Talking to a Bot: A Structured Evaluation of AI-Moderated Interviews

Prepared by **Lauren McCluskey**, Principal of Responsive Research, Inc. | **April 1, 2026**

A rigorous, multi-cohort evaluation of AI-moderated qualitative interviews across 166 sessions — examining participant experience, response depth, and analytical usability to define where AI moderation excels and where human judgment remains irreplaceable.

WHITEPAPER AI MODERATION QUALITATIVE RESEARCH

Executive Summary

AI moderation is rapidly expanding the scale and accessibility of qualitative research. Yet critical questions remain about what is gained — and what may be lost — when conversations are mediated by AI rather than a human moderator. This study was designed to examine those questions rigorously, across a structured sample of 167 total sessions, evaluating AI-moderated interviews from multiple vantage points simultaneously.

The evaluation included **n=101 panel participants** recruited through the AI platform's panel, **n=28 qualitative recruits** sourced through traditional human recruitment methods, and a **bonus cohort of n=37 qualitative researchers** drawn from an insights community — analyzed separately as a meta-evaluation layer. Together, these three cohorts provide a layered, multi-perspective view of the AI moderation experience that few prior studies have attempted.

Participant Engagement

How engagement and output quality differ based on who is recruited and how.

The Participant Experience

Comfort, ease, and willingness to disclose in an AI-mediated setting.

Depth & Quality

Level of elaboration, nuance, and reflective richness in responses.

Analytical Usability

Clarity, interpretability, and downstream usefulness of AI outputs.

The findings reveal a consistent and important pattern: participants experienced AI moderation as comfortable, natural, and easy to engage with. Qualitative researchers, however, identified meaningful constraints — particularly around probing depth, emotional nuance, and the risk of meaning compression at scale. The central thesis of this report is not that AI moderation fails, but that it performs differently across different research objectives, and that understanding those differences is essential to deploying it responsibly and effectively.

Study Design

How This Study Was Conducted

This evaluation was fielded in October 2025, with data collection completed within a three-to-four day window from study open to close — itself a testament to the speed advantage of AI-moderated research. The study employed AI-moderated interviews using a platform that enabled both text and voice modality, giving participants agency over how they engaged. Each participant was exposed monadically to three distinct concepts, with structured probing parameters applied consistently across the sample.

Fieldwork Timeline

October 2025 3–4 days, study open to close

Average LOI

~20-25 minutes per session

Modality

Text and/or voice enabled — participant choice

Exposure Design

Monadic — three concepts per participant

Method Structure

Interviews were conducted using a structured flow with controlled probing parameters. This design choice enabled consistency and comparability across sessions, but it also — by design — limited the adaptive flexibility that characterizes skilled human moderation. The controlled probing structure is therefore both a methodological asset and a lens through which findings should be interpreted.

Comparative analysis was conducted across participant types, allowing the research to isolate the effect of sample source on output quality. This comparative layer is one of the study's most distinctive contributions: rather than evaluating AI moderation in isolation, it contextualizes performance relative to recruitment method and participant orientation.

Sample Design

Three Cohorts, Three Perspectives

The study's three-cohort design is foundational to its analytical value. Each cohort brings a distinct orientation, incentive structure, and behavioral profile to the data — and understanding those differences is essential to interpreting what the AI moderation platform captured, and what it did not.

1

Panel Participants (n=101)

Recruited through the AI platform's panel at a **$3 incentive**. These participants approach research tasks with efficiency in mind — they are accustomed to structured, task-oriented interactions and tend to provide concise, on-prompt responses. Their behavior is well-suited to pattern detection at scale.

2

Pure Qual Recruits (n=28)

Recruited through traditional human qualitative methods at a **$30 incentive**. These participants were selected for their ability to engage conversationally and reflectively. They bring a different behavioral orientation — more verbose, more expressive, and more likely to surface the kind of layered, nuanced input that characterizes deep qualitative work.

3

Qual Researchers — Bonus Cohort (n=37)

Experienced researchers recruited from the qualitative insights community, with **no monetary incentive**. This cohort served as a meta-evaluation layer, assessing the AI moderation experience through a professional lens — evaluating interaction quality, probing effectiveness, and analytical usability rather than simply reporting their own experience.

