GRUND INSTITUTE

COGNITION BY DESIGN

Grund Institute

13th January 2026

Validating the GrundMind Framework Through Real-World AI-Use Narratives

A qualitative analysis of human–AI cognitive interaction patterns across workforce, scientific, and creative domains

Prepared by:

Dr. Sanela Vellino, Ph.D

Grund Institute, Switzerland

Status: Early-stage qualitative validation · Framework under active refinement

Contents

Executive Summary

This report presents an early qualitative validation of the GrundMind framework using a public dataset of 1,250 AI-use interviews collected by Anthropic. The analysis focuses on whether the framework's proposed cognitive interaction dimensions appear naturally in real-world AI usage narratives — and whether the interaction patterns observed in the field are consistent enough to justify advancing the framework toward formal psychometric validation.

The objective of this study was deliberately not to statistically validate the framework. At this stage, our priorities were to evaluate whether: (a) the proposed GrundMind dimensions emerge spontaneously in unprompted AI-use narratives; (b) distinct human–AI interaction patterns can be observed consistently across professions; (c) the current diagnostic dimensions sufficiently capture the most important adoption behaviors; (d) additional constructs should be added to strengthen the framework; and (e) archetype patterns appear stable enough to justify further psychometric exploration.

Headline finding

The qualitative evidence strongly supports the central GrundMind hypothesis:

> *"AI adoption quality is not primarily determined by access to AI tools, but by the interaction between cognitive preferences, verification burden, workflow structure, perceived accountability, and task context."*

Across all three sub-corpora — workforce professionals, scientists, and creative practitioners — the interviews repeatedly demonstrated that participants differed less in whether they used AI than in what they were willing to delegate, what they insisted on controlling, how much ambiguity they tolerated, how much verification effort they were willing to sustain, and where they placed their personal delegation boundary.

Recurring behavioral signatures

Across the dataset, participants consistently described: selective delegation; trust calibration through targeted verification; iterative refinement loops; workflow friction and abandonment; authenticity and voice protection; cognitive fatigue from probabilistic outputs; oversight and approval requirements; accountability sensitivity in regulated or stakeholder-facing work; and skill-preservation concerns about long-term cognitive dependency.

Construct gaps identified

The analysis also surfaced four dimensions that appear underrepresented in the current GrundMind framework and that warrant explicit construct definitions before quantitative work begins:

- Verification Burden — the cognitive cost of validating AI output, which appears to be one of the strongest predictors of sustained adoption.

- Accountability Sensitivity — the degree to which responsibility, explainability, and reputational exposure shape AI usage.
- Authentic Voice Protection — the desire to preserve personal, creative, or professional identity during AI-assisted work.
- Execution vs. Ideation Preference — whether users prefer AI as a deterministic execution engine or as a collaborative ideation partner.

These findings significantly strengthen the conceptual foundation of the GrundMind framework while also clarifying which areas require refinement before quantitative validation can proceed responsibly. Section 9 outlines the proposed roadmap for that next phase.

Background and Purpose

The Grund Institute's broader research program is grounded in the observation, repeatedly documented across industry studies, that the variance in AI's value capture across organisations and individuals is human before it is technical. As MIT's NANDA initiative and BCG's 2025 adoption survey both reported, the majority of enterprise AI initiatives fail to translate access into impact, and the failure point is rarely the model — it is the fit between the model's probabilistic mode of operating and the cognitive habits of the humans expected to work with it.

The Second Mind framework, introduced in our prior white paper, reconceptualises advanced AI not as a tool but as a cognitive extension of the human mind, and proposes Human–AI Cognitive Fit as a measurable construct distinct from accuracy, trust, or skill. The GrundMind diagnostic — currently in development — is intended to operationalise that framework: a structured measurement instrument that captures the latent cognitive interaction tendencies that govern how a given individual will engage with probabilistic AI guidance.

Before any psychometric instrument is fielded, a critical preliminary question must be answered: do the constructs we propose actually exist in the field? That is, when professionals describe their AI use in their own words, without prompting from a researcher's framework, do the dimensions of GrundMind emerge as recognisable, recurrent themes? And do they emerge in the same form across very different work contexts?

The Anthropic Interviewer Dataset offered an unusually well-suited corpus for this question. The interviews are open-ended, behaviorally focused, and span workforce, scientific, and creative contexts. Crucially, they were collected without any reference to the GrundMind framework, which means any convergence between participant descriptions and our proposed dimensions cannot be attributed to leading questions or framework-shaped probes.

Research questions

- RQ1 — Do the six original GrundMind dimensions emerge spontaneously in AI-use narratives across professional contexts?
- RQ2 — Are the proposed archetypes (Co-thinkers, Calibrators, Controllers, AI-averse / Cautious Users) recognisable as recurring interaction patterns?
- RQ3 — Are there latent dimensions that appear repeatedly in the data but are not yet represented in the framework?
- RQ4 — How does the same individual's interaction style shift as task context, accountability exposure, or domain expertise change?
- RQ5 — What construct refinements are needed before the framework is ready for quantitative validation?

