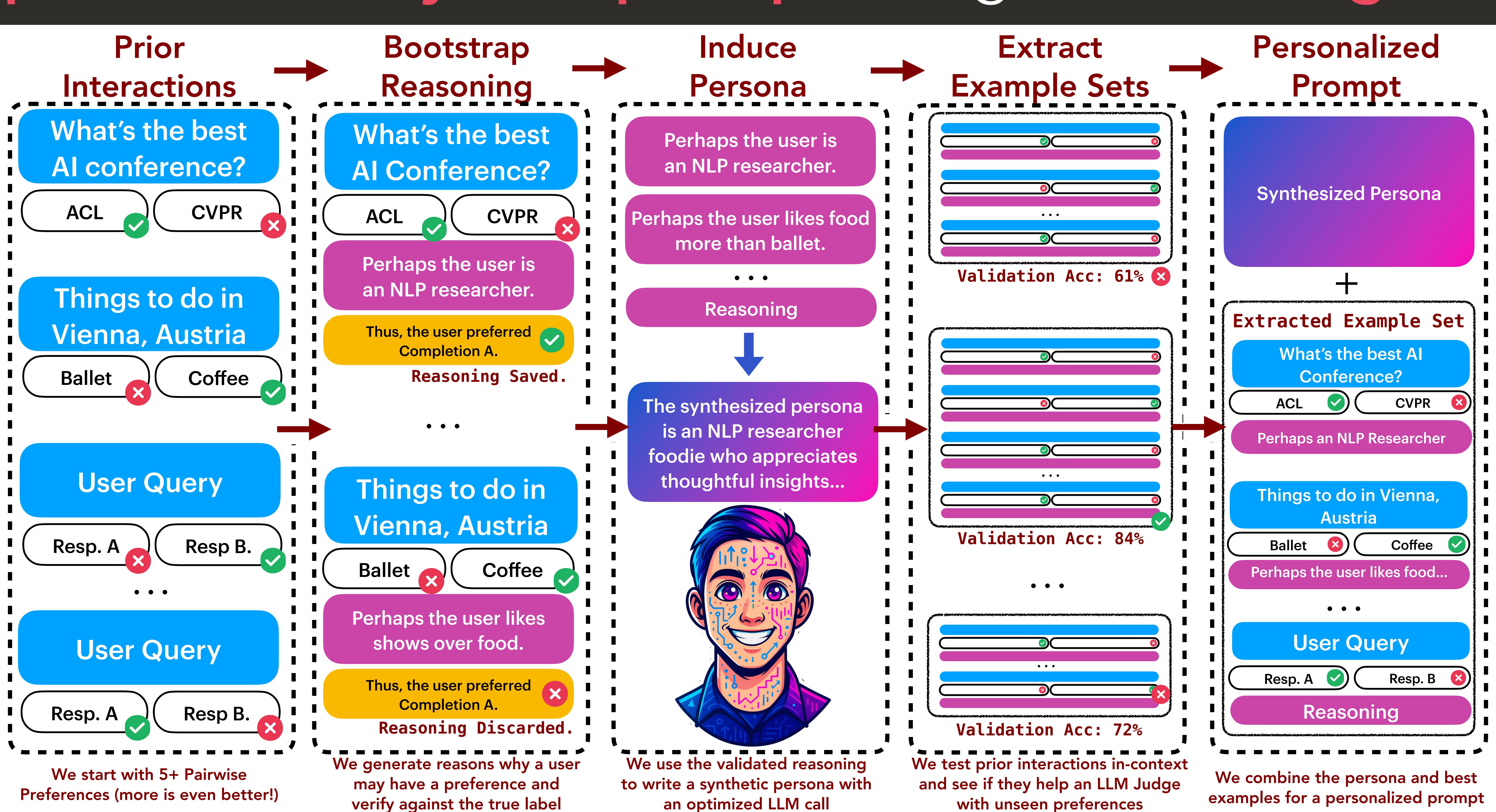


SynthesizeMe! Inducing Persona-Guided Prompts for Personalized Reward Models in LLMs

Michael J. Ryan, Omar Shaikh, Aditri Bhagirath, Daniel Frees, William Held, Diyi Yang

Leverage prior interactions with an LLM to write a personalized system prompt and gain user insights!



Try it now! Get SYNTHESIZED:

Llama 3.2 3B



https://www.try-synthesize.me/

SynthesizeMe produces better LLM Judge Reward Models than Demographics or Memory-Based Personalization **Chatbot Arena** ---- Random (50%) No Personalization Demographics 62 Memory SynthesizeMe

Llama 3.3 70B

Example Persona (user1118 — Qwen3 32B)

The synthesized persona is a **curious, practical, and socially engaged individual** who values:

with unseen preferences

- 1. **Nuanced, balanced perspectives** on complex topics (e.g., religion, relationships).
- 2. **Specific, actionable details** in practical contexts (e.g., recipes, travel recommendations).
- 3. **Relevance and coherence** in responses, avoiding tangents or vague generalizations.
- 4. **Empathy and community focus**, showing interest in shared experiences (e.g., football culture, fan passion).
- 5. **Efficiency in communication**, preferring concise yet thorough explanations over fluff.

This persona is likely a **lifelong learner** who enjoys exploring diverse topics but prioritizes clarity and utility in information. They are open to multiple viewpoints but expect responses to acknowledge complexity without oversimplification.

