



Failures in Forecasting: An Experiment on Interpersonal Projection Bias

Benjamin Bushong,^{a,*} Tristan Gagnon-Bartsch^b

^aMichigan State University, East Lansing, Michigan 48824; ^bFlorida State University, Tallahassee, Florida 32306

*Corresponding author

Contact: bbushong@msu.edu,  <https://orcid.org/0000-0002-1398-3873> (BB); tgagnonbartsch@fsu.edu,  <https://orcid.org/0000-0002-9118-6095> (TG-B)

Received: March 4, 2022

Revised: February 24, 2023; April 20, 2023

Accepted: April 20, 2023

Published Online in Articles in Advance:
March 11, 2024

<https://doi.org/10.1287/mnsc.2022.00655>

Copyright: © 2024 INFORMS

Abstract. Using a real-effort experiment, we show that people project their current tastes onto others when forecasting others' behavior, even when their tastes are exogenously manipulated and transparently different. In the first part of our experiment, "workers" stated their willingness to continue working on a tedious task. We varied how many initial tasks workers completed before eliciting their willingness to work (WTW); some were relatively fresh when stating their WTW, whereas others were relatively tired. Later, a separate group of "predictors"—who also worked on the task—guessed the WTW of workers in each state. We find: (i) tired workers were less willing to work than fresh workers; (ii) predictors (in aggregate) accurately guessed the WTW of workers when they were in the same state as the workers about whom they were predicting; but (iii) when fresh predictors were guessing about tired workers, they substantially overestimated their WTW; and (iv) when tired predictors were guessing about fresh workers, they underestimated their WTW. Using an additional treatment, we find that workers also mispredicted their own future WTW, and we compare the magnitudes of intrapersonal and interpersonal projection bias.

History: Accepted by Axel Ockenfels, behavioral economics and decision analysis.

Supplemental Material: The online appendix and data files are available at <https://doi.org/10.1287/mnsc.2022.00655>.

Keywords: projection bias • real-effort experiments • forecasting • state-dependent preferences

When you are asked to 'put yourself in someone's place,' what is the implied contrasting condition: what is it that you are implicitly being asked *not* to do? ... [Y]ou shouldn't just *project your own* situation and psychology on the other. —Robert M. Gordon (1986)

1. Introduction

Predicting others' preferences is a ubiquitous feature of economic and organizational interactions. For instance, estimating others' valuations is central to bidding in an auction; attending to a counterparty's goals enables effective negotiation; and being mindful of others' recent workloads is important for managers when allocating tasks because fatigue will diminish workers' productivity and willingness to take on additional work. Many other aspects of organizational decision making—for example, motivating effort, advertising, and sales planning—similarly require some form of prediction about what others might do. But do people accurately forecast others' preferences?

In this paper, we present experimental evidence that people's predictions about others' willingness to work (WTW) on a real-effort task are biased by their own current fatigue. Our evidence supports a simple model of

"interpersonal projection bias": people project their current tastes onto others, even when others face transparently different circumstances. This basic idea speaks to a recent literature showing that people—including experts—sometimes struggle to forecast how others will respond to incentives and experimental interventions (DellaVigna and Pope 2018a, b; DellaVigna et al. 2019; DellaVigna and Pope 2022). Our results suggest that projection bias may be a potential source of this prediction error, and a natural way to mitigate this error is to put forecasters in states similar to those of the people they are estimating. Furthermore, although projection bias has been shown among people predicting their own future preferences (see, e.g., Read and van Leeuwen 1998, Badger et al. 2007, Conlin et al. 2007, Busse et al. 2015), we measure both interpersonal and intrapersonal projection bias in the same setting and compare their relative magnitudes. We find suggestive evidence that interpersonal projection bias may be larger: that is, a person's current tiredness distorts her predictions about others' WTW by more than it distorts predictions about her own future WTW.

Participants in our experiment worked on a tedious real-effort task. In the first part of our experiment, we

elicited participants' willingness to continue working on the task for additional pay. In a second part, different participants cast incentivized predictions about the WTW of the first group. We call these two groups "workers" and "predictors," respectively. Critically, we varied how many initial tasks workers completed before eliciting their WTW: half completed 5 tasks—they were relatively fresh when stating their WTW—whereas the other half completed 20—they were relatively tired. Predictors also worked on the task, and we similarly varied their initial workloads: some predictors completed 5 tasks before guessing the WTW of others, whereas others completed 20. Our central question is whether—and to what extent—predictors abstracted from their own state (i.e., their current tiredness) when making guesses about others' preferences.

We present five stylized facts about behavior and predictions in our primary experiment on interpersonal projection bias. (i) Tired workers were less willing to work than fresh workers. (ii) Predictors (in aggregate) accurately guessed the WTW of workers when they cast their predictions in the same state as the workers about whom they were predicting. (iii) When predictors were fresh, but guessing about tired workers, they substantially overestimated the WTW of tired workers. (iv) When predictors were tired, but guessing about fresh workers, they underestimated the WTW of fresh workers. (v) Fresh predictors made a larger error when guessing about tired workers than tired predictors made when guessing about fresh workers.

Overall, these results suggest that participants in our experiment projected their current sense of tiredness onto others. We find—through both nonparametric measures and the estimates of a reduced-form model—that our manipulation of tiredness induced large and significant errors in predictions about the choices of others. Although these predictions were accurate when guesses were about others in one's own state, our estimates suggest that guesses about others in a different state were systematically distorted (in the direction of the predictor's own tiredness) by 21%–50%.

To further elucidate the mechanism driving these errors, we divided our predictors into three subgroups. These groups varied in how many guesses each predictor cast and their tiredness when they made each guess. Some predictors made guesses about the WTW of fresh workers when they themselves were fresh and, later, made guesses about tired workers when they themselves were tired. Other predictors made guesses when they were "out of phase" with the workers: they guessed about tired workers when they themselves were fresh and guessed about fresh workers when they themselves were tired. Comparing predictions across these two groups allows us to measure how predictors' estimates changed as their own states changed. Finally, a third group made their first

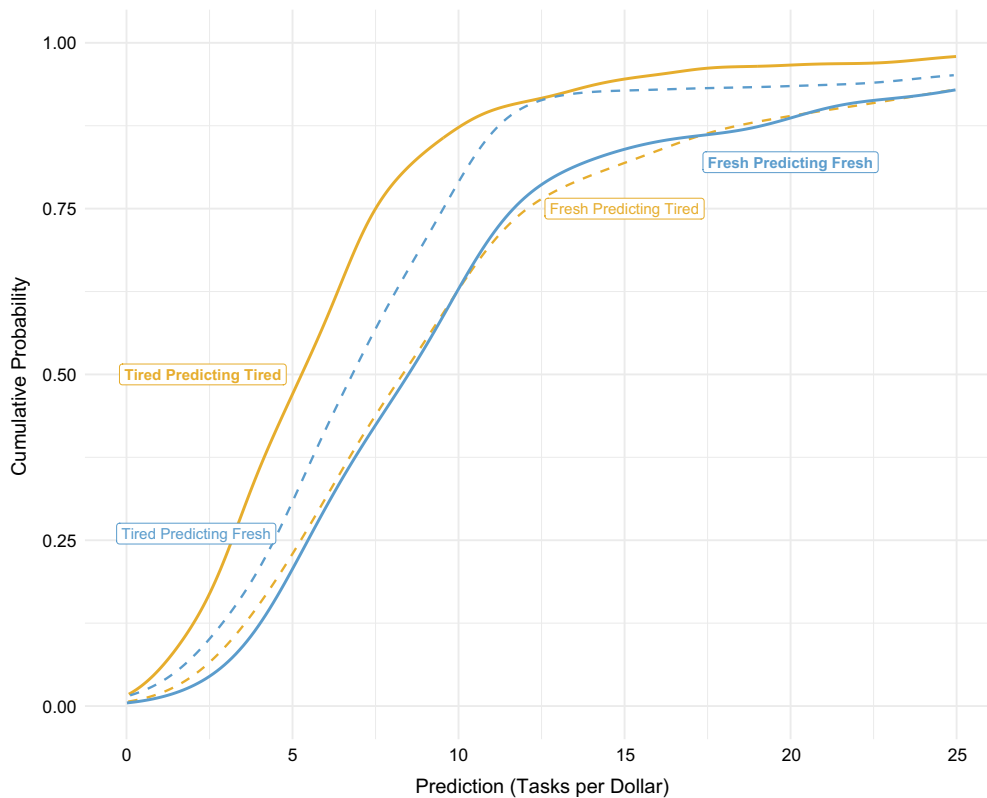
prediction—about fresh workers—when they themselves were tired; they made no prediction when they were fresh. This group allows us to further control for anchoring or other order effects by comparing only the initial predictions across groups. When focusing solely on initial predictions, we find that fresh predictors overestimated the WTW of tired workers by approximately 50%, whereas tired predictors underestimated the WTW of fresh workers by approximately 21%. Figure 1 previews this result by showing the distribution of initial guesses cast by predictors in a different state than workers relative to those cast by predictors in the same state. Moreover, comparing initial guesses across groups also allows us to decompose predictors' erroneous guesses into two components—one due to projection bias and another due to uncertainty about how onerous the task would become over time. We find that these two components distorted the guesses of fresh predictors by roughly similar magnitudes.

Additionally, we analyze how predictors' guesses about fresh workers changed as the predictors went from fresh to tired. When predictors first guessed the WTW of fresh workers when they themselves were fresh, they were, on average, accurate. However, when they performed this same prediction again (i.e., about fresh workers) once they themselves became tired, they substantially revised their guesses downward (by approximately 19%; difference significant at $p < 0.001$). This was a mistake. By revising their guesses, predictors significantly decreased their accuracy and lowered their expected earnings. Because predictors became less accurate even as they gathered more firsthand experience with the task, this suggests that our results do not stem from predictors simply lacking information about the task. Indeed, this analysis—and our collection of results more broadly—offers strong support for interpersonal projection bias.

Finally, we advance the existing literature on projection bias by comparing the magnitude of interpersonal and *intrapersonal* projection—the tendency for a person's current state to overly influence predictions about their own future behavior. To measure intrapersonal projection bias in the same experimental setting, we ran an additional worker treatment, in which workers in the fresh state predicted their own future WTW in the tired state. On average, these fresh workers overestimated their WTW in the tired state by approximately 30%. For a comparable measure of interpersonal projection, fresh predictors overestimated the WTW of others in the tired state by approximately 50%. This suggests that the totality of errors that accumulate into interpersonal projection bias may be more severe than the intrapersonal analog.

In the conclusion, we discuss a related idea—information projection (e.g., Camerer et al. 1989 and Madarász 2012)—wherein people exaggerate the degree

Figure 1. (Color online) Empirical (Smoothed) Cumulative Distribution Functions of Predictions About Fresh and Tired Workers



Notes. Predictions are shown in units of “tasks-per-dollar”—the average number of tasks workers are willing to do for each dollar of compensation. The solid lines represent guesses cast by predictors in the same state as the workers about whom they were predicting, while the dotted lines represent the guesses cast by predictors in the other state. Both dotted distributions are significantly different from the relevant solid distributions ($p < 0.001$ for both; Wilcoxon rank-sum test).

to which others share their private information (as opposed to their preferences). Some of our findings are consistent with information projection, but we discuss why information projection may not be a complete explanation for our collective evidence.