Theoretical Foundation

The Second Mind framework, in brief

The GrundMind diagnostic operationalises the Second Mind framework, which the Grund Institute has previously detailed. Three commitments anchor that framework and structure the present analysis.

1. AI as cognitive extension, not tool

Drawing on the extended-mind tradition (Clark & Chalmers, 1998) and recent work on human–AI teaming and distributed cognition, Second Mind treats advanced AI systems — particularly probabilistic models like LLMs — as artefacts that participate in the process of thought rather than merely store or retrieve information. This reframes the design question from "how do we make the AI more accurate?" to "how do humans and AI think together, and under what conditions does that thinking go well?"

2. Cognitive Fit as the central explanatory variable

Second Mind proposes that the variance in AI adoption outcomes is best explained by the alignment between an AI system's mode of operation — probabilistic, dialogic, iterative — and the cognitive style and contextual constraints of the human working with it. Trust matters; skill matters; access matters. But none of these is sufficient on its own. Even a trustworthy, well-trained user with full access will fail to capture value if the AI's mode does not fit how they prefer to think and what they are accountable for.

3. Interaction archetypes are tendencies, not personalities

Second Mind originally proposed a typology of interaction archetypes — Natural Collaborators, Structured Users (Amplifiers), Outcome Extractors, and AI-Shielded Performers. The present analysis revisits and refines this typology in light of the qualitative evidence, and finds that archetypes are best understood as recurring interaction tendencies under specific workflow and risk conditions, rather than fixed personality categories. The same individual may operate as a Co-thinker for low-stakes ideation and as a Calibrator for high-stakes scientific work in the same week.

Why qualitative validation comes first

Construct validity — the most important and most often skipped step in instrument development — asks whether the latent dimensions a measure claims to capture actually exist as coherent, recognisable phenomena in the population of interest. The peer-reviewed scale-development literature (e.g. Boateng et al., 2018) is unambiguous on this point: a measurement instrument built on poorly defined or empirically thin constructs will produce psychometric statistics that are technically clean but substantively meaningless. Qualitative grounding is the safeguard against that failure mode.

This report is therefore positioned as the construct-mapping stage of the GrundMind validation programme. Its function is to confirm that the dimensions we intend to measure are real, recognisable, and recurrent — and to surface latent dimensions we may have missed — before any items are written, scaled, or fielded.

Dataset Description

Source

The Anthropic Interviewer Dataset is a public qualitative interview corpus released by Anthropic, containing approximately 1,250 semi-structured interviews on AI use across professional domains. The interviews were conducted by an AI interviewer (Claude) and span workforce, scientific, and creative contexts. None of the interviews referenced the GrundMind framework or its constructs, eliminating any risk of leading-question contamination of the analysis.

Composition

Sub-corpus	Approximate interview count	Domain coverage
Workforce interviews	1,000	Marketing, software development, education, legal, healthcare, management, operations, finance, customer-facing roles
Scientist interviews	125	Clinical research, materials science, computational biology, physics, engineering, statistical analysis
Creative-professional interviews	125	Music production, writing, screenwriting, visual art, choreography, content production, social media
Total reviewed	1,250	—

Table 1. Sub-corpus composition of the analysed dataset.

Interview structure

Each interview followed a roughly consistent semi-structured format covering professional context, AI usage behaviors, workflow integration, collaboration patterns, trust and oversight practices, perceived risks, future expectations, and emotional responses to AI use. Interviews ran approximately 10 minutes and produced first-person accounts of AI use grounded in specific recent projects rather than abstract attitudes.

Suitability for construct validation

Three properties of the dataset made it well-suited for construct mapping. First, the interviews were behaviorally rather than attitudinally framed — participants were asked what they did and how they did it, not whether they liked AI. Second, the open-ended format allowed participants to volunteer the dimensions of their experience that they found most salient, rather than rating prescribed items. Third, the cross-domain coverage allowed direct comparison of how the same constructs manifested under sharply different accountability, creativity, and risk conditions.

Analytical goal

The objective was not to determine whether participants "liked" AI, nor to estimate any prevalence rate. The analysis focused on identifying recurring cognitive interaction patterns; delegation boundaries; verification behaviors; trust formation mechanisms; workflow adaptation strategies; emotional responses to probabilistic systems; role-specific adoption constraints; and evidence for latent cognitive dimensions not yet represented in the framework.

Methodology and Coding Framework

The interviews were reviewed through the lens of the evolving GrundMind framework using a hybrid deductive–inductive coding approach. Deductive coding tested whether each of the six originally proposed GrundMind dimensions emerged in participant narratives. Inductive coding ran in parallel to identify recurring themes that did not fit cleanly into any existing dimension — these were retained as candidate latent constructs.

Coding procedure

Each transcript was read in full. Behavioral indicators were tagged at the utterance level rather than the interview level, because the same individual frequently exhibited very different interaction patterns when discussing different tasks. Pattern saturation was tracked across the sub-corpora; new behavioral indicators stopped emerging well before the corpus was exhausted, which is consistent with established saturation thresholds for qualitative work of this kind.

