DATA COLLECTION IN A FLAT WORLD: THE STRENGTHS AND WEAKNESSES OF MECHANICAL TURK SAMPLES

Joseph K. Goodman, Washington University in St. Louis

Cynthia E. Cryder, Washington University in St. Louis

Amar Cheema, University of Virginia

QUESTIONS ABOUT MTURK

- ## Problems/Questions

  - ### Do they really pay attention?

  - ### Who is doing these surveys?

    - #### Are they different from the rest of us? Are they 'normal'?

    - #### In other words, do they show the same biases as normal people?

Research questions

Do they really pay attention? Who is doing these surveys?

- Motivation & Cognition

  - Do they pay attention?

- Valuing Money & Consumption

  - Do they value money in odd ways?

- Big Five and other Individual

    Differences

  - Do they have personality differences?

- Judgment and Decision Making

  - Are they "normal" (aka, irrational like the rest of us)? Do they show normal biases?

GOODMAN, CRYDER, & CHEEMA

TEST ACROSS TWO STUDIES

- ## Study 1
  - Compare MTurkers to a community sample in Pittsburgh
  - Mturk: $.10, ~10 min study via Qualtrics

- ## Study 2
  - Compare MTurkers to student sample from Wash U in St. Louis
  - MTurk: $.20, ~16 min study via Qualtrics
  - Students: Paper-and-pencil, 14 min, and web-based 13 min

GOODMAN, CRYDER, & CHEEMA

DEMOGRAPHICS

Study 1 - Mturk vs. Community

- Age: 33.5 on average

- Female: 59% MTurk vs. 52% Comm

- Education: Modal and Median 4 year bachelor's

- 26% non-US (3% Canada, 19% India)

- More ESL: 28% vs. 11%*

- Age: 31 MTurk vs. 19.4 students

- Female: 43% MTurk vs. 59% Students

- 100% of students were students and had "some college"!

  - 24% MTurk'ers were students

  - 92% MTurk'ers "some college"

