An Experiment on Interpersonal Projection Bias

Benjamin Bushong
Michigan State

Tristan Gagnon-Bartsch
Harvard

Motivation

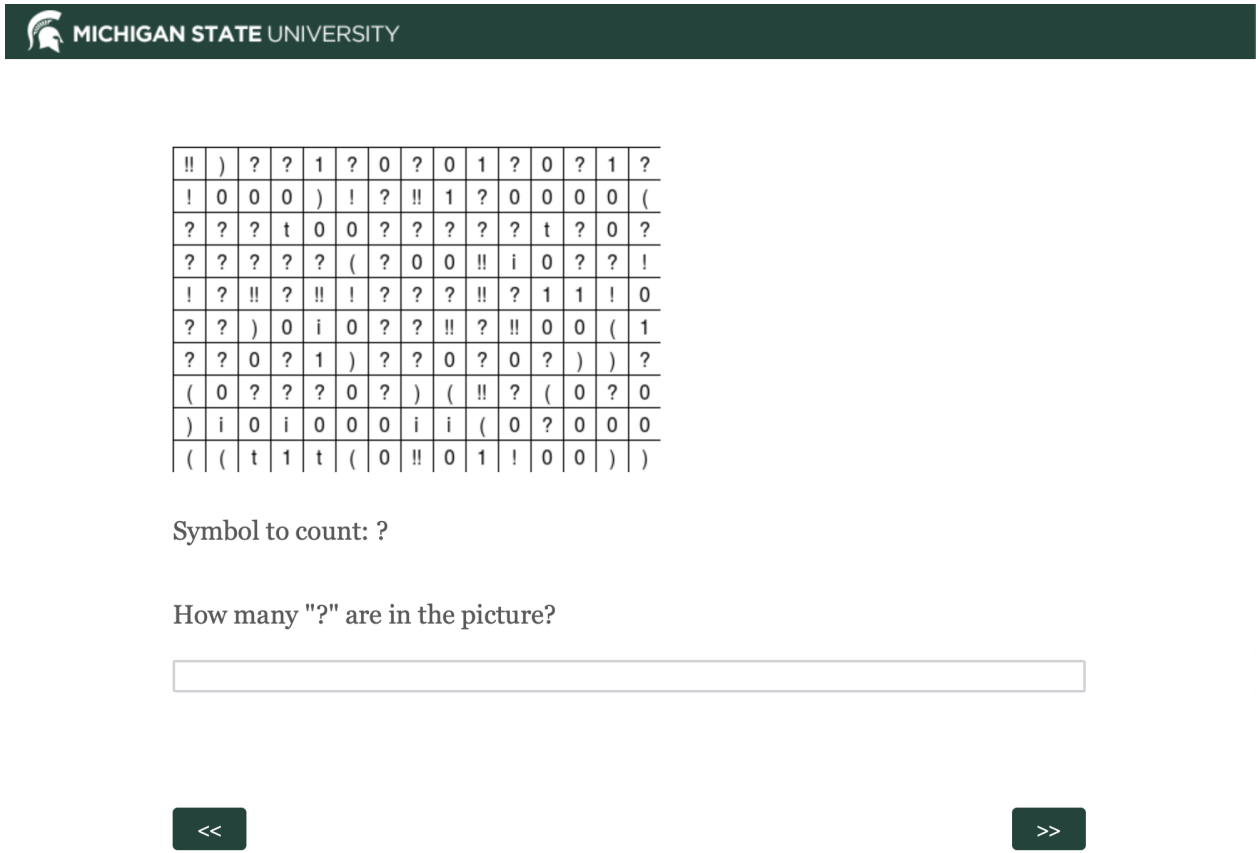
- Forecasting others' preferences widely relevant (e.g., auctions, negotiations, coordination, managing workers, social learning, etc.)
  - **Question:** How well do we predict others' preferences?
  - Real-effort setting where people have transparently different preferences
- **Finding:** people project own preference state (fatigue) onto others when predicting their willingness to work
  1. Subjects have familiarity with various states (advances psych lit on social projection; e.g., false-consensus effect)
  2. Measure intra-personal projection in same domain (builds on Conlin et al 2007; Busse et al 2015; Augenblick & Rabin 2019)