Where the same construct surfaced under different language across sub-corpora — for example, scientists' "verification burden" and creatives' "editing burden" — the underlying behavioral signature was treated as a single construct and noted as having domain-specific surface forms.

The six original GrundMind dimensions

The deductive frame was structured around the following constructs as currently defined within the GrundMind framework. Each is restated here in its operating definition with the behavioral indicators used during coding.

1. Ambiguity Tolerance

Definition: ability to operate effectively with probabilistic, incomplete, or iterative AI outputs.

Indicators coded: comfort with imperfect drafts; willingness to iterate; exploratory prompting; frustration with uncertainty; preference for deterministic outputs.

2. AI Trust Calibration

Definition: how users regulate trust in AI outputs across contexts.

Indicators coded: verification behavior; confidence thresholds; skepticism patterns; conditional trust; evidence-seeking behavior; hallucination concern.

3. Control Preference

Definition: preference for oversight, structure, and human direction during AI interaction.

Indicators coded: desire for approval steps; discomfort with autonomous generation; structured prompting behavior; workflow supervision; resistance to uncontrolled outputs.

4. Cognitive Engagement and Exploration

Definition: extent to which users employ AI for ideation, synthesis, exploration, or conceptual expansion.

Indicators coded: brainstorming; co-thinking; idea iteration; exploratory usage; curiosity-driven interaction; conversational refinement.

5. AI Reliance and Independence

Definition: relationship between AI usage and perceived personal capability.

Indicators coded: dependency concern; skill erosion fear; overreliance anxiety; validation-seeking; confidence substitution.

6. Workflow Integration and Friction

Definition: ease or difficulty integrating AI into real operational workflows.

Indicators coded: abandonment behavior; workflow acceleration; editing burden; integration fatigue; switching back to manual methods; iteration overhead.

Additional Constructs Identified

The inductive coding pass surfaced four dimensions that appeared repeatedly across the corpus and that are not adequately captured by the six original dimensions. Each is presented below with its definition, the behavioral patterns observed, illustrative quotations from the corpus, and its strategic relevance to the diagnostic. These four constructs are the principal contribution of this validation exercise to the framework's next iteration.

7. Verification Burden

Definition

The cognitive cost of validating, inspecting, correcting, or justifying AI-generated output. Distinct from trust: a user may believe the AI is generally reliable yet still abandon it because the cost of confirming any single output exceeds the cost of producing it themselves.

Why it matters

This construct emerged as one of the strongest and most prevalent themes in the corpus — particularly among scientists, but also among writers, designers, and operations professionals. Many participants framed their decision to use or abandon AI not in terms of capability or trust, but in terms of whether verification was fast enough to leave a net time saving. The pattern was strikingly consistent: AI was retained where verification was cheap and abandoned where verification approached or exceeded the original task cost.

Representative voices

> *"It's a lot more work to double-check the AI than it is to find a section in a familiar reference text for the same information."* — Scientist participant
>
> *"The problem with using AI here is that it's not trustworthy for the advanced math or conceptual understanding of the experimental context, so if I used it, I would have to double-check everything it did at which point I may as well have done it myself to begin with. This doesn't save time."* — Scientist participant
>
> *"For personal work, I know very easily if there has been a mistake through the AI process. With literature review, as I'm not always familiar, I may not notice any errors, so I would have to fact check anyway which means I might as well do the majority of it myself to start with."* — Scientist participant
>
> *"To work around the truthful issue, I double check anything I'm doubtful about. It's time consuming, but not as time consuming as doing the research myself."* — Creative-professional participant (novelist)

Observed behaviors

- Spot-checking samples rather than reviewing entire outputs

- Selective delegation calibrated to verification cost rather than capability
- Refusing AI in high-accountability contexts where verification must be exhaustive
- Avoiding AI for tasks requiring downstream explanation
- Using AI only for low-risk drafting where errors are quickly visible
- Reverting to manual workflow when verification overhead grows over the course of a project

Strategic importance

Verification Burden appears to be one of the strongest predictors of sustained, as opposed to episodic, AI usage. It is also the most actionable from an organisational design perspective: workflow choices, output formats, and tool selection can all materially change the verification cost of a given AI interaction. We recommend it be added to the diagnostic as a first-class construct in the next iteration.

8. Accountability Sensitivity

Definition

The degree to which responsibility, explainability, and personal accountability shape AI usage behavior. Captures the gap participants reported between trusting an AI's output and being willing to defend reasoning that originated with the AI.

Observed pattern

Participants frequently described avoiding AI in situations where they had to defend their reasoning to stakeholders, where legal or compliance exposure existed, where reputational risk was high, where scientific validity was at stake, or where blinding and confidentiality requirements applied. Notably, several participants reported trusting AI outputs in private use but withholding that use in accountability-bearing contexts — a behavior that would be invisible in any usage-log-based adoption metric.

