. Two events can occur:

(a) The individual receives the bonus, $b_t = 1$. This happens with probability

$$p_{t-1} \Pr(b_t = 1|EM) + (1 - p_{t-1}) \Pr(b_t = 1|LM) = p_{t-1}(\alpha e_t - \beta) + (1 - p_{t-1})\gamma$$

and yields

$$p_t = \frac{p_{t-1}(\alpha e_t - \beta)}{p_{t-1}(\alpha e_t - \beta) + (1 - p_{t-1})\gamma}. \quad (3)$$

(b) The individual does not receive a bonus, $b_t = 0$. This happens with probability

$$p_{t-1}(1 + \beta - \alpha e_t) + (1 - p_{t-1})(1 - \gamma)$$

and yields

$$p_t = \frac{p_{t-1}(1 + \beta - \alpha e_t)}{p_{t-1}(1 + \beta - \alpha e_t) + (1 - p_{t-1})(1 - \gamma)}. \quad (4)$$

Now suppose the individual chooses effort level e'_t . Two events can occur: (c) The individual does not receive a bonus, $b_t = 0$. This happens with probability $p_{t-1}(1 + \beta - \alpha e'_t) + (1 - p_{t-1})(1 - \gamma)$. Substituting in $e'_t = 2e^* - e_t$ together with $\alpha e^* = \beta + \gamma$ and $\gamma = 1/2$ gives

$$p_{t-1}(\alpha e_t - \beta) + (1 - p_{t-1})\gamma$$

and yields

$$p_t = \frac{p_{t-1}(\alpha e_t - \beta)}{p_{t-1}(\alpha e_t - \beta) + (1 - p_{t-1})\gamma}. \quad (5)$$

Note that this outcome exactly coincides with that of (a).

(d) The individual receives the bonus, $b_t = 1$. This happens with probability $p_{t-1}(\alpha e'_t - \beta) + (1 - p_{t-1})\gamma$. Again substituting in for e'_t gives

$$p_{t-1}(1 + \beta - \alpha e_t) + (1 - p_{t-1})(1 - \gamma)$$

and yields

$$p_t = \frac{p_{t-1}(1 + \beta - \alpha e_t)}{p_{t-1}(1 + \beta - \alpha e_t) + (1 - p_{t-1})(1 - \gamma)}. \quad (6)$$

Note that this outcome exactly coincides with that of (b).

We have shown that the distribution over values of p_t is exactly the same whether the individual chooses effort level e_t or e'_t (because outcomes (a) and (c) are equivalent and (b) and (d) are equivalent). Given the objective of maximizing payoff in period $g > t$ the individual is, therefore, indifferent between choosing effort e_t or e'_t . We can repeat the same argument for periods $t + 1, t + 2, \dots$ up to period g .

Choosing extremes: We now want to show that to maximize payoff in period $g > t$ the individual should choose either E_L or E_H . In particular,

she should maximize $|E_M - e_t|$. For specificity, we first consider effort levels $e'_t > e_t > E_M$. We will show that the individual's expected payoff in period g is higher, *ceteris paribus*, if she chooses effort level e'_t in period $t < g$ compared to e_t .

We consider, in turn, the case where the individual is of type LM and then type EM. If she is type LM then her payoff is maximized in period g if she chooses effort level $e_g = E_L$. We have seen that she will do that if $p_{g-1}\alpha B < 1$. Her payoff is, thus, increasing in the $\Pr(p_{g-1} < \frac{1}{\alpha B})$. Similarly if the individual is of type EM her payoff is increasing in the probability $\Pr(p_{g-1} > \frac{1}{\alpha B})$. We, therefore, need to show that choosing effort level $e'_t > e_t$ increases the expected probability that p_{g-1} is the right side of $1/\alpha B$.

