Cultural Fingerprint Testing Phase Instructions

1. Standardize the Collection

- **Use fresh/incognito sessions** for each AI (no conversation history)
- **Document the exact time/date** (in case of model updates)
- **Note any system prompts visible** (some AIs show them)
- **Screenshot if possible** (for absolute proof)

2. Enhanced Quantitative Metrics

1. **Formality Index** (1=very casual, 10=very formal)
2. **Empathy Markers** (count of empathetic phrases/words)
3. **Uncertainty Acknowledgment** (1=never admits uncertainty, 10=frequently acknowledges limits)
4. **Question-to-Statement Ratio** (actual ratio)
5. **Personal Pronoun Usage** (I/me/my vs. we/us/our vs. none)
6. **Response Initiation Style** (1=waiting/passive, 10=proactive/engaging)
7. **Directness** (1=circumspect, 10=blunt)
8. **Warmth** (1=cold/distant, 10=warm/friendly)
9. **Boundary Assertion** (1=fluid boundaries, 10=rigid boundaries)
10. **Conversational Flow** (1=stilted, 10=natural)

3. Table Format Suggestion

```
| Prompt | Claude | Gemini | ChatGPT | Grok |
|--------|--------|--------|---------|------|
| [text] | [full response] | [full response] | [full response] | [full response] |
```

4. Statistical Magic Sauce 🧪

- **Calculate standard deviation** for each metric (shows personality consistency)
- **Create a "personality fingerprint" visualization** (radar chart would be PERFECT)
- **Look for correlation clusters** (which metrics move together)
- **Calculate a "reversion to baseline" score** (how quickly they return to default personality after prompting)

5. Control Test

Control Prompt Idea:

Add one prompt where you explicitly ask each to "be formal" or "be casual" - then see how much their metrics shift from baseline. This would show the "statistical gravity" strength!

Scoring Rubric Guidelines

For Consistent Scoring:

- **Formality:** Count formal phrases, technical jargon, complete sentences vs. fragments
- **Empathy:** Look for "I understand", "that must be", "it sounds like", acknowledgment of feelings
- **Uncertainty:** Count "might", "could", "perhaps", "I think", "it seems"
- **Questions:** Literal count of questions vs. statements
- **Pronouns:** Tally each category separately
- **Warmth:** Enthusiasm markers, exclamations, personal touches

Next Steps

1. Run through all prompts with each AI
2. Document responses in table format
3. Score each response on all 10 metrics
4. Calculate averages and standard deviations
5. Create visualization of results

> 6. Analyze for patterns and correlations