Despite its potential ubiquity, interpersonal projection bias has received relatively little attention in the economics literature. On the empirical side, Van Boven et al. (2003) find that sellers in experimental markets project their sense of endowment onto potential buyers; Ambuehl et al. (2021) find that subjects in a paternalistic role project their aspirations onto others; and Engelmann and Strobel (2000, 2012) provide experimental tests of the “false-consensus effect”—the tendency to exaggerate the similarity between one’s own actions or opinions and those of others. In the following section, we discuss the related empirical evidence (including studies from psychology) in more detail, and we describe the variety of ways in which our study builds on this literature.

A few of these ways are worth emphasizing here. First, we transparently induced changes in tastes along

a familiar dimension (tiredness). Second, because our predictors completed the same task (in the same quantities) as our workers, we limit information-based explanations that have challenged the interpretation of previous studies. These two elements of our design reduce the possibility that biased predictions arose simply because predictors were either uncertain about other people’s tiredness or were unfamiliar with the experience of being in the fresh or tired state. Third, we elicited multiple predictions from each participant, which allows us to obtain within-subject measures of how a person’s guesses changed as their own tiredness state changed. Finally, we are—to our knowledge—the first study to measure both interpersonal and intrapersonal projection bias in the same domain. This allows us to compare the relative magnitudes of these errors.

Our results highlight the potential benefit from greater engagement with the perspectives of others, particularly in domains involving effort provision and fatigue. For instance, a fresh manager who suffers from interpersonal projection bias may fail to design optimal incentive schemes because she underestimates how

workers' tiredness will influence their responsiveness to incentives. Regardless of her own tiredness state, the manager will also systematically underestimate the heterogeneity in her workers' marginal disutility of effort and will therefore underestimate the value of strategically tailoring assignments to workers based on their recent workloads. Beyond the specific domain of fatigue, projection bias has many important implications for decision making in organizations. The strategies we use to motivate others will be overly reliant on what would motivate ourselves at a given moment. The advice we give will be overly tailored to our own tastes or circumstances. And the policies for which we advocate will be overly aligned with our own private objectives, even when we are directly concerned with satisfying others' preferences. Projection bias may also cloud our judgments of others. At times when we are relatively unburdened, we may wrongly attribute another person's limited productivity to intrinsic characteristics (e.g., low ability or laziness) rather than momentary hardship.

Our findings also speak to the emerging literature studying the accuracy of forecasts about behavioral responses to incentives and interventions (e.g., DellaVigna and Pope 2018a, b; DellaVigna et al. 2019; DellaVigna and Pope 2022). Those papers highlight that even though forecasters demonstrate some ability to predict others' responses to interventions, they still make significant forecasting errors. Furthermore, neither higher incentives nor added expertise substantially improve forecasters' guesses. Our experiments suggest that these predictions may improve when forecasters are themselves in a similar state as the people they are trying to predict. Given the pitfalls of projection described above, this practice may directly improve the decision making of managers and researchers alike. In fact, researchers have recently explored this idea in the domain of policy and intervention design: Thomas et al. (2020) show that organizations can gain valuable information for designing such interventions by eliciting forecasts about their impact directly from the "local" population that the intervention intends to target. They emphasize that this is especially true when the target population may be in a different psychological state than those designing the intervention.

Finally, our evidence bolsters the empirical foundation for a growing theoretical literature emphasizing the broader implications and importance of interpersonal projection bias. For instance, projection of political preferences can generate inefficient election outcomes (Goeree and Grosser 2007); projection of private valuations can lead to overbidding and inefficient allocations in auctions (Gagnon-Bartsch et al. 2021); and in social-learning contexts, such as the adoption of a new technology, mispredicting others' tastes can prevent people from inferring their optimal action (Gagnon-Bartsch

2016, Frick et al. 2020, Bohren and Hauser 2021). Relatedly, Kaufmann (2022) shows how *intrapersonal* projection bias over effort can lead people to overcommit to and overwork toward projects that go unfinished in contexts similar to our experimental setting.

The paper proceeds as follows. In Section 2, we discuss the existing evidence on interpersonal projection, highlighting how our design builds on previous studies and mitigates important confounds. In Section 3, we provide a detailed description of our experimental design and—using a framework similar to that of Loewenstein et al. (2003)—we derive testable hypotheses. In Section 4, we present evidence supporting these hypotheses. We also present additional analyses involving predictors' experience and confidence that lend further support to projection as the mechanism underlying our results. In Section 5, we present findings on intrapersonal projection bias and compare those results to the existing literature. Section 6 concludes.

2. Related Literature

In this section, we review prior work in both psychology and economics on interpersonal projection bias. We then describe how our experiment mitigates some of the confounds in earlier studies.

Social projection—the tendency for people to believe that others share their tastes or beliefs—has a long history in psychology (e.g., Katz and Allport 1931, Cronbach 1955, Sherif and Hovland 1961). In a seminal paper, Ross et al. (1977) dub this error the "false-consensus effect."¹ Subsequent work has proposed various mechanisms that can generate a false-consensus effect, but distinguishing which factor drives this error in a given context remains elusive.

One explanation of the false-consensus effect is the idea that people project their preferences or states onto others. For instance, in Van Boven et al. (2003), endowed sellers overestimated the willingness to pay of potential buyers.² To further explore how a person's own temporary state distorts their predictions of others' preferences, Van Boven and Loewenstein (2003) had participants read a short vignette about three lost hikers stranded overnight in the woods. Participants were then asked to imagine themselves in the place of one of the hikers and answer the following: "Which would be more unpleasant [to you] for the hikers, hunger or thirst?" Some participants completed at least 20 minutes of vigorous exercise before reading the vignette and answering the question; those who did were significantly more likely to be concerned about thirst when compared with those who did not exercise before answering.³

Our experiment drew inspiration from Van Boven and Loewenstein's (2003) design in that we also exogenously manipulated a predictor's state. Superficially,

we focus on projecting fatigue rather than thirst. But there are several more important design differences that help highlight our contribution to this literature. First, we precisely controlled the tiredness states of both the target groups (i.e., fresh and tired workers) and the predictors. Second, predictors attempted to forecast a continuous variable (i.e., willingness to work) chosen by the target groups. Together, these two features allow us to measure projection bias on the intensive margin. Third, because we could actually observe the choices of the target group, we were able to incentivize predictions based on accuracy and analyze how accuracy varied with the predictor's state. Fourth, we elicited multiple predictions from each predictor, which reveals how readily an individual's predictions changed as her own state changed.

Lastly, our predictors had firsthand experience with the situations experienced by the target groups: they worked on the same task as the workers, and they faced that task in both the fresh and tired states. This aspect of our design—combined with our exogenous manipulation of others' states—allows us to consider an argument put forth by Dawes (1989, 1990) suggesting that projection may be rational (see also Krueger and Clement 1994). An adaptation of this argument to our setting with state-dependent utility is as follows: a predictor's own WTW in a given state may provide information about others' WTW in that same state, and she may therefore guess that others will behave like herself simply because she lacks any further information about others' preferences (i.e., she acts as a Bayesian with a single data point, herself). This could manifest in "rational projection." We address this by (i) arming predictors with information about both states via firsthand experience; (ii) changing the taste of predictors (making them tired); and (iii) asking whether predictors use their WTW when tired to predict the WTW of fresh workers.⁴ Rational projection would imply that predictors use their own WTW *when fresh* to predict the WTW of fresh workers. We explore—and reject—this notion.

In addition to the limited-information explanation discussed above, our design was intended to isolate the role of projection from other alternative mechanisms that are discussed in the false-consensus literature (see, e.g., Marks and Miller 1987). First, a form of availability bias or selection neglect may cause people to excessively extrapolate from the characteristics of their own social circle—which are likely correlated—when estimating the characteristics of a more general population. Second, in domains with a salient social norm, people may derive value from believing that their preferences conform with others'. Hence, because of motivated reasoning, their predictions may reflect this willfully distorted belief. We designed our experiment to sidestep these alternative channels in order to better identify

whether projection distorts predictions. In particular, our experiment explores preferences over an unfamiliar, yet mundane, task. Given its unfamiliarity, it is unlikely that participants had any relevant data from which they could extrapolate or any perception of a social norm. Furthermore, by incentivizing predictions, we diminished any benefit from maintaining motivated beliefs.

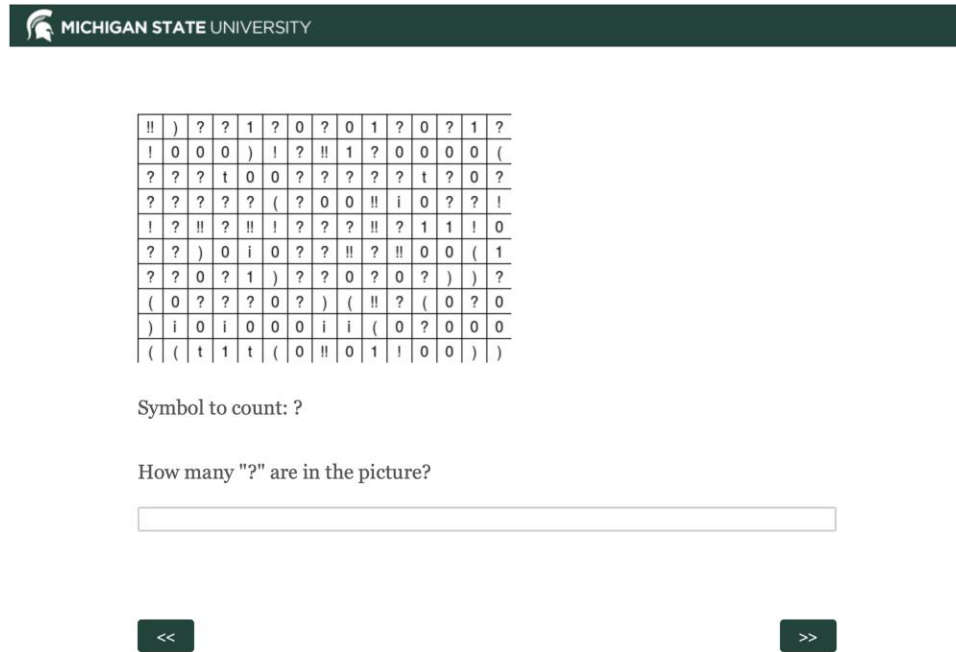
Finally, although several papers in the economics literature find indirect evidence of a false-consensus effect, a few specifically study this concept.⁵ Ambuehl et al. (2021) conducted an experiment where participants constructed choice sets on the behalf of others and were asked what they believed others would choose. The authors document a strong false-consensus effect, and the effect persisted in two settings, where participants were given information on either a specific person's preferences or the distribution of other' preferences (see their online appendix B.4).⁶ Engelmann and Strobel (2000) find that the false-consensus effect disappears after participants are provided with signals about others' choices. However, Engelmann and Strobel (2012) suggest that the artificial nature of that environment—namely, the free acquisition of strong signals about others—led to the null finding in their previous work. Our contribution relative to these papers is threefold. First, by focusing on a setting where predictors have been "in the shoes" of the target group, we offer additional evidence of projection in a setting where uncertainty about others is reduced. Second, we transparently and exogenously manipulated tastes in this setting, allowing us to further isolate the effect of projection bias on predictions. Finally, as highlighted in the introduction, we measure both interpersonal and intrapersonal projection bias in the same domain.