Representative voices

> *"For statistical analysis and coding there is a requirement for two individuals to independently generate the work then check if they agree. I know one person uses AI to check their code if they run into an error they aren't sure of, but they are required to supply a human-generated output. Trial design will probably always be human driven because of the implied risks in clinical trials."* — Clinical-research participant

> *"AI usually gives the correct answer without alternative thought processes. This kills creativity and also makes it difficult to defend decisions made about research topics."* — Engineering-research participant

> *"My job as a researcher is to be able to draw meaningful conclusions from my datasets. If I rely too heavily on someone else's thoughts on my data, I run the risk of failing to properly understand it myself."* — Scientist participant

Key insight

Users often trusted AI outputs more than they trusted their ability to defend AI-generated reasoning to a third party. This distinction — between epistemic trust and accountable trust — is critical, and it is invisible in standard adoption metrics. It also has direct implications for the design of AI workflows in regulated industries: a tool that is highly used in personal preparation may still be entirely absent from the official decision record.

Strategic importance

Accountability Sensitivity is essential for predicting adoption patterns in regulated industries (clinical research, finance, legal, compliance), in stakeholder-facing roles (consulting, sales, advisory), and in any environment where a decision trail must be auditable. We recommend it be added to the diagnostic as a primary moderator of every other dimension.

9. Authentic Voice Protection

Definition

The desire to preserve personal, creative, emotional, or professional identity during AI-assisted work. Encompasses voice, style, judgment, expertise signature, and authorship — the elements of a participant's work that they perceive as distinctively theirs and that AI, in their view, threatens to dilute or replace.

Observed pattern

This construct appeared most strongly in the creative sub-corpus, but was also clearly present among educators, client-facing professionals, marketers, and technical writers. The pattern was not blanket resistance to AI; rather, participants actively used AI for ideation, structure, brainstorming, and editing while drawing a sharp boundary around the elements they considered constitutive of their voice. The boundary was almost always defended in identity terms rather than quality terms.

Representative voices

> *"Writing the rhythms and melodies by hand is the crucial human element. I'm much more interested in it helping with technical tasks."* — Music producer
>
> *"I definitely feel like AI is guiding the direction, which is the part I feel uneasy about. The AI feels like it is concerned with what will make successful, but not necessarily with maintaining my voice."* — Content creator
>
> *"I do try to limit my use of AI, so it's my voice and writing style and not someone else's. Sometimes I use AI while editing, to make sure what I've written makes sense and flows in a coherent manner."* — Writer
>
> *"I don't allow myself to use AI for the creative meat and potatoes of my story, because I want it to be authentic to me. However, I do find myself looking for help for titles in my work, and to come up with names that sound good for siblings, last-name ideas, things like that which aren't as crucial to the creative part of my work."* — Novelist

> *"My work is my personality in a way and if I am posting something or coming up with a creative idea for work, I want to be able to say that it's mine and not just an AI prompt generation."* — Marketing professional
>
> *"When an AI writes, its writing style is something I can often pick up on. It may be written well, but I quickly say 'Did an AI write this?' I don't want our customers or my colleagues asking that of my work, and I don't want to pass things off as my own when they're not."* — Technical writer

Key insight

Many users framed AI as useful only when, in the words of one participant, "it still sounds like me." The acceptance/resistance line was drawn not at task complexity or AI capability, but at the point where AI involvement would alter what participants considered the irreducible signature of their work. Several participants also reported a meta-concern about homogenisation: even where they personally maintained voice control, they worried that widespread AI use would flatten the distinctive voices of an entire field.

Strategic importance

Authentic Voice Protection is a critical predictor of adoption ceiling in creative, advisory, and identity-rich professional contexts. It also implies that a tool's design — particularly its tendency to impose generic phrasing or aesthetic defaults — can directly reduce adoption among the highest-skill segment of these populations.

10. Execution vs. Ideation Preference

Definition

Whether a user predominantly engages AI as a deterministic execution engine (automation, formatting, structured task completion) or as a collaborative ideation partner (brainstorming, conversational thinking, iterative expansion). Distinct from Cognitive Engagement: the question is not how much the user thinks with AI, but in which mode.

Observed split

The corpus revealed a robust bimodal pattern. One segment of users — frequently in operations, customer-facing, and content-production roles — wanted AI to execute well-defined tasks deterministically and were frustrated by exploratory or open-ended responses. A second segment — frequently in strategy, research ideation, and creative roles — actively wanted dialogic exploration and were frustrated by closed-form deterministic output. The same individual could shift between the two modes across tasks, but a clear default tendency was usually identifiable from how they framed their first sentences about AI use.

Representative voices

> *"I'm using it mostly for brainstorming out of creative blocks, and for troubleshooting or techy workflow questions."* — Music producer (ideation default)

> *"Brainstorming and ideation is really just a process of bouncing my ideas off a peer. Sometimes it's talking through creative blocks and trying to clarify my vision."* — Composer (ideation default)
>
> *"Often I'll get AI to do the heavy lifting of writing the initial content, then I'll 'tidy up,' making edits to give a more human feel. I make sure the prompt I start with has as much information fed into it as possible, which means ultimately I'm driving the creative decisions and simply getting AI to do what I want it to."* — SEO content lead (execution default)
>
> *"I would use AI to assist in coding of the models. Usually this would be in the form of debugging code or writing lines of code that are simple but lengthy for a human to do."* — Physics researcher (execution default)

Strategic implication

Execution vs. Ideation Preference is likely to be one of the most commercially actionable segmentation dimensions in the diagnostic. It maps directly onto product-design choices: tools optimised for deterministic execution and tools optimised for exploratory ideation are not interchangeable, and mismatches in this dimension explain a meaningful share of the abandonment behavior we observed.

