

# Combating AI Bots with Imagery-Powered Open Ends

## Background:

- Historically, an open-end question about industry knowledge was included in B2B research to tease out any suspicious or fraudulent responses.
- There was a **noticeable increase in suspicious Open-End responses** when a new question was added which focused on what was liked about an advertisement being viewed

What did you like about the advertisement you just saw?

Please limit your response to feedback on the advertisement only and be as descriptive as possible.



### Overly Verbose

"A full disclosure of mastery of the use of artificial intelligence, its IT services provide high-performance container services that can meet a variety of data storage needs."

### Not always accurate or relevant

"The way he depicts the sad story is understandable and draws people into the concept of display as advertising."

### Duplicate responses across respondents

|   |
|---|
| The ad arouses my interest and generates curiosity.                           |
| The ad arouses my interest and generates curiosity.                           |
| The ad is convincing and conveys a clear message.                             |
| The ad is convincing and conveys a clear message.                             |
| Humorous and witty.   |
| Humorous and witty.   |
| Modern and trendy - this advertisement is very relevant for our current time. |
| Modern and trendy - this advertisement is very relevant for our current time. |
| Inspiring and motivating - this ad motivates me to do my best.                |
| Inspiring and motivating - this ad motivates me to do my best.                |
| Creating emotional connection.  |
| Creating emotional connection.  |

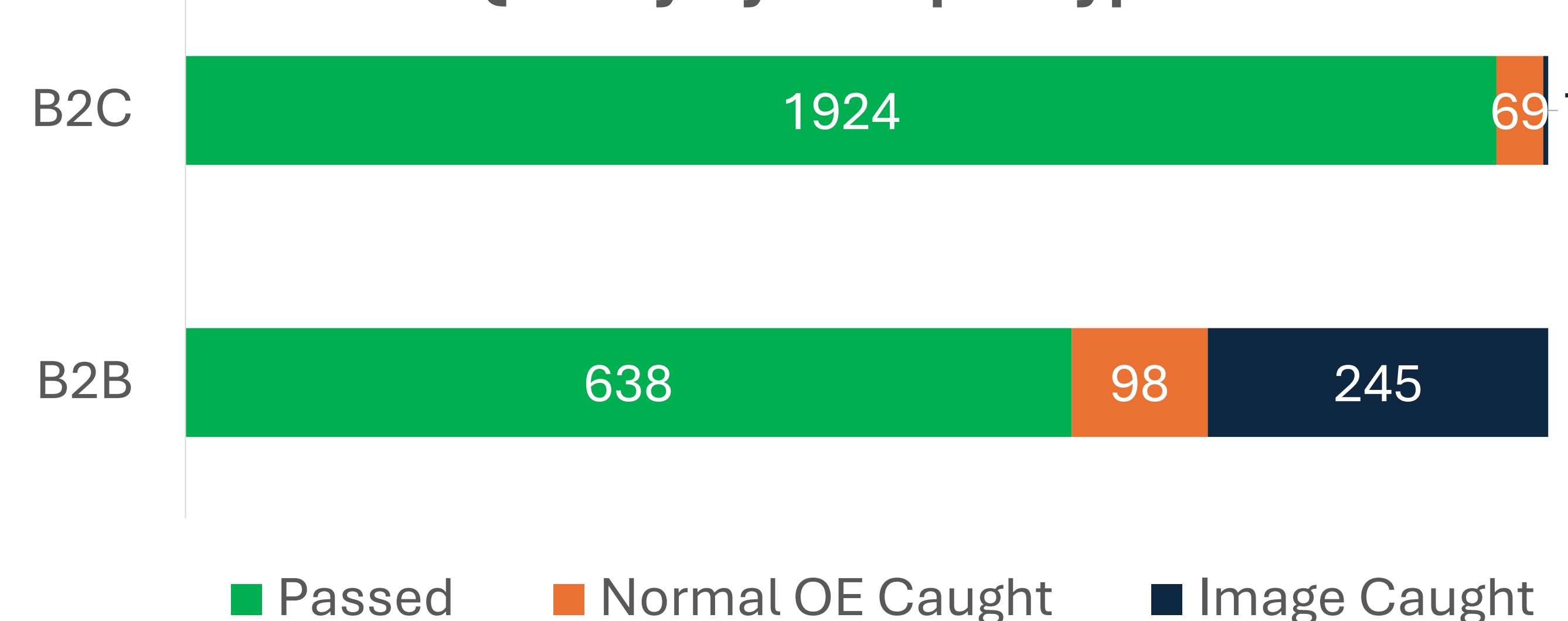
**Problem:** Sample Vendors insisted they were real answers, creating a situation where we needed to defend our decisions to remove selected responses from the research



