

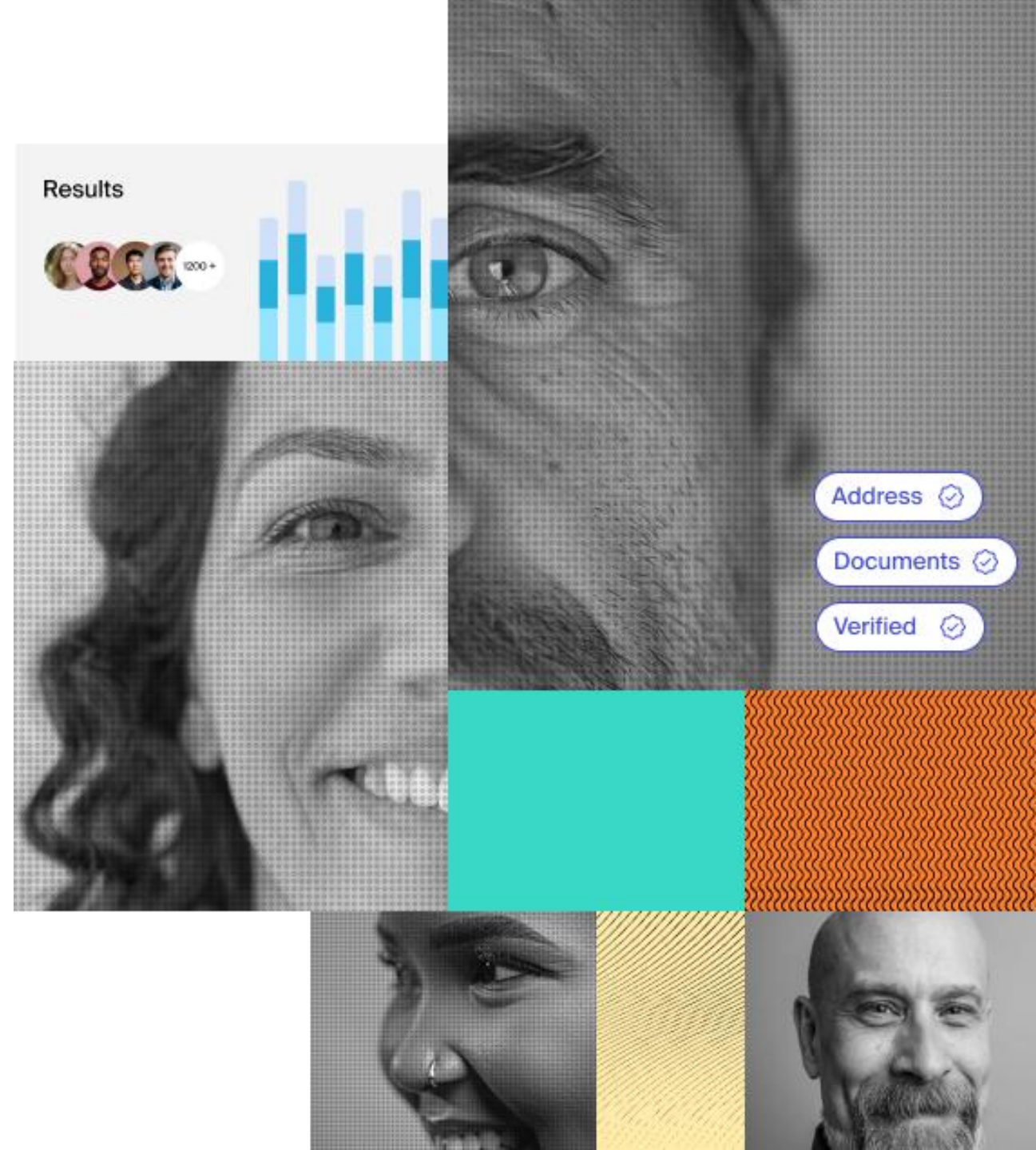


Putting academic research to work: How to apply the latest from academia to elevate your survey research

Ben Leff, Co-Founder & CEO, Verasight

ben@verasight.io


July 2025



Overview



Founded by academic researchers, Verasight works with 100s of universities, media, and government organizations on their survey research



Verasight builds a proprietary panel with our multi-mode methodology, which directly recruits and verifies respondents for our clients



This presentation will share the latest research that we leverage to collect high quality data with our academic clients



Key Themes

1. Importance of respondent verification for consumer & B2B samples
2. A warning about attention checks in an age of AI/bots
3. Major limitations to synthetic samples – beware of the hype

Respondent Verification - Voter Files

What is a voter file?

