

Speech-Based Quality Analysis in COA Administration: Profiling Clinician Behaviors in MADRS Interviews



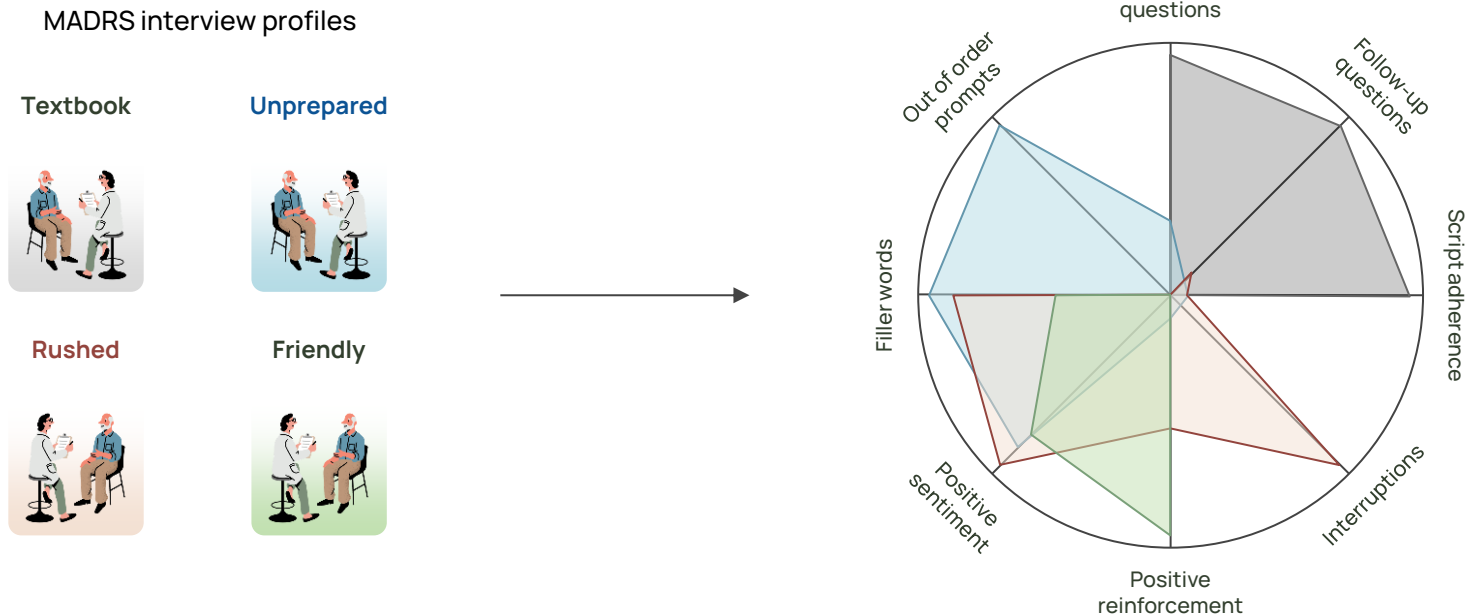
Clinical interview

Can we automate secondary review?

Reviewing *a//*interviews in near real-time

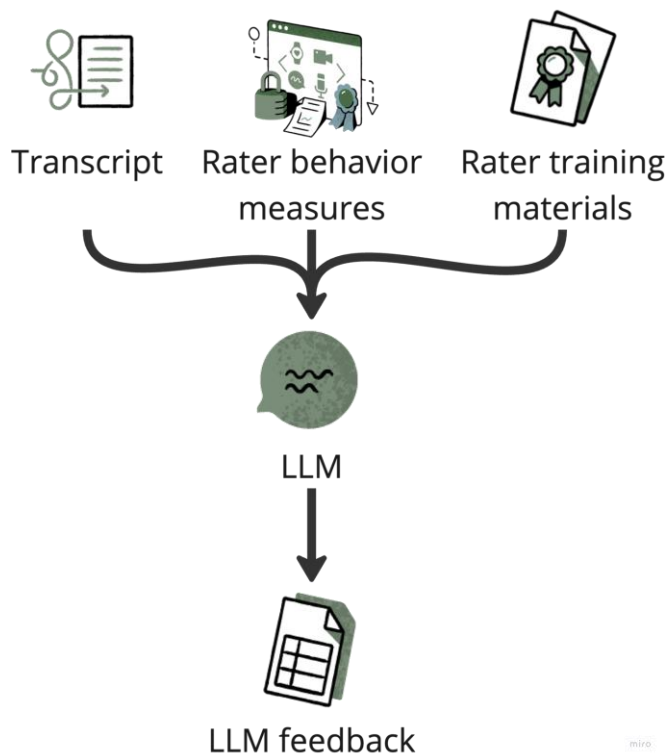
Measuring rater behavior

Starting by quantifying behaviors not captured in language



These measures are incorporated into the input of the large language model

Secondary review using an LLM



Structured, balanced

"The clinician conducting the interview demonstrates a **strong adherence to the structured guidelines** outlined in the MADRS SIGMA manual ... appropriately **follows up to clarify the severity and frequency of symptoms**, demonstrating a balance between ... **high standard of interview quality and reliability.**"

Disorganized and passive

"... However, there are some areas where further refinement could enhance the interview. The clinician sometimes **uses filler phrases** ("ok," "um," "all right") which could detract from a focused exploration of symptoms. They occasionally seem to **overlook or underexplore follow-up prompts on key symptoms**, such as confirming ..."

Dominant, reinforcing

"... the clinician appears **impatient and skips over certain questions**, as seen in their decision to skip exploring concentration difficulties due to a subjective judgment ... their **manner occasionally imposes an overly simplified perspective on complex symptoms**, limiting the opportunity for nuanced responses."

Positive, validating

"aligns with MADRS SIGMA's guidelines ... However, the clinician occasionally **risks over-empathizing**, which could lead to subjective reinforcement. They also tend to use phrases like "that's great" and "good job," which, while encouraging, could **unintentionally minimize the severity of symptoms by focusing on positive interpretations.**"