This situation is a statistical game, in the sense of Blackwell and Girshick (1979). The individual has the choice of performing experiment e'_t or experiment e_t . Note that

$$\Pr(b_t|EM) - \Pr(b_t|LM) = (\alpha e - \gamma - \beta)(2b_t - 1).$$

is increasing in e . Choosing effort level e'_t allows, therefore, a more accurate belief p_t and, therefore, p_{g-1} than choosing e_t . This means that experiment e'_t is more informative than experiment e_t (see Definition 12.2.1 of Blackwell and Girshick 1979). From this we can conclude that the expected payoff in period g is maximized, *ceteris paribus*, by choosing E_H , or E_L in period t . An analogous argument treats the case $e'_t < e_t < E_M$.

We are now in a position to prove the Proposition. Suppose that in period 1 the individual chooses E_L if $p_0 < 1/\alpha B$ and chooses E_H otherwise. We know that this maximizes payoff in period 1. Moreover, we know that, *ceteris paribus*, it maximizes expected payoff in period 2, 3, ..., T . The individual cannot, therefore, do any better. This argument can be repeated in period 2. Specifically, suppose that in period 2 the individual chooses E_L if $p_1 < 1/\alpha B$ and chooses E_H otherwise. We know that this maximizes payoff in period 2. Moreover, we know that, *ceteris paribus*, it maximizes expected payoff in period 3, 4, ..., T . Again, therefore, the individual cannot do any better. Repeating this argument for periods 3 to T gives the desired result. ■

Proof of Proposition 2

From equation (2) we can see that

$$\frac{dX^*}{dp_A} = C \left(-\frac{1}{1-p_A} - \frac{1}{p_A} \right)$$

where $C > 0$ is a constant that depends on γ, β, α and E_H . Thus $dX^*/dp_A < 0$ and $X^*(p_A)$ is a decreasing function of p_A .

Suppose that the manager is type EM and so $p_A > p_0$. Then $X^*(p_A) < X^*(p_0)$ and so the manager has a lower threshold than the benchmark threshold. If $X < X^*(p_A)$ then $\bar{e}(b, p_A) = \bar{e}(b, p_0) = E_L$. If $X > X^*(p_0)$ then $\bar{e}(b, p_A) = \bar{e}(b, p_0) = E_H$. If $X \in (X^*(p_A), X^*(p_0))$ then $\bar{e}(b, p_A) = E_H > \bar{e}(b, p_0) = E_L$. Thus, for any X we have $\bar{e}(b, p_A) > \bar{e}(b, p_0)$ giving the desired result. A similar logic treats the case of $p_A < p_0$. ■

Proof of Proposition 3

Suppose that the only information Alice has about Bob is the number of bonuses X that Bob received from period \bar{t} to T . Suppose also that Bob knows his own type and has a belief responsive strategy where he chooses $e^* < E_h \leq E_H$ if type EM. Then it is optimal for Alice to use the benchmark threshold

$$X^*(p_A = p_0, E_h) = \frac{M(\ln(1-\gamma) - \ln(1+\beta-\alpha E_h)) + \ln(1-p_A) - \ln(p_A)}{\ln(1-\gamma) - \ln(1+\beta-\alpha E_h) + \ln(\alpha E_h - \beta) - \ln \gamma}. \quad (7)$$

We first consider how the optimal threshold changes if Bob uses a belief responsive strategy. Setting $p_A = 0.5$ and $\gamma = 0.5$ we get that

$$\frac{dX^*(\frac{1}{2}, E_h)}{dE_h} \propto (\alpha E_h - \beta)(\ln(\alpha E_h - \beta) - \ln(1 + \beta - \alpha E_h)) - (\ln(1 - \gamma) - \ln(1 + \beta - \alpha E_h)). \quad (8)$$

Set $x = \alpha E_h - \beta$. Then equation (8) can be written,

$$\frac{dX^*(p_A, E_h)}{dE_h} \propto x \ln(x) + (1-x) \ln(1-x) - \ln(0.5) \geq 0 \quad (9)$$

where the equality is strict for all values of x other than 0.5. This tells that if Bob has a belief responsive strategy with $E_h < E_H$ then the benchmark

threshold is lower, $X^*(p_0, E_h) < X^*(p_0, E_h)$, than if Bob used the optimal strategy.