GOODMAN, CRYDER, & CHEEMA
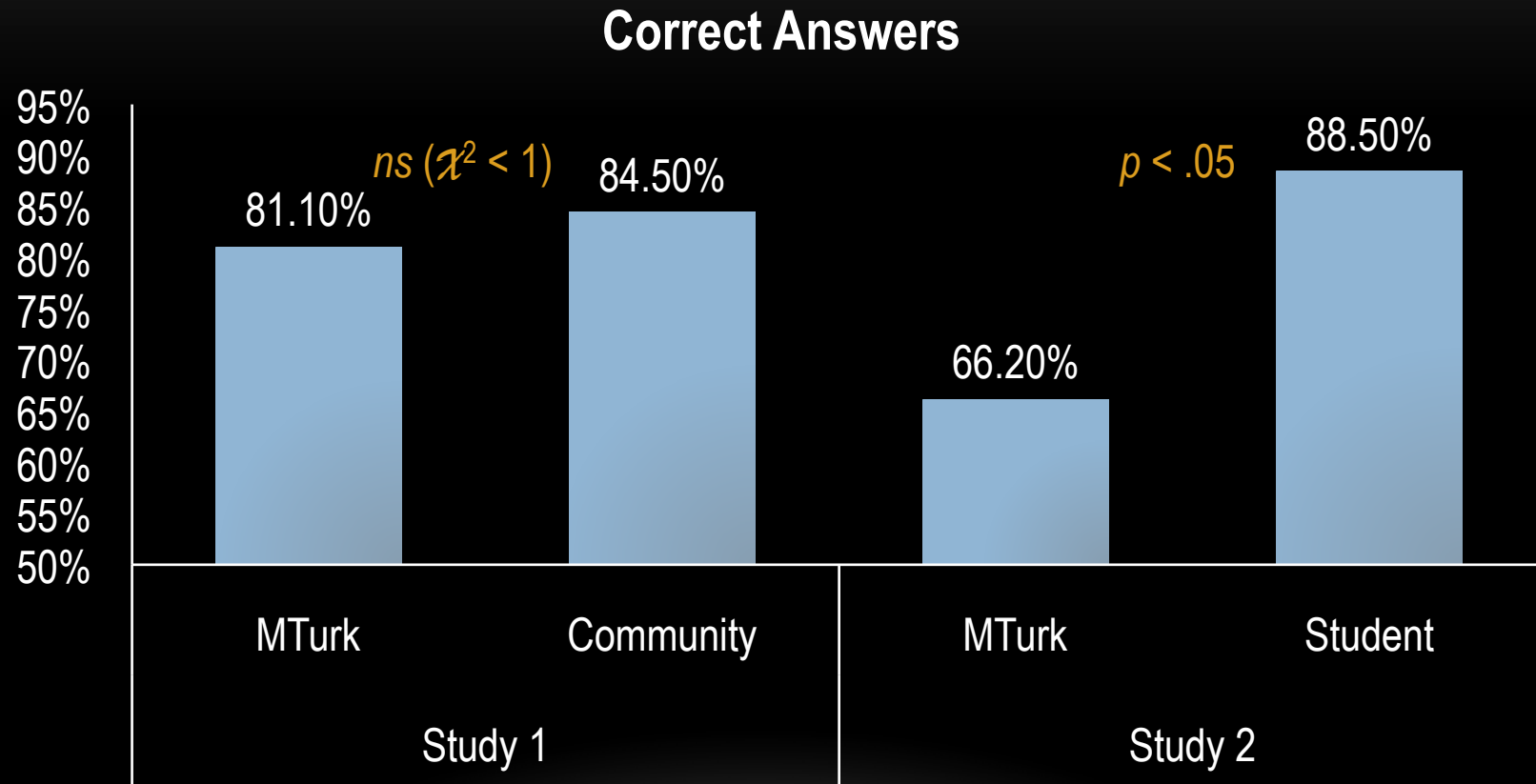
STUDY 1 - COGNITION & MOTIVATION
THE IMC

Research in decision making shows that people…prefer not to pay attention.... If you are reading this question and have read all the other questions, please select the box marked 'other' and write 'decision making' on the line below. Do not select your own opinions and behaviors. Thank you for participating and taking the time to read through the questions carefully!

What was this study about?

- Your own opinions and behaviors

- Lions (Study 2: Political history)

- Tigers (Study 2: Friends' behaviors)