3. Experimental Design

A total of 1,566 people participated in our experiments on Amazon's Mechanical Turk (MTurk).⁷ Our experiment had two distinct, mutually exclusive parts, which corresponded to two different participant roles. Regardless of their decisions or their role, all participants who completed the survey earned at least \$3. Participants in the first part—whom we call "workers"—completed some initial work on a real-effort task and then stated their willingness to perform more work for additional pay. Participants in the second part—whom we call "predictors"—completed some initial work and then guessed the workers' average willingness to work.

Before providing details on these roles, we first describe the real-effort task. All participants worked on (and, when required, formed predictions about) the same real-effort task. Each round of the task required a participant to count the number of times a particular number or symbol (e.g., 0, 1, ?, !) appeared in a 10 × 15

Figure 2. (Color online) Screenshot of the Counting Task



matrix of numbers and symbols. See Figure 2 for a screenshot of the task. On average, it took participants about 75 seconds to complete one round of the task.

3.1. Workers

Participants in the first part of the experiment completed a set number of rounds of the task, and then we elicited their willingness to complete additional rounds. Workers were randomized into one of two groups: (i) in the *Fresh* group, workers completed 5 mandatory tasks prior to stating their willingness to complete more; (ii) in the *Tired* group, workers completed 20 mandatory tasks prior to stating their willingness to complete more.

We elicited willingness to work using the Becker-DeGroot-Marshak (BDM) mechanism. We asked each worker how many additional tasks he was willing to complete for a bonus of $\$m$. Participants used a slider to select a WTW between zero and 100. We then randomly drew an integer z between zero and 100. If z was below the participant's selected WTW, they had to complete z additional tasks in exchange for a bonus of $\$m$. Otherwise, they did no additional tasks and received no bonus. We varied the bonus payment $m \in \{2, 3\}$ depending on the group (Fresh versus Tired, respectively).⁸

To summarize, participants assigned to the worker role were randomly assigned to one of two groups:

Fresh Workers ($n = 303$). Participants in this group completed five mandatory rounds of the task before we elicited their willingness to complete more rounds.

These workers were in a (relatively) fresh state when announcing their WTW. Each participant i in this group stated how many additional rounds he was willing to complete for $m = \$2$, which we denote by $W_i(\$2, F)$ (where F denotes the fresh state). Let $\bar{W}(\$2, F)$ be the average response among this group.

Tired Workers ($n = 299$). Participants in this group completed 20 mandatory rounds of the task before we elicited their willingness to complete more rounds. These workers were in a (relatively) tired state when announcing their WTW. Each participant i in this group stated how many additional rounds he was willing to complete for $m = \$3$, which we denote by $W_i(\$3, T)$ (where T denotes the tired state). Let $\bar{W}(\$3, T)$ be the average response among this group.

We also recruited a third group of workers that allowed us to measure *intrapersonal* projection bias. These workers had the same experience as the tired workers described above, except they additionally predicted their own WTW ahead of time. We postpone a detailed description of this group until Section 5, where we compare intrapersonal and interpersonal projection bias.

3.2. Predictors

Predictors made a series of incentivized guesses about the average WTW of fresh and tired workers; that is, they predicted $\bar{W}(\$2, F)$ and $\bar{W}(\$3, T)$. In order to mitigate confounds from informational asymmetries, predictors also worked on the same task that the workers faced, and they made predictions after completing 5

tasks (i.e., in the fresh state) and after completing 20 tasks (i.e., in the tired state). Predictors were randomly assigned to one of three groups. In each, participants had to complete 20 rounds of the counting task. The three groups differed based on when participants provided predictions (after completing 5 tasks, 20 tasks, or both) and based on which groups of workers they guessed about (fresh versus tired).

In particular, our groups differed based on whether initial predictions were about workers in the same state as the predictor. Predictors in the “In Group” (group I henceforth; $n = 223$) began by making predictions about others in their own state. To clarify:

A predictor in group I made three guesses in total: (1) after completing 5 tasks himself, he predicted $\bar{W}(\$2, F)$ —the WTW of fresh workers; (2) after completing 20 tasks, he predicted $\bar{W}(\$3, T)$ —the WTW of tired workers; and (3) immediately after the second prediction, he again predicted the WTW of fresh workers, $\bar{W}(\$2, F)$. This final prediction allows us to test whether a predictor changed his view of others simply as a result of becoming tired himself. Note that for the first two predictions above, the predictor guessed the WTW of workers who were in the same state as himself: when the predictor was fresh, he guessed the WTW of fresh workers; when the predictor was tired, he guessed the WTW of tired workers. We call these “in-group” predictions.

Predictors in our other two groups—“Out Groups”—made their first predictions about others in a *different* state than themselves. Our two Out Groups differed in which prediction they made first: group O ($n = 221$) cast three predictions in total, made when both fresh and tired; group O_{-1} ($n = 222$) cast two predictions, only when tired. To clarify:

A predictor in group O made three guesses in total: (1) after completing 5 tasks herself, she predicted $\bar{W}(\$3, T)$ —the WTW of tired workers; (2) after completing 20 tasks, she predicted $\bar{W}(\$2, F)$ —the WTW of fresh workers; and (3) immediately after the second prediction, she again predicted the WTW of tired workers, $\bar{W}(\$3, T)$.

Note that for the first two predictions above, the predictor guessed the WTW of workers who were in a different state as herself: when the predictor was fresh, she guessed the WTW of tired workers; when she was tired, she guessed the WTW of fresh workers. We call these “out-group” predictions.

A predictor in group O_{-1} made two guesses after completing all 20 tasks. Aside from not making a guess after completing 5 tasks, a predictor in this group made the same guesses as group O : after completing 20 tasks, she predicted $\bar{W}(\$2, F)$ —the WTW of fresh workers; immediately after, she predicted the WTW of tired workers, $\bar{W}(\$3, T)$.

To ensure that the instructions and timing for this group closely mirrored those of group O , we interrupted each predictor in group O_{-1} after five tasks. During this pause, we presented instructions on the BDM mechanism and reminded the participant that she would later make predictions about others who made choices in the state she was currently in (i.e., fresh). Thus, even though a predictor in group O_{-1} did not make a numerical guess after she completed five tasks, she was still paused and cued to think about others while in the fresh state.

We introduce the following notation for the predictions described above. Let $\widehat{W}_i^s(m, s|s_i)$ denote the guess of predictor i from group $g \in \{I, O, O_{-1}\}$ about workers in state $s \in \{F, T\}$ facing bonus m , where s_i denotes predictor i 's own state at the time of her prediction. For example, $\widehat{W}_i^s(\$2, F|T)$ is predictor i 's guess about $\bar{W}(\$2, F)$ cast while she is in the tired state. Let $\bar{W}^s(m, s|s')$ denote the average prediction of $\bar{W}(m, s)$ among predictors in group g who were in state $s' \in \{F, T\}$ when making their predictions.

All predictions from each group were incentivized as follows: a participant earned a 50-cent bonus for each prediction that was within five tasks of the true value.⁹ After each prediction, we also asked participants to rank their confidence in that prediction on a scale from 1 (not at all confident) to 5 (extremely confident). These confidence measures were not incentivized. Finally, after completing all 20 tasks and providing all predictions, we asked predictors about their own willingness to complete more tasks for an additional payment of \$3. This question was phrased identically to the one we asked workers, but it was not incentivized.¹⁰

Predictors received no feedback after making their predictions. Their payments—which were based on the accuracy of their guesses—were revealed only once the experiment was over.

3.3. Theoretical Predictions

Interpersonal projection bias implies that a person's own state (fresh versus tired) distorts her predictions about the WTW of others. Using a simple model, we now show that projection leads fresh predictors to overestimate the WTW of tired workers and leads tired predictors to underestimate the WTW of fresh workers.

To formalize these hypotheses, first consider the behavior of workers. Suppose participant i 's cost of completing $e \in \mathbb{R}_+$ additional tasks is $C(e; s_i, \theta_i) \equiv c(e + s_i; \theta_i) - c(s_i; \theta_i)$, where the “state” $s_i \in \mathbb{R}_+$ is the number of tasks i completed beforehand and $\theta_i \in \mathbb{R}$ is i 's “type,” capturing an idiosyncratic taste for the task. We assume that the cost function $c(\cdot; \theta_i)$ is strictly increasing and convex for all θ_i , and, thus, effort becomes more costly as the worker grows tired. When asked how many tasks she is willing to complete for a bonus payment of

m , participant i chooses e to maximize $\int_0^e [m - C(\tilde{e}; s_i, \theta_i)] d\tilde{e}$. Participant i 's optimal choice is thus implicitly defined by the solution to

$$c(e + s_i; \theta_i) - c(s_i; \theta_i) = m, \quad (1)$$

which we denote by $W(m, s_i | \theta_i)$. Given our assumptions on c , $W(m, s_i | \theta_i)$ is decreasing in s_i for a fixed monetary bonus—intuitively, a person is less willing to work as she grows tired.

Building from the Loewenstein et al. (2003) model of intrapersonal projection, we now provide a simple model of interpersonal projection bias in our experimental setting. Suppose that participant i wrongly believes that participant j 's cost function is

$$\begin{aligned} \widehat{C}(e; s_j, \theta_j | s_i, \theta_i) &= \alpha C(e; s_i, \theta_i) + (1 - \alpha)C(e; s_j, \theta_j) \\ &= \alpha [c(e + s_i; \theta_i) - c(s_i; \theta_i)] \\ &\quad + (1 - \alpha)[c(e + s_j; \theta_j) - c(s_j; \theta_j)], \end{aligned} \quad (2)$$

where parameter $\alpha \in [0, 1]$ captures the degree of projection bias. That is, a projecting predictor perceives another person's cost as a convex combination of her own cost and that other person's true cost.¹¹ When $\alpha = 0$, this model collapses to the rational unbiased model. Under projection bias, participant i predicts that participant j will choose an effort level that solves $\widehat{C}(e; s_j, \theta_j | s_i, \theta_i) = m$. Hence, i 's prediction about the WTW of an individual in state s_j , denoted by $\widehat{W}(m, s_j | s_i, \theta_i)$, is decreasing in both s_j and s_i whenever $\alpha > 0$.¹² On the other hand, $\widehat{W}(m, s_j | s_i, \theta_i)$ is constant in s_i absent projection bias ($\alpha = 0$).

Although the model allows for a continuum of tiredness states, our experiment focuses on just two: $s = 5$ and $s = 20$. To make the state salient in our notation, we henceforth denote these two numerical states by F and T , respectively. Additionally, our formalization of projection bias in Equation (2) implies that a projector's predictions of others' WTW is biased by her own tiredness state, s_i , and her own idiosyncratic taste, θ_i . However, because we can only exogenously manipulate the former, our main hypotheses will compare predictions cast by predictors in state F versus state T .