Evaluation Framework

Five Domains of Assessment

Performance was assessed across five structured domains, each designed to capture a distinct dimension of AI moderation capability. Together, these domains provide a comprehensive picture of what AI moderation does well, where it encounters limitations, and how those limitations manifest differently depending on the evaluator's perspective.

Comfort & Disclosure

Willingness to share openly, including on personal or sensitive topics. Measured through participant self-report and researcher observation.

Interaction Quality

The naturalness versus structure of the AI-moderated conversation — how conversational the exchange felt from both participant and researcher perspectives.

Depth of Response

Level of elaboration, nuance, and reflective richness. Assessed relative to sample type and compared against what human moderation typically yields.

Probe Effectiveness

The AI's ability to follow up meaningfully and deepen responses — the degree to which probing extended or enhanced what participants initially offered.

Analytical Usability

Clarity and downstream usefulness of AI-generated outputs — whether themes and summaries accurately reflected the richness of underlying participant input.

Sample Structure & Behavioral Context

Sample Design Shapes Output — Not Just Data Quality, But Data Type

One of the most important findings of this study is that sample design is not merely a methodological consideration — it is a determinant of the type of data generated. The differences observed between Panel and Pure Qual participants are not simply demographic or attitudinal; they are behavioral, and they directly influence the character of the AI moderation output produced. Understanding this distinction is essential to interpreting any AI-moderated research project.

Sample Type	n	Incentive	Observed Behavior
Panel Participants	101	$3	Concise, efficient, task-oriented responses
Pure Qual Recruits	28	$30	Verbose, reflective, conversational responses
Qual Researchers	24	None	Critical, evaluative, assessment-focused

Key Takeaways

- Sample is not neutral — it actively shapes output character
- Incentive level influences depth, effort, and verbosity
- Panel vs. Pure Qual differences are behavioral, not just demographic
- Predictably, the researcher cohort evaluates differently than participants experience

Bottom Line: AI performance must be interpreted through the lens of who is responding — not just how the system operates. The same platform, the same probing logic, and the same topic can yield meaningfully different outputs depending solely on who sits on the other side of the conversation.

Participants Felt Comfortable. Researchers Saw the Gaps.

Perhaps the most striking finding in this study is the divergence between how participants experienced AI moderation and how researchers evaluated it. These are not two perspectives on the same phenomenon — they are assessments of fundamentally different things. Participants evaluated the experience of engaging with AI. Researchers evaluated the analytical depth of what that engagement produced.

Participant Experience

- High levels of comfort engaging with AI
- Willingness to share personal and sensitive experiences
- Described as easy, intuitive, and pleasant
- Modality choice (text or voice) enhanced comfort
- Minimal concern about the absence of a human presence

Participants across both cohorts reported a consistently positive engagement experience. The ability to choose between text and voice modality appeared to be a meaningful comfort lever — participants could interact on their own terms, which likely reduced friction and inhibition.

Researcher Evaluation

- Interaction perceived as linear and structured ('too survey-like')
- Limited sense of genuine conversational flow
- Probing often did not extend depth meaningfully
- Emotional nuance surfaced but remained underdeveloped

Experienced qualitative researchers, by contrast, evaluated the same interactions through the lens of analytical yield. They noted that while the platform produced clean, structured outputs, it consistently fell short on the dimensions that define high-quality qualitative research — adaptive probing, emotional development, and context retention.

Key Distinction: A positive participant experience is a necessary condition for good qualitative research — but it is not sufficient. Comfort creates the conditions for disclosure; probing depth and interpretive skill determine what is done with that disclosure.

Participant vs. Researcher Lens

Experience Metrics vs. Analytical Depth Metrics

The gap between participant-reported experience and researcher-assessed analytical quality is not a contradiction — it reflects two legitimate but distinct evaluative frameworks operating simultaneously. This distinction has significant implications for how AI moderation performance is reported and interpreted by platform providers and research buyers alike.