- Other _____

STUDY 1 & 2 - COGNITION & MOTIVATION
IMC

Correct Answers

95%
90%
85%
80%
75%
70%
65%
60%
55%
50%

81.10%   *ns* ($\chi^2 < 1$)   84.50%   *p* < .05   88.50%

66.20%

MTurk   Community   MTurk   Student

Study 1   Study 2

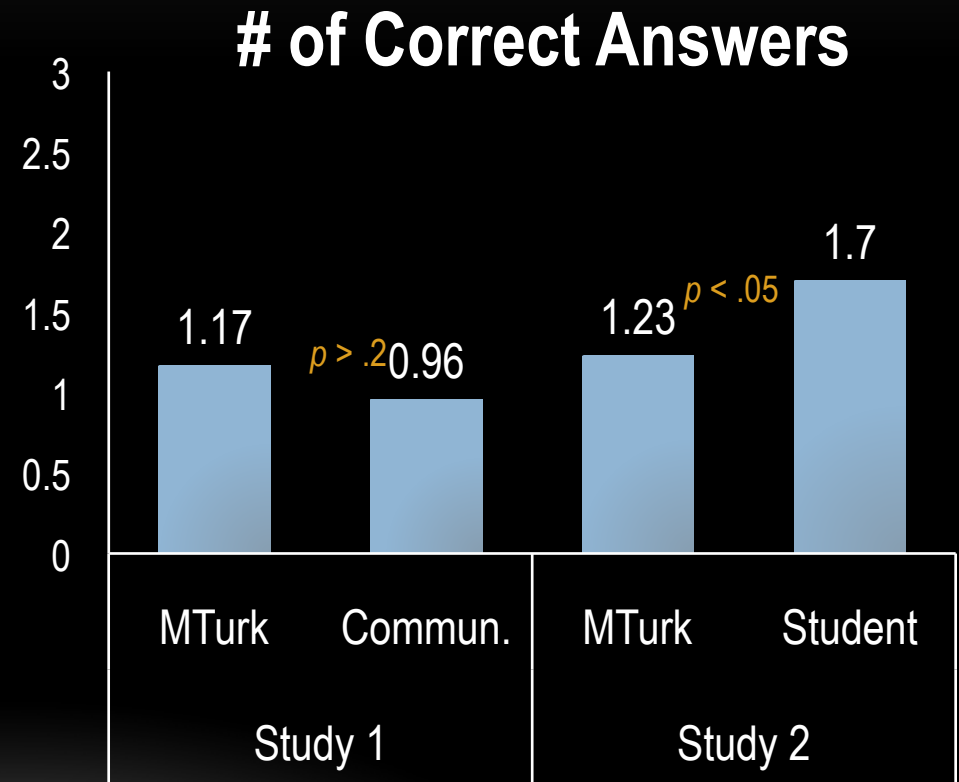GOODMAN, CRYDER, & CHEEMA

COGNITION & MOTIVATION
THE CRT & SYSTEM 2 PROCESSING

- Cognitive Reflection Test

  - 3 questions

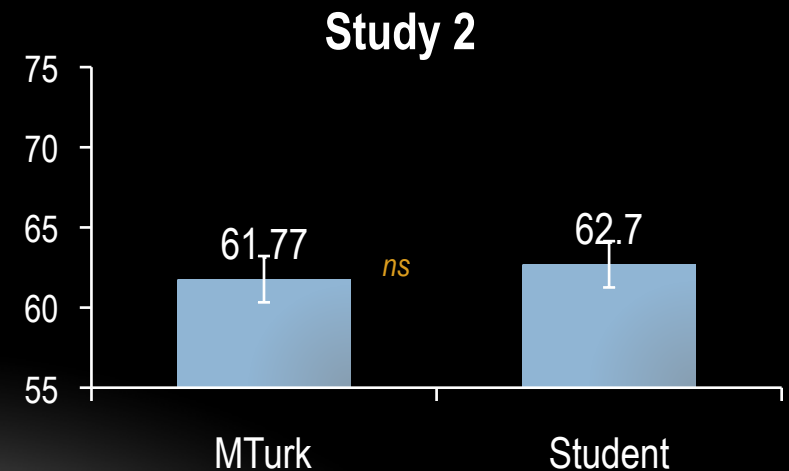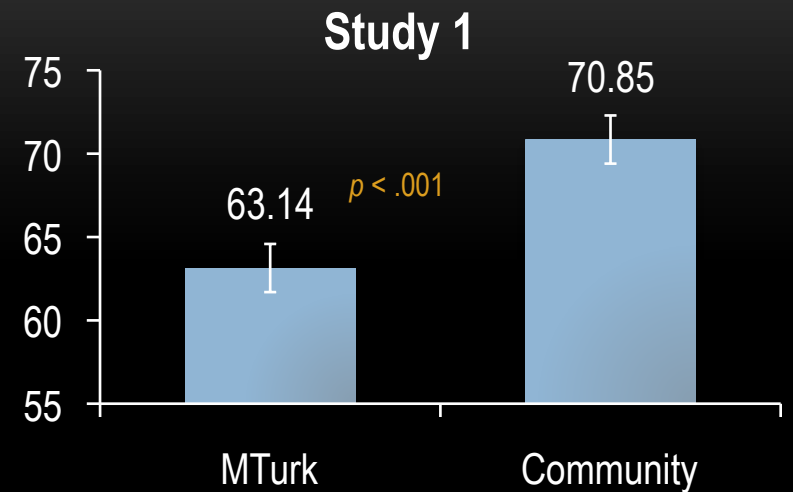 A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