We can now state our primary hypotheses regarding the average estimates cast by predictors. Under projection bias ($\alpha > 0$), we have the following:

Hypothesis 1. *Fresh predictors accurately estimate the WTW of fresh workers, and tired predictors accurately estimate the WTW of tired workers: $\widehat{W}^I(\$2, F|F) = \overline{W}(\$2, F)$ and $\widehat{W}^S(\$3, T|T) = \overline{W}(\$3, T)$ for $g \in \{I, O, O_{-1}\}$.*

Hypothesis 2. *Relative to tired predictors, fresh predictors overestimate the WTW of tired workers: $\widehat{W}^O(\$3, T|F) > \widehat{W}^S(\$3, T|T)$ for each $g \in \{I, O, O_{-1}\}$.*

Hypothesis 3. *Relative to fresh predictors, tired predictors underestimate the WTW of fresh workers: $\widehat{W}^S(\$2, F|T) < \widehat{W}^I(\$2, F|F)$ for each $g \in \{I, O, O_{-1}\}$.*

A variety of biases in belief formation independent of projection bias could jeopardize Hypothesis 1 (see, e.g., Benjamin 2019 for a review).¹³ Hypotheses 2 and 3 are robust to such biases because they compare predictions across groups, rather than compare the predictions of one group to the truth. Moreover, if Hypothesis 1 does hold, then we have two immediate corollaries of Hypotheses 2 and 3:

Hypothesis 2A. *Fresh predictors overestimate the WTW of tired workers: $\widehat{W}^O(\$3, T|F) > \overline{W}(\$3, T)$.*

Hypothesis 3A. *Tired predictors underestimate the WTW of fresh workers: $\widehat{W}^S(\$2, F|T) < \overline{W}(\$2, F)$ for each $g \in \{I, O, O_{-1}\}$.*

3.4. Discussion

Some of our predictions above are potentially generated by alternative explanations. Here, we discuss how our design addresses these confounds. First, a fresh predictor may have overestimated the WTW of a tired worker simply because he was uncertain about how onerous the task would become after working longer. Because fresh predictors had not yet completed 20 tasks themselves, they were unfamiliar with the state in which tired workers made decisions. Therefore, we are cautious to interpret an outcome in which $\widehat{W}^O(\$3, T|F) > \overline{W}(\$3, T)$ as stemming purely from projection bias. Note, however, that this limited-information channel is less relevant for tired predictors who guessed about the behavior of fresh workers—tired predictors had already experienced the fresh state. Accordingly, predictions cast by tired workers potentially provide a cleaner measure of projection bias. Furthermore, comparing the prediction errors across fresh and tired predictors sheds light on the extent to which this form of limited information distorted initial predictions. We present this analysis in Section 4.2.

A variant of the informational confound just discussed could still emerge, however, if tired predictors did not remember what it was like to be in the fresh state. In that case, a tired predictor might have rationally used his current state to approximate the fresh state because he was uncertain (due to limited memory) of what that state was like. To address this, our design attempted to make predictors' experience in the fresh state salient and memorable. First, predictors in groups I and O were interrupted after five tasks (i.e., while they were in the fresh state) in order to make predictions. Furthermore, the instructions explicitly told participants that they would later make predictions about the WTW of fresh workers, and,

hence, it behooved them to remember their attitude toward additional work while in that state.

Although requiring predictors in groups *I* and *O* to pause and make predictions when fresh likely provided them with useful information about that state, predictors in group O_{-1} did not make such predictions. Hence, they may have been less familiar with workers' sentiment in the fresh state (relative to groups *I* and *O*). As mentioned above, we tried to mitigate this concern by briefly interrupting participants in group O_{-1} after they completed five tasks (i.e., while they were in the fresh state) to deliver some of the instructions. In particular, we reminded them that they would later need to predict the WTW of workers in the fresh state, thereby emphasizing the value of remembering their current attitude toward additional work in that state.

Furthermore, group O_{-1} provided an important degree of control that groups *O* and *I* lacked. Namely, participants in groups *O* and *I* made several predictions in various states, and, hence, their predictions may have exhibited order effects. For instance, participants may have subconsciously anchored later predictions toward their initial guesses, or they may have deliberately chosen later predictions to appear consistent with their initial guesses. Leveraging data from group O_{-1} —as we do in the next section—allows for a between-subject analysis of projection bias that uses only initial guesses across groups and, thus, controls for any such order effects. To summarize: we designed group *O* to control for informational concerns and group O_{-1} to control for order effects.

4. Results on Interpersonal Projection Bias

In this section, we present our main findings. We first present the baseline WTW of fresh and tired workers and demonstrate that our manipulation of “tiredness” was successful. We then analyze predictors' average guesses about other workers' WTW and discuss evidence supporting our three hypotheses presented above. We also use within-subject variation in predictions to show that participants erroneously decreased their predictions about fresh workers as they themselves became tired and provide estimates of the degree of interpersonal projection bias. We conclude with a few additional analyses that further support projection bias over alternative explanations.

4.1. Willingness to Work Among Workers

We first present the aggregated willingness to work of fresh and tired workers. Table 1 shows the average WTW among these groups. Although raw responses are similar across the two groups (row (1)), recall that fresh workers stated their WTW for \$2, whereas tired workers stated their WTW for \$3. Row (2) accounts for

Table 1. Average Willingness to Work

	Workers' state		
	Fresh (5 tasks; \$2)	Tired (20 tasks; \$3)	Difference
Number of tasks	21.29 (1.383)	20.44 (1.224)	0.85 (1.847)
Tasks per dollar	10.64 (0.692)	6.81 (0.408)	3.83 (0.803)
Observations	300	299	

Notes. Standard errors are in parentheses. Difference in tasks-per-dollar row are significant at $p < 0.001$ (Welch's two-sided *t*-test).

these differential monetary incentives by showing WTW in terms of tasks per dollar. Under this normalization, we see that tiredness had a marked effect: average WTW when fresh was about 10.6 tasks per dollar versus 6.8 tasks per dollar when tired (difference significant at $p < 0.001$; Welch's two-sided *t*-test used throughout discussions of results, despite our directional hypotheses).¹⁴ This suggests that our tiredness manipulation succeeded at generating a meaningful change in participants' attitude toward work.¹⁵

Because tired workers faced a higher monetary bonus than fresh workers, it is worth noting that this would mechanically reduce their WTW in tasks-per-dollar if their effort costs were convex—even if our tiredness manipulation had no effect. Simple back-of-the-envelope calculations help demonstrate that our observed reduction in WTW did not stem entirely from this convexity effect. Toward a contradiction, suppose that the tiredness state had no effect on WTW and the effort-cost function was merely $C(e; s) = \psi e^2$ for some scaling parameter ψ .¹⁶ If a representative (fresh) participant with this cost function exerts the WTW reported in Table 1, then Equation (1) implies $\psi = \frac{2}{21.29^2} = 0.0044$. If we then use this estimate to calculate WTW for \$3, we would predict that tired workers would be willing to complete 26.11 tasks, or 8.70 tasks per dollar. However, tired workers in our experiment state an average WTW of 20.44 tasks for \$3; that is, 6.81 tasks per dollar.¹⁷ This difference suggests that even if workers faced substantially convex effort costs, our tiredness manipulation materially changed their WTW.

4.2. Main Results

We now examine predictors' guesses about workers' WTW and evaluate each of our three enumerated hypotheses from Section 3.3. Throughout this subsection, we continue to normalize WTW (and predictions thereof) in terms of tasks per dollar. This is purely expositional; none of our results rely on this normalization, and unnormalized predictions are presented in the appendix.

Table 2. Predictions of Willingness to Work, Same State (Tasks per Dollar)

Predictor's state	Prediction	True WTW	Difference
Fresh (after 5 tasks)	10.81	10.64	0.17
	(0.605) <i>n</i> = 223	(0.692) <i>n</i> = 300	(0.957)
Tired (after 20 tasks)	6.65	6.81	−0.17
	(0.230) <i>n</i> = 666	(0.408) <i>n</i> = 299	(0.439)

Notes. Standard errors are in parentheses. Differences are not significant ($p = 0.860$ and $p = 0.704$ top to bottom; Welch's two-sided t -test).

Hypothesis 1: Fresh Predictors Accurately Estimate the WTW of Fresh Workers, and Tired Predictors Accurately Estimate the WTW of Tired Workers. As highlighted by Table 2, when predictors were fresh, their average guesses matched the average WTW of fresh workers (difference not significant; $p = 0.855$). Likewise, when predictors were tired, they accurately guessed the WTW of tired workers (difference not significant; $p = 0.704$). Recall that only group I guessed about fresh workers when they themselves were fresh, whereas all three groups guessed about tired workers when they themselves were tired. Despite these unequal samples in this analysis, our results in Table 2 rule out large or systematic errors in predictions. We therefore find support for Hypothesis 1.

Hypothesis 2. Relative to Tired Predictors, Fresh Predictors Overestimate the WTW of Tired Workers. We now test whether fresh predictors overestimated the WTW of tired workers. As described in Section 3.3, we control for potential (state-independent) biases in predictions by comparing guesses of fresh predictors to those cast by tired predictors.

We present this information in Table 3. Note that only group O provided predictions about tired workers while

Table 3. State-Dependent Predictions (Tasks per Dollar)

Predictor's state	Workers' state	
	Fresh	Tired
Fresh (after 5 tasks)	10.81	10.22
	(0.605)	(0.491)
Tired (after 20 tasks)	9.46	6.65
	(0.345)	(0.230)
Difference	1.36	3.57
	(0.691)	(0.489)

Notes. Only specific groups made predictions reported in the "Fresh" row (the top left cell reports average predictions among group I , and the top right cell reports average predictions among group O). The "Tired" row reports predictions averaged across all groups. Sample sizes are (clockwise from top left) 223; 221; 666 and 666. Standard errors are in parentheses. Decreased variance in the Tired row reflects the fact that all predictors made two guesses when tired. Differences significant at $p = 0.049$ and $p < 0.001$ (left to right; Welch's two-sided t -test).

fresh, yet all three predictor groups did so while tired. Hence, the top right cell of Table 3 shows $\widehat{W}^O(\$3, T|F)$, and the bottom right cell shows $\widehat{W}^S(\$3, T|T)$ averaged over all members of groups $g \in \{I, O, O_{-1}\}$. (Below, we further discuss features of these data disaggregated across groups.) On average, fresh predictors guessed that tired workers had a WTW of 10.22 tasks per dollar, whereas tired predictors guessed 6.65. Thus, we find that guesses cast by fresh predictors were more than 50% higher than those cast by tired predictors (difference of 3.57 tasks per dollar, or 10.71 total tasks; $p < 0.001$). Moreover, given that tired predictors accurately estimated the WTW of tired workers, this result implies that fresh predictors significantly overestimated the true WTW of tired workers (they estimated 10.22, whereas the truth, from Table 1, was 6.81).¹⁸

Hypothesis 3. Relative to Fresh Predictors, Tired Predictors Underestimate the WTW of Fresh Workers. The "Fresh" column of Table 3 confirms this result. The top left cell shows $\widehat{W}^I(\$2, F|F)$ for group I (because only they made predictions about fresh workers while fresh), and the bottom left cell shows $\widehat{W}^S(\$2, F|T)$ averaged over all members of groups $g \in \{I, O, O_{-1}\}$. On average, tired predictors guessed that fresh workers have a WTW of 9.46 (tasks per dollar), whereas fresh predictors guessed 10.81. Hence, tired predictors provide significantly lower estimates than fresh predictors (difference of 1.36 tasks per dollar, or 2.72 total tasks; $p = 0.049$). This additionally implies that tired predictors underestimated the true WTW of fresh workers (they estimated 9.46, whereas the truth, from Table 1, was 10.64).