Dimension	Participants (Panel + Pure Qual)	Qual Researchers
Comfort sharing	High	—
Ease of use	High	—
Enjoyment	High	—
Conversational feel	Moderate–High	Low
Naturalness	High	Moderate
Depth of insight	Not explicitly evaluated	Moderate
Probe effectiveness	Not evaluated	Limited

It is worth noting that participants did not report the experience as sterile or transactional. The perception of transactionality is a researcher-imposed frame — a professional assessment of what the interaction did not achieve analytically, not a reflection of how participants felt while in it. This distinction matters: it means AI moderation can deliver an authentic, comfortable participant experience even when its analytical outputs are constrained.

Bottom Line: AI performs well on experience metrics, but is more constrained on analytical depth metrics. These two dimensions should not be conflated in platform evaluation or research design decisions.

Recruitment Effects

Who You Recruit Shapes What You Get

The behavioral differences between Panel and Pure Qual participants were among the most consistent and consequential findings in this study. These differences did not emerge from the AI moderation platform itself — they were introduced by the sample design before a single interview began. This finding has important implications for how AI moderation projects are scoped, recruited, and interpreted.

Panel Participants

- More concise, task-oriented responses
- Tend to answer the question as asked — no more
- Lower levels of spontaneous elaboration
- Efficient, structured input well-suited to pattern detection

Panel participants are a strong fit for AI moderation at scale. Their behavioral orientation aligns naturally with the platform's strengths — consistent structure, efficient data collection, and rapid thematic synthesis across large samples.

Pure Qual Recruits

- More verbose and expressive responses
- Greater reflection and storytelling tendencies
- Higher levels of spontaneous elaboration
- More layered and nuanced input overall

Pure Qual participants generate the kind of rich, narrative input that AI moderation is best positioned to capture — but may not fully develop. Their responses surface more of what researchers look for, even when the probing does not actively elicit it. The depth originates with the participant, not the platform.

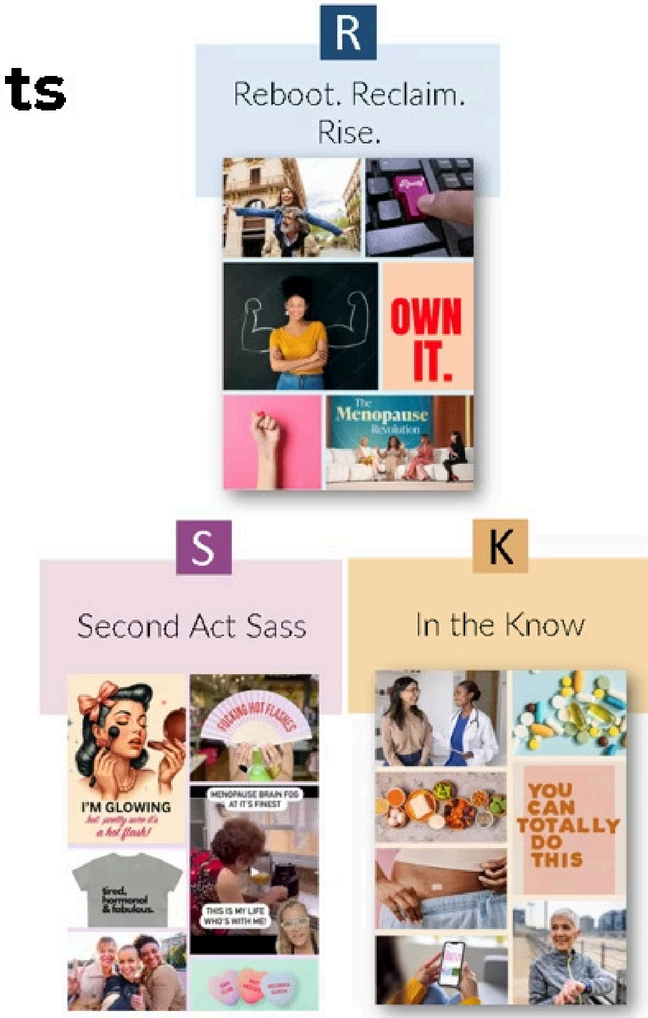
Both cohorts were exposed to the same lead concept — confirming that observed differences in output quality are attributable to sample source and recruitment approach, not topic or stimulus. **Pure Qual participants generate richer narratives. Panel participants enable scalable pattern detection.** Both have a role; the research design should make that choice deliberately.