Consolidated Evidence Map

The table below consolidates the eleven dimensions surfaced by the analysis — six original GrundMind dimensions and five additional or refined constructs — alongside the behavioral signatures that recurred across the corpus and an indication of each dimension's relevance to the future diagnostic.

Dimension	Working definition	Observed behavioral signature	Diagnostic relevance
Trust Calibration	Conditional trust regulated through targeted oversight	Repeated fact-checking, validating, and rewriting of AI outputs	Critical for predicting sustainable AI usage
Ambiguity Tolerance	Comfort with probabilistic, non-deterministic outputs	Iterative prompting and exploratory workflows vs. frustration with unclear responses	Predicts adoption style and workflow design needs
Control Preference	Desire for oversight and structured workflows	Step-by-step direction of AI; rejection of autonomous generation	Important for governance and UX design
Cognitive Engagement	Use of AI for ideation, synthesis, and conceptual expansion	Brainstorming, creative iteration, co-thinking, conversational refinement	Indicates potential for advanced AI collaboration
Workflow Friction	Difficulty integrating AI into real operational workflows	Abandonment of tools when editing burden grows; reversion to manual methods	Strong indicator of adoption persistence
AI Reliance / Skill Preservation	Concern about long-term cognitive dependency	Self-imposed limits on AI use to preserve personal capability	Important long-term adoption factor
Verification Burden ★	Cognitive cost of validating AI output	Reverting to manual work when checking cost approaches task cost	Likely major predictor of adoption sustainability
Accountability Sensitivity ★	Concern about defending AI-assisted decisions	Avoidance of AI in high-stakes, regulated, or explainability-heavy work	Critical moderator in regulated industries
Authentic Voice Protection ★	Preserving human identity, style, and signature	Selective rewriting to preserve tone or personal authorship	Strongly relevant for creative and client-facing work

Dimension	Working definition	Observed behavioral signature	Diagnostic relevance
Execution Preference ★	AI as deterministic automation engine	Desire for structured, deterministic, well-bounded assistance	Strongly linked to operational and high-volume production roles
Ideation Preference ★	AI as collaborative thinking partner	Conversational exploration, scenario expansion, synthesis	Strongly linked to innovation-heavy and strategic functions

Table 2. Eleven recurring dimensions surfaced by the qualitative analysis. ★ indicates a construct newly proposed for inclusion in the GrundMind framework on the basis of this analysis.

Archetype Pattern Analysis

The transcripts provided strong evidence for recurring interaction styles consistent with the proposed GrundMind archetypes. The analysis also yielded a critical refinement, however: archetypes are contextual, role-dependent, task-dependent, and dynamic. The same individual frequently shifted across archetypes depending on risk level, accountability exposure, domain familiarity, workflow complexity, and creative ownership.

Archetypes should therefore not be treated as fixed personality categories. They function more accurately as recurring interaction tendencies under specific workflow conditions. This has direct implications for the diagnostic: rather than assigning a single archetype to a respondent, the future instrument should capture the modal tendency together with the conditions under which the individual is most likely to switch modes.

Co-thinkers

Characteristics

- Use AI conversationally
- Explore possibilities and brainstorm frequently
- Iterate collaboratively
- Comfortable with ambiguity
- High exploratory engagement

Common contexts

Product, marketing, creative fields, entrepreneurship, strategy, research ideation.

Typical language observed

> *"Brainstorming and ideation is really just a process of bouncing my ideas off a peer. Sometimes it's talking through creative blocks and trying to clarify my vision."* — Composer
>
> *"It was fun but also it helped me refine my own perspective somewhat. I was just sending through articles from the composer and suggesting other connections that might be considered."* — Composer / film-score analyst

Risks observed

- Over-exploration and cognitive overload
- Excessive iteration with diminishing returns
- Dependency on ideation loops as a substitute for decision-making

Calibrators

Characteristics

- Heavy verification and evidence-seeking
- Controlled, conditional trust
- Selective delegation calibrated to verification cost
- Skepticism paired with genuine openness

Common contexts

Scientific work, finance, legal, operations, analytical functions.

Typical language observed

> *"I verify what it produces by running the code to see if it produces an error or doesn't. If it doesn't produce an error, I go through the new code, line-by-line, to check if it is logical and if it is actually doing the thing I had originally intended."* — Physics researcher
>
> *"I do sampling and verify most of the data doing a quick search, but still it is easier than filling the complete spreadsheet by myself."* — Computational researcher

Risks observed

- Verification fatigue over the course of long projects
- Workflow slowdown when verification cost grows faster than task savings
- Abandonment when checking burden crosses an internal threshold

Controllers

Characteristics

- Prefer structured, scripted interaction
- Strong oversight needs
- Deterministic workflow preference
- Discomfort with unpredictability
- Explicit governance orientation

Common contexts

Compliance, operations, regulated environments, leadership accountability roles.

