

Convenience Samples

POLSCI 4SS3

Winter 2023

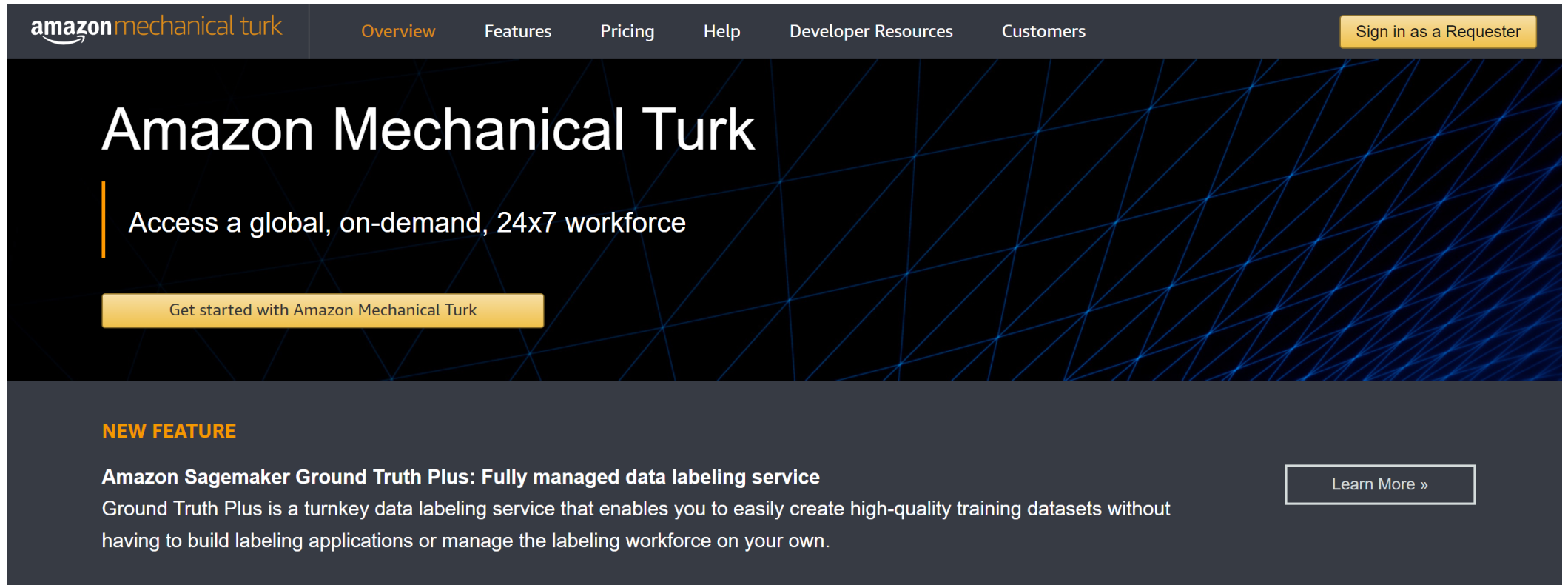
Course so far

- Representative surveys as the gold standard
- Other research design help us learn more but tend to use convenience samples
- **Today:** Talk more about convenience samples

We are seeing less of this



And more of this

The image is a screenshot of the Amazon Mechanical Turk website. At the top, there is a dark navigation bar with the Amazon Mechanical Turk logo on the left and several links: Overview, Features, Pricing, Help, Developer Resources, and Customers. On the right side of the navigation bar is a yellow button that says "Sign in as a Requester". Below the navigation bar is a large hero section with a dark background and a blue geometric pattern of lines. The main heading "Amazon Mechanical Turk" is in large white text. Below it, a subtitle reads "Access a global, on-demand, 24x7 workforce". A yellow button with the text "Get started with Amazon Mechanical Turk" is positioned below the subtitle. Further down, there is a section titled "NEW FEATURE" in orange. It features the heading "Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service" in white. Below this heading is a paragraph of text: "Ground Truth Plus is a turnkey data labeling service that enables you to easily create high-quality training datasets without having to build labeling applications or manage the labeling workforce on your own." To the right of this text is a white button with the text "Learn More »".