Core Finding: While Panel and Pure Qual participants demonstrated clear differences in response style and depth, **it is important to note that they often aligned on the same leading concept in this study.**

Panel and traditional qual participants netted out in the same place

	R	S	K
Total scores of 8-10	65%	59%	55%
Total scores of 1-3	11%	14%	12%
Panel average	7.5	7.2	7.1
Traditional average	7.8	7.0	7.6
Panel preferred	36%	24%	32%
Traditional preferred	43%	25%	32%
Panel least liked	15%	16%	24%
Traditional least liked	18%	21%	25%

Traditional Qual = 28
Panel = 101



Responsive Research

Concept R leads across both cohorts, but alignment in outcomes does not equal equivalence in insight.

> **Core Finding:** AI can help you identify what works.
> Sample design determines how well you understand why it works.

Depth Generation

Depth Is Not Evenly Distributed

One of the most important structural findings of this evaluation is that depth of response is not a product of the AI moderation platform — it is primarily a function of who is in the conversation. This has profound implications for how AI moderation is positioned, scoped, and evaluated. When depth appears in AI-moderated data, it should be attributed to participant articulation, not probe quality.

Panel Responses

Skew toward **surface and moderate depth**. Responses are on-prompt, efficient, and consistent — but rarely extend beyond what was directly asked. Elaboration is the exception, not the norm.

Pure Qual Responses

Show a meaningfully **higher proportion of deep, reflective input**. Participants volunteer context, storytelling, and nuance — generating the kind of layered data that qualitative analysis requires.

Depth Driver

Depth is **primarily participant-driven — not probe-driven**. AI probing maintained consistency and structure, but did not reliably unlock or elevate responses that were not already trending toward elaboration.

Core Finding: AI captures structure consistently. Depth depends on who is in the conversation. Platform evaluation that does not account for sample source will systematically misattribute the source of analytical value in AI-moderated data.

Depth Generation Model

A Framework for Understanding When Depth Emerges

The below 2×2 framework captures the interaction between probing quality and participant input, illustrating why depth in AI-moderated research is conditional rather than guaranteed. Understanding this framework is essential for research designers who want to set appropriate expectations for AI moderation outputs and avoid misreading the absence of depth as a platform failure when it may be a sample design issue.

Condition	Strong Participant Input	Weak Participant Input
Strong Probing (Human Moderator)	Deep, layered insight — the gold standard of qualitative depth	Moderate insight — rescued and developed through skilled adaptive probing
Limited Probing (AI Moderation)	Moderate–High insight — captured when participants naturally provide it	Low insight — neither input nor probing generates meaningful depth

Key Insights from the Framework

- AI Moderation captures depth when participants naturally provide it
- AI Moderation does not consistently elevate or unlock weaker input
- Skilled human moderators actively *create* depth through adaptive, thoughtful probing
- AI is more effective at capturing than generating insight

Depth Emerges From:

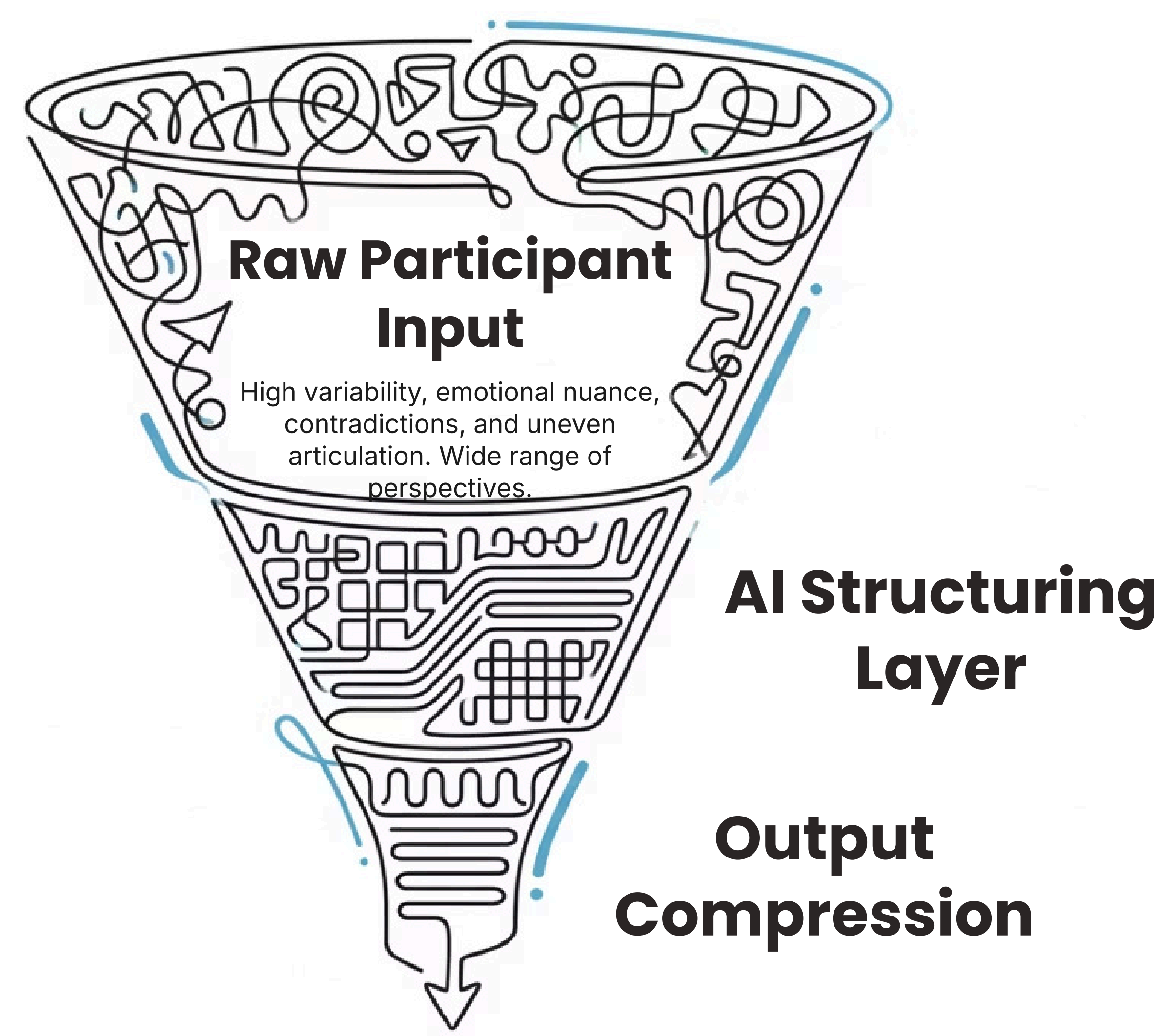
- **Participant mindset** — orientation toward reflection vs. task completion
- **Articulation ability** — verbal fluency and narrative capacity
- **Motivation** — incentive, topic relevance, and engagement level

Bottom Line: AI can scale responses — but it does not yet reliably create deeper meaning. The platform amplifies what participants bring; it does not substitute for what they lack.

The Flattening Effect

AI Scales Conversations — But Can Compress Meaning

Among the most consequential dynamics identified in this evaluation is what we term the "flattening effect" — the tendency of AI moderation, at scale, to reduce the variance and texture of participant input as it moves through standardized questioning, thematic clustering, and output generation. This is not a flaw unique to any single platform; it is a structural feature of how AI moderation processes and synthesizes unstructured human expression.



What Gets Lost in Processing

- Variance is progressively reduced through each processing layer
- Signal becomes clearer — but less textured and differentiated
- Edge cases and subtle distinctions may be systematically diminished
- Distinct individual experiences risk compression into generalized themes

The Core Tension

Scale and consistency are genuine strengths of AI moderation. But they come at a cost: the same mechanisms that make AI output clean and actionable also reduce the qualitative texture that gives insights their strategic and emotional resonance.

