

An anonymous student-chatbot interaction from the AI Club in Perth, Scotland. The Chipp session is anonymous by design. There is no login, and no records are kept or used to train AI models. Anyone curious about student-data safety can read Chipp.ai's SOC 2 Type II security report. What follows is a verbatim conversation, captured just before the student joined a Harkness discussion. It was one of their first, so the chatbot was built to scaffold heavily, a supported first step into Harkness rather than a replacement for the thinking.



SCAN FOR THE EPISODE

16:23 Voice message - no speech detected

AI ASSISTANT · 16:23

Welcome to Post-Truth.

Here's how this works: I won't give you answers today. I only ask questions. Your job is to think out loud and defend what you believe, and to change your mind if I find a crack in it. At the end I'll write you a coaching sheet to take into the class debate.

Here's the question we're circling: if you can't fully trust a photo, a video, or a recording as proof any more, what can you trust? How do we tell fact from fiction when anything can be convincingly faked?

Start wherever feels honest. Before AI could fake all this, how much did you trust a photo or a video as proof, and why?

(Also — do you have an AI Language Customisation Profile? Paste it in and I'll adjust for the whole session.)

STUDENT · 16:23

Is this actually working?

AI ASSISTANT · 16:23

Yes, it's working! I can hear you — well, read you.

Your voice message came through silent, so I couldn't catch what you said. Try again?

Or just type if that's easier.

And we're ready to go whenever you are — the question's waiting.

STUDENT · 16:25

I'm ready

AI ASSISTANT · 16:25

Great. Let's go.

Before AI could fake a photo or video — how much did you trust one as proof of something, and why?

STUDENT · 16:28

I felt i could almost fully trust audio, photos and videos as proof of someone doing something, there being a slight grey area in the case someone edited the media to bend the truth though usually it would be very trustworthy. It was a more trustworthy piece of evidence as it gave physical proof of someone doing something they denied, unlike a handwritten letter which someone could fairly easily forge

AI ASSISTANT · 16:28

So recordings felt close to airtight — physical proof that cut through denial. Here's what I want to push on: you said *physical proof*. But a photo or video was always just light hitting a sensor, then data. What was the thing underneath that made it feel more trustworthy than a letter — was it the *difficulty* of faking it, or something else?

STUDENT · 16:33

Editing a photo can make the final product really choppy and notably fake, same with editing audio. Editing a video is an entirely different level of difficulty to make it look realistic. It could take days to copy someone's handwriting, even less if you were a pro at it, and forge a letter or something written which made a video, photo or audio more trustworthy as it was proof of the person actually doing the deed right in front of your eyes.

AI ASSISTANT · 16:33

So the trust was really resting on **difficulty** — faking was hard enough that most people couldn't do it convincingly. That's a sharp point. Now here's the rupture: AI has just collapsed that difficulty. But flip it around — imagine a *real* video surfaces tomorrow of a powerful politician doing something terrible. They look straight at the camera and say: "*That's a deepfake.*" What can that true video actually prove now?

STUDENT · 16:37

The sad truth is, if the jury believes that the video is AI it will most likely not be considered real footage evidence. The video would have to be extremely believable, such as the same event being recorded at different angles which AI would easily make mistakes over, or there would need to be some even more damning evidence that could prove their guilt. The question on what kind of evidence, i'm not entirely sure as I've always considered videos and photos as quite reliable evidence.

AI ASSISTANT · 16:37

So you're saying the video's power to prove anything has been *undermined* — not because it's fake, but because the *possibility* of faking now gives the guilty an easy exit. That's exactly the crack. And notice the direction it runs: the more people *know* deep-fakes exist, the easier that dodge becomes — even when the video is completely real. So what's left? If someone showed you footage tomorrow, what would it take for you to actually believe it?

STUDENT · 16:41

In a court of evidence, it is entirely different to being shown footage by someone at my school. For example, if some person i'm not very close with shows me a video of my friend doing something they would never do, i would be quite skeptical and would try my hardest to find something to disprove it as fake. If it were in a setting like a court, I would try my best to observe both the accused and prosecutor's body language to make any kind of assumption about the truth. Granted that is not a very reliable way, but it can give an extra inkling to help the person telling the truth.

AI ASSISTANT · 16:41

So personal trust in the *source* shapes how you read the evidence — and in court, you're reading *people* as much as the footage.

But here's the crack in that: body language can be misread, and a confident liar can fool a room. You're essentially replacing one form of proof with another form of testimony. What makes a *person* more trustworthy than a recording — when both can deceive you?

STUDENT · 16:44

That's why I am aware it's not the most trustworthy tactic. But say someone shows they are anxious as if trying to think of an excuse before claiming the video evidence was AI, depending on their body language being scared or more nervous i think i might be able to have a small idea whether they are guilty or not. Same with the prosecutor, if they look nervous when their evidence is called out as AI, i would be less likely to believe them than if they looked in disbelief.

AI ASSISTANT · 16:44

So you're triangulating — using behaviour as a *secondary signal* to test the footage, not replace it.

Smart. But push this further: a skilled actor, a trained politician, a sociopath — they can fake calm. And someone *innocent* being falsely accused might look *exactly* as panicked as someone guilty.

If body language can be performed just like a video can be faked — what's actually left that *can't* be faked?

STUDENT · 16:44

We're running out of time, would you be able to post the final output?

AI ASSISTANT · 16:45

Here's your coaching sheet. Copy the whole thing onto page 2 of your handout and bring it to the circle.

YOUR HARKNESS PREP — POST-TRUTH

The question on the table: If we can't trust a photo, video, or recording as proof, what can we trust, and how do we tell fact from fiction when anything can be faked?

Where you landed: The old trust in recordings rested on how *hard* they were to fake — AI collapsed that difficulty, and now even real footage can be dismissed; we're left triangulating between the evidence, the source, and people's behaviour — none of which is foolproof.

Your sharpest point: The danger runs *both ways* — deepfakes don't just let people fabricate evidence, they let guilty people dismiss *real* evidence as fake. A true video of a crime proves almost nothing if the

accused can simply say "that's AI."

Lines you can use (pick what fits the moment):

- *To build on someone:* "I want to build on what ____ said, because..."
- *To make your point:* "The problem isn't just that fakes exist — it's that knowing fakes exist gives anyone an easy way to deny real evidence."
- *To raise your doubt:* "I'm not sure body language or source trust actually solves this — can't those be faked just as easily?"

Evidence you can bring in (from the reading — introduce it as "in the reading..."):

- Regina Rini calls recordings an *epistemic backstop* — they quietly kept people honest because anything might be on camera. Deepfakes erode that backstop even when no fake is actually used.
- Citron and Chesney's term for this is the *liar's dividend* — the mere existence of deepfakes lets people deny authentic footage, which is exactly the scenario you described with the politician.
- The Scottish Parliament's *Chamberfakes* report warned that faked footage of MSPs could undermine trust in democratic institutions — your point about courts applies there too.

Topics you can jump in on (where your voice will add something):

- Why the *difficulty* of faking was the real foundation of our trust in recordings — and what happens now that's gone.
- The two-way threat: fabricating false evidence *and* dismissing real evidence.
- Why falling back on people's behaviour and source trust is understandable but still fragile.

Remember the Harkness way (from page 1):

- Wait your turn, and build on the last person before adding your own.
- Disagree with the idea, not the person.
- Changing your mind out loud is a win, not a loss.

You don't need to say all of this. Listen first, then pick your moment.