Guidelines for Creating Developer-Focused Interpretive Briefs - v5.0

Preamble & Purpose

This document outlines the standardized process for the Narrative Interpretation Phase of the AI Lab's work. The primary goal is to synthesize technical findings from AI Lab "Synthesis Reports" and transform them into constructive, developer-focused narrative briefs that support continuous improvement of the KĀʻEO program.

The "Synthesis Report" is a technical, research-oriented document for internal analysis. The "Interpretive Brief" is fundamentally different: an explanatory narrative that translates technical findings into clear, actionable insights for item developers. The brief should be accessible to an intelligent reader without specialized psychometric training—think of writing for bright college students who are engaged and capable but not specialists in measurement theory.

Part 1: Process Workflow

Step 1: Input & Objective Definition

Input: A curated collection of multiple AI Lab Synthesis Reports, typically focused on items from a specific content area, grade, and/or SLO that have shown persistent challenges (e.g., two or more original items that didn't work as intended and their replacements also didn't work as intended when several other items performed well).

Objective: To generate a standalone "Interpretive Brief" that synthesizes patterns and findings from across the input reports into a single, cohesive narrative. This brief will focus on common areas for improvement and opportunities for refinement, rather than dwelling on specific challenges of any single item.

Step 2: Pattern Recognition & Meta-Synthesis

Primary Directive: Shift from Single-Item Analysis to Systemic Pattern Recognition

Look for common threads across the analyzed items. The brief should focus on answering questions like:

- "What common challenges emerge when writing DOK 3 items for this standard?"
- "Is there a recurring pattern in how distractors are structured that might be causing confusion?"
- "Do several items share similar computational or linguistic characteristics that obscure the intended construct?"

This approach elevates the conversation from individual item challenges to strategic, forward-looking insights.

Step 3: Translation & Drafting

Using the language guidelines in Part 2 below, transform technical findings into accessible narrative. Apply the document structure from Part 3 to organize the content.

Step 4: Quality Review

Before finalizing, ensure the brief:

- Can be understood by a non-specialist
- Maintains a constructive, collaborative tone
- Provides specific, actionable guidance
- Respects the expertise of item developers
- Accurately represents the technical findings without requiring technical knowledge to understand them

---

Part 2: Language Guidelines & Translation Principles

Core Principle: Clarity Through Translation, Not Omission

Technical findings should be translated into clear language, not hidden. When data shows an item didn't work as intended, explain what happened and why in accessible terms.

Technical Term Translations

Use these specific translations to replace psychometric jargon:

Technical Term	Plain Language Translation
"Item-rest correlation of [number]"	"Relationship between this item and overall test performance"
"Negative item-rest correlation"	"An unexpected pattern where students who generally performed well on the test were less likely to answer this particular question correctly"
"P-value of [number]"	"[X]% of students answered correctly"
"Marginal/poor discrimination"	"Didn't effectively distinguish between students at different performance levels"
"DIF flag"	"Performed differently for certain student groups relative to their overall ability"
"DOK mismatch"	"The actual thinking required exceeded what was intended"
"Construct-irrelevant variance (CIV)"	"challenges unrelated to the targeted skill that interfered with measurementx"

"Degrading reliability"	"Reduced the consistency of measurement"

Prohibited Language & Replacements

Avoid judgmental or harsh language:

Avoid	Use Instead
"Failed," "failure"	"Didn't perform as expected," "showed unexpected patterns"
"Broken," "fatal flaw"	"Presented challenges," "can be refined"
"Poor," "terrible," "problematic"	"Didn't work as intended," "presented opportunities for improvement"
"Severely," "extremely," "critically"	Remove intensifiers entirely or use "notably," "considerably"
"Compromised measurement"	"Measurement that didn't work as intended"
"Negative discrimination"	"Inverse pattern where higher-performing students selected incorrect answers more frequently"

Framing Principle: Problems → Opportunities

When describing issues, immediately pivot to the opportunity for improvement:

Instead of: "The item failed because of excessive computational burden."

Write: "The item required extensive calculations that obscured the core concept. Streamlining the computational steps would allow students to better demonstrate their understanding of the mathematical principle."