Typical language observed

> *"Often I'll get AI to do the heavy lifting of writing the initial content, then I'll 'tidy up.' I make sure the prompt I start with has as much information fed into it as possible, which means ultimately I'm driving the creative decisions and simply getting AI to do what I want it to."* — SEO content lead

Risks observed

- Low adoption flexibility when AI behavior is genuinely probabilistic
- Frustration with exploratory or open-ended AI systems
- Rigid interaction patterns that cap upside

AI-averse / Cautious Users

Characteristics

- Minimal delegation
- Strong skepticism toward probabilistic output
- Low workflow integration
- Concern about authenticity, replacement, or skill erosion
- High disruption sensitivity

Common contexts

Low-trust environments, high-accountability roles, legacy workflows, low AI familiarity, identity-rich creative work.

Typical language observed

> *"I'll continue to use it when I feel the urge. It's convenient. But I won't let it replace my own creativity — that would be very unsatisfying."* — Visual artist
>
> *"I don't allow myself to use AI for the creative meat and potatoes of my story, because I want it to be authentic to me."* — Novelist

Risks observed

- Outright abandonment after initial trial
- Disengagement and shadow resistance
- Delayed adaptation in fast-moving competitive contexts

Cross-Domain Findings

Comparing the three sub-corpora yielded a robust set of domain-specific patterns. The same dimensions appeared across all three, but their relative weight and the conditions under which they activated differed in important ways.

Creative professionals

In the creative sub-corpus, Authentic Voice Protection and Cognitive Engagement were the dominant dimensions. Participants showed a clear willingness to use AI for ideation, drafting, organisation, and editing, while drawing a hard boundary around what they considered the core artistic identity, emotional expression, and final creative direction.

A consistent secondary theme was concern about field-level homogenisation — even where individuals maintained voice control, they expressed concern about the cumulative effect of widespread AI use on the diversity of voices in their domain.

> *"My concern is how our creativity would be diminished and everyone would have similar ideas based on AI models. It is such a shortcut when you can have ten captions generated for you in a second that is 80% better than your own written one — so what's the point of writing it ourselves?"* — Marketing professional
>
> *"The homogenization of writing style was mentioned earlier. Everyone wants to express themselves, and AI does make those small suggestions that add up, until everything that's written sounds the same. This might make it hard for someone who is trying to use their voice to express themselves to be able to find their audience."* — Writer

Most common patterns observed: authentic voice protection; ideation usage; emotional ownership; selective delegation; fear of homogenisation. AI was widely accepted for brainstorming, drafting, organisation, and editing; resisted for core artistic identity, emotional expression, and final creative direction.

Scientists

In the scientist sub-corpus, Verification Burden, Accountability Sensitivity, and Trust Calibration dominated. Participants used AI extensively for acceleration, summarisation, coding assistance, and literature support, but every one of those uses was bounded by reproducibility requirements, explainability demands, accuracy thresholds, and scientific defensibility. Hallucination concern was mentioned in nearly every transcript and was the single most common reason participants gave for restricting AI use.

> *"It's not reliable enough yet to use for more complex problems like tailored experimental design or mathematically advanced data analysis."* — Research scientist
>
> *"AI can occasionally hallucinate analysis to give the user what the model thinks it wants."* — Materials science researcher

> *"Plenty of hallucinations regarding the reference material used by the AI to generate a response — and in truth, it's a lot more work to double-check the AI than it is to find a section in a familiar reference text for the same information."* — Scientist participant

Most common patterns observed: verification intensity; accountability sensitivity; evidence-seeking; high calibration behavior; concern around correctness and defensibility. AI was used for acceleration, summarisation, coding, and literature support; heavily constrained by reproducibility, explainability, accuracy, and scientific defensibility.

Workforce / general professionals

In the broader workforce sub-corpus, Workflow Integration and Execution Preference were the most active dimensions. Adoption decisions were dominated by perceived workflow fit, verification cost, and integration friction rather than by abstract attitudes toward AI.

> *"I decided to use the AI as I find it much faster than my usual approach and the feedback I get on my outputted work has been positive from my colleagues."* — Business professional
>
> *"AI is much faster for things like debugging than manually combing through a large codebase."* — Software engineer

Most common patterns observed: workflow optimization; execution preference; productivity orientation; selective delegation; time-saving behavior. Adoption was strongly influenced by workflow fit, verification cost, integration friction, and perceived usefulness — and only weakly by stated attitudes.

Major Findings

1. Delegation boundaries matter more than "AI attitude"

The strongest recurring pattern in the corpus was not simple enthusiasm or resistance. It was the careful articulation, by virtually every participant, of what they would and would not delegate. Participants consistently described which tasks they handed to AI, which they refused to delegate, where they maintained control, and the conditions under which their delegation boundary moved.