We next consider the threshold Alice will use if she has own experience bias. Recall that

$$\frac{dX^*(p_A, E_h)}{dp_A} = C \left(-\frac{1}{1-p_A} - \frac{1}{p_A} \right) < 0.$$

Thus, for any $E_h < E_H$ there exists $p_A > p_0$ such that

$$|X^*(p_0, E_h) - X^*(p_A, E_H)| < |X^*(p_0, E_h) - X^*(p_0, E_H)|.$$

It follows that Alice's estimates will be more accurate with belief p_A than unbiased prior p_0 . If $p_A < p_0$ then estimates will be less accurate than with unbiased prior p_0 .

References

- [1] Akerlof, G. A., & Shiller, R. J. (2010). *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism* Princeton University Press
- [2] Anderson, L., Fryer, R., & Holt, C. (2006). Discrimination: experimental evidence from psychology and economics. *Handbook on the Economics of Discrimination*, 97-118.
- [3] Arceo-Gomez, E. O., & Campos-Vazquez, R. M. (2014). Race and marriage in the labor market: A discrimination correspondence study in a developing country. *American Economic Review*, 104(5), 376-380.
- [4] Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments?. *Labour Economics*, 30, 32-40.
- [5] Bertrand, M., Chugh, D. & Mullainathan, S. (2005). Implicit discrimination. *American Economic Review*, 95(2), 94-98.
- [6] Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, 94(4), 991-1013.

- [7] Blackwell, D. A., & Girshick, M. A. (1979/1954). *Theory of games and statistical decisions*. John Wiley & Sons.
- [8] Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about gender. *American Economic Review*, 109(3), 739-73.
- [9] Breitmoser, Y. (2019). Knowing me, imagining you: Projection and overbidding in auctions. *Games and Economic Behavior*, 113, 423-447.
- [10] Carlsson, M., & Rooth, D. O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics*, 14(4), 716-729.
- [11] Castillo, M., & Petrie, R. (2010). Discrimination in the lab: Does information trump appearance? *Games and Economic Behavior*, 68(1), 50-59..
- [12] Christensen-Szalanski, J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational behavior and human decision processes*, 48(1), 147-168.
- [13] Charness, G., & Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab?. In *Handbook of Labor Economics*, Elsevier.
- [14] Dickinson, D. L., & Oaxaca, R. L. (2014). Wages, Employment, and Statistical Discrimination: Evidence from the Laboratory. *Economic Inquiry*, 52(4), 1380-1391.
- [15] Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). 'The overconfidence effect in social prediction'. *Journal of Personality and Social Psychology*, 58(4), 568.
- [16] Engelmann, D., & Strobel, M. (2012). Deconstruction and reconstruction of an anomaly. *Games and Economic Behavior*, 76(2), 678-689.
- [17] Ewens, M., Tomlin, B., & Wang, L. C. (2014). Statistical discrimination or prejudice? A large sample field experiment. *Review of Economics and Statistics*, 96(1), 119-134.
- [18] Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: A survey, in In J. Benhabib, M. Jackson and A.

Bisin (Eds) *Handbook of Social Economics*, Vol. 1A, The Netherlands: North-Holland.

- [19] Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.
- [20] Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1), 35-42.
- [21] Guryan, J., & Charles, K. K. (2013). Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots. *The Economic Journal*, 123(572), F417-F432.
- [22] Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 566.
- [23] Krueger J, Clement RW. The truly false consensus effect: an ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*. 1994 Oct;67(4):596.
- [24] Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, 96(1), 222-235.
- [25] Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, 21(3), 262-283.
- [26] Neumann, N., Bckenholt, U., & Sinha, A. (2016). A meta-analysis of extremeness aversion. *Journal of Consumer Psychology*, 26(2), 193-212.
- [27] Pager, D., Western, B., & Bonikowski, B. (2009). Discrimination in a low-wage labor market a field experiment. *American Sociological Review*, 74(5), 777-799.
- [28] Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3), 775-816.
- [29] Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair womens careers in science. *Proceedings of the National Academy of Sciences*, 111(12), 4403-4408.

- [30] Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3), 279-301.
- [31] Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2), 207-232.
- [32] Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.