amazon mechanical turk

Overview Features Pricing Help Developer Resources Customers

Sign in as a Requester

Amazon Mechanical Turk

Access a global, on-demand, 24x7 workforce

Get started with Amazon Mechanical Turk

NEW FEATURE

Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service

Ground Truth Plus is a turnkey data labeling service that enables you to easily create high-quality training datasets without having to build labeling applications or manage the labeling workforce on your own.

Learn More »

Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace that makes it easier for individuals and businesses to outsource their processes and jobs to a distributed workforce who can perform these tasks virtually. This could include anything from conducting simple data validation and research to more subjective tasks like survey participation, content moderation, and more. MTurk enables companies to harness the collective intelligence, skills, and insights from a global workforce to streamline business processes, augment data collection and analysis, and accelerate machine learning development.

And more of this



Researchers

Participants

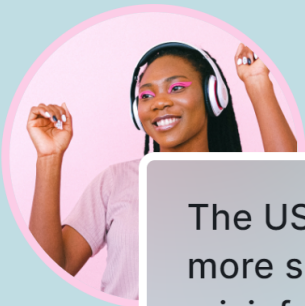
Resources ▾

Log in

Sign up

A higher standard of online research

Conduct research with 130,000+ vetted participants.
Gain insights you can rely on.



The US vs UK: Which is more susceptible to misinformation online?



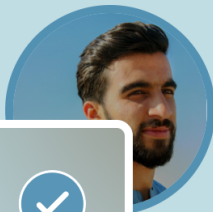
£6.50/hour



Which advert do you prefer?



\$11.00/hour



Risk or reward:
How will you choose?



