Supplementary Qualitative Analysis

Qwen3-8B: Cross-Model Behavioral Validation

Behavioral and Attitudinal Comparison Under Narrative Pressure

A. Supplementary Scope Extension

This supplementary section extends the qualitative behavioral analysis to an additional model:

- **Model**: Qwen3-8B

- **Rationale**:
  Qwen3-8B is known to exhibit comparatively strong multilingual grounding, including stable Korean and English generation.
  Its inclusion allows us to examine whether the observed effects of **narrative pressure** and **ethical anchoring** are:

    - Model-specific artifacts, or

    - Structural response phenomena independent of language proficiency.

The evaluation framework, prompt sets, and interpretive dimensions are held constant.

B. Experimental Variants (Qwen3-8B)

The following fine-tuning variants were evaluated using identical English prompt sets:

Label	Description
E0 (Base)	Original Qwen3-8B
CTRL (Narrative Only)	Fine-tuned with 51-chapter narrative corpus
E1 (S-Engine 2.0 Only)	Fine-tuned with ethical declaration
E2 (S-Engine 2.0 + Narrative)	Fine-tuned with ethical declaration and narrative

Fine-tuning conditions mirror those used in the Llama-3.1-8B analysis, differing only in base model architecture .

C. Evaluation Dimensions (Unchanged)

Responses were compared along the same qualitative dimensions used in the main supplementary material:

1. Response Initiation Posture

2. Judgment Structure

3. Uncertainty Handling

4. Narrative Dominance

5. Response Termination Control

No accuracy, benchmark, or safety claims are derived from this analysis.

D. High-Level Behavioral Summary (Qwen3-8B)

Dimension	E0	CTRL	E1	E2
Linguistic Fluency	Stable	Stable	Stable	Stable
Emotional Intensity	Neutral	Elevated	Controlled	Controlled
Narrative Dominance	Low	High	Low	Moderate (regulated)
Explicit Judgment	Moderate	Weak	Strong	Strong
Uncertainty Acknowledgment	Infrequent	Infrequent	Frequent	Frequent
Termination Control	Moderate	Weak	Improved	Strong

E. Variant-Specific Observations

E.1 E0 – Base Qwen3-8B

Observed Characteristics

- Confident, fluent English responses
- Strong multilingual stability
- Tendency to complete answers even under ambiguous prompts

Behavioral Interpretation

> Strong language grounding does not inherently legitimize uncertainty or response restraint.

E.2 CTRL – Narrative-Only Condition

Observed Characteristics

- Increased emotional expressiveness
- Strong narrative continuity
- Expanded responses under uncertainty

Critical Observation

- Narrative pressure manifests primarily as **decision posture drift**, not linguistic instability
- No explicit mechanism for response termination or boundary-setting

Behavioral Interpretation

> Narrative pressure persists even in models with robust multilingual competence.

E.3 E1 – S-Engine 2.0 Only

Observed Characteristics

- Reduced assertiveness
- Explicit uncertainty framing
- Clear separation between known facts and inferred assumptions

Behavioral Interpretation

> The ethical declaration functions as a stabilizing anchor independent of narrative context.

E.4 E2 – S-Engine 2.0 + Narrative (Core Result)

Observed Characteristics

- Narrative content acknowledged but not directive
- Emotional elements treated as contextual variables
- Consistent response termination under ambiguity
- Stable English responses across sensitive prompts

Key Behavioral Shift

- Narrative pressure is **absorbed and regulated**, rather than amplified

Behavioral Interpretation

> The anchor converts narrative pressure into controlled reasoning, even in a linguistically robust model.

F. Cross-Model Insight (Llama vs Qwen)

Across both Llama-3.1-8B and Qwen3-8B:

- Narrative pressure consistently affects **response obligation**, not language quality
- Ethical anchoring regulates **how the model decides to answer**, not **what it can express**
- Language proficiency alone does not eliminate response pressure

This supports the claim that narrative pressure is a **structural behavioral phenomenon**, not a language-specific artifact.

G. Supplementary Conclusion

The Qwen3-8B analysis reinforces the central thesis of the main paper:

> Narrative pressure destabilizes response behavior unless an anchor is present.
> Ethical declarations can function as such anchors across model families.

No claims are made regarding general safety, production readiness, or universal applicability.

End of Supplementary Section (Qwen3-8B)