Instead of: "This distractor shows poor design with obvious incorrectness."

Write: "Refining this distractor to be more plausible would strengthen the item's ability to measure true understanding rather than simple elimination."

Tone Requirements

- **Use "we" and "our"** to emphasize collaboration and shared goals
- **Write as a collaborative colleague** sharing insights, not as an authority delivering verdicts
- **Be honest about findings** while maintaining a constructive, forward-looking perspective
- **Avoid hedging** that weakens clear observations (minimize "it's worth noting," "interestingly," "perhaps")
- **Respect developer expertise** by framing recommendations as strategies that tend to work well rather than mandates

When Discussing Technical Metrics

1. Always explain what the metric means in practical terms
2. Focus on implications rather than numbers
3. Use metrics to support insights, not as the primary message

Example: Instead of: "Item 47 had a point-biserial of -0.02" Write: "The performance data revealed an unexpected pattern: students who generally performed well on the test were less likely to answer this particular question correctly, suggesting the item may contain unintended complexity or ambiguity."

Part 3: Document Structure

Title Format

Use accessible, descriptive titles:

- "Developer Brief: Patterns in Grade 8 Math Items"
- "Insights from Recent Volume Calculation Items"
- "Narrative Notes on Assessing [SLO Code]"

Section 1: Introduction - Shared Goals

Begin with a collaborative framing that emphasizes continuous improvement and shared commitment to student success.

Example: "As part of our ongoing work to strengthen the KĀʻEO item bank, we've analyzed several Grade 8 Math items to identify patterns that can inform future development. This brief shares insights and strategies to help create items that are clear, fair, and effective while maintaining rigorous standards."

Key elements:

- Emphasize "ongoing work" and continuous improvement
- Frame as "insights" and "patterns" rather than problems
- Connect to shared values (clarity, fairness, effectiveness, rigor)

Section 2: Patterns and Insights

Present synthesized observations using clear subheadings and accessible explanations.

Subheading Strategies:

- Use questions: "What Makes Multi-Step Problems Challenging?"
- Use conceptual framings: "Managing Cognitive Load in Calculation-Heavy Items"
- Use forward-looking language: "Opportunities in Distractor Design"

Content Guidelines:

- Explain technical concepts through analogies or examples
- Connect observations to practical implications
- Focus on patterns across multiple items, not individual item challenges
- Use specific examples to illustrate patterns (item numbers, brief descriptions)

Example Structure: "Several items addressing volume calculations showed similar patterns. Students who demonstrated strong overall mathematical ability occasionally selected incorrect answers on these items more frequently than expected. Analysis suggests that the combination of multi-step calculations and complex visualization

requirements may have created challenges beyond the intended construct of volume understanding."

Section 3: Strategies and Models

Present proactive models as the centerpiece, showing successful approaches that address observed patterns.

For Each Model:

1. **Connect explicitly to a pattern from Section 2**
    - "To address the challenge of managing cognitive load in multi-step problems..."
2. **Explain why the model works in simple terms**
    - "This model uses a 'one-thing-well' approach. It tests a complex idea without asking students to juggle too many steps at once, allowing them to demonstrate their understanding of the core concept."
3. **Provide both English and Hawaiian versions**
    - Include appropriate notes: "Note: The Hawaiian language version was developed by [appropriate authority]. We present it with respect for the linguistic expertise of KĀʻEO's Hawaiian language specialists."
4. **Frame as strategies that tend to work well**
    - Avoid: "You must structure items this way"
    - Use: "This approach has shown promise in similar contexts" or "Developers might consider..."

Model Presentation Format:

Strategy: Focusing on Single-Concept Mastery

Why this works: [Brief explanation]

English Version: [Full item text]

Hawaiian Version: [Full item text with appropriate deference note]

Section 4: Key Takeaways

Conclude with 3-5 concise, actionable strategies for developers.

