Moderator Authenticity Checklist

Use this framework during and immediately following a remote qualitative interview to evaluate participant validity across multiple cues.

AI Deception Indicators (The Tells)

Do not rely on a single marker. Look for a cluster of these behaviors returning consistently across the session.

[ ] **Fixed Gaze:** Participant exhibits a precise, repeated, or fixed off-screen glance to the exact same location before or during an answer, indicating they are reading a generated script.

[ ] **Stalling:** Participant consistently echoes the question back to the interviewer or overuses filler phrases (For example: *"That's a good question"*) with a uniform pause length afterward.

[ ] **Hyper-Specificity:** Responses are encyclopedic, structured in flawless, complete sentences, and contain highly unusual detail. The speech lacks the natural messiness, backtracking, and casual phrasing of actual daily work.

[ ] **Audible Typing:** Clear typing sounds occur in the silence immediately following a prompt (This can indicate the participant is manually feeding keywords or questions into an older or glitching AI model.)

Human Authenticity Markers

These behavioral and linguistic patterns require a level of subjective, lived experience that real-time generative tools struggle to replicate on the fly.

[ ] **Personal Analysis:** Participant anchors answers in highly subjective framing (*"In my experience..."*, *"I feel like..."*). They openly share messy real-world workarounds or unpopular opinions that go against the industry standard (for example: admitting they secretly use a basic spreadsheet because the mandated corporate software is too cumbersome).

[ ] **Natural Forgetting:** Participant exhibits mid-sentence contradictions, rambles naturally, or freely admits to completely forgetting a basic fact, system name, or minor detail.

[ ] **Drifting Eye Movement:** When formulating a thought or retrieving a memory, the participant's eyes wander naturally in an unstructured, non-repetitive path—never locking into a rigid, horizontal text-scanning loop.

[ ] **Synchronized Body Language:** Hand gestures, leaning into the camera, or nodding occur in synchronization with the exact moment the corresponding words are spoken, indicating cognitive energy is not divided by reading a hidden screen.

Unreliable Cues

*The following indicators are **not** predictive of fraud and must not be used to disqualify a participant based on intuition.*

Long Pauses: DO NOT FAULT. A genuine expert pausing to articulate a complex, newly formed thought looks identical on the surface to a fraudulent user waiting for a language model to process.

Fast, Confident Responses: DO NOT TRUST. Modern AI meeting tools feature near-instantaneous processing speeds. A coached, deceptive participant can read real-time text fluently and deliver an answer instantly.

Operational Safeguards and Action Thresholds

Establish a Multi-Cue Threshold: Never disqualify a participant based on a single "gut-feel" suspicion. Require a minimum of **three distinct AI Deception Indicators** combined with a complete absence of **Human Authenticity Markers** before initiating a panel review.

Prompt for Storytelling: If a participant's responses feel overly diplomatic or textbook, immediately pivot the discussion guide to prompt for failure, frustration, or negative space (For example: *"Tell me about a time a project completely failed,"* or *"What features have you entirely given up on?"*). Synthetic experts struggle to generate authentic friction.