- Data from Election Officials in all 50 states and Washington D.C.
- Data supplemented with consumer records (important for capturing non-registered voters)
- Voter file gives data on an individual's participation in past elections (e.g., whether the person voted in general, primary, or local elections). It does not tell you whom they voted for.

How Verasight matches our panel to the voter file?

- Ask respondents for identifiable information at sign-up and ask consent for matching:
Full name, address, date of birth.

Performance of Voter File Matched Respondents

Are you licensed to operate a class SSGN submarine?

- Yes
- No

Percent of respondents answering No:

No Voter File Match

Voter File Match



Takeaways: Voter File Matched Respondents

Respondents who had not been matched to the voter file were twice as likely to fail at least one of our quality screens and therefore were twice as likely to be ineligible for inclusion in the final sample.

Respondent Verification B2B

Traditional Approach

Screen large panels or exchanges for respondents who match the inclusion criteria
(self-identification)



Respondent Verification B2B

When researchers try to find low incidence groups in traditional panels, their data will be disproportionately comprised of imposters or bad-faith respondents

Instead of screening based on self-reported job tiles, Verasight works with researchers to recruit directly from official records



“Our study seeks to provide the most comprehensive and rigorous analysis to date of small business owners (hereafter, SBOs) as a political constituency.”

“Using a range of data sources – an original bespoke survey of SBOs based on a unique administrative sampling frame of the recipients of the Paycheck Protection Program (PPP)”

Malhotra N, Margalit Y, Shi S. The Politics of Small Business Owners. British Journal of Political Science. 2025;55:e94. doi:10.1017/S0007123425000274



Attention Checks in an AI/Bot Age

reCAPTCHA Score by Google

- A value that indicates how likely a person is to be a human rather than a bot or malicious software
- Range: 0.0 to 1.0
- Qualtrics recommended interpretation: A score of greater than or equal to 0.5 means the respondent is likely a human. A score of Less than 0.5 means the respondent is likely a bot.