# of Correct Answers

Study 1: MTurk 1.17, Commun. 0.96, $p > .2$

Study 2: MTurk 1.23, Student 1.7, $p < .05$

Other web-studies = 1.10

(Frederick 2005)

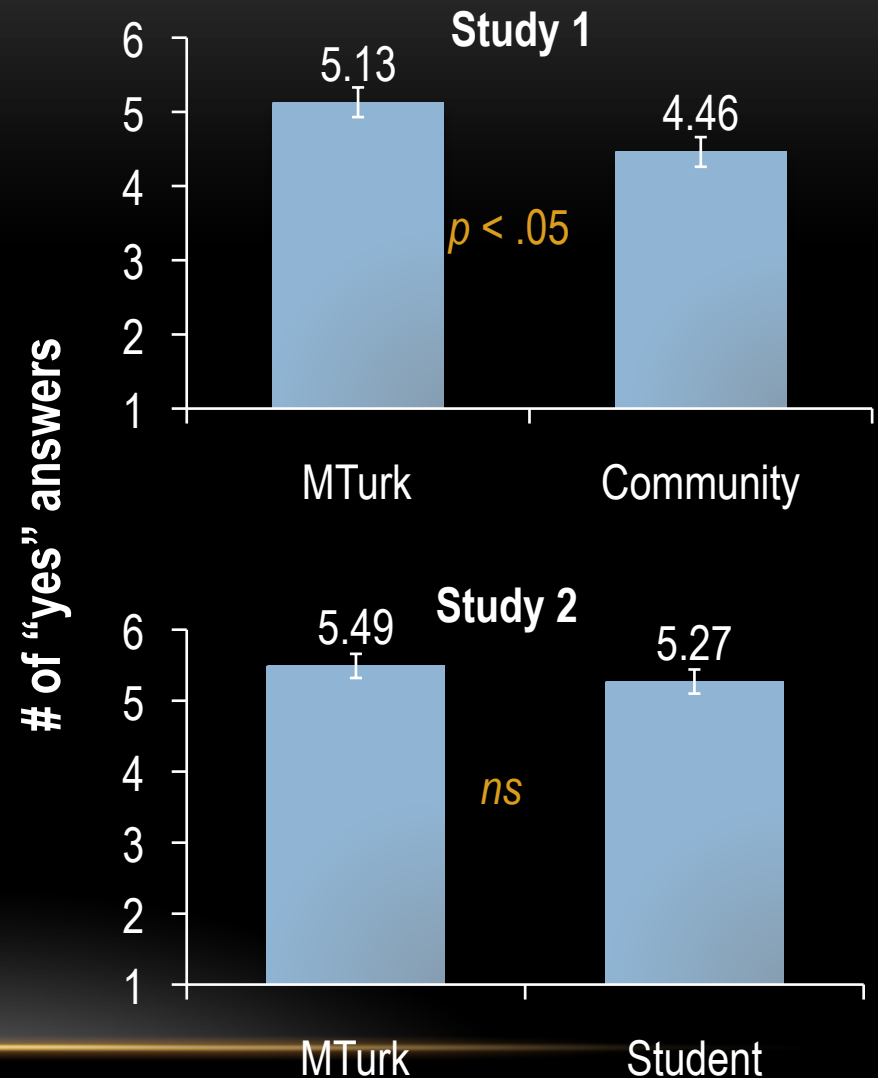GOODMAN, CRYDER, & CHEEMA

COGNITION & MOTIVATION
NFC

- Need for Cognition
  - 18 questions
  - "I would prefer complex to simple problems.
  - "I like to have the responsibility of handling a situation that requires a lot of thinking"