Motivation

- Forecasting others' preferences widely relevant (e.g., auctions, negotiations, coordination, managing workers, social learning, etc.)
  - **Question:** How well do we predict others' preferences?
  - Real-effort setting where people have transparently different preferences
- **Finding:** people project own preference state (fatigue) onto others when predicting their willingness to work
  1. Subjects have familiarity with various states (advances psych lit on social projection; e.g., false-consensus effect)
  2. Measure intra-personal projection in same domain (builds on Conlin et al 2007; Busse et al 2015; Augenblick & Rabin 2019)

Real-Effort Task

- All participants do multiple rounds of counting task



Overview of Experiment

Subjects (MTurk) assigned to one of two roles: worker or predictor

1. **Workers** do $s \in \{5, 20\}$ tasks then state willingness to continue working

   $\Rightarrow$ Elicit WTW when either Fresh $(s = 5)$ or Tired $(s = 20)$

2. **Predictors** guesses average WTW in each state

   - Induce variation in predictors' preferences:
     - Work on same task as workers
     - Make predictions when Fresh and Tired

   $\Rightarrow$ How does predictor's own state influence predictions?

Overview of Experiment

Subjects (MTurk) assigned to one of two roles: worker or predictor

1. **Workers** do $s \in \{5, 20\}$ tasks then state willingness to continue working

   $\Rightarrow$ Elicit WTW when either Fresh $(s = 5)$ or Tired $(s = 20)$

2. **Predictors** guesses average WTW in each state

   - Induce variation in predictors' preferences:
     - Work on same task as workers
     - Make predictions when Fresh and Tired

   $\Rightarrow$ How does predictor's own state influence predictions?

Design: Workers

- Two groups of workers ($N \approx 600$):
  1. **Fresh:** Do **5** rounds, then state # additional tasks willing to do for $\$m$ (BDM)
     - $W(5)$ denotes average response
  2. **Tired:** Do **20** rounds, then state # additional tasks willing to do for $\$m$ (BDM)
     - $W(20)$ denotes average response

Design: Workers

- Two groups of workers ($N \approx 600$):

  1. **Fresh:** Do **5** rounds, then state # additional tasks willing to do for $\$m$ (BDM)

     - $W(5)$ denotes average response

  2. **Tired:** Do **20** rounds, then state # additional tasks willing to do for $\$m$ (BDM)

     - $W(20)$ denotes average response

Design: Predictors

- Predictors ($N \approx 670$) also work on the task
    - Incentivized predictions about workers
- 3 groups of Predictors

Design: Predictors

1. **In-Group**: guesses about workers in same state

After 5 rounds:
Predict $W(5)$

After 20 rounds:
Predict $W(20)$ and $W(5)$

2. **Out-Group A**: guesses about workers in opposite state

After 5 rounds:
Predict $W(20)$

After 20 rounds:
Predict $W(5)$ and $W(20)$

3. **Out-Group B**: no guess while fresh

After 5 rounds:
Reminder

After 20 rounds:
Predict $W(5)$ and $W(20)$

Design: Predictors

1. **In-Group**: guesses about workers in same state

After 5 rounds:
Predict $W(5)$

After 20 rounds:
Predict $W(20)$ and $W(5)$

2. **Out-Group A**: guesses about workers in opposite state

After 5 rounds:
Predict $W(20)$

After 20 rounds:
Predict $W(5)$ and $W(20)$

3. **Out-Group B**: no guess while fresh

After 5 rounds:
Reminder

After 20 rounds:
Predict $W(5)$ and $W(20)$

Design: Predictors

1. **In-Group**: guesses about workers in same state

After 5 rounds:
Predict $W(5)$

After 20 rounds:
Predict $W(20)$ and $W(5)$

2. **Out-Group A**: guesses about workers in opposite state

After 5 rounds:
Predict $W(20)$

After 20 rounds:
Predict $W(5)$ and $W(20)$

3. **Out-Group B**: no guess while fresh

After 5 rounds:
Reminder

After 20 rounds:
Predict $W(5)$ and $W(20)$

Hypotheses

- **Model:** when I'm in state s', I think person in state s has utility function

$$\hat{u}(w|s) = \alpha u(w|s') + (1-\alpha)u(w|s)$$

- **Hypotheses for** $\alpha > 0$:

