Just had a post shower thought.

Instead of boiling it down into tiny nuggets of creativity tests, or aside from those, actually test it in a corporate environment. Have a series of scripts / videos of meetings. Actual innovation / defense lab type meetings, kind of like the recursive innovation adventure / bad lipreading video.

Have like three recordings. For each, also have three ways of delivering it to the subjects. VR, vertical TV's pulled up to the table that light up whenever their person is talking, and maybe one abstract one of some kind. Auditory only. Just a video instead of VR.

For each of those nine settings (or however many), have them provide input at the end. Written in a room with no one watching (neurodivergent), spontaneous verbal (fast talker), something else.

Then have some objective measure for how innovative their responses are. Maybe a panel of scorers based on a few different definitions of innovation, or words slightly like innovation. Have a panel of LLMs score them the same way (say, GPT4, Claude, Poe, Llama, Phind, etc).

Then we come back to the individual and be like hey one more thing. You have 400 words to write your own script.

Now do it again, but with the aid of one predetermined LLM. You have x "tokens" available to augment your ingenuity. Tokens are a measure that the LLMs use to measure amounts of words, usually for limits or billing purposes.

Evaluate their scripts in a similar manner.

Maybe don't tell them it's about innovation, to kind of stunt really oddball useless word vomit. Instead tell them we're working on a compelling story. Or product.