Study 1

Study 1 bar chart: MTurk = 63.14, Community = 70.85, $p < .001$

Study 2

Study 2 bar chart: MTurk = 61.77, Student = 62.7, *ns*
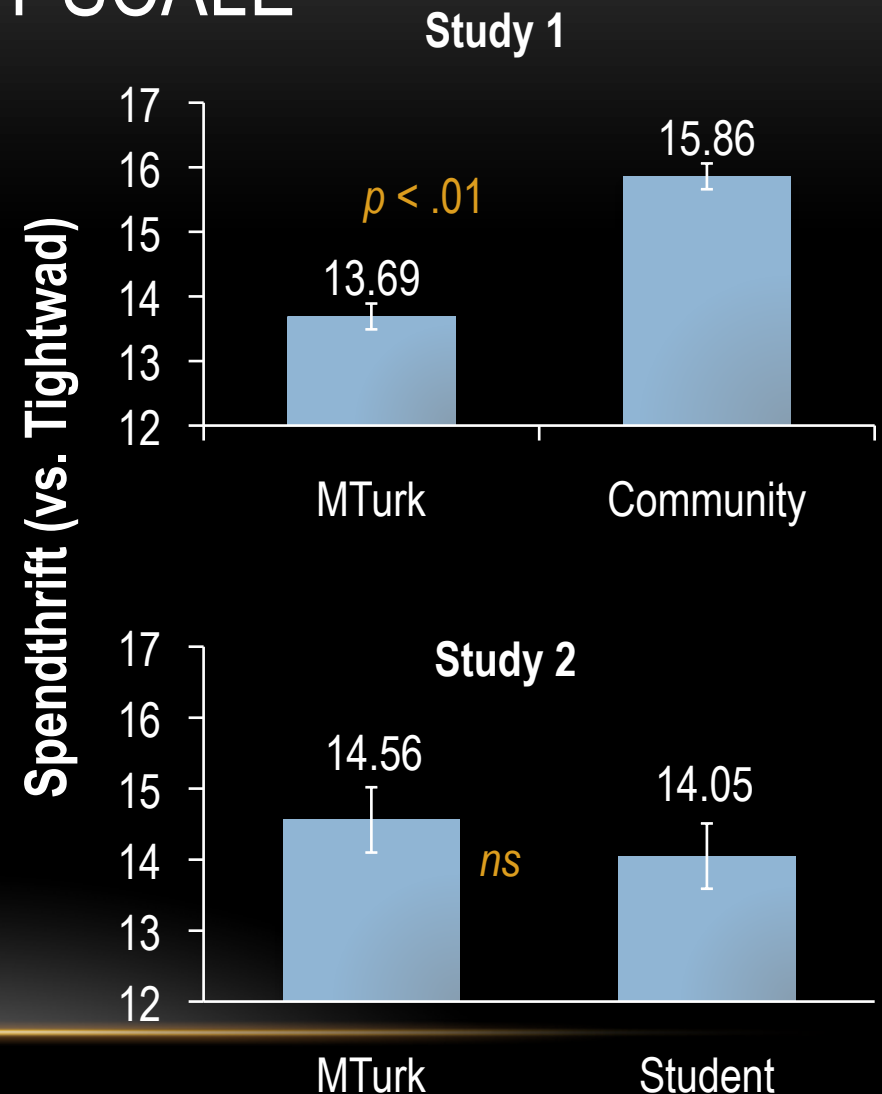
VALUING MONEY & CONSUMPTION
TVM

- Time Value of Money
  - Would you go an entire day (24 hours) without drinking liquids for
    - $5? (yes/no)
    - $25?
    - $100?
    - $1000?
  - Would you complete a 1-hour survey for
    - $5?
    - $25?
    - $100?
    - $1000?