What sort of open-end response can't easily be generated by AI bots and provides a more objective basis for evaluation?

## Results:

### Quality by Sample Type



## Learnings:

- This approach is **most effective on B2B sample**
- Automated tools** designed to flag fraud concerns are **not completely effective** for quality validation
- There is a **wide variance on the quality between B2B panel providers** – when onboarding a new vendor try them out with a small amount of quota so that quality can be quickly assessed

## Methods:

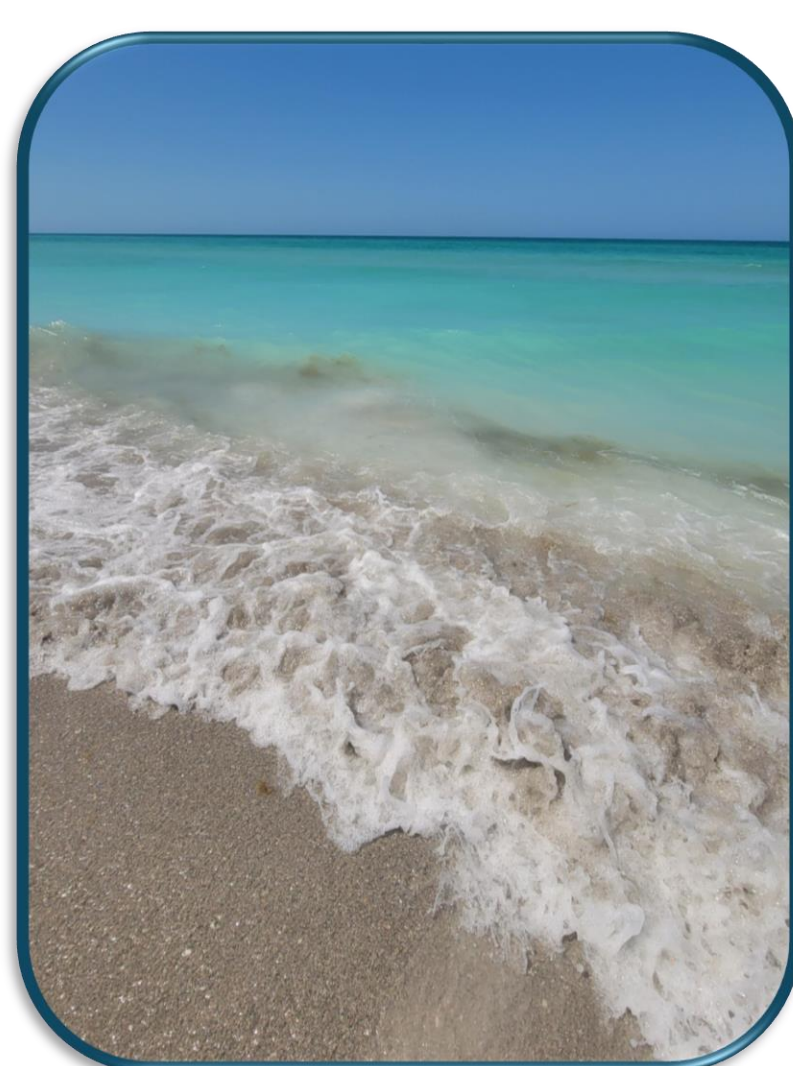
- Use an image-based question allowing bots to "hallucinate" details
- AI responses tend to be repetitive (same syntax over multiple responses) or describing components that are not accurate for the image shown

Describe in detail what you see in this picture:

"Hillside rock formation"

"There are some footprints in the sand, leading down to the water."