Without active researcher interpretation, nuance can be lost not in the conversation — but in the synthesis.

AI can scale conversations — but it can also flatten what makes them meaningful. Active human interpretation is the safeguard against meaning compression.

Comparative Capability

Different Strengths. Different Roles.

Rather than framing AI and human moderation as competitors for the same research task, this evaluation supports a more nuanced and strategically useful framing: AI and human moderation have distinct capability profiles that make each better suited to different research objectives. The following comparison is not a performance scorecard — it is a design guide for research practitioners who want to deploy each modality where it performs best.

Capability	AI Moderation	Human Moderation
Scale	High	Low
Speed	High	Moderate
Participant Comfort	High	High
Consistency	High	Moderate
Probing Depth	Moderate–Low	High
Emotional Insight	Moderate	High
Narrative Development	Low	High
Synthesis & Meaning-Making	Low	High

The pattern in this comparison is clear and consistent: AI moderation excels at the structural, scalable dimensions of research — speed, consistency, and participant-facing comfort. Human moderation retains a decisive advantage wherever depth, meaning-making, and narrative development are required. These are not incremental differences — they reflect fundamentally different capabilities shaped by fundamentally different mechanisms.

> **AI scales conversations. Humans build meaning.** These are complementary strengths — not competing ones.

Strategic Fit

Fit-for-Purpose, Not One-Size-Fits-All

The most important strategic reframing to emerge from this evaluation is the shift from asking "should we use AI moderation?" to asking "where does AI moderation perform best?" This is a design question, not a technology question. The answer depends on the research objective, the depth of insight required, the sensitivity of the topic, and the ultimate use of the findings.

High AI Fit

- Concept screening at scale
- Message testing with structured stimuli
- Early-stage directional feedback
- Precursor to quantitative work
- Rapid pattern detection across large samples
- Precursor to deeper Qual work (to handpick articulate participants)

Hybrid Opportunities

- Sensitive topics where AI comfort enables disclosure
- Exploratory learning with human synthesis overlay
- Iterative concept development across rapid cycles
- Large-sample screening followed by human deep-dives

Human-Led Priority

- Deep exploratory research on complex topics
- Strategic positioning and brand narrative work
- Emotional journey mapping
- Complex decision dynamics and behavioral deep-dives

AI performs best when structure matters more than depth — when the goal is to collect, organize, and detect patterns across many voices efficiently. Human moderation performs best when meaning matters more than scale — when the goal is to understand, interpret, and develop insight from fewer but richer interactions. The most sophisticated research designs will integrate both.

The Evolving Researcher Role

From Asking Questions to Protecting Meaning

As AI moderation scales conversations, the role of the qualitative researcher does not diminish — it transforms. The traditional moderator role centered on relationship-building, adaptive questioning, and in-the-moment interpretive judgment. In an AI moderation context, those skills must be redirected toward a different but equally critical function: ensuring that meaning survives the processing pipeline.

1

Traditional Role

- Ask questions
- Probe for depth
- Guide the conversation
- Build rapport in real time

2

Evolving Role in AI Context

- Interpret AI outputs critically
- Reconstruct nuance that was compressed
- Identify what is missing or underdeveloped
- Protect against oversimplification
- Advocate for the participant's voice in analysis

This shift demands a higher order of analytical skill — not less expertise, but different expertise. Researchers who work effectively with AI-moderated data must be comfortable reading between the lines of structured outputs, recognizing when themes are too clean to be fully accurate, and interrogating synthesis artifacts rather than accepting them at face value. The researcher becomes a meaning guardian, not just an analyst.

Core Principle: As AI scales conversations, the researcher must ensure meaning survives. The platform generates the transcript; the researcher protects its integrity.

Key Findings Summary

Six Things This Study Confirms

Across 167 sessions, three cohorts, and five evaluation domains, this study yields six findings that are consistent, consequential, and actionable for qualitative researchers and insight professionals evaluating AI moderation as a research method.