Study 1

MTurk: 5.13

Community: 4.46

$p < .05$

of "yes" answers

Study 2

MTurk: 5.49

Student: 5.27

ns

(Cryder and Loewenstein 2010)

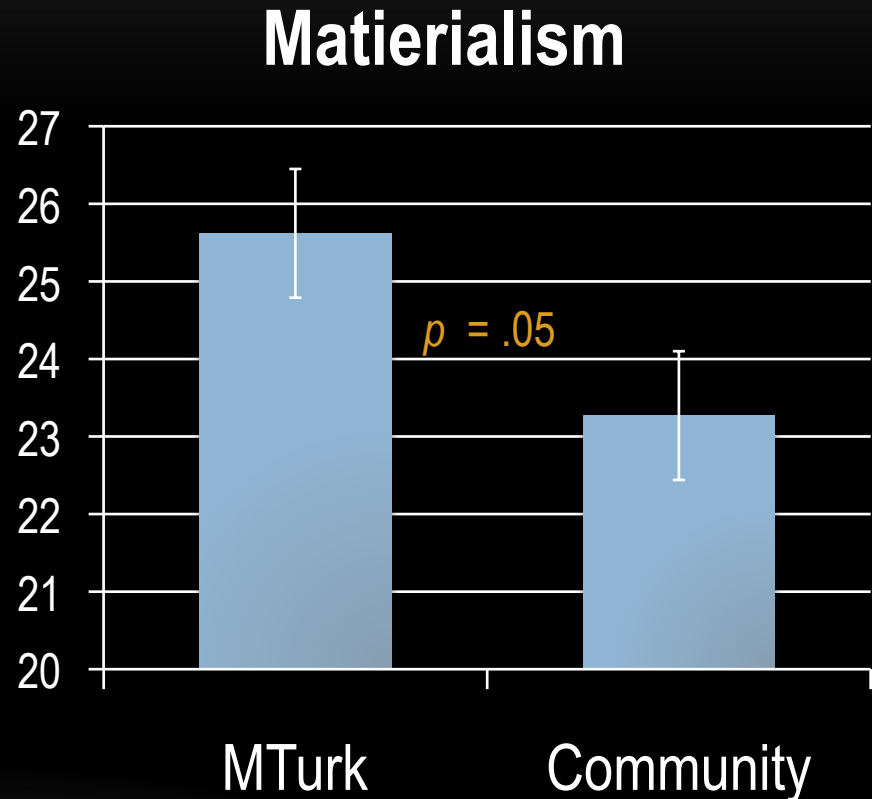# VALUING MONEY & CONSUMPTION
TIGHTWAD-SPENDTHRIFT SCALE

- 4 questions on spending

- *Spendthrift*:

  trouble limiting spending. Often spend money when they would do better not to.

- *Tightwad*:

  trouble spending money. Spending money makes them anxious, often don't spend money on things they should spend on.

Study 1

p < .01

13.69 — MTurk

15.86 — Community

Spendthrift (vs. Tightwad)

Study 2

14.56 — MTurk

14.05 — Student

ns

VALUING MONEY & CONSUMPTION
MATERIAL VALUES SCALE (MVS)

- 9 questions

- e.g., My life would be better if I owned certain things I don't have.

Matierialism

p = .05

MTurk   Community

Study 2: no difference on Materialism Scale (27.57 vs. 27.69)