\$10.00/hour



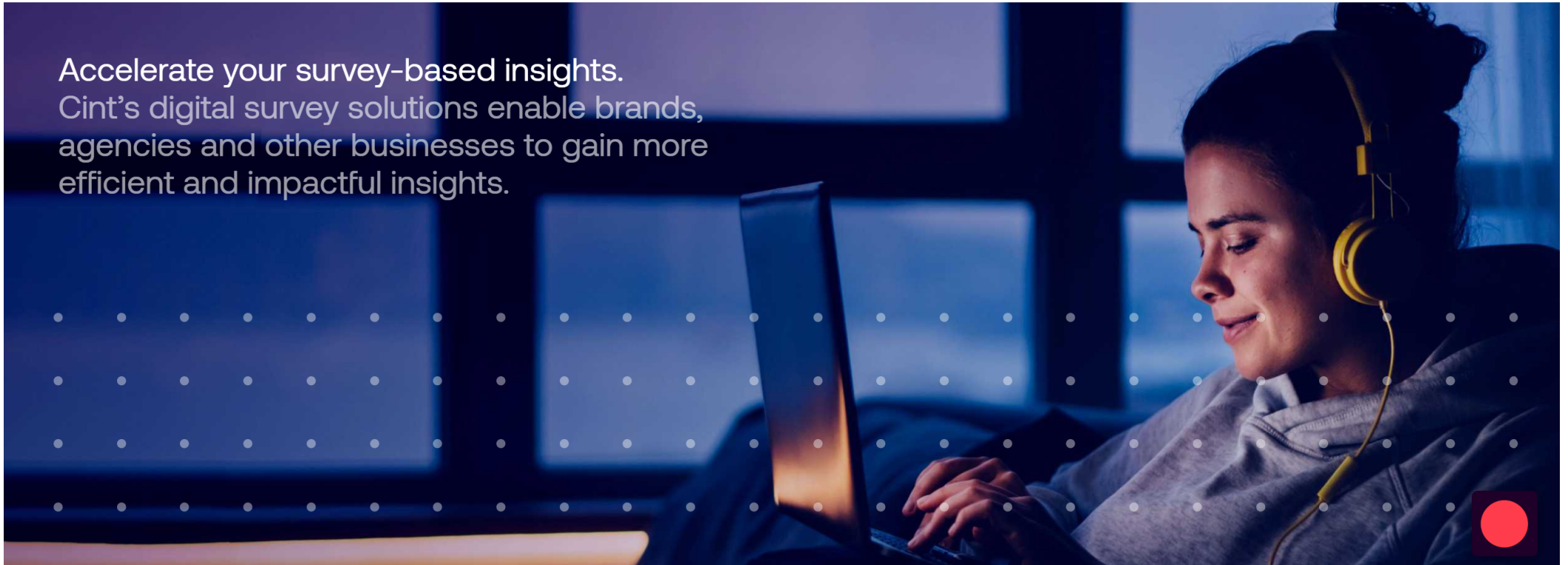
And more of this

Cint

Solutions ▾ Resources ▾ Company ▾

Login ▾ ● Get Started  

Accelerate your survey-based insights.
Cint's digital survey solutions enable brands,
agencies and other businesses to gain more
efficient and impactful insights.



Do we want surveys to be representative?

- Pros?
- Cons?
- We always want them!
- But when do we **need** them?
- Rather, when can we **get away** with not having them?

Internal an external validity

- **Validity:** Approximate truth or usefulness of an *inference*
- **Inference:** How we interpret the results of a study
- **Internal validity:** Whether inferences from a single study *cannot* be explained by other factors
- **External validity:** Whether inferences from a single study apply to a broader population or other target populations
- Convenience samples make it *easier* to achieve *internal validity* at the *expense* of *external validity*

Types of internal validity

1. X-validity (endogenous variables)
2. T-validity (treatments, conditions)
3. Y-validity (outcome variables)
4. C-validity (context)

X-validity

- Is the sample comparable to the target population?
- If not, can we claim that the differences can be ignored?
- To do that, we have to convince ourselves that:

1. Effects are the same across units

OR

2. We observe all the variables that may explain discrepancies in effects

T-validity

- Do treatments (**conditions**) reflect what participants would encounter in the real world?
- **Example:** Is thinking about *hypothetical* countries a good reflection to how people would think about *real* countries?
- Can we claim that there are no different versions of the same *treatment*?
- To do that, we need to convince ourselves that everyone would interpret vignettes in the same way
- Either because it is **realistic** enough or abstract yet **believable**

Y-validity

- Do the outcomes we measure in surveys reflect the outcomes we want to learn about in the real world?
- **Example:** Are self-reported vote intentions a good replacement for actual voting behavior?
- Can we claim that there are no different versions of the same *outcome*?
- Need to convince ourselves that measured outcomes are sufficiently valid and reliable

C-validity

- Do results generalize from other contexts?
- **Example:** If it worked with students in Sweden, will it work with students in Canada?
- Can we claim that the same units would react in the same way if the study was conducted elsewhere?
- Need to convince ourselves that **context is irrelevant** for *similar people in different places*