1 Participants Report a Consistently Positive Experience

Comfort, ease, and willingness to disclose were high across both participant cohorts. AI moderation does not create a barrier to engagement — for most participants, it removes one.

2 Transactionality Is a Researcher Frame, Not a Participant Experience

The perception that AI moderation feels "sterile" or "transactional" originates from the researcher evaluative lens, not from participant self-report. This distinction matters for how platform limitations are characterized and communicated.

3 Sample Source Shapes Insight Quality — Even When Outcomes Align

Panel and Pure Qual participants often converged on the same leading concept in this study, suggesting that AI-moderated research can reliably identify directional winners across sample types.

However, the *quality of insight behind those choices differed meaningfully.* Pure Qual participants provided more deeply articulated and diagnostically useful responses, offering greater clarity on why concepts resonated and how they could be improved.

4 AI Can Capture Depth — But Doesn't Consistently Create It

When participants naturally provide rich, reflective input, AI moderation captures it effectively. But the platform does not reliably unlock depth from participants who are not already inclined to provide it.

5 Outputs Require Active Interpretation

AI-generated themes and summaries are a starting point, not an endpoint. Preserving meaning requires a researcher who can read outputs critically and reconstruct what the processing pipeline may have compressed or generalized.

6 AI Is a Complement, Not a Replacement

The evidence does not support AI moderation as a universal substitute for human-led qualitative research. It supports AI moderation as a powerful, fit-for-purpose tool that performs best in specific research contexts alongside — not instead of — human expertise.

Study Limitations

Interpreting These Findings in Context

This study was designed with rigor, but findings should be interpreted within the boundaries of its design. The limitations described below do not undermine the central conclusions — but they do define the conditions under which those conclusions apply, and they point toward important areas for future investigation as AI moderation continues to evolve rapidly.

Single Platform

This evaluation reflects the performance of a single AI moderation platform. Findings may not generalize to other platforms with different probing architectures, language model backends, or interaction designs. Comparative cross-platform research would strengthen the evidence base significantly.

Specific Use Case

All interviews were focused on a specific topic (menopause). Platform performance on sensitive health topics may differ from performance on product testing, brand research, or behavioral inquiry. Use case is a meaningful variable in evaluating AI moderation capability.

Controlled Probing Parameters

The structured, controlled probing design was a deliberate methodological choice that enabled consistency and comparability. However, it also constrained the AI's probing flexibility. Performance under more adaptive probing conditions may differ from what was observed here.

Non-Generalizable Sample

The sample is qualitative and non-generalizable in the statistical sense. Findings represent patterns and observations from a specific set of participants evaluated at a specific point in time — not population-level estimates of AI moderation performance.

Point-in-Time Snapshot

AI moderation technology is advancing rapidly. The findings reported here reflect platform capability as observed in October 2025. Given the pace of development in this space, findings should be treated as a current-state baseline rather than a definitive or durable assessment.

Implications for Practice

What This Means for Qualitative Researchers

The findings of this evaluation have direct, practical implications for qualitative researchers, insight professionals, and research buyers who are evaluating or already deploying AI-moderated methods. These implications span research design, sample strategy, output interpretation, and the positioning of AI moderation within broader research portfolios.

01

Design for the Sample, Not Just the Platform

Recruitment strategy should be a first-order design decision in AI moderation projects. If the research objective requires depth, nuance, and narrative richness, Pure Qual recruitment methods will generate meaningfully better inputs — regardless of platform capability. Panel recruitment is appropriate when scale and efficiency are the priority.

02

Match the Method to the Objective

Not all research questions are equally well-served by AI moderation. Use the fit-for-purpose framework to identify where AI moderation belongs in the research design — and where it should be complemented or replaced by human moderation. Resist the temptation to deploy AI moderation universally simply because it is available.

03

Treat AI Outputs as a Starting Point

AI-generated themes, summaries, and synthesis outputs should be approached as an analytical input, not a final deliverable. Researchers must be prepared to interrogate outputs, identify compression artifacts, and reconstruct the nuance and texture that the processing pipeline may have reduced. This is not optional — it is the core analytical task.