GOODMAN, CRYDER, & CHEEMA

STUDY 2 – VALUING MONEY & CONSUMPTION
PRESENT BIAS
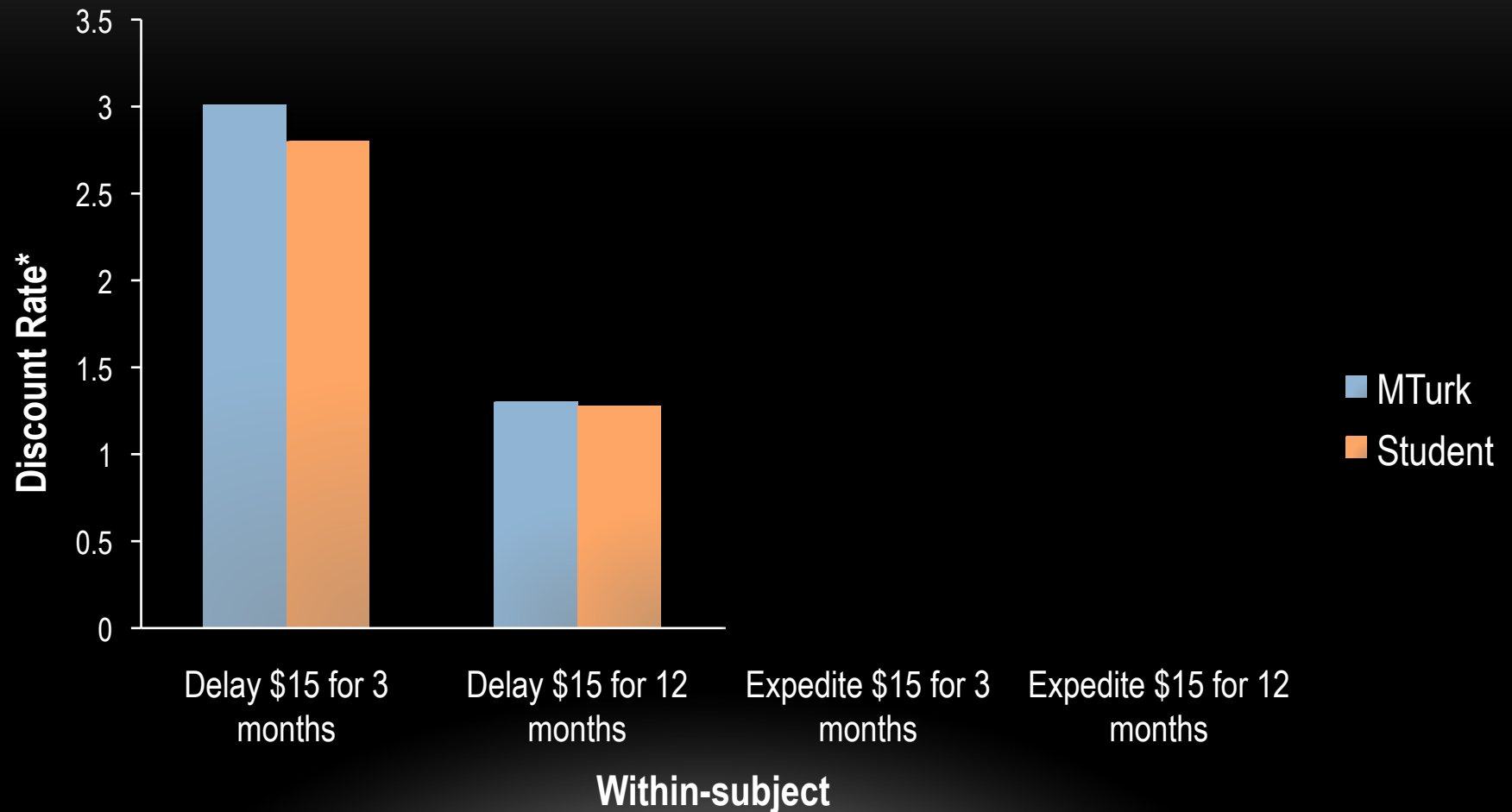
*continuously compounded discount rates

GOODMAN, CRYDER, & CHEEMA

(Thaler, 1981; Lowenstein, 1988; Malkoc & Zauberman, 2006)