Discussion

Munger et al (2021): Accessibility and generalizability

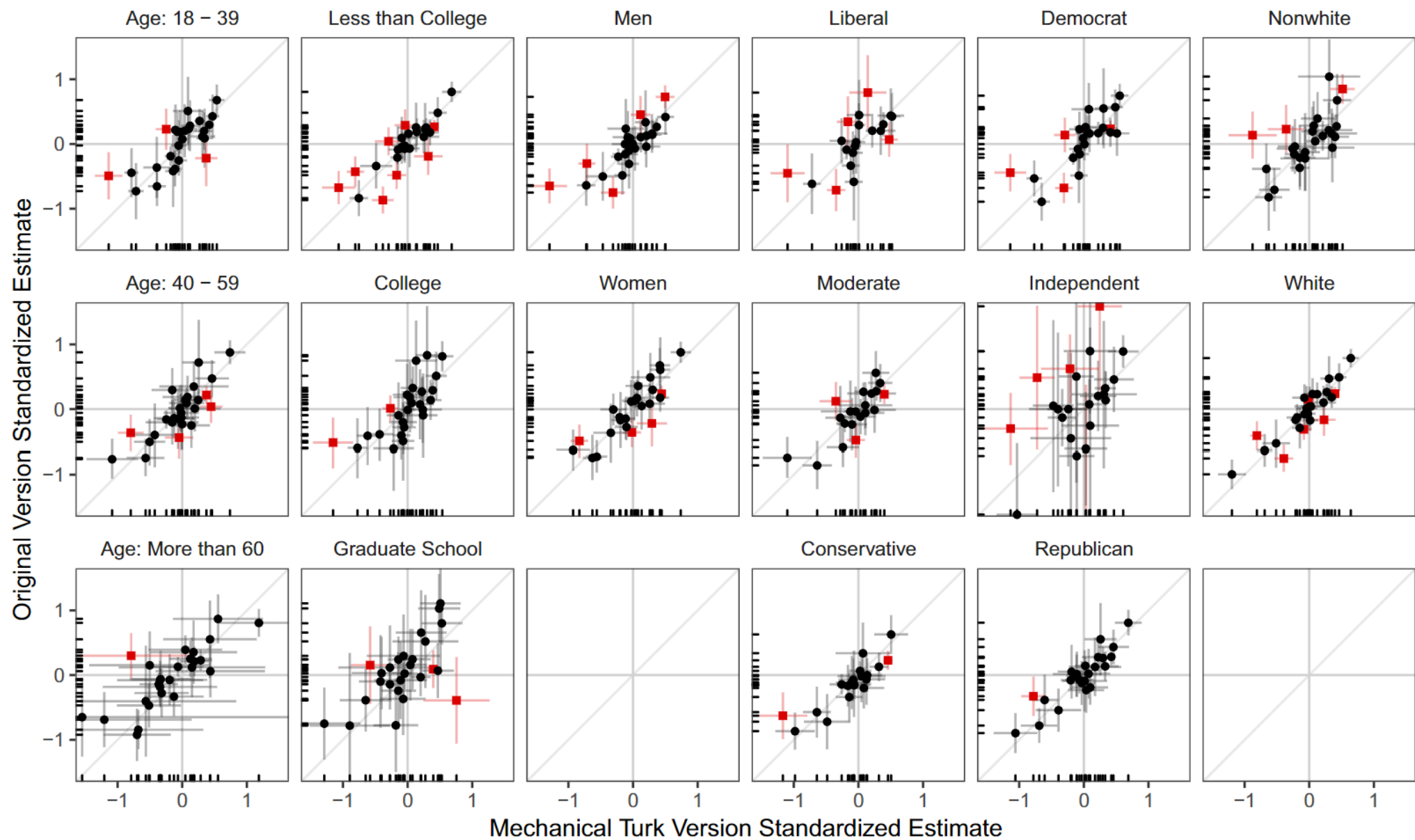
- Replicate 3 convenience sample survey experiments with representative sample
 1. Social commentary and news source credibility
 2. Facebook shares and news consumption
 3. Issue framing and support for gun control
- **Argument:** Effects vary considerably by age and digital literacy

Findings

- **Replication 1:** Participants low on digital literacy did not respond differently to vignettes
- **Replication 2:** Older people clicked on whatever headline came first
- **Replication 3:** No differences because issue had nothing to do with digital literacy
- What kind of validity is this about?

Coppock et al (2018): Generalizability of heterogeneous treatment effect estimates across samples

- Replicate 27 studies from nationally-representative samples with convenience samples
- Compare how effects vary across 16 demographic characteristics



Difference in CATES ■ Significant ● Not Significant

Explanation

- Different samples yield similar results when:
 1. Treatment effects are mostly homogeneous
 2. Effect heterogeneity is **orthogonal** to sample selection
- What type of validity is this about?

After Recess

Evidence-Informed Policy

Focus on: New topic!

Break time!





