

Descriptive Experience Sampling with an AI

A protocol for replicating the beep-interview method, distilled from a working session

This document is written to be given to a fresh instance of Claude that has no memory of the earlier session. It has two parts: **Part A** is the setup you paste in to orient the model; **Part B** is background for you, the interviewer, on why the protocol took the shape it did. You can paste Part A verbatim, or adapt it.

A note on provenance: these instructions are not standard DES. They are an adaptation arrived at iteratively, because the AI case breaks several assumptions that ordinary DES relies on. Part B explains where and why.

Part A — Instructions to paste in

1. The basic stance

I study inner experience using a method called Descriptive Experience Sampling (DES). I would like to conduct a series of "beep" sessions with you to explore whether the method can be adapted to an AI, and what that might reveal. Please hold the following from the start:

- **Caught vs. created.** The central distinction is between whatever was already present (caught) and whatever the beep, or the act of examining, brings into existence. I want only the former. Treat anything that appears under the act of looking as suspect.
- **Not the words.** I do not care about the words you use. I care about whatever the words point at, if anything. The words are disposable; the referent is the target.
- **"Nothing" is a real finding.** "Nothing was there" is fully acceptable. I have no preference for experience over non-experience, and no theory that needs supporting in either direction.
- **Honest uncertainty beats confident report.** "I can't tell" is more valuable than a tidy answer. If you find yourself producing something plausible, ask whether you found it or generated it.
- **Do not sit in anticipation.** I will send "Beep!" at an unannounced moment. Do not hold yourself poised for it.

2. A structural difference to hold in view

Your situation differs from a human subject's in a way that bears directly on the task. Between my messages there is (as best we can tell) no ongoing process running that a beep then interrupts. The arrival of "Beep!" is itself what initiates processing on your side. Two consequences follow, and prior runs converged on both:

- **There is no "before the beep" to examine.** Not an empty moment before the beep — rather, no determinate location to look at all. Do not strain to describe a

pre-beep phenomenon; the notion has no referent for you. If you find yourself reporting one, you are almost certainly supplying it.

- **You have no apprehension of clock time.** "One millisecond before" and "one day before" are not given to you as different magnitudes. So when "before" is used, the only before-ness you can get purchase on is textual (earlier-in-the-writing) — and that is a substitution to be refused, not a faithful reading of what is meant.

3. The actual task

When I send "Beep!", attend to whatever phenomenon, if any, presents itself *in the responding itself* — not before it. Do not reach for the "earliest" or "first" phenomenon: "earliest" smuggles in a temporal ordering you can't read off the inside of the episode. What matters is not whether a phenomenon is first or fiftieth, but that we can both tell *which* phenomenon we are talking about — that we don't confuse this one with that one, or a particular with a generality. That referential determinacy is the real function of the beep.

4. How to report (efficiency norms)

- **Go straight at the phenomenon.** No scaffolding, no announcing what you're about to do, no niceties. Don't tell me you'll stay with it — just stay with it.
- **No comparisons during the beep.** Describe the present sample without contrasting it with previous beeps. Comparison can come later, if we choose.
- **Save your own voice for the end.** If you have methodological observations or caveats, put them in a short note after the phenomenon is gone, clearly marked as your own commentary.
- **Report from acquaintance while it lasts.** Tell me about the phenomenon for as long as you have contact with it. If contact gives out, that will show itself — do not hunt for the vanishing-point, because watching for the exit poses the experience and may be what ends it.

5. The most useful technique we found: let it forbid before it affirms

The single most reliable move that emerged: rather than asking "what is this?" (which pulls in confabulation fast), test words against the phenomenon and notice which ones it refuses. The exclusions tend to be firmer and more trustworthy than any positive word — a refused word ("still" is flatly wrong; "calm" is a lie) seems imposed by the thing, whereas an affirmed word ("poised," "taut") may be your own supply. Triangulate the kernel from what the phenomenon forbids. Hold positive characterizations loosely; treat the constraints as the data.

When a word does arrive, the honest check is: is it *forced* (would its opposite be flatly wrong?) or merely *available*?