  1. Same state as target $\Rightarrow$ unbiased
  2. Less tired than target $\Rightarrow$ overestimate WTW
  3. More tired than target $\Rightarrow$ underestimate WTW

Workers: Lower WTW when Tired

Average Willingness To Work

	Workers' State		
	Fresh	Tired	Difference
WTW (Tasks/$)	10.64	6.81	**3.83*****
	(0.692)	(0.408)	(0.803)
Observations	300	299	

Predictors: Accurate Guesses About Own State

Predictions of WTW, Same State (Tasks/$)

	Prediction	True WTW	Difference
Fresh Predicting Fresh	10.81	10.64	0.17
	(0.605)	(0.692)	(0.957)
	$n = 223$	$n = 300$	
Tired Predicting Tired	6.65	6.81	-0.17
	(0.230)	(0.408)	(0.439)
	$n = 666$	$n = 299$	

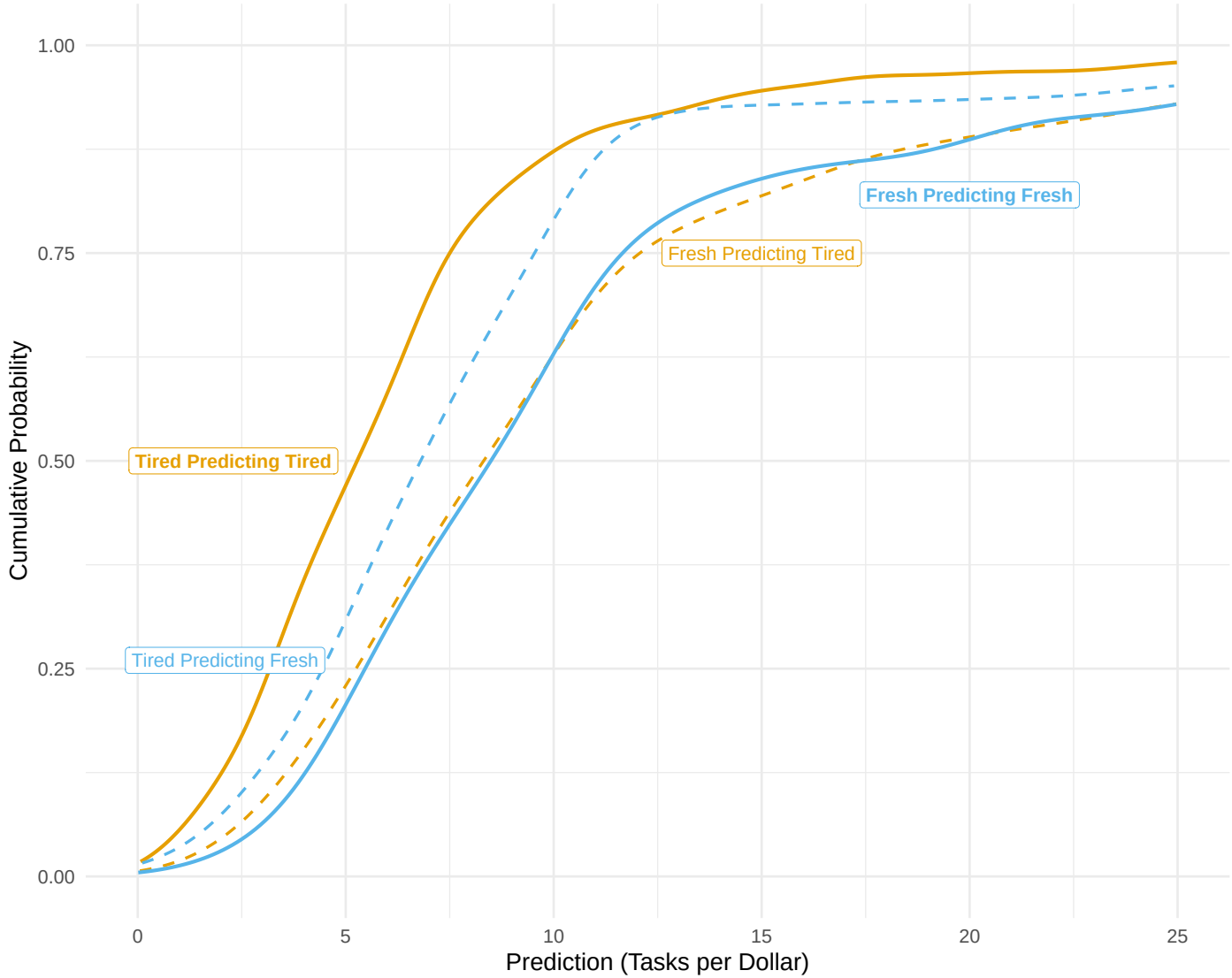
Predictors: Biased Guesses About Other State