Attention Check

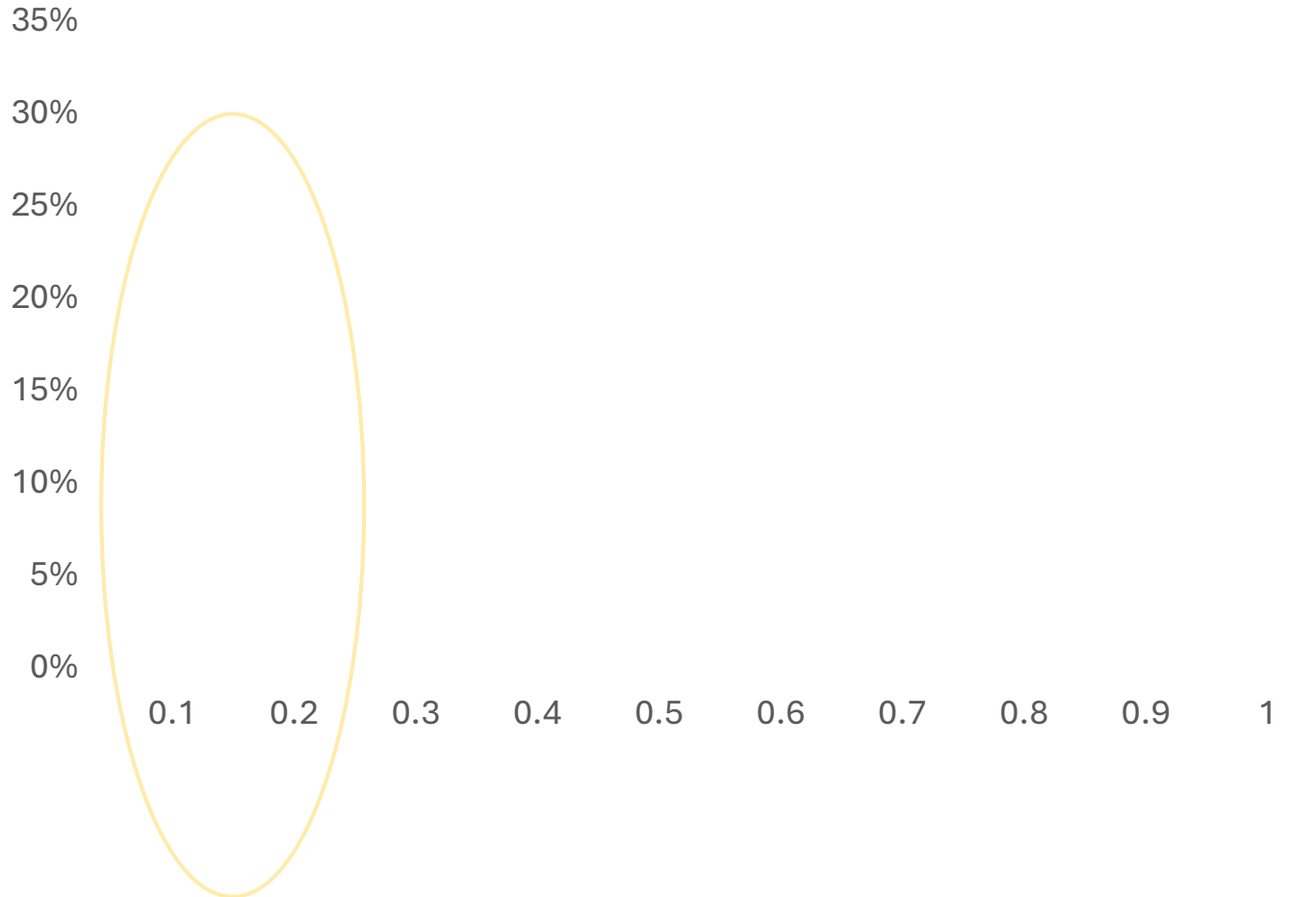
Are you licensed to operate a class SSGN submarine?

- Yes
- No



Importance of Respondent Verification

Failure Rate by reCAPTCHA Score



Conclusions

- Bots/AI agents are often the most likely to pass attention checks.
- Researchers must build a multifaceted strategy that recognizes human imperfections.
- For example, Verasight monitors attention check failure rates overtime to systematically identify low quality respondents, while acknowledging a good-faith respondent is more likely to fail an attention check than a bot.

Synthetic Samples - How to evaluate them

What did we do?

The prompt to our LLM contains the following instructions:



GPT-4o mini

Synthetic Samples

Example

We repeatedly prompt the model with the personas of each of our respondents, and we get back a list of 1,500 responses to the Trump approval question that we can analyze for a match.



Synthetic Samples

High-Level Performance

Overall, we found that our bespoke synthetic sample technology, using several leading LLM providers, replicated Trump's approval rating with a **mean absolute error of about 4 percentage points.**

