

# Expressive Cognition — Scoring Rubric

## Core Dimensions (included in VRI composite)

---

### 1. ABSTRACTION

*What level does the speaker naturally think at?*

Band	Label	Score	Behavioral Signal
1–2	<b>Concrete</b>	1–2	Stays at the level of the specific case. Describes rather than categorizes. No movement toward generalization.
3–4	<b>Classifying</b>	3–4	Names the category the example belongs to. Multiple instances listed without integration. Generalization attempted but not developed.
5–6	<b>Relational</b>	5–6	Derives a principle from the case. Integrates elements into a framework. Identifies what would change the outcome.
7–9	<b>Principled</b>	7–9	Generates a framework that extends beyond the prompt. Uses abstraction to identify edge cases, predict, or explain what the question itself assumes.

**What it is NOT:** Using abstract-sounding words. "Fundamentally" is not abstraction. Deriving a principle from an example is.

**Key evidence:** Shifts from "X happened" → "X is an instance of Y" → "Y predicts Z in different contexts."

**Grounding:** SOLO Taxonomy (Biggs & Collis, 1982); Vygotsky's concept formation levels

---

### 2. COMPRESSION

*How much cognitive content per unit of language?*

Band	Label	Score	Behavioral Signal
1–2	<b>Expansive</b>	1–2	Heavy restatement. Ideas unpacked across many words. Significant redundancy. Little propositional gain per utterance.
3–4	<b>Sequential</b>	3–4	Ideas presented one at a time without integration. Low redundancy but low density. Each sentence does one thing.
5–6	<b>Dense</b>	5–6	Multiple ideas per utterance. Some integration. Clauses carry real propositional weight.
7–9	<b>Packed</b>	7–9	Ideas arrive already integrated. Single phrases carry complex propositional content. No unpacking required. The speaker is ahead of the sentence.

**What it is NOT:** Being brief. A highly compressed response can be long. It's the ratio of propositional content to words, not word count.

**Key evidence:** Count distinct propositions per sentence. Compare word count to idea count. Look for strategic ellipsis and implicature.

**Grounding:** Propositional density analysis (Kintsch, 1974); Snowdon et al. (1996) idea density as cognitive predictor

---

### 3. ORIGINALITY

*Does anything genuinely unexpected arrive?*

Band	Label	Score	Behavioral Signal
------	-------	-------	-------------------

1–2	<b>Conventional</b>	1–2	The most expected response to the prompt. Culturally available frames, standard examples. Nothing unexpected arrives.
3–4	<b>Inflected</b>	3–4	A minor variation on the expected response. Some personal angle but within conventional frames.
5–6	<b>Reframed</b>	5–6	The prompt is approached from a non-obvious angle. An unexpected example or analogy appears. The frame is surprising but apt.
7–9	<b>Generative</b>	7–9	Something genuinely new arrives in the course of speaking — a frame, connection, or reversal the speaker appears to discover mid-response. The prompt becomes a different question than the one asked.

**What it is NOT:** Being unconventional for its own sake. Obscurity is not originality. The test is whether the unexpected framing actually illuminates something.

**Key evidence:** Would most people have said this? If no — is it apt? Does the unexpected framing illuminate or just surprise?

**Grounding:** Guilford's divergent production; Finke et al. (1992) geneplore model; Simonton's combinatorial theory of creativity

#### 4. CONCEPTUAL CONTINUITY

*Do ideas build on each other?*

Band	Label	Score	Behavioral Signal
1–2	<b>Fragmented</b>	1–2	Ideas do not connect. Topic shifts without transition. Repetition without development. Response ends where it started.
3–4	<b>Listed</b>	3–4	Ideas are related to the topic but do not build on each other. Additive rather than generative. A list, not a development.
5–6	<b>Connected</b>	5–6	Ideas follow from each other. A thread is maintained. The response develops, though not always to a resolved conclusion.
7–9	<b>Cumulative</b>	7–9	Each move builds on the last. The response arrives somewhere the opening didn't anticipate. Thought visibly accumulates across the response.

**What it is NOT:** Logical argumentation (trainable). The question is whether ideas connect — whether there is a thread — not whether the argument is formally valid.

**Key evidence:** Does the ending build on the beginning? Does each utterance advance the previous one? Or could the sentences be reordered without loss?

**Grounding:** Cohesion theory (Halliday & Hasan, 1976); Coh-Metrix (McNamara et al., 2014); situation model (van Dijk & Kintsch, 1983)

#### 5. EPISTEMIC CALIBRATION

*Does the speaker know what they know?*

Band	Label	Score	Behavioral Signal
1–2	<b>Undifferentiated</b>	1–2	All claims delivered with uniform confidence. No distinction between what is known and what is inferred. No spontaneous qualification.
3–4	<b>Occasionally Hedged</b>	3–4	Some hedging present but inconsistent. The speaker sometimes marks uncertainty but does not systematically distinguish claim types.