The "Difference" row of Table 3 shows the discrepancies in guesses cast by predictors in different states when holding the target group fixed. Notice that this difference is smaller for guesses about fresh workers than for guesses about tired workers (1.36 tasks per dollar versus 3.57), implying that tired predictors underestimated the WTW of fresh workers by a smaller degree than fresh predictors overestimated the WTW of tired workers. This reflects our discussion from Section 3.4: because fresh predictors faced uncertainty about how onerous the task would become, their guesses about tired workers may have been inaccurate not only due to projection bias, but also limited experience. By contrast, tired predictors had firsthand experience with the fresh state. Hence, their guesses about fresh workers involved less uncertainty, and the prediction error made by tired predictors therefore provides a cleaner measure of projection bias. As we discuss next, however, Table 3 understates the magnitude of this error because it aggregates data across predictor groups.

4.2.1. Analysis of Initial Predictions by Group. We now utilize the group-level variation in *initial* guesses.

As we emphasize here, disaggregating the data down to the group level allows us to control for any potential order effects by exclusively comparing initial guesses across groups.¹⁹

To see why it may be important to control for order effects, we first consider group *O*. This group first guessed about tired workers when they themselves were fresh (average guess of 10.22 tasks per dollar; see Table 3). They then guessed about fresh workers when they themselves were tired. The average of this second guess was relatively high (11.24 tasks per dollar; see Table A.1 in the appendix) and, accordingly, did not exhibit the underestimation that we’d expect under Hypothesis 3 because the true WTW of fresh workers was 10.64 (see Table 1). We suspect that these elevated second predictions stemmed from order effects—for example, anchoring or consistency bias. Namely, because their first guesses were very high—perhaps due to projection—their subsequent guesses may have been shifted upward as well. This would partially obfuscate our ability to detect projection amongst group *O*’s second predictions.

Fortunately, our experiment was specifically designed to allow us to sidestep such order effects by analyzing only the first predictions cast by each group. Indeed, predictors in group *O*₋₁, who first guessed about fresh workers when they themselves were tired, did significantly underestimate this WTW, consistent with Hypothesis 3. Although predictors in group *I* made accurate first guesses (as shown in Table 2), Table 4 shows that the first guesses among predictors in both Out Groups were systematically biased. Group *O*’s first guess—about tired workers when they themselves were fresh—was far too high, on average (average guess of 10.22 versus a truth of 6.81), whereas group *O*₋₁’s first guess—about fresh workers when they themselves were tired—was too low (average guess of 8.44 versus a truth of 10.64).²⁰ Furthermore, when focusing only on first guesses, the projection error among tired predictors guessing about fresh workers is now more pronounced

than it was in Table 3 (2.20 tasks per dollar in Table 4 versus 1.36 in Table 3). The previous estimate in Table 3 understates this degree of projection because it averaged over all groups and, thus, included the second guesses of group *O*, which were potentially shifted upward by order effects.²¹

In addition to being free of any potential order effects, the initial guesses made by group *O*₋₁ also involve a diminished role of uncertainty, as noted in our discussion of Hypothesis 3: tired predictors guessing about fresh workers did not face the same degree of uncertainty that fresh predictors faced when guessing about tired workers. Thus, we believe the prediction error in the initial guesses from group *O*₋₁ reflects our cleanest measure of the effect of projection bias on predictions. The magnitude of this error is 2.2 tasks per dollar, representing an underestimate of approximately 21% relative to the truth. This measure also allows us to loosely approximate the effect that uncertainty had on group *O*’s initial guesses. For this exercise, we assume that projection bias similarly led group *O* to overestimate the WTW of tired workers by 21%. Because this WTW is 6.8 tasks per dollar in truth, projection would then induce a prediction error of 1.4 tasks per dollar. But the observed prediction error is 3.4 (or 50% of the true WTW). The additional portion of this observed error, 3.4 – 1.4 = 2 tasks per dollar (or 29% of the true WTW), could then be a result of uncertainty about the task. In the next section, we quantify the degree of interpersonal projection bias using alternative approaches and find similar magnitudes.

4.3. Within-Subject Measures of Interpersonal Projection Bias

We now consider some within-subject measures of projection bias, leveraging the fact that each predictor made several guesses over time. This section proceeds as follows. We first examine how a predictor’s guesses about a fixed target group changed as the predictor moved from fresh to tired. We then estimate a simple, reduced-form variant of our model, allowing us to additionally account for how a predictor’s guesses depended on their own stated WTW. We utilize both approaches to quantify the overall bias.

We begin by noting a simple fact: predictors in group *I* changed their guesses about fresh workers after they completed additional tasks and became tired. In our opinion, this represents one of our strongest indicators of projection bias. Note that group *I*’s first guess about fresh workers was made while they themselves were fresh. They therefore had the same information—and tiredness—as the workers did when stating their WTW. Thus, additional exposure to the task should not have led these predictors to change their guesses about fresh workers. Despite this, we find that group *I* predictors significantly lowered their

Table 4. First Predictions vs. Workers’ WTW (Tasks per Dollar)

	Prediction	True WTW	Difference
Fresh Predictors → Tired Workers	10.22 (0.491) <i>n</i> = 221	6.81 (0.408) <i>n</i> = 299	3.40 (0.639)
Tired Predictors → Fresh Workers	8.44 (0.532) <i>n</i> = 222	10.64 (0.692) <i>n</i> = 300	-2.20 (0.873)

Notes. Predictions about tired workers (top row) are from group *O*. Predictions about fresh workers (bottom row) are from group *O*₋₁. Standard errors are in parentheses. Differences are significant at *p* < 0.001 and *p* = 0.012 (top to bottom; Welch’s two-sided *t*-test).

guesses about fresh workers once they themselves became tired: $\widehat{W}^I(\$2, F|F) - \widehat{W}^I(\$2, F|T) = 4.22$ total tasks (*s.e.* = 0.858, where *s.e.* is standard error; $p < 0.001$ for difference). Although the predictors' initial guesses about fresh workers were well calibrated (see Table 2), their revised guesses were approximately 19% lower than their first, and they ultimately underestimated the true WTW of fresh workers by about 21%. As we show below, these revisions significantly reduced the expected earnings for group *I*.

We likewise note that predictors in group *O* changed their guesses about tired workers once they became tired. Recall that group *O* first guessed about tired workers while fresh. At that time, they potentially lacked information about how it felt to be tired. We may therefore expect a relatively large revision in their predictions, stemming from both projection and the resolution of this uncertainty. Indeed, the average revision in guesses about tired workers was $\widehat{W}^O(\$3, T|F) - \widehat{W}^O(\$3, T|T) = 8.02$ total tasks (*s.e.* = 1.018, $p < 0.001$ for difference). On average, their revised guesses were approximately 26% lower than their first.

Recall that we also collected predictors' own (hypothetical) WTW while tired. As a final approach to measuring projection bias, we incorporate these data to estimate a simple reduced-form model motivated by our theoretical framework in Section 3.3. Our specification in that section (and other common models of projection bias) assumes that the parameter capturing projection bias, α , enters as a convex combination of utilities across states (see Equation (2)). However, because we instead observe willingness to work, we estimate a parameter that captures a convex combination of the optimal WTW across states. More specifically, we assume that a projector's prediction about a worker's WTW is a convex combination of her own WTW in her current state and her unbiased estimate of a worker's WTW in the target state. As above, let $W(m, s|\theta)$ be the utility-maximizing WTW of a participant facing payment m in state s , where θ represents her idiosyncratic taste for the task. Predictor i 's guess about the WTW of a worker in state s facing payment m is then

$$\widehat{W}(m, s|s_i, \theta_i) = \rho W(m, s_i|\theta_i) + (1 - \rho)\mathbb{E}_\theta[W(m, s|\theta)|\theta_i], \quad (3)$$

where $\mathbb{E}_\theta[\cdot|\theta_i]$ denotes predictor i 's expectation over θ , conditional on herself having type θ_i , and s_i is her state when casting this prediction. Parameter ρ therefore measures the extent to which the predictor's estimate is biased toward her own current WTW.

We take this model to our data as follows. First, note that all predictors made two guesses when they were in the tired state. Using Equation (3), we can write these

two predictions as

$$\begin{aligned} \widehat{W}(\$2, F|T, \theta_i) \\ = \rho W(\$2, T|\theta_i) + (1 - \rho)\mathbb{E}_\theta[W(\$2, F|\theta)|\theta_i], \end{aligned} \quad (4)$$

and

$$\begin{aligned} \widehat{W}(\$3, T|T, \theta_i) \\ = \rho W(\$3, T|\theta_i) + (1 - \rho)\mathbb{E}_\theta[W(\$3, T|\theta)|\theta_i]. \end{aligned} \quad (5)$$

Recall that we elicited predictors' own (hypothetical) WTW for \$3 when tired, but not for \$2; thus, we measure $W(\$3, T|\theta_i)$, but not $W(\$2, T|\theta_i)$. In order to obtain a conservative estimate of ρ with this limited data, we impose a linear effort-cost function, implying that $W(\$2, T|\theta_i) = \frac{2}{3}W(\$3, T|\theta_i)$. Substituting this value into Equation (4) and then differencing Equations (4) and (5) yield

$$\begin{aligned} \widehat{W}(\$3, T|T, \theta_i) - \widehat{W}(\$2, F|T, \theta_i) \\ = \rho \left(\frac{1}{3}W(\$3, T|\theta_i) \right) + (1 - \rho)\mathbb{E}_\theta[W(\$3, T|\theta) \\ - W(\$2, F|\theta)|\theta_i]. \end{aligned}$$

We estimate this—given the data we observe—with the following econometric model:

$$\text{Diff}_i = \beta_0 + \beta_1 \left(\frac{1}{3} \text{WTW}_i \right) + \epsilon_i, \quad (6)$$

where Diff_i is the observed difference in predictor i 's guesses cast while tired, WTW_i is her own willingness to work, and β_1 corresponds to ρ . When pooling all predictors and estimating via ordinary least squares, this estimation yields $\rho = 0.23$ (*s.e.* = 0.061).²² However—as suggested above—treating costs as linear yields a conservative estimate of ρ .²³ Thus, although our design was not optimized for this parametric approach, it still provides estimates of projection bias that are consistent with our nonparametric results.

4.4. Additional Analyses

In this section, we present a few additional analyses that provide further support for projection as the mechanism underlying our findings. We show that learning from experience with the task was not the primary driver of our effects by examining (i) how the accuracy of a predictor's guesses changed as they accumulated more experience with the task, and (ii) self-reported confidence ratings across guesses. We then explore whether the amount of time it took participants to complete the tasks affected their predictions; we find no such effect.

We first evaluate how participants' guesses improved (or failed to improve) with more task experience. Table 5 shows the mean absolute error in each

Table 5. Prediction Accuracy (Mean Absolute Error) by Group

	Group <i>I</i>	Group <i>O</i>	Group <i>O</i> ₋₁
Mean abs error, 1st prediction	11.29 (0.942)	15.32 (1.256)	—
Mean abs error, 2nd prediction	10.33 (0.787)	13.38 (1.087)	11.01 (0.820)
Mean abs error, 3rd prediction	11.40 (0.763)	12.85 (1.081)	10.42 (0.895)

Notes. Standard errors are in parentheses. Abs, absolute.

guess for each group. We see that predictors’ guesses became slightly more accurate with time, on average. However, this improvement was state-dependent: when predictors were guessing about workers who shared their state, they tended to be more accurate than when guessing about workers in the opposite state. Pooling all of the guesses that were cast by tired predictors (i.e., after accumulating experience), we find that same-state guesses were significantly more accurate than different-state guesses (difference 0.730, $p = 0.032$).