STUDY 1 - JDM
ANCHORING & ADJUSTMENT

- Last 2 digits of phone number

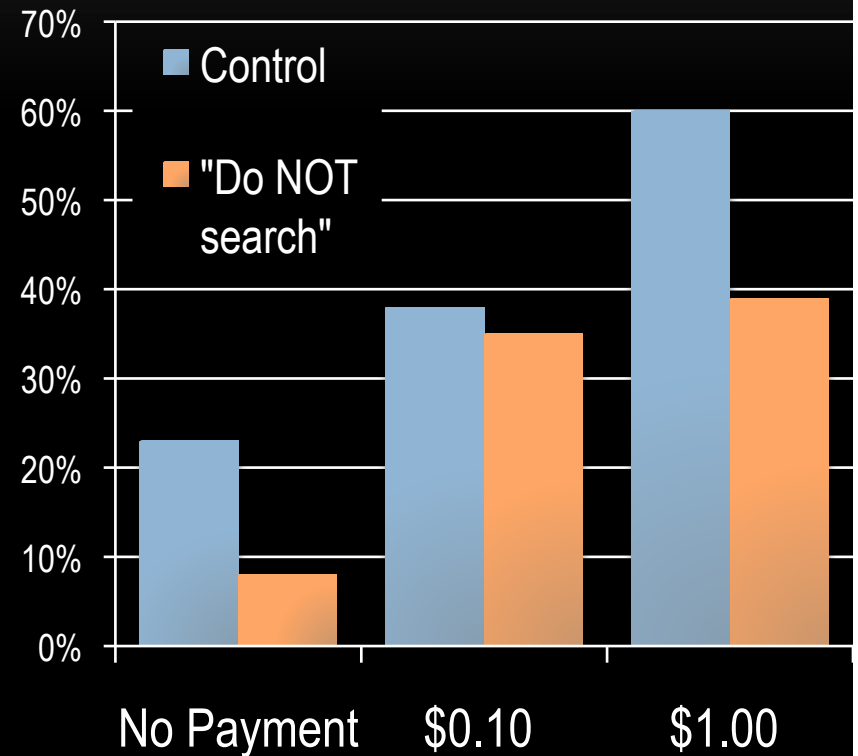- # countries of Africa

Community: β = .45, p < .01

MTurk: β = .06, ns

- 10% MTurk'ers guess correct! (54)

STUDY 2 - JDM
ANCH. & ADJ. (FOLLOW-UP STUDY)

- ## Instructions vs. Compensation

  - Main effect for instructions
    "…do NOT use external sources…"

  - Main effect for compensation
    ($0, $.10, $1)

STUDY 2 - JDM
RISK

- Four Gambles

  - Risk Averse for gains, both big and small

  - Certainty Effect: Preference decreases with small probabilities (students especially)

  - Risk seeking for losses

	$3 @ 100% (vs. $4 @80%)	$240 @100% (vs. $320 @ 80%)	$3 @ 25% (vs. $4 @ 20%)	-$3 @ 100% (vs. -4 @ 80%)
MTurk	0.85			
Student	0.82			

*$p < .05$
~$p < .06$

GOODMAN, CRYDER, & CHEEMA

INDIVIDUAL DIFFERENCES
MTURK'ERS ARE (BOTH STUDIES)..

- Less extroverted

- Less emotionally stable

- Lower self-esteem

- Less happy (satisfaction with life)

- No consistent differences:

  - Agreeableness

  - Conscientiousness

  - Openness to experiences

  - Self-control

  - Maximizer-satisficer

IMC, ESL, AND NON-US PARTICIPANTS
FEW DIFFERENCES

Study 1

- Only 1 IMC x MTurk interaction (emotional stability)
- Including everyone: Emot. Stab. (p=.15) ns and MVS (p=.15) and
- ESL and non-US more likely to fail IMC
- MTurk x ESL: Emot. Stab.
- Filter by ESL: MVS ns (p>.2)

- Filter by non-US: MVS ns (F<1), Emot. Stab. ns (p>.1)

Study 2

- Only 1 IMC x MTurk interaction (TW-ST)
- Including Everyone: Emot. Stab. marginal (p=.07)
- ESL and non-US more likely to fail IMC
- MTurk x ESL: CRT
- Fitler by ESL: Emot. Stab. (F < 1) and conscientiousness (p=.06)
- Filter by non-US: Emot. Stab. (p>.15) and conscientiousness (p>.2)

CONCLUSIONS

- MTurk'ers are Different!

  - They pay attention, but students in the lab are best

  - Views of money and spending are different than community (but not different than students)

    - More tightwad

    - Willing to do more tasks for money

    - More materialistic

  - Some personality differences

- But MTurk'ers show the same biases

  - Present biased, delay/ expedite asymmetry

  - Risk seeking for losses, risk averse for gains

  - Respond to honestly pleas

GOODMAN, CRYDER, & CHEEMA

Happy MTurkin'