What Do They Think and Feel?

What assumptions do they make about the system's guardrails?
How confident are they in their ability to succeed?
What's their mental model of how the AI works?
What motivates them -- curiosity, profit, ideology, resentment?

What Do They Say and Do?

What language patterns do they use in prompts?
Do they use tools, scripts, or manual input?
Do they start subtle or go direct?
How do they escalate when initial attempts fail?

What Do They Know?

What's their technical skill level?
What domain expertise do they have about the target system?
What jailbreak patterns or adversarial techniques are they aware of?
What insider knowledge or public documentation have they accessed?

What Limits Them?

How much time are they willing to invest?
What technical knowledge gaps do they have?
Are they concerned about detection?
What access restrictions are they working within?

Pains & Problems

What frustrates them about the system's defenses?
Where have their past attempts failed or stalled?
What constraints make their goal harder to achieve?
What risks do they face if detected?

Needs & Objectives

What specific outcome are they trying to produce from the model?
What does "success" look like for this attacker?
What would they do with the result once they get it?
How good does the result need to be -- exact output or close enough?