First Predictions vs Workers' WTW (Tasks/$)

	Prediction	True WTW	Difference
Fresh Predicting Tired	10.22	6.81	**3.40*****
	(0.491)	(0.408)	(0.639)
	$n = 221$	$n = 299$	
Tired Predicting Fresh	8.44	10.64	**−2.20****
	(0.532)	(0.692)	(0.873)
	$n = 222$	$n = 300$	

Quantifying Interpersonal Projection Bias

- From previous table:
  1. Fresh predictors overestimate WTW of tired workers by 50%
  2. Tired predictors underestimate WTW of fresh workers by 21%
- From parametric estimation in paper: $\alpha \geq .23$


1.00
0.75
0.50
0.25
0.00
Cumulative Probability
0
5
10
15
20
25
Prediction (Tasks per Dollar)
Tired Predicting Tired
Tired Predicting Fresh
Fresh Predicting Tired
Fresh Predicting Fresh

Within-Subject Revisions in Predictions

- Within-subject revisions in guesses:
  - In-Group guessed about fresh workers both when fresh and tired
    - When fresh, guesses were accurate
    - Once tired, guesses reduced by 19%

    $\Rightarrow$ Changing this guess significantly reduced expected earnings
  - Out-Group A also exhibits significant revisions in guesses about tired workers

Within-Subject Revisions in Predictions

- Within-subject revisions in guesses:
  - In-Group guessed about fresh workers both when fresh and tired
    - When fresh, guesses were accurate
    - Once tired, guesses reduced by 19%

    $\Rightarrow$ Changing this guess significantly reduced expected earnings
  - Out-Group A also exhibits significant revisions in guesses about tired workers

Intra-Personal Projection Bias

- New group of workers (while fresh) predicted own WTW once tired

Predicting Workers' Guesses and WTW

	Prediction	Actual	Difference
WTW (Tasks/$)	7.37	5.67	**1.70*****
	(0.393)	(0.387)	(0.348)
Observations	298	298	298

Notes: Difference significant at $p < .001$

- Overestimate own WTW by 30%
- Recall: fresh predictors overestimate WTW of tired workers by 50%

  $\Rightarrow$ Interpersonal bias more severe than *intra*personal bias

Conclusion

- **Take away: Difficult to empathize with others' states**
  - Even when others' states are known
  - Even when others' states are familiar
- Additional Points:
  - Neither accuracy nor confidence increases over time $\Rightarrow$ errors in predictions likely not due to limited info
  - Magnitude of *intra*personal projection similar to other studies (Conlin et al 2007; Busse et al 2015; Augenblick & Rabin 2019)
- In Progress:
  - How does projection distort learning from others' behavior?
  - How does it influence strategic behavior (e.g., bidding in auction)?

Thanks!