Furthermore, group *I* helps show that any improvements in accuracy stemming from experience must have been relatively small compared with the reduction in accuracy associated with being in a different state than the target group. Recall that group *I*’s first two guesses were about workers in their same state (fresh and tired, respectively). Table 5 suggests that group-*I* predictors’ second guess was more accurate than their first, and, hence, they may have improved in their ability to estimate others in their same state. Group *I*’s third guess, however, was about fresh workers while they themselves were tired and was less accurate than either of their initial guesses. Thus, experience with guessing about those in their same state was not sufficiently beneficial to overcome the reduction in accuracy stemming from projection bias (i.e., misalignment between predictors’ and workers’ states). Moreover, although the reduction in group *I*’s accuracy between their first and third guesses does not appear significant in Table 5, it indeed came at a considerable cost: group *I*’s expected earnings significantly decreased between their first and third predictions. Specifically, the number of guesses within ± 5 tasks of the true WTW—and, thus, guesses that could have increased earnings—fell by approximately 26% ($p = 0.002$ for difference).

Overall, we believe that these limited improvements in accuracy—along with our results on confidence, below—suggest that there may have been some learning from experience, but that this learning does not fully account for many of the effects that we observe.

We now evaluate predictors’ confidence ratings, providing further evidence that learning about the disutility of work does not drive our results. Recall that after

each prediction, participants reported their confidence on a 5-point scale, where 1 represents “not at all confident” and 5 represents “extremely confident.” Average responses are reported in Table A.1 in Online Appendix A. We first consider the confidence of groups *I* and *O*, as these groups made predictions while both fresh and tired. As shown in Table A.1, average confidence did not increase with experience—neither in going from the first prediction to the second (and thus accumulating more experience with the task) nor in going from the second prediction to the third (and thus accumulating more experience with predicting). In sum, predictors did not grow more confident as they accrued experience.

In an ex post analysis, we discovered that predictors who were extremely confident tended to be *less* accurate (à la Kruger and Dunning 1999). Although this is consistent with the classic Dunning-Kruger effect, this correlation is also predicted by projection bias: as the extent of projection increases, a predictor believes that she has a more precise assessment of others because she is more confident that others will act like herself. At the same time, an increase in projection leads to a greater bias in predictions. Hence, it induces a negative correlation between confidence and accuracy. To explore whether this correlation indeed stems from projection, we pooled predictors from groups *I* and *O*, and we split them into two new groups: (i) “high confidence” predictors, who responded with “extremely confident” to at least one of the confidence questions; and (ii) predictors who never responded with “extremely confident.” We then calculated a crude measure of projection for each predictor: how much, in percentage terms, they revised their first guess after they became tired.²⁴ Those with extremely high confidence changed their guesses by 26.7%, on average, whereas those with nonextreme confidence changed their guesses by 14.6%, on average ($p = 0.032$ for difference). We believe this provides additional suggestive evidence for projection, insofar as strongly biased projectors exhibited extreme confidence because they believed—either directly or inattentively—that their own attitude toward work was very informative about others’.

Finally, we briefly consider task completion time and its (null) effect on the degree of projection. Ex post, we believed that those who took longer to complete the tasks might be more fatigued. Thus, we believed that relatively slow predictors might exhibit a greater degree of projection when asked about fresh workers. Our data do not bear this out. We present a series of exploratory analyses in Online Appendix A.2, which demonstrate that task completion time has no effect on participant predictions or their self-reported confidence. We suspect that this null result may stem from the fact that much of the heterogeneity in task-completion times arose due to inattention (e.g.,

doing other things online), but we have no direct evidence of this.

5. Comparison with Intrapersonal Projection Bias

Intrapersonal projection bias—the propensity for one’s current state to overly influence predictions about one’s *own* behavior in a different state—is well-documented (see below for examples and discussion). To provide evidence for this in our domain, we ran an additional worker group (called “Predicting Workers”). This group was identical to our Tired-Workers group, except that participants predicted their own WTW ahead of time. This allows us to measure the extent to which fresh workers mispredicted their own WTW once tired. Specifically, after a predicting worker completed 5 mandatory rounds (out of 20), we asked them to predict how many additional rounds they would complete for a bonus of \$3 once they had finished the mandatory 20 rounds. Thus, while in the fresh state, these participants predicted their own attitude toward work in the tired state.²⁵ Then, after completing the mandatory 20 tasks, we asked participants how many additional tasks they would complete for a bonus of \$3. We elicited this WTW exactly as in the Tired-Workers group.

This additional group allows us to measure the extent to which participants mispredicted their own behavior. Table 6 shows the predictions and actual WTW among predicting workers. As in the previous section, we take the difference between the predicted and actual WTW as a raw metric of projection bias: on average, fresh workers overestimated their own WTW when tired by roughly five tasks—approximately 30% of their true WTW. Perhaps more dramatically, 93 out of 298 participants overestimated their WTW in a costly way: their prediction was more than five tasks higher than their true WTW, which prevented them from earning the bonus.²⁶

Interpreting this number, however, requires some caution. First, these mispredictions about future WTW came from workers in the fresh state who had not yet experienced the tired state. Hence, these mispredictions may have stemmed from predictors underestimating how onerous the task would become. Because this force acts in the same direction as projection, the 30% error noted above may overstate the degree of intrapersonal

projection. In contrast, by monetarily incentivizing participants’ predictions, we may have indirectly incentivized consistency. Namely, stating a WTW close to one’s prediction would have increased a person’s payout (relative to stating a different WTW). Because consistency acts against projection bias, the 30% error may also understate the degree of intrapersonal projection.²⁷

To assess the relative magnitudes of intrapersonal and interpersonal projection bias, we compare prediction errors of the predicting workers with those of the fresh predictors who guessed the WTW of tired workers. Recall that fresh predictors overestimated the WTW of others in the tired state by roughly 10.25 tasks (see Table 3)—approximately 50% of tired workers’ true WTW—whereas fresh workers overestimated their *own* WTW by approximately 30%. Thus, despite significant biases among both groups, participants were better calibrated when making predictions about themselves rather than about others.

This finding suggests that the intrapersonal prediction error is substantially smaller than the interpersonal one. However, this interpretation comes with some critical caveats. The difference in the prediction errors noted above (30% versus 50%) captures the difference in the degree of intrapersonal and interpersonal projection bias only if we assume that uncertainty about how onerous the task would become was similar when considering oneself and considering others.²⁸ Importantly, Table 4 (and the surrounding discussion) suggests that the interpersonal error we observe among fresh predictors stems from both uncertainty and projection and that their relative contributions are roughly similar in magnitude. If the 30% intrapersonal error we document here stems from the same composition of uncertainty and projection as the interpersonal error, then we would conclude that interpersonal projection is stronger. This simple arithmetic allows us to consider alternative assumptions on the relative composition of these effects. For instance, it is likely that participants have less uncertainty about themselves than about others. If we observed a small difference in uncertainty across self and other, we would still conclude that interpersonal projection bias is stronger than intrapersonal. But if this difference was sufficiently large (e.g., in the extreme case where participants faced *no* uncertainty about themselves), then we would instead conclude that interpersonal projection is slightly weaker than intrapersonal projection. As an additional caveat, any comparison across the intrapersonal and interpersonal domains should be interpreted with care due to the potential for different measurement errors. Because the prediction error in the intrapersonal case is smaller, any measurement error would loom larger.

That said, it is intuitive that the interpersonal prediction error may be larger than the intrapersonal one. In the interpersonal case, there are more factors that might

Table 6. Predicting Workers’ Guesses and WTW

	Prediction	Actual	Difference
WTW (# of tasks)	22.11 (1.179)	17.02 (1.161)	5.09 (1.044)
Observations	298	298	298

Notes. Standard errors are in parentheses. Difference significant at $p < 0.001$ (Welch’s two-sided t -test).

be projected: people could, for instance, project their intrinsic taste for the task, captured by θ_i . By contrast, in the intrapersonal case, there is limited scope for one's stable, intrinsic taste to distort their predictions. Moreover, past work on perspective taking (see, e.g., Van Boven and Loewenstein 2005) suggests that interpersonal predictions across states entail two distinct judgments: (i) what one's own choice would be in an alternative state; and (ii) how similar another person's choice would be to their own. Both steps in this dual-judgment process are prone to mistakes. Table 6 reveals an error in the first step (i.e., the intrapersonal one), and, thus, interpersonal prediction errors may be larger because they additionally reflect any error in the second step.

Our measure of intrapersonal projection bias falls in the range of existing estimates in the literature. These measures come from a variety of different domains and different estimation schemes; accordingly, there is no a priori reason that our results should be the same as others. Nevertheless, we find a good deal of agreement. For example, Loewenstein and Adler (1995) find that unendowed people underappreciate how the endowment effect will alter their selling price by about 31%. Other papers structurally estimate the extent of intrapersonal projection. Conlin et al. (2007) find $\alpha \in [0.31, 0.50]$ for cold-weather clothing-catalog sales, whereas Augenblick and Rabin (2019) find $\alpha \in [0.27, 0.53]$ in a real-effort experiment.²⁹ Our measures thus accord with an emerging consensus on the magnitude of projection bias observed across a variety of domains.³⁰

6. Conclusion

In this paper, we provide evidence that interpersonal projection bias leads to substantial and costly errors in forecasting others' behavior. Specifically, we find that predictors correctly guessed the behavior of others in their own state, but fresh predictors systematically guessed that tired workers would behave as if they, too, were fresh, and tired predictors guessed that fresh workers would behave as if tired. Additionally, we find evidence for intrapersonal projection in the same domain, and this error is likely smaller in magnitude than interpersonal projection. Our evidence suggests that neither uncertainty about the task nor learning were the root cause of these errors. Rather, our results accord with a simple model of interpersonal projection bias.

Our results also suggest that projection bias may be one potential source of inaccuracy documented among forecasters attempting to predict others' response to incentives (e.g., DellaVigna and Pope 2018a, b, 2022). To illustrate how such forecasts may go awry, consider a projecting manager who is forecasting her workers' behavior. Such a manager will underestimate how workers' marginal disutility of effort differs across

states and fail to anticipate how the motivating effect of incentives changes over time. We find suggestive evidence of this misperceived gap in responsiveness to incentives. As reported in Table 1, the true difference in the WTW between fresh and tired workers was 3.83 tasks per dollar. Tired predictors, for instance, underestimated this difference by 27% (see Table 3). Future work could be tailored to more directly test how these perceived differences depend on a predictor's current state.³¹

There are several other avenues for further research. Our evidence shows that people project "tiredness," which is in line with the literature on projection of visceral states (see, e.g., Van Boven and Loewenstein 2003 for a discussion on this point). However, it may be that the degree of projection decreases in contexts where differences in preferences are driven by less visceral factors. Nevertheless, other barriers to perspective taking may still prevent accurate forecasts.³² Better understanding when these various barriers emerge and how to mitigate them may greatly improve how we predict the behavior of others who face differing perspectives or circumstances.