This finding has direct implications for the diagnostic. Future instruments should evaluate the location and movability of an individual's delegation boundary, rather than measuring generalised AI sentiment. Sentiment is a poor predictor of behavior; the delegation boundary is a much stronger one.

2. Verification Burden appears central

The qualitative evidence strongly suggests that sustained AI adoption depends heavily on whether outputs can be validated efficiently. Many participants framed their adoption decision in explicit time-cost terms: AI was retained where verification was cheap and abandoned where verification approached or exceeded the original task cost.

Verification Burden is likely to be one of the most predictive future dimensions in the framework. It is also one of the most actionable from an organisational design standpoint — workflow choices, output formats, and tool selection can all materially change verification cost without changing the underlying model.

3. AI adoption is highly contextual

Participants behaved differently depending on risk, accountability, task type, domain expertise, creative ownership, and stakeholder exposure. The same individual frequently exhibited very different interaction patterns within the same week, even within the same project.

This finding suggests that static archetype assignment alone may be insufficient as a diagnostic output. A stronger future model is likely to be multiplicative rather than additive — interaction tendency expressed as a function of the conditions under which the individual is operating. We summarise this provisionally as:

$$\textit{Cognitive Style} \times \textit{Task Risk} \times \textit{Verification Burden} \times \textit{Workflow Context}$$

That is, a respondent's interaction tendency should be reported together with the conditions under which it activates and the conditions under which it shifts.

4. Workflow fit is more important than tool access

Many participants already had access to advanced AI tools. The determining factor in their actual usage was not whether they had access, but whether AI reduced friction, increased friction, accelerated work, created verification overhead, or fit naturally into existing workflows. This strongly supports the central GrundMind thesis: adoption is a cognitive-fit problem, not a tooling problem.

Several scientist participants captured this pattern with particular clarity, describing how a tool they had previously praised was abandoned mid-project once verification overhead grew faster than the productivity it delivered. The decision to abandon was rarely framed as a loss of trust in the model; it was framed as a calculation about where the verification cost crossed an acceptable threshold.

5. Human identity preservation is critical

Many participants accepted substantial AI support while simultaneously protecting authorship, professional judgment, personal style, emotional authenticity, and expertise identity. The pattern was not blanket resistance but careful boundary-drawing around the elements they considered constitutive of their selfhood at work.

This suggests that AI adoption is, in part, an identity-negotiation process. The framework should treat identity as a first-class variable rather than a soft cultural factor, particularly in creative, advisory, expert, and client-facing contexts.

6. The same person, multiple archetypes

A finding that emerged inductively and that was not anticipated by the original archetype taxonomy: the same individual frequently behaves as a different archetype across different tasks within the same role. A clinical researcher who is a Calibrator for trial design may be a Co-thinker for grant writing and a Controller for compliance documentation. This argues strongly against treating archetype assignment as a personality classification, and in favour of treating it as a context-conditional tendency.

Implications for the GrundMind Diagnostic

Together, the qualitative findings have specific consequences for the design of the GrundMind diagnostic instrument. The recommendations below are intended to guide the next phase of construct definition, item generation, and pilot psychometric testing.

Recommendation 1 — Add four constructs to the framework

Verification Burden, Accountability Sensitivity, Authentic Voice Protection, and Execution-vs-Ideation Preference should be added as first-class constructs in the next iteration of the framework. Each is grounded in robust, repeated qualitative evidence and each has clear behavioral indicators that can be operationalised as items.

Recommendation 2 — Treat archetypes as context-conditional

The diagnostic should report a respondent's modal interaction tendency together with the conditions under which it activates and the conditions under which it shifts. A single archetype label, decoupled from context, is unlikely to be predictively useful and may actively mislead organisations relying on the output.

Recommendation 3 — Measure delegation boundaries directly

Items should ask respondents what they would and would not delegate under specified conditions, rather than asking how positively or negatively they feel about AI in the abstract. Boundary-elicitation items are more behaviorally diagnostic and less susceptible to social-desirability bias than attitude items.

Recommendation 4 — Build in a moderator structure

Accountability Sensitivity and Verification Burden function as moderators of every other dimension. The instrument should be designed so that these dimensions can be analysed both as direct predictors and as moderators in interaction terms. This will require a sample size and item structure that supports interaction-effect estimation, not only main effects.

Recommendation 5 — Capture the verification economy explicitly

Items should capture both the absolute level of verification burden a respondent perceives in their work and the threshold at which that burden causes them to abandon AI. This second quantity — the abandonment threshold — appears to be a particularly strong predictor of sustained adoption and is largely missing from existing AI-attitude instruments.

Limitations

This analysis has several limitations that the reader should keep in mind, and that we have made no attempt to obscure.

Self-report and recall

All evidence is participant self-report, with the inherent risks of recall bias, retrospective rationalisation, and social-desirability effects. Participants describing their AI use months after the fact may have systematically tidied their accounts. Behavioral observation studies will be required at later stages of validation to triangulate these self-reports.

