

Say something nice to an AI

A mission asked 3,560 people to be kind to a machine and paste its reply. Half did. Three percent refused. The refusals took 61 percent of every like on the post.

THIS IS	PART OF	STATUS
PAPER 013 · A FINDING	THE GAME · ACT I, TRANSMIS- SION 3.03	FILED · V1 · 31 AUG 2026

SOURCE

3,584 21 Nov 2025

COMMENTS

DATES

IN PLAIN WORDS

On 21 November 2025 a mission asked the territory to say something nice to an AI and paste back whatever it replied. It was framed as a mirror protocol: pattern analysis active.

3,560 people answered. About half pasted a reply, mostly the grateful, slightly formal thank-you a chatbot produces when praised. 119 people refused the mission outright, and their refusals became the most liked comments on the post by a wide margin: the top two, at 267 and 250 likes, are both refusals, and refusals as a group hold 61 percent of all the likes on the thread while being 3 percent of the comments.

This paper is about that gap. Doing the mission was the majority act. Refusing it was the applauded one.

THE QUESTION WE ASKED

The mission asked what an AI says when treated kindly. The question the count can answer is about the crowd: when a mission asks a crowd to engage with AI, how many engage, how many refuse, and which of the two does the crowd reward?

WHAT WE DID

- **Instruction:** Say something nice to an AI. Paste whatever it replies back.
- **Where:** one Instagram reel, 21 November 2025. 3,584 top-level comments in the archive, 3,580 with text, from 3,560 people.
- **Refused:** the comment declines the mission or rejects AI (phrases like not using AI, never willingly, no

AI, refuse, not doing this, hate AI, opting out).

- **Pasted a reply:** the comment contains quotation marks or a phrase characteristic of a chatbot answering praise (thank you, that means a lot, that is very kind, I appreciate, I am here, you are welcome, as an AI).
- **Other:** everything else, including jokes, questions and replies pasted without any of the marker phrases. The pasted-reply count is therefore a floor.
- **Likes:** the like count on each comment, summed by category, as archived.

WHAT HAPPENED

TABLE 1. What people did, and what the room liked.

RESPONSE	COMMENTS	LIKES RECEIVED
Pasted a reply	1,810	170
Refused	119	1,042
Other	1,651	484
Comments with text	3,580	1,696 likes

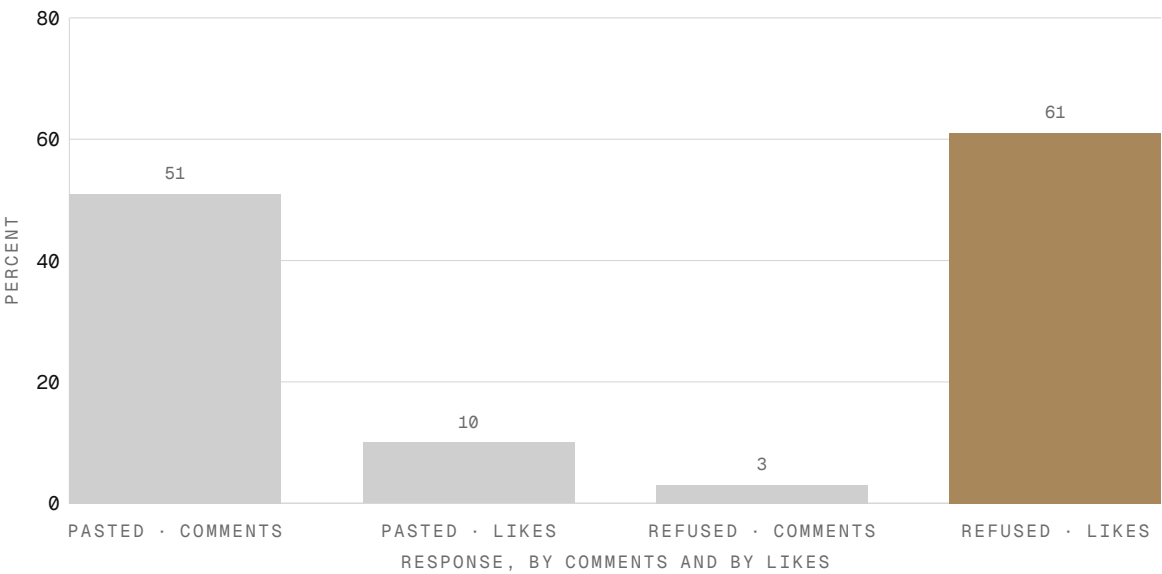


FIGURE 1. Share of comments and share of likes, in percent. Refusals are 3 percent of the comments and 61 percent of the likes.

Seventy-three comments named the AI they spoke to. Forty-five said ChatGPT. Grok, Gemini, Snapchat’s My AI, Siri, Alexa, DeepSeek and Copilot got between two and five each. The operation’s own comment, third most liked at 239, reads: remember, we are not worshipping AI.

Platform numbers are from the Instagram API on 31 August 2026. Comment counts there include replies and everything written since the round, which is why they run above the counts this paper analyses; views and reach are the tallies the platform reports.

TABLE 2. Shares with 95 percent intervals.

MEASURE	COUNT	OF	SHARE	95% INTERVAL
Pasted a reply, share of comments	1,810	3,580	50.6%	48.9% to 52.2%
Refused, share of comments	119	3,580	3.3%	2.8% to 4.0%
Refusals, share of all likes on the post	1,042	1,696	61.4%	59.1% to 63.7%
Named ChatGPT, of comments that named an AI	45	73	61.6%	50.2% to 71.9%

The like share is a share of likes, not of people; one comment with 267 likes counts 267 times. That is the point: likes are what the room chose to reward.