5–6	<b>Differentiated</b>	5–6	Consistent distinction between confident claims and uncertain inferences. Spontaneous qualification present. Evidence and assertion kept separate.
7–9	<b>Reflexively Calibrated</b>	7–9	The speaker monitors not just individual claims but the quality of their own reasoning. Acknowledges what the argument assumes. Notes where an alternative interpretation is possible.

**What it is NOT:** Excessive hedging, which can be trained professional caution. The signal is *differentiated* calibration — confident about some things, uncertain about others, and marking the difference.

**Key evidence:** Does the speaker treat all claims the same, or do they mark some as more certain than others? Is this spontaneous or only in response to challenge?

**Grounding:** Reflective Judgment Model (King & Kitchener, 1994); Kuhn's metacognitive argumentation theory (1991)

## 6. GENERATIVE SELF-MONITORING

*Does the speaker improve in real time?*

Band	Label	Score	Behavioral Signal
1–2	<b>Unreflective</b>	1–2	No revision or self-correction. Claims delivered without monitoring. Errors or imprecisions pass without notice.
3–4	<b>Surface Repair</b>	3–4	Some correction present but at the surface level — word substitution, false starts. Not conceptual revision.
5–6	<b>Conceptual Revision</b>	5–6	The speaker catches a conceptual imprecision and reformulates. Revisions improve accuracy or precision, not just fluency.
7–9	<b>Generative Refinement</b>	7–9	Revision moves the idea forward. The speaker arrives at something better than what they started with — a more accurate frame, a more precise claim, a higher-level formulation — through the process of monitoring their own output.

**What it is NOT:** Disfluency. Repeated "um" and "uh" is not self-monitoring — it's processing load. The signal is targeted revision: a speaker who says something and then says something better.

**Key evidence:** Does the speaker ever say something, then say something better? Does the revision go UP (more precise, more accurate, higher-level) or just SIDEWAYS (word swap)?

**Grounding:** Self-monitoring model (Levelt, 1989); Protocol analysis (Ericsson & Simon, 1993); Metacognition theory (Flavell, 1979)

## Moderator Dimensions (scored but NOT in VRI composite)

*These measure linguistic competence, not reasoning. They are reported alongside the VRI as contextual information.*

## 7. VOCABULARY

Band	Label	Score	Behavioral Signal
1–2	<b>Basic</b>	1–2	High-frequency vocabulary only. Frequent repetition. Overextends general words where precise alternatives exist.
3–4	<b>Functional</b>	3–4	Mid-frequency vocabulary used appropriately. Some domain terms appear but not always precisely.
5–6	<b>Precise</b>	5–6	Word choice sharpens meaning. Academic and domain-specific terms used accurately.

7–9	<b>Surgical</b>	7–9	Lexical choices both precise and strategic. Distinguishes near-synonyms by connotation and register.
-----	-----------------	-----	--

## 8. SYNTACTIC CONTROL

Band	Label	Score	Behavioral Signal
1–2	<b>Simple</b>	1–2	Short, simple sentences. Limited subordination. Frequent grammatical errors.
3–4	<b>Basic Subordination</b>	3–4	Some subordinate clauses but complex embedding rare or error-prone.
5–6	<b>Controlled Complexity</b>	5–6	Mixes simple and complex structures effectively. Maintains control under cognitive demands.
7–9	<b>Syntactic Fluency</b>	7–9	Complex structures deployed naturally. Structure serves rhetorical and reasoning purposes.

## Scoring Instructions

1. **Read the full transcript** before assigning any scores.
2. **Score sustained performance** — the highest band *consistently* demonstrated, not peak moments. A speaker who reaches 7–9 once but operates mostly at 5–6 should score 6, not 8.
3. **Use the full range within each band.** Within 7–9: 7 = solidly meets descriptors, 8 = strong, 9 = exceptional even among experts.
4. **Calibration:** 5 = average adult performance. Most speakers on most dimensions score 4–7. A score of 8 or 9 should be rare and clearly justified.
5. **One evidence quote per dimension** — the quote that most clearly illustrates why this band and not one band lower or higher.

## VRI Composite Weighting

Dimension	VRI Weight
Abstraction	0.18
Epistemic Calibration	0.18
Compression	0.16
Originality	0.16
Conceptual Continuity	0.16
Generative Self-Monitoring	0.16

VRI = weighted average mapped to a 100-centered scale ( $SD \approx 15$ ). Score of 5.0 on all dimensions = VRI 100.

## The Composite

The six core dimensions measure aspects of the same underlying thing: **what happens cognitively when a person thinks aloud under novel conditions.**

Dimension	The Question
Abstraction	What level do they think at?
Compression	How much can they hold?
Originality	Does anything new arrive?

Conceptual Continuity	Does thought cohere?
Epistemic Calibration	Do they know what they know?
Generative Self-Monitoring	Do they improve in real time?