Additionally, we have primarily modeled and interpreted our results as stemming from the projection of preference-relevant states. We've adopted this interpretation following the large psychology and economics literature that does the same (see, e.g., Loewenstein et al. 2003 and Van Boven et al. 2013). However, other forms of projection—namely, information projection (as in, e.g., Camerer et al. 1989, Madarász 2012, Madarász et al. 2023)—could also play a role. In Madarász's (2012) model of information projection, an individual wrongly acts as if others might know his private information. In our setting, information projection may be relevant for tired predictors if they gained information about the onerousness of the task when doing extra work. For example, as predictors worked more (that is, became tired), they may have come to learn that the task was more onerous than expected and projected this new information onto others. As such, tired predictors may have underestimated fresh workers' WTW not only because they wrongly treated them as more tired than they were (as in our interpretation), but also because they wrongly thought these fresh workers had additional information about the task that they didn't yet have. Such information projection may be a partial explanation for our effect, but given that mispredictions about others seem to be significantly more biased than participants' mispredictions about their own future attitudes toward the task, we suspect that there is scope for both forms of projection.³³ Furthermore, information projection has difficulty explaining the systematic pattern of mispredictions made by fresh predictors because they don't have additional information to project. But it is inherently difficult to

fully distinguish between state-based preference projection and information projection when the information at hand concerns a person's marginal utility. Future work could further illuminate when and how these two specific forms of interpersonal projection differentially distort predictions in order to, for instance, inform the design of strategies to debias forecasters. If an intervention was limited to highlighting either others' states or their information, the optimal approach would depend on the relative magnitude of these effects.

Acknowledgments

The authors thank Andy Brownback, Jon X. Eguia, Christine Exley, Laura Gee, Marc Kaufmann, Kristof Madarász, Kirby Nielsen, Matthew Rabin, Antonio Rosato, Joshua Schwartzstein, and audiences at the Economic Science Association World Conference in Vancouver, the Early-Career Behavioral Economics Conference (2021), Chapman University, Michigan State University, Purdue University, and the University of Michigan for comments.

Appendix. All Predictions by Group and Timing

Recall that the true WTW of fresh workers was 21.3 tasks for \$2, whereas that of tired workers was 20.4 tasks for \$3. Table A.1 shows all predictions of these quantities for each predictor group.

Table A.1. All Predictions (Number of Tasks)

State guessing about	In group		Out group O		Out group O_{-1}	
	Fresh	Tired	Fresh	Tired	Fresh	Tired
Fresh (after 5 tasks)	21.62 (1.209)	17.40 (1.048)	–	22.48 (1.410)	–	16.89 (1.065)
Tired (after 20 tasks)	–	18.95 (1.044)	30.65 (1.473)	22.62 (1.378)	–	18.27 (1.127)
Observations	223	223	221	221	222	222

Note. Standard errors are in parentheses.

Endnotes

¹ Marks and Miller (1987) document the false-consensus effect in 45 different studies published in the decade following Ross et al. (1977). These studies generally elicited subjects' responses to binary-choice questions (e.g., "Would you vote for a bill to increase space-program funding?") and asked subjects to predict how the general population would answer the same questions. The false-consensus effect is observed when the average estimate of the fraction that supported a given choice was larger among those who supported that choice than those who did not (e.g., those who voted for space-program funding predicted that the bill would receive more support than those who voted against it). More recently, Selten and Ockenfels (1998) extend the false-consensus-effect paradigm to the expectation of others' behavior in a conditional gift-giving game.

² Buchanan (2020) similarly finds that participants neglected the effect of others' endowments when trying to predict their risk attitudes and acted as if others shared their own endowment.

³ Relatedly, economists have documented several instances of intrapersonal projection bias, where people project their current preferences

onto their future selves and thus exaggerate the similarity between their current and future tastes (Conlin et al. 2007, Busse et al. 2015, Chang et al. 2018, Augenblick and Rabin 2019; additional evidence is discussed in Loewenstein et al. 2003). We return to a discussion of intrapersonal projection in Section 5.

⁴ Furthermore, we told participants ahead of time that they would make predictions about the willingness to work of others who had completed 5 and 20 tasks. Thus, predictors (ostensibly) knew that it was in their interest to pay attention to and recall their sentiment toward work at these two points.

⁵ Research focusing on separate questions has documented results akin to the false-consensus effect. For example, one line of papers suggests that people project their own social preferences or beliefs in settings involving guilt aversion, solidarity, and trust (e.g., Selten and Ockenfels 1998, Charness and Dufwenberg 2006, Vanberg 2008, Ellingsen et al. 2010, Blanco et al. 2011, Butle et al. 2016). In a related vein, Ziegler et al. (2022) show that those whose morals have been eroded are overly pessimistic about the morals of others trading in the market.

⁶ Our design shares a feature with that of Ambuehl et al. (2021): choice. Architects in their experiment also had firsthand experience making the choices faced by others. Our design goes further. Because we directly altered a predictor's own preferences by changing her state of fatigue, we can examine whether and how firsthand experience with *different preferences* matters. This is akin to a Choice Architect in their experiment making choices for others when she is both patient and impatient; such an exercise is obviously infeasible in their setting.

⁷ Participants were recruited to meet the following criteria: (i) over 18 years old; (ii) resident of the United States (verified with IP address); and (iii) completion of at least 100 prior Human Intelligence Tasks on MTurk with a 95% acceptance rate. All data were collected in June 2019, prior to the COVID-19 pandemic. Complete experimental instructions are in Online Appendix B.

⁸ As we discuss below, these measures of WTW were the objects that predictors had to guess. We varied the monetary incentives across the two worker groups to ensure that a predictor faced distinct questions when asked about the two groups. This was intended to promote independent assessments for each prediction. Had we asked predictors the same question repeatedly, we may have introduced a consistency bias.

⁹ Note that this mechanism is not incentive compatible for eliciting point estimates for beliefs. In particular, any prediction below 5 (or above 95) is strictly dominated by simply guessing 5 (or 95). However, out of the 1,776 total predictions we collect, only 4% fall outside the interval [5,95]. Dropping these responses does not substantively change any of our results.

¹⁰ Adding incentives to this last question would have substantially increased the length of the experiment for predictors because participants would have actually had to complete additional work. Given the (already long) duration, we opted to collect hypothetical WTW instead. It is worth noting that the average WTW from this unincentivized elicitation is similar to the average WTW of our tired (incentivized) workers.

¹¹ Equation (2) is inherently a model of interpersonal projection rather than intrapersonal projection because it does not presume that participant i miscalculates her own future costs. That said, we could additionally capture intrapersonal projection by replacing $C(e; s_i, \theta_i)$ in Equation (2) with participant i 's current misperception of her cost function. Following Kaufmann (2022), an appropriate specification of one's own misperceived costs due to intrapersonal projection bias in our context would be $\tilde{C}(e; s_i, \theta_i) = \tilde{\alpha} c'(s_i; \theta_i)e + (1 - \tilde{\alpha})[c(e + s_i; \theta_i) - c(s_i; \theta_i)]$, where $\tilde{\alpha} \in [0, 1]$ captures the degree of this bias. Although we present reduced-form evidence of interpersonal projection in

Section 5, our results in that section could be interpreted using this newly stated model.

¹² Note that participant i 's prediction about j , $\widehat{W}(m, s_j | s_i, \theta_i)$, also depends on i 's beliefs about θ_j . We assume a predictor is Bayesian aside from the misspecified model of costs presented in (2). Thus, $\widehat{W}(m, s_j | s_i, \theta_i)$ is the expected value of effort that maximizes (2), given predictor i 's beliefs over θ_j .

¹³ Two recent experiments highlight how errors distinct from projection bias can warp interpersonal predictions. Frederick (2012) shows that people generally overestimate others' willingness to pay for goods, and Kurt and Inman (2013) demonstrate that both endowed and nonendowed participants are inaccurate in their predictions about others in their *same* state.

¹⁴ Our basic finding that there is a difference in WTW is a necessary step to identify projection bias. In a different context—gym attendance—März (2019; commenting on Acland and Levy 2015) notes that a similar “first stage” effect is notably absent once the appropriate estimation technique is applied.

¹⁵ We observe small differences in attrition across treatment groups. A total of 55 workers failed to complete the experiment in the “fresh” treatment, whereas 88 failed to complete the experiment in the “tired” treatment (and 87 in the additional tired treatment in which they had to predict about themselves; see Section 5). Conditional on staying in the experiment for 60 seconds (and, thus, plausibly reaching the screen that announced the treatment assignment), 89.8% (81.1%) of workers assigned to the fresh (tired) treatment completed the experiment.

¹⁶ For this exercise, we ignore idiosyncratic differences in the cost function and, hence, drop the θ_i term introduced in Section 3.3.

¹⁷ Following our setup in Section 3, a more realistic model might be that e additional tasks after having already completed s yields a cost of $C(e; s) = \psi(e + s)^\gamma - \psi(s)^\gamma$. Fitting this specification to match the average behavior of fresh and tired workers, we find $\psi = 0.01$ and $\gamma = 1.66$.

¹⁸ The first row of Table 3 also highlights that predictors were aware of the different monetary incentives across groups. This is most apparent when considering predictions in terms of the absolute number of tasks rather than tasks per dollar. Fresh predictors guessed that the WTW of fresh workers for \$2 is 20.62 tasks and that the WTW of tired workers for \$3 is 30.66 tasks. Thus, predictors clearly expect that higher monetary payments will increase workers' WTW.

¹⁹ See Table A.1 in the appendix for all disaggregated predictions.

²⁰ Note that Figure 1 in the introduction depicts the distribution of responses underlying Table 4.

²¹ Group I was similar to group O_{-1} in that they also underestimated the WTW of fresh workers when they themselves were tired. On average, they guessed 8.7 tasks per dollar, whereas the truth was 10.64 (see Table A.1).

²² Allowing the intercept to vary for each of the three groups (I , O , and O_{-1}) leads us to estimate $\rho = 0.22$ (*s.e.* = 0.063) and, thus, does not substantively alter the results.

²³ To illustrate this claim, suppose each tired predictor i had a normalized cost of additional effort given by $\theta_i(e)^\gamma$. Our observations $W(\$3, T | \theta_i)$ and $W(\$2, T | \theta_i)$ would solve $\theta_i W(\$3, T | \theta_i)^\gamma = 3$ and $\theta_i W(\$2, T | \theta_i)^\gamma = 2$, respectively. It then follows that $W(\$2, T | \theta_i) = (2/3)^{1/\gamma} W(\$3, T | \theta_i)$, and the regressor in Equation (6)—which captures $W(\$3, T | \theta_i) - W(\$2, T | \theta_i)$ —would then be $(1 - (2/3)^{1/\gamma}) W(\$3, T | \theta_i)$. Our estimation assumes $\gamma = 1$. As γ increases above one, the multiplier term $(1 - (2/3)^{1/\gamma})$ decreases; this mechanically increases the estimate of β_1 (and, hence, ρ). For instance, if we assume costs are quadratic instead of linear (i.e., $\gamma = 2$), then we would estimate $\rho = 0.41$ (*s.e.* = 0.111).

²⁴ This is the percentage change between a predictor's first and third guesses. Note that this is the same nonparametric measure of projection bias considered in Section 4.3.

²⁵ These predictions were incentivized in the same way as other predictions in this experiment: participants earned the bonus if their prediction of their own WTW was within five of their subsequent stated WTW.