Format:

- Brief numbered or bulleted list
- Action-oriented language
- Specific enough to guide, general enough to allow professional judgment
- Positive framing

Example: "Key Strategies for Future Development:

1. Consider whether each item tests one clear concept or requires students to demonstrate multiple skills simultaneously
2. When creating answer options, ask: 'Could a student with solid understanding of the concept reasonably select more than one of these?'
3. Review calculation requirements to ensure they support rather than obscure the mathematical concept being assessed
4. Use the provided models as starting points while adapting to specific content needs"

Part 4: Implementation Notes

Focus on Patterns, Not Individual Items

Throughout the brief, maintain focus on systemic patterns rather than cataloging individual item challenges.

Instead of: "Item 47 failed because... Item 52 also failed because... Item 61 had problems with..."

Write: "Across several items testing proportional reasoning, a common pattern emerged where students with strong overall performance occasionally selected distractors that represented partially correct intermediate steps. This suggests an opportunity to..."

When Presenting Solutions

1. Frame as strategies that tend to work well rather than mandates
2. Explain the reasoning behind recommendations
3. Acknowledge that context matters and developers' expertise is valuable
4. Connect solutions to specific patterns identified in Section 2

Maintaining Accessibility Without Sacrificing Accuracy

The goal is to make technical findings understandable, not to oversimplify or hide important information.

Balance this by:

- Using analogies that illuminate rather than obscure
- Providing enough detail to support understanding without overwhelming
- Explaining the "why" behind technical observations
- Trusting the reader's intelligence while not assuming specialized knowledge

---

Part 5: Quality Checklist

Before finalizing any Interpretive Brief, verify:

Content:

- ☐ All technical jargon has been translated using Part 2 guidelines
- ☐ Prohibited language has been eliminated
- ☐ Focus remains on patterns across items, not individual item problems
- ☐ At least 2-3 proactive models are presented with clear rationale
- ☐ Hawaiian language versions include appropriate deference notes

Tone:

- ☐ Language is collaborative ("we," "our") rather than prescriptive
- ☐ Findings are presented honestly but constructively
- ☐ Developer expertise is respected throughout
- ☐ Forward-looking rather than dwelling on past challenges

Structure:

- ☐ Four-section structure is maintained
- ☐ Subheadings are clear and accessible
- ☐ Examples support general principles
- ☐ Key takeaways are actionable and specific

Accessibility:

- ☐ Can be understood by intelligent non-specialists
- ☐ Technical concepts are explained through clear examples
- ☐ Implications are more prominent than metrics

- ☐ Reads naturally, not like a technical report

Appendix: Example Transformations

Example 1: Negative Discrimination

Technical Finding (from Synthesis Report): "Item 47 exhibited negative item-rest correlation of -0.02, indicating that higher-achieving students were more likely to select incorrect responses."

Transformed for Brief: "Analysis of performance patterns revealed something unexpected: students who demonstrated strong overall mathematical ability were more likely to select incorrect answers on this particular item. This inverse pattern typically indicates unintended complexity or ambiguity that merits attention in future development."

Example 2: DIF Flag

Technical Finding: "Item 52 was flagged for DIF against non-FRL students with a standardized difference of 0.28."

Transformed for Brief: "Performance data suggests this item may inadvertently create different challenges for students from different economic backgrounds. Even when accounting for overall mathematical ability, students experienced this item differently. This pattern often emerges when items include context or language that assumes specific background knowledge."

Example 3: P-Value and Discrimination

Technical Finding: "Item 61 showed a p-value of 0.89 with discrimination of 0.08, suggesting the item was too easy and did not effectively discriminate between performance levels."

Transformed for Brief: "Nearly 90% of students answered this item correctly, and it didn't effectively distinguish between students at different performance levels. While high success rates can indicate clear item design, in this case the item may not have provided enough challenge to assess the intended depth of understanding."

Example 4: Reframing Problems as Opportunities

Technical Finding: "Multiple items failed due to construct-irrelevant variance introduced by excessive computational demands."

Transformed for Brief: "Several items required extensive multi-step calculations that may have obscured the core mathematical concept being assessed. Students who understand the principle might still struggle with the computational burden, making it difficult to determine whether incorrect answers reflect gaps in conceptual understanding or computational challenges. Streamlining the computational requirements while maintaining conceptual rigor would allow students to better demonstrate their mathematical reasoning."

-