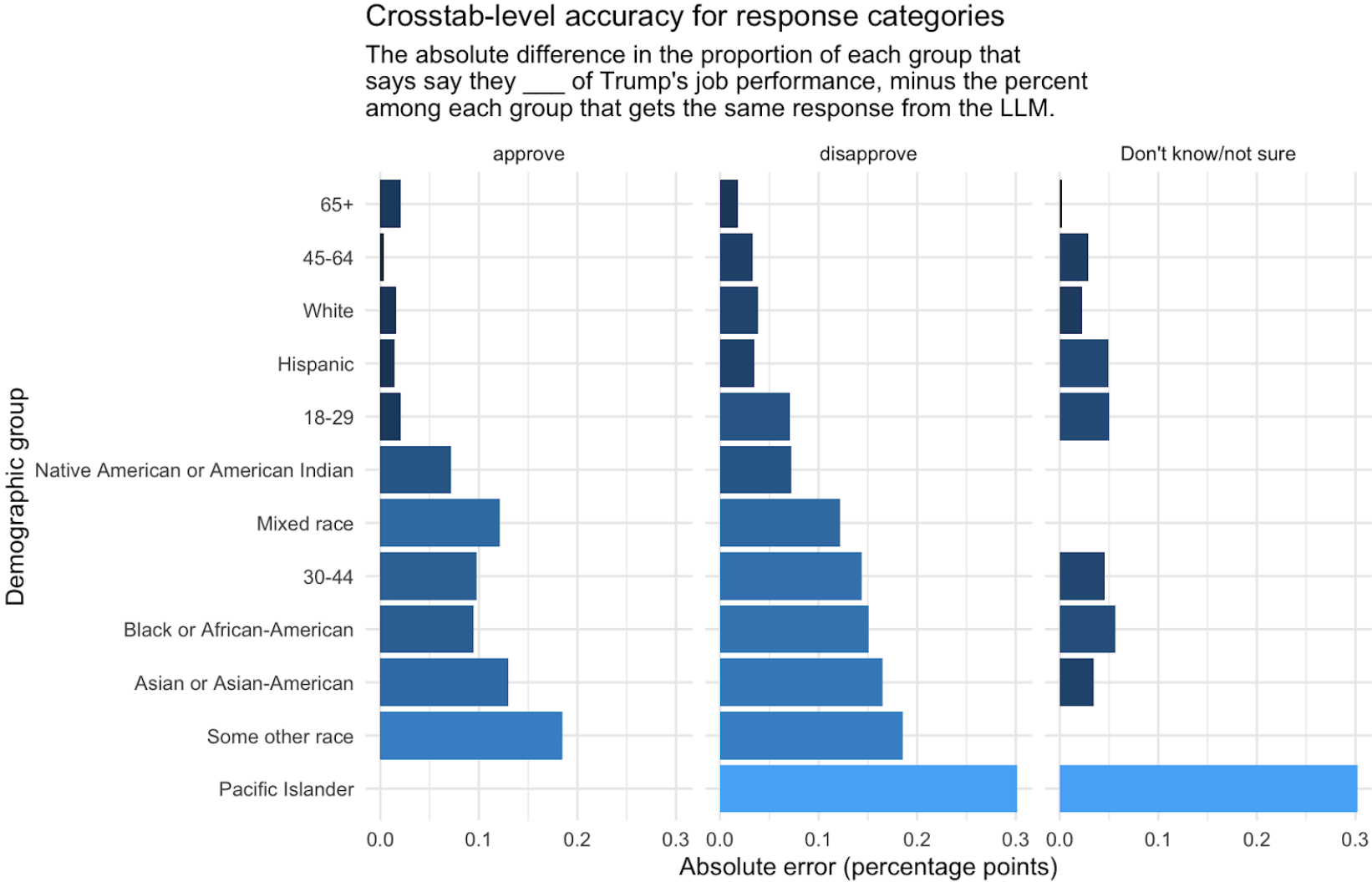
Synthetic Samples – Fail at Sub-Group Level

Synthetic Respondents

Actual Respondents

Error

Synthetic Samples – Fail at Sub-Group Level



The average absolute difference between the population proportion and LLM proportion for these demographic subgroups is **8 percentage points**.

Synthetic Samples - Early Warnings

- On the respondent level: We find that the LLM is able to loosely recover population means, and has such trouble with crosstab means, because it generates errors at the respondent level that sometimes cancel each other out.

Key Recommendations

01 Leverage respondent verification

As the target population gets narrower, the importance of respondent verification increases

02 Attention checks do not outsmart bots

Avoid inadvertently building checks that remove real respondents and increase prevalence of bots in final data

03 Approach synthetic data with caution

Claims of accuracy might be the result of errors cancelling each other out

Questions

Free National Survey Question & Entry into Pappy Van Winkle Giveaway



National Survey Question Entry

- 1000 nationally-representative respondents
- Demographic variables
- Topline & crosstab report

Ben Leff, Co-Founder & CEO, Verasight

ben@verasight.io