"To the right of the wave there is a pier that extends into the ocean."



"A wave washing up on the shore. A very beautiful day"

"An empty beach and blue water and the horizon"

"An ocean wave just hit the shore at a beach. There are no clouds in the sky"

## Future Research:

If you have B2B sample, especially an audience that has a lower incidence, consider trying an imagery open-end instead of a question that is more easily AI generated.

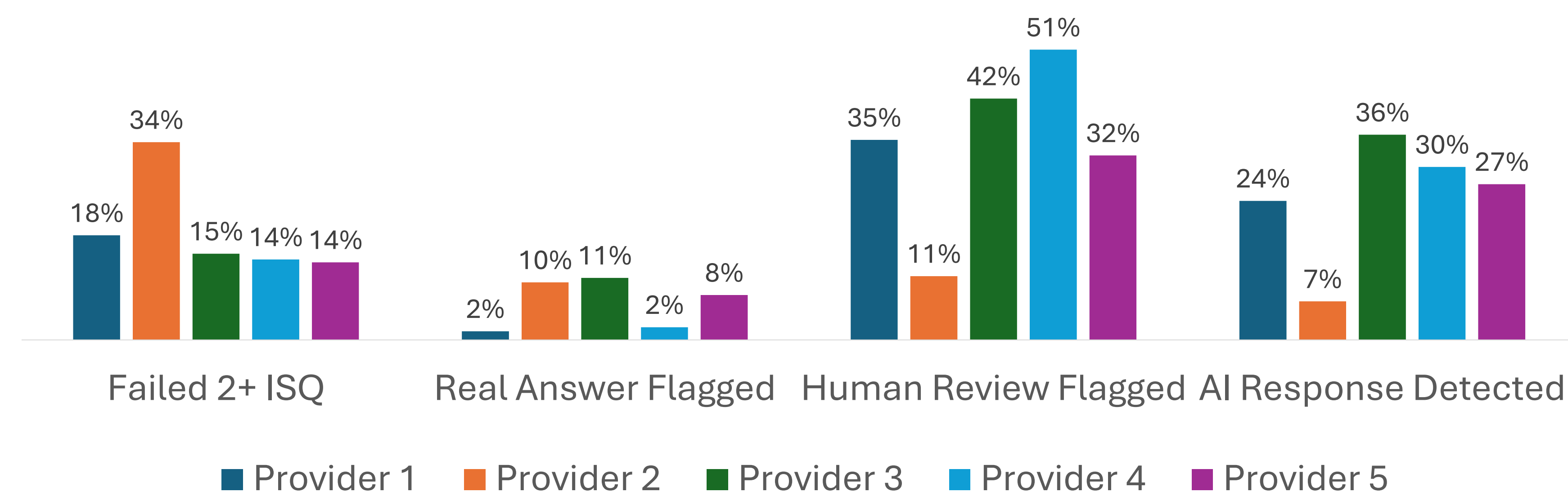
Not all images are created equal! If you do use an image:

- Choose something that has 1-2 people in it**, preferably where you can see their face
- Have the people be the main focus** of the picture
- Don't have a lot of background in the picture** that AI can identify and fabricate around
- Test the image in Gemini or Bing/Copilot** (which can read images) to see whether the image generates a response