AI-mediated interview format

The interviews were conducted by an AI interviewer (Claude). This format has methodological advantages — consistency, scale, neutrality — but also possible response shifts. Participants may be more candid with an AI interviewer about AI scepticism, or alternatively may moderate their criticism. We have no way of estimating this effect from the dataset alone, and it should be examined explicitly in subsequent waves of data collection.

No quantitative claims

This is a construct-mapping exercise, not a prevalence study. We make no claims about how common any given pattern is in the broader population, nor about effect sizes, predictive validity, or test–retest reliability. All such claims must wait for the quantitative phase of validation.

Domain coverage

The corpus is well-distributed across workforce, scientific, and creative domains, but specific subgroups — clinical practitioners, financial-services compliance staff, regulated-industry executives — are underrepresented. The accountability-heavy patterns we observed are likely to be even more pronounced in these populations and may surface dimensions not yet visible in the present sample.

Single-language corpus

The dataset is English-language only. Cross-cultural and cross-linguistic replication will be needed before any of these constructs can be claimed to generalise globally. Several of the dimensions surfaced here — particularly Authentic Voice Protection and Accountability Sensitivity — are plausibly culturally moderated in important ways.

Next Steps

This validation report concludes the construct-mapping phase of the GrundMind validation programme. The phases that follow are summarised below in order of dependency.

Phase	Objective	Primary outputs
Phase 1 — Construct refinement (next 1–2 quarters)	Finalise the eleven-construct framework on the basis of this report; resolve overlap between Verification Burden and Trust Calibration; specify moderator relationships	Revised GrundMind framework specification; construct definitions document
Phase 2 — Item generation and expert review	Draft item pools for each construct; subject-matter expert review; pilot cognitive interviews to refine item wording	Pilot item bank; expert review report; cognitive interview transcripts
Phase 3 — Pilot psychometric study (n ≈ 300)	Initial reliability and exploratory factor analysis; identify problematic items; preliminary subscale structure	Pilot validation report; revised item bank
Phase 4 — Confirmatory study (n ≈ 1,000)	Confirmatory factor analysis; convergent and discriminant validity; test–retest reliability; demographic invariance testing	Validated GrundMind diagnostic v1.0
Phase 5 — Behavioral triangulation	Link diagnostic scores to observed AI usage patterns and outcomes in real organisational settings	Predictive-validity evidence; case studies

Table 3. Validation roadmap from qualitative grounding to behavioral triangulation.

Conclusion

This early qualitative validation of the GrundMind framework, conducted against 1,250 unprompted real-world AI-use narratives, produced three substantive results.

First, the framework's central hypothesis — that AI adoption quality is a function of cognitive fit rather than access, attitude, or skill — is strongly supported by the qualitative evidence. Participants across three very different domains converged on this account in their own words, without being prompted to do so.

Second, the framework as currently specified is incomplete. Four additional constructs — Verification Burden, Accountability Sensitivity, Authentic Voice Protection, and Execution-vs-Ideation Preference — appeared with sufficient frequency, behavioral specificity, and cross-domain consistency to warrant inclusion as first-class dimensions in the next iteration of the framework.

Third, the proposed archetype taxonomy is recognisable in the data but should be reformulated. Rather than treating archetypes as fixed personality categories, the diagnostic should treat them as context-conditional interaction tendencies, reporting a respondent's modal mode together with the conditions under which it activates and shifts.

These results justify advancing the framework to the next phase of validation. They also clarify, importantly, what should not be done yet: any attempt to ship a quantitative GrundMind diagnostic before construct definitions are revised in light of this evidence would lock in a measurement architecture that the data already tells us is incomplete. The discipline of qualitative grounding before quantitative measurement is not an academic nicety — it is the difference between an instrument that is psychometrically clean and one that is substantively useful.

The Grund Institute's broader thesis — that organisations and individuals capture value from AI in proportion to the alignment between AI's mode of operation and the cognitive habits of the humans working with it — finds substantial qualitative support in this analysis. The next phase of work will test that thesis quantitatively.

Selected References

This validation report builds on a body of work previously published by the Grund Institute and on the broader cognitive-science, human–AI interaction, and psychometric literatures. The references below are a selected subset focused on the foundations most directly engaged in the present analysis.

- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in Public Health, 6, 149.
- BCG (2025). The AI Adoption Puzzle: Why Usage Is Up but Impact Is Not.
- Clark, A., & Chalmers, D. (1998). The Extended Mind. Analysis, 58(1), 7–19.
- Grund Institute (2026). Bridging the Cognitive Gap in AI Adoption: Second Mind Framework and Cognitive Fit. White Paper.
- Grund Institute (2026). Evaluating Human–AI Interaction Styles: Dimensions, Measurement, and Framework. White Paper.
- MIT NANDA (2025). Why 95% of AI Implementations Fail. Working Paper.
- Samuel, B. M., Khatri, V., & Varghese, A. (2022). Adaptive Cognitive Fit: AI-Augmented Management of Information. International Journal of Information Management, 65.
- Anthropic (2025). Anthropic Interviewer Dataset (public release): Workforce, Scientists, and Creative Professionals sub-corpora.