04

Position AI Moderation as a Complement

The most powerful research designs will integrate AI and human moderation strategically — using AI for scale, speed, and early-stage pattern detection, and human moderation for depth, emotional insight, and strategic meaning-making. This is a design philosophy, not a compromise: it maximizes the strengths of both modalities.

Closing Perspective

The Question Is Not "AI or Human?" — It's "Where Does Each Perform Best?"

This evaluation began with a deceptively simple question: what happens when conversations are mediated by AI rather than a human moderator? The answer, as this study demonstrates, is neither straightforwardly positive nor negative — it is conditional. AI moderation performs with genuine strength in specific contexts and reveals meaningful constraints in others. The appropriate response to this finding is not skepticism about AI moderation, nor uncritical enthusiasm. It is strategic precision.

> The researcher's job is not to choose between AI and human moderation. It is to understand each well enough to deploy both wisely — and to protect the integrity of meaning at every stage of the process.

Platforms like Glaut are expanding what is possible in qualitative research — enabling conversations at a scale and speed that were previously unimaginable, and doing so in a way that participants find comfortable and engaging. That is a genuine and significant contribution to the field. The findings of this study are not an argument against AI moderation; they are a guide to using it well.

As AI moderation technology continues to advance — as probing logic becomes more adaptive, as synthesis becomes more nuanced, as the gap between AI-captured and AI-created depth narrows — the framework established in this evaluation will need to be revisited. The point-in-time findings reported here will evolve. What will not evolve is the underlying principle: **that the quality of qualitative insight depends not just on the technology that mediates it, but on the human judgment that designs, deploys, and interprets it.**

For Research Designers

Use sample source and research objective as primary design variables. AI moderation's output quality is a function of **who is in the conversation**, not just how the platform operates.

For Analysts

Approach AI outputs as inputs. Reconstruct nuance, interrogate themes, and protect meaning at the synthesis stage. **The flattening effect is real — active interpretation is the antidote**.

For Research Buyers

Evaluate AI moderation on fit-for-purpose terms, not universal performance claims. The right question is not "how good is AI moderation?" but "**is AI moderation right for this objective?**"

About This Study

Methodology & Acknowledgments

This research was a collaborative effort made possible by the contributions of many individuals and partners. I would like to thank my research collaborator, Vivian Harris of VMH Qualitative, for her thoughtful partnership throughout the study design and interpretation phases. I am also grateful to Glaut and their support team for enabling the AI-moderated interviews that formed the foundation of this research, and to QualBids for expert qualitative recruitment support. Additional thanks go to Bryant Brown Healthcare for design support and to Bello for providing the live qualitative meeting platform used in researcher debrief sessions. While this study benefited from the expertise and support of these partners, all interpretations, conclusions, and perspectives presented here are those of the author.

Study Specifications

- **Total Sample:** n=166 sessions
  - **Panel Participants:** n=101 ($3 incentive)
  - **Pure Qual Recruits:** n=28 ($30 incentive)
  - **Qual Researchers (Bonus Cohort):** n=37 (no monetary incentive)
- **Fieldwork:** October 2025
- **Duration:** 3–4 days
- **Average LOI:** ~24 minutes
- **Modality:** Text and/or voice (participant choice)
- **Topic:** Menopause (sensitive health context)

About Responsive Research, Inc.

Responsive Research, Inc. is a Los Angeles-based qualitative research consultancy specializing in evaluative research, methodological innovation, and the integration of emerging technologies into qualitative practice. Lauren McCluskey brings extensive experience in both traditional and technology-mediated qualitative methods, with a focus on preserving analytical rigor as the research landscape evolves. She can be reached at lauren@responsive.rocks.

Prepared for Glaut

Glaut is an AI moderation platform enabling qualitative research at scale. This report represents an independent evaluation intended to advance evidence-based practice in AI-assisted qualitative research.

This document was prepared as a professional whitepaper for Glaut. All findings reflect the independent analysis of Lauren McCluskey of Responsive Research, Inc. and are based on data collected in October 2025. Findings should be interpreted as a point-in-time assessment of a rapidly evolving technology landscape.