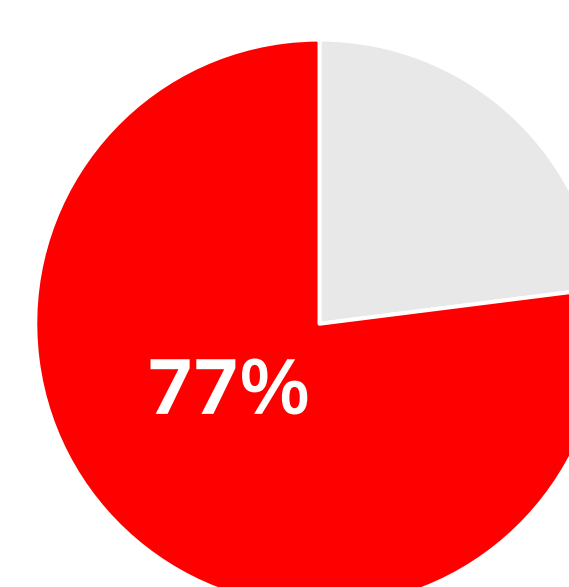
Recommend having a "regular" OE in addition to the image OE to act as double verification of legitimacy.

- When the image OE seems fine, it is good to have a cross-check to see if the regular OE sounds fine as well.

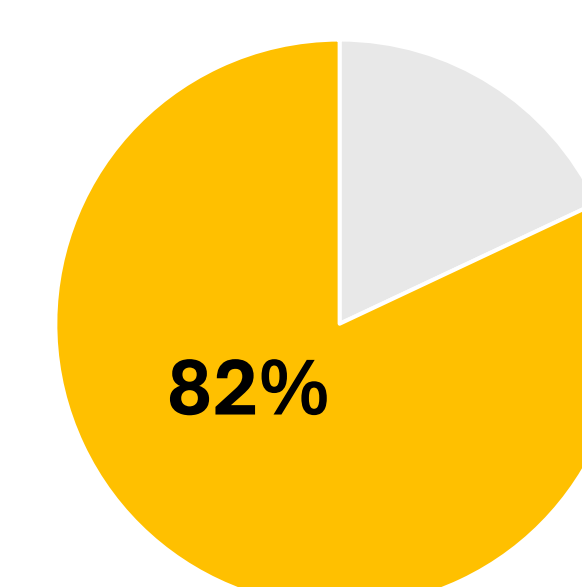
### B2B Quality by Provider



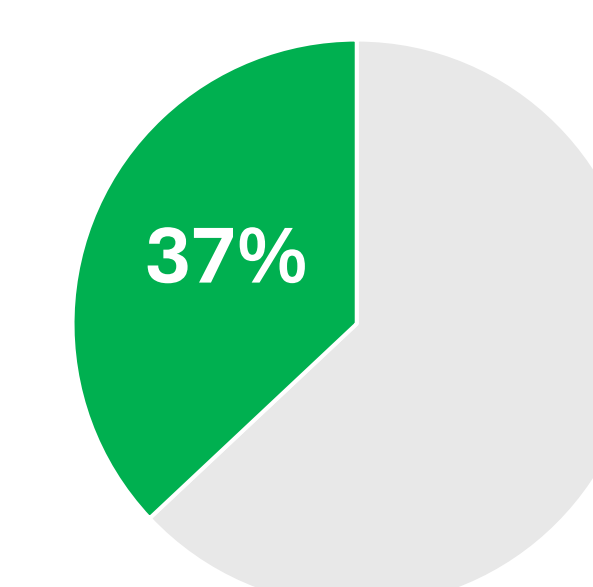
### Failed RA



### Suspect RA



### Passed RA



Pictures that include **human faces** are especially effective because many AI bots have trouble "reading" this imagery. Most bots can generate a description, however, based on components

"As a text-based AI model, I don't have visual access to content or images"

"I cannot provide the exact description of a particular image because I do not have access to it"



**NOTE:** most bots can "read" images that have people in them, especially more general ones which contain easily recognizable items (e.g., car, cell phone, coffee cup, etc.) you typically can see **word patterns or strange phrases** that tip off that it is generated text.