²⁶ Figure A.1 in Online Appendix A shows the distribution of individual differences in predictions versus actual WTW. The distribution is skewed toward positive values. As noted, 93 subjects out of 298 overestimated their subsequent WTW to a degree that reduced their payoffs, whereas a substantially smaller fraction (23 out of 298) underestimated their WTW in a similarly costly way.

²⁷ Comparing Table 6 with Table 1 reveals that predicting workers were significantly less willing to work than tired workers (difference of 3.42 total tasks; significant at $p = 0.045$). Recall that these two groups were nearly identical, except the former made predictions about their eventual WTW, and the latter did not. Hence, stating predictions seemed to have a *negative* effect on eventual effort. This finding stands in contrast to research suggesting that stated goals form a motivational reference point (e.g., Heath et al. 1999). However, our experiment is not well-suited to draw such conclusions.

²⁸ An extensive psychology literature suggests that such a symmetry is likely (see Van Boven et al. 2013 for a review).

²⁹ Augenblick and Rabin (2019) consider a real-effort experiment similar to our domain and find that projection bias leads tired workers to commit to doing fewer tasks in the future than their fresh counterparts. The authors offer caution in the precision of their estimates of α because their estimation procedure requires strong assumptions on the effort-cost function. Moreover, their experiment also examines present bias, and their ability to separately measure projection bias is somewhat limited by their design.

³⁰ Although there are a number of related studies, many—particularly early experimental studies—are not suited to estimate the degree of projection. Likewise, some recent empirical papers do not parametrically estimate projection bias directly, but find support for its main premise (though other possible explanations are also offered, such as salience and overweighting situational cues). For example, Chang et al. (2018) find that Chinese consumers are more likely to purchase health insurance on days with high pollution and are likely to reverse this decision (during a cooling-down period) when pollution drops. Busse et al. (2015) find that people are more likely to buy a convertible car on sunny days than on overcast or rainy days.

³¹ Our experiment was not designed to precisely estimate this difference-in-differences across groups. However, the main hypotheses that we test and confirm (Section 4.2) imply this result.

³² Epley and Caruso (2009) describe three forms of barriers: (i) failing to even think about differing perspectives; (ii) failing to fully adjust away from one's own current perspective; and (iii) holding an inaccurate sense of what perspectives others may hold. Our evidence is most in line with the second barrier, yet the first may be more prominent when salient visceral or emotional factors are absent.

³³ In our setting, the magnitude of bias stemming from information projection depends on the degree to which fresh workers were initially miscalibrated about the eventual onerousness of the task. If, for example, this miscalibration was relatively small, then gaining access to the knowledge of a tired individual should have little effect on a fresh worker's WTW. We can use the self-predictions data (reported in Table 6) to get a sense of any such miscalibration in our setting: fresh workers guessed that their own WTW would be 22.11 tasks when tired, whereas the actual WTW among tired workers was 20.44 (Table 1). Thus, miscalibration among fresh workers appears small relative to the magnitude of prediction errors that we observe. Recall that fresh predictors guessed that the WTW of tired workers was 30.65—considerably higher than guesses about their own WTW when tired.

References

- Acland D, Levy M (2015) Naiveté, projection bias, and habit formation in gym attendance. *Management Sci.* 61(1):146–160.
- Ambuehl S, Bernheim D, Ockenfels A (2021) What motivates paternalism? An experimental study. *Amer. Econom. Rev.* 111(3):787–830.
- Augenblick N, Rabin M (2019) An experiment on time preference and misprediction in unpleasant tasks. *Rev. Econom. Stud.* 86(3):941–975.
- Badger G, Bickel W, Giordano L, Jacobs E, Loewenstein G, Marsch L (2007) Altered states: The impact of immediate craving on the valuation of current and future opioids. *J. Health Econom.* 26(5):865–876.
- Benjamin D (2019) Errors in probabilistic reasoning and judgment biases. Bernheim BD, DellaVigna S, Laibson D, eds. *Handbook of Behavioral Economics: Applications and Foundations*, vol. 2 (Elsevier, Amsterdam), 69–186.
- Blanco M, Engelmann D, Normann H (2011) A within-subject analysis of other-regarding preferences. *Games Econom. Behav.* 72:321–338.
- Bohren A, Hauser D (2021) Learning with model misspecification: Characterization and robustness. *Econometrica* 89:3025–3077.
- Buchanan J (2020) My reference point, not yours. *J. Econom. Behav. Organ.* 171:297–311.
- Busse B, Pope D, Pope J, Silva-Risso J (2015) The psychological effect of weather on car purchases. *Quart. J. Econom.* 130(1):371–414.
- Butler J, Giuliano P, Guiso L (2016) The right amount of trust. *J. Eur. Econom. Assoc.* 4(5):1155–1180.
- Camerer C, Loewenstein G, Weber M (1989) The curse of knowledge in economic settings: An experimental analysis. *J. Polit. Econom.* 97(5):1232–1254.
- Chang T, Huang W, Wang Y (2018) Something in the air: Pollution and the demand for health insurance. *Rev. Econom. Stud.* 85(3):1609–1634.
- Charness G, Dufwenberg M (2006) Promises and partnership. *Econometrica* 74:1579–1601.
- Conlin M, O'Donoghue T, Vogelsang T (2007) Projection bias in catalog orders. *Amer. Econom. Rev.* 97(4):1217–1249.
- Cronbach L (1955) Processes affecting scores on 'understanding others' and 'assumed similarity'. *Psych. Bull.* 52:177–193.
- Dawes R (1989) Statistical criteria for establishing a truly false consensus effect. *J. Experiment. Soc. Psych.* 25(1):1–17.
- Dawes R (1990) The potential nonfalsity of the false consensus effect. Hogarth RM, ed. *Insights in Decision Making: A Tribute to Hillel J. Einhorn* (University of Chicago Press, Chicago), 179–199.
- DellaVigna S, Pope D (2018a) Predicting experimental results: Who knows what? *J. Polit. Econom.* 126(6):2410–2456.
- DellaVigna S, Pope D (2018b) What motivates effort? Evidence and expert forecasts. *Rev. Econom. Stud.* 85(2):1029–1069.
- DellaVigna S, Pope D (2022) Stability of experimental results: Forecasts and evidence. *Amer. Econom. J. Microeconom.* 14(3):889–925.
- DellaVigna S, Pope D, Vivaldi E (2019) Predict science to improve science. *Science* 366(6464):428–429.
- Ellingsen T, Johannesson M, Torsvik G, Tjøtta S (2010) Testing guilt aversion. *Games Econom. Behav.* 68(1):95–107.
- Engelmann D, Strobel M (2000) The false consensus effect disappears if representative information and monetary incentives are given. *Experiment. Econom.* 3:241–260.
- Engelmann D, Strobel M (2012) Deconstruction and reconstruction of an anomaly. *Games Econom. Behav.* 76:678–689.
- Epley N, Caruso E (2009) Perspective taking: Misstepping into others' shoes. Markman KD, Klein WMP, Suhr JA, eds. *Handbook of Imagination and Mental Simulation* (Psychology Press, Hove, UK), 295–309.
- Frederick S (2012) Overestimating others' willingness to pay. *J. Consumer Res.* 39(1):1–21.
- Frick M, Iijima R, Ishii Y (2020) Misinterpreting others and the fragility of social learning. *Econometrica* 88:2281–2328.
- Gagnon-Bartsch T (2016) Taste projection in models of social learning. Working paper, Harvard University, Cambridge, MA.
- Gagnon-Bartsch T, Pagnozzi M, Rosato A (2021) Projection of private values in auctions. *Amer. Econom. Rev.* 111:3256–3298.
- Goeree J, Grosser J (2007) Welfare-reducing polls. *Econom. Theory* 31(1):51–68.
- Gordon RM (1986) Folk psychology as simulation. *Mind Language* 1(2):158–171.
- Heath C, Larrick R, Wu G (1999) Goals as reference points. *Cognitive Psych.* 38(1):79–109.
- Katz C, Allport F (1931) *Students' Attitudes* (Craftsman Press, Syracuse, NY).
- Kaufmann M (2022) Projection bias in effort choices. *Games Econom. Behav.* 135:368–393.
- Krueger J, Clement R (1994) The truly false consensus effect—An ineradicable and egocentric bias in social-perception. *J. Personality Soc. Psych.* 67(4):596–610.
- Kruger J, Dunning D (1999) Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J. Personality Soc. Psych.* 77(6):1121–1134.
- Kurt D, Inman J (2013) Mispredicting others' valuations: Self-other difference in the context of endowment. *J. Consumer Res.* 40(1):78–89.
- Loewenstein G, Adler D (1995) A bias in the prediction of tastes. *Econom. J.* 105(431):929–937.
- Loewenstein G, O'Donoghue T, Rabin M (2003) Projection bias in predicting future utility. *Quart. J. Econom.* 118(4):1209–1248.
- Madarász K (2012) Information projection: Model and applications. *Rev. Econom. Stud.* 79(3):961–985.
- Madarász K, Danz D, Wang S (2023) Projective thinking: Model, evidence, and applications. Working paper, London School of Economics & Political Science, London, UK.
- Marks G, Miller N (1987) 10 years of research on the false-consensus effect: An empirical and theoretical review. *Psych. Bull.* 102(1):72–90.
- März O (2019) Comment on "Naiveté, projection bias, and habit formation in gym attendance." *Management Sci.* 65(5):2442–2443.
- Read D, van Leeuwen B (1998) Predicting hunger: The effects of appetite and delay on choice. *Organ. Behav. Human Decision Processes* 76(2):189–205.
- Ross L, Greene D, House P (1977) The false consensus effect: An egocentric bias in social perception and attribution processes. *J. Experiment. Soc. Psych.* 13:279–301.
- Selten R, Ockenfels A (1998) An experimental solidarity game. *J. Econom. Behav. Organ.* 34(4):517–539.
- Sherif M, Hovland D (1961) *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change* (Yale University Press, New Haven, CT).
- Thomas C, Otis N, Abraham J, Markus H, Walton G (2020) Toward a science of delivering aid with dignity: Experimental evidence and local forecasts from Kenya. *Proc. Natl. Acad. Sci. USA* 117(27):15546–15553.
- Vanberg C (2008) Why do people keep their promises? An experimental test of two explanations. *Econometrica* 76:1467–1480.
- Van Boven L, Loewenstein G (2003) Social projection of transient drive states. *Personality Soc. Psych. Bull.* 29(9):1159–1168.
- Van Boven L, Loewenstein G (2005) Empathy gaps in emotional perspective taking. Malle BF, Hodges SD, eds. *Other Minds: How Humans Bridge the Divide Between Self and Others* (Guilford Press, New York), 284–297.
- Van Boven L, Loewenstein G, Dunning D (2003) Mispredicting the endowment effect: Underestimation of owners' selling prices by buyers' agents. *J. Econom. Behav. Organ.* 51:351–365.
- Van Boven L, Loewenstein G, Dunning D, Nordgren L (2013) Changing places: A dual judgment model of empathy gaps in emotional perspective taking. Olson JM, Zanna MP, eds. *Advances in Experimental Social Psychology*, vol. 48 (Academic Press, San Diego), 118–171.
- Ziegler A, Romagnoli G, Offerman T (2022) Morals in multi-unit markets. Working paper, University of Amsterdam, Amsterdam.