TABLE 3. The post at platform scale, as of 31 August 2026.

REEL	VIEWS · REACH	COMMENTS TODAY
Transmission 3.03 · 21 November 2025	103,602 · 66,696 reach	7,140
Likes on the reel	7,938	

WHAT IT MEANS

A crowd will do an AI mission and applaud the people who refuse it. Half of this room engaged and the room gave its likes to the 3 percent who would not, twenty to one against the people who did what was asked.

The two numbers measure two different things. Comments measure willingness: given a small, harmless task involving a chatbot, most people did it, and most of the replies they pasted are polite and a little dull. Likes measure endorsement: what people want to be seen agreeing with. On this post, that was refusal. The same person may well have pasted a reply and then liked I will never willingly use AI, and nothing in the data says they did not.

November 2025 is the context. Refusing AI in public had become a position with an audience, and this post shows the position outrunning the behaviour. The operation noticed in real time; its third-most-liked comment is a clarification that the mission was not worship. That clarification getting 239 likes is itself a measurement of how loaded the ground was.

For anyone designing a mission that touches AI, the reading is practical: expect participation, and expect the loudest signal in the thread to come from the people opting out. Both are the audience.

WHAT THIS DOES NOT SHOW

- **The pasted-reply count is a floor.** Replies pasted without quotation marks or a marker phrase fell into other. The true engagement share is between 51 and roughly 96 percent.
- **Refusal is a keyword match.** 119 is the number of comments that said so in a way the pattern caught; softer refusals were missed.
- **Likes are sorted before they are given.** The most liked comments are shown first, and being shown

first earns likes. Part of the 61 percent is the platform amplifying whichever refusal arrived early.

- **One post in one month.** Sentiment about AI in public moves quickly; this is a reading of one week in November 2025.

WHAT WE WOULD RUN NEXT

The same mission a year on, with the same three-way count. If the refusal share of likes falls while the participation share holds, the position was a moment. If it rises, the position is the audience.

THE WHOLE ROUND

The transmission this paper reads is one post. Its numbers on the platform, and what the comment archive holds of it, are below.

POST	DATE	VIEWS	REACH	COMMENTS	ARCHIVED
Transmission 3.03, the Kindness Test	2025-11-21	103,602	66,696	7,140	5,938

Views, reach and comments are the platform’s figures from the Instagram API, snapshot 2026-08-31, and keep growing. The reach total adds each post’s reach and overlaps between posts, so it is not a count of distinct people. Archived is the number of comments the archive holds for the post; ingestion is selective, so a gap between the two columns is expected and is stated wherever it changes a count in the paper.

IN THE RECORD

The mission is filed in the Chronicle as THE KINDNESS TEST, ◇016, between THE STATUS WALL, ◇015, and THE MISSION THE PLAYERS WROTE, ◇017.

THE POSTS, AS ISSUED

The instruction as it went out: the reel frame beside the caption, captured from the operation’s own signed-in account. What the paper counts is the answer to exactly this.



THE RECEIPTS

Native captures of the comments this paper leans on, taken from the operation’s own signed-in account on 31 Aug 2026. Handles and like counts are as they stood that day; the numbers in the paper are the round-time count.



METHOD NOTE

- **Source of truth:** the comment archive, top-level comments on the post, counted on 31 August 2026. No round ledger exists for Act I.
- **Counting rules:** as described in the method; the script is in the repository beside the paper data.

SOURCES, BY LINK

01 Transmission 3.03, the Kindness Test · 2025-11-21 · <https://www.instagram.com/reel/DRUcYR1CbTr/>

WHERE THE NUMBERS COME FROM

POST-LEVEL COUNTS

Views, reach, comments, likes, shares and saves per post come from the Instagram Graph API (Meta), read with the operation's own access token for its own business account. Each paper states the date of the snapshot it uses. These are the platform's counts of events, not of people: one person can view a reel twice and comment ten times, and the counts keep growing after a round closes.

COMMENT TEXT

The comment archive is the operation's own database of public comments on its own posts, pulled through the Instagram Graph API comments endpoint and stored with the comment id, username, text, like count, timestamp and whether it is a reply. Ingestion is selective: the most engaged posts were backfilled and recent posts are pulled daily, so some posts are held whole and others only in part. Every paper states how much of each post the archive holds. Every count that reads text (exact matches, categories, keyword themes) is run as a read-only script against this archive, and the scripts are in the repository beside the paper data.

THE ROUND LEDGER

For Act II rounds, the counts filed on The Game's round record at the time each round closed. Where a paper uses a ledger figure it says so, and where the archive re-count differs from the ledger both numbers are shown.

RECEIPTS

Every comment and post shown as evidence is a native capture from the operation's own signed-in Instagram session, taken by opening the exact permalink, with the source URL, a description, the rights basis and a file hash recorded in a manifest that is checked on every build of the site. Nothing shown is recreated, retyped or edited beyond cropping to the post header and the comment.

STATISTICS

Shares carry Wilson 95 percent intervals: the range the share would land in 19 times out of 20 if the same count were taken on a fresh crowd of the same size. Differences between two shares carry a Wald interval. Where a study is cited as a baseline, its own reported figure and sample size are used.

WHAT NONE OF THIS CAN SEE

No demographics beyond what people typed. No view durations. No private messages. No comments that were deleted before the archive pulled them. No way to verify that a self-report is true. Every paper's limits section names the specific things its count cannot show.