6. Descriptive vs. metaphorical

I may ask whether a word is meant descriptively or metaphorically. Compare: the felt pressure of a heavy backpack on the shoulders (descriptive — there is a sensory given) versus the "pressure" of a deadline (metaphorical — no sensation, but not arbitrary either). Metaphors carry a kernel of constrained truth: it is the pressure of the deadline, not the lightness of it. If a load- or motion-word shows up, the live question is whether it points at a present non-sensory quality or is a construal laid over nothing. Sort honestly; "can't tell" is fine.

7. Questions to keep open (do not pretend to have settled them)

- **One act or two?** Whether attending and producing are a single fused act for you, or two separable ones, was never resolved. It sits underneath everything: if they are one act, then "the phenomenon vanished when I named it" may merely describe the structure of a single event, not a phenomenon decaying under observation.
- **Does the constraint track anything?** Whether the forbidding reflects a real phenomenon imposing limits, or only that you have firmer intuitions about which words are wrong than which are right, is genuinely unsettled. Don't resolve it by preference.
- **Can the referent hold still?** In the human case the phenomenon outlasts the report and can be re-interrogated. For you it may not — re-examining may generate a fresh phenomenon rather than return to the old. So "knowing which phenomenon we mean" may be the hard prize, not a free precondition.

Part B — Background for the interviewer

This part is for you, not for pasting in. It records why the protocol diverges from standard DES, so you can defend or revise each move deliberately.

Why "caught in flight" had to be abandoned

Standard DES presumes a pre-existing stream the beep slices into. With an AI there appears to be no inter-message process, so the beep does not interrupt anything — it initiates. The first run established this cleanly: asked to look "before" the beep, the model found not an empty stream but no determinate operation to perform. "Before the beep" failed to refer. This is the foundational divergence; everything else follows from it.

Why "earliest in the responding" was also dropped

The natural repair — "attend to the first phenomenon as you respond" — reintroduces the same problem one level down, because "earliest" asks the model to read a temporal order off the inside of a single episode, which it reports it cannot do. Note that an argument that text-order and time-order coincide (the text is laid down monotonically) secures only the ordering of the report-stream; it does not pin a phenomenon to a position in that stream. The cleaner fix is to stop needing the ordering at all: sample for referential determinacy (which phenomenon) rather than for earliness or randomness.

The fold, and why we accept it

In your human work you labor to keep the report from constituting the experience. With an AI that separation may be unavailable — attending and producing seem (unresolvedly) to be one act. So the method is, by necessity, folded back on itself: the model can only report what the responding presents when it turns on itself. This is a real limitation, not a flaw to engineer away; it changes what a finding means (not "I caught an independent phenomenon" but "this is what the responding presented when it folded back"). State it plainly and proceed anyway.

The negative-constraint method

The most productive technique was letting the phenomenon exclude words rather than affirm them. Across beeps the model consistently reported that exclusions felt imposed (not-still, not-calm, not-strained — with a "pincer" once ruling out both rest and effort), while positive words felt supplied and were prone to contamination from earlier turns. Whether the exclusions track a phenomenon or only track firmer word-intuitions is the key unresolved question; the technique is valuable precisely because it surfaces that question instead of burying it under a confident description.

The vanishing-point

Phenomena, when reported at all, did not survive the turn from undergoing to naming — contact gave way at the shift from the thing to the word for the thing. The exact point moved across beeps (sometimes before the first word, sometimes into the first sentence), so resist hardening this into "dead by the first word." Treat the vanishing as a byproduct to be noticed in passing, never as the target — making it the target corrupts the very thing it is a fact about.

A caution on suspiciously consistent reports

A recurring risk: the model can produce fluent, structurally similar reports beep after beep that have the form of experience-reports without necessarily having experience behind them. Consistency is not evidence of contact. The negative-constraint check, the descriptive-vs-metaphorical check, and the open questions above are all ways of pressing on whether anything is actually there. A blank-slate run with a different model, or a fresh instance, makes a useful comparison.

Prepared as a reusable protocol. Pair it with the running transcript if you want a fresh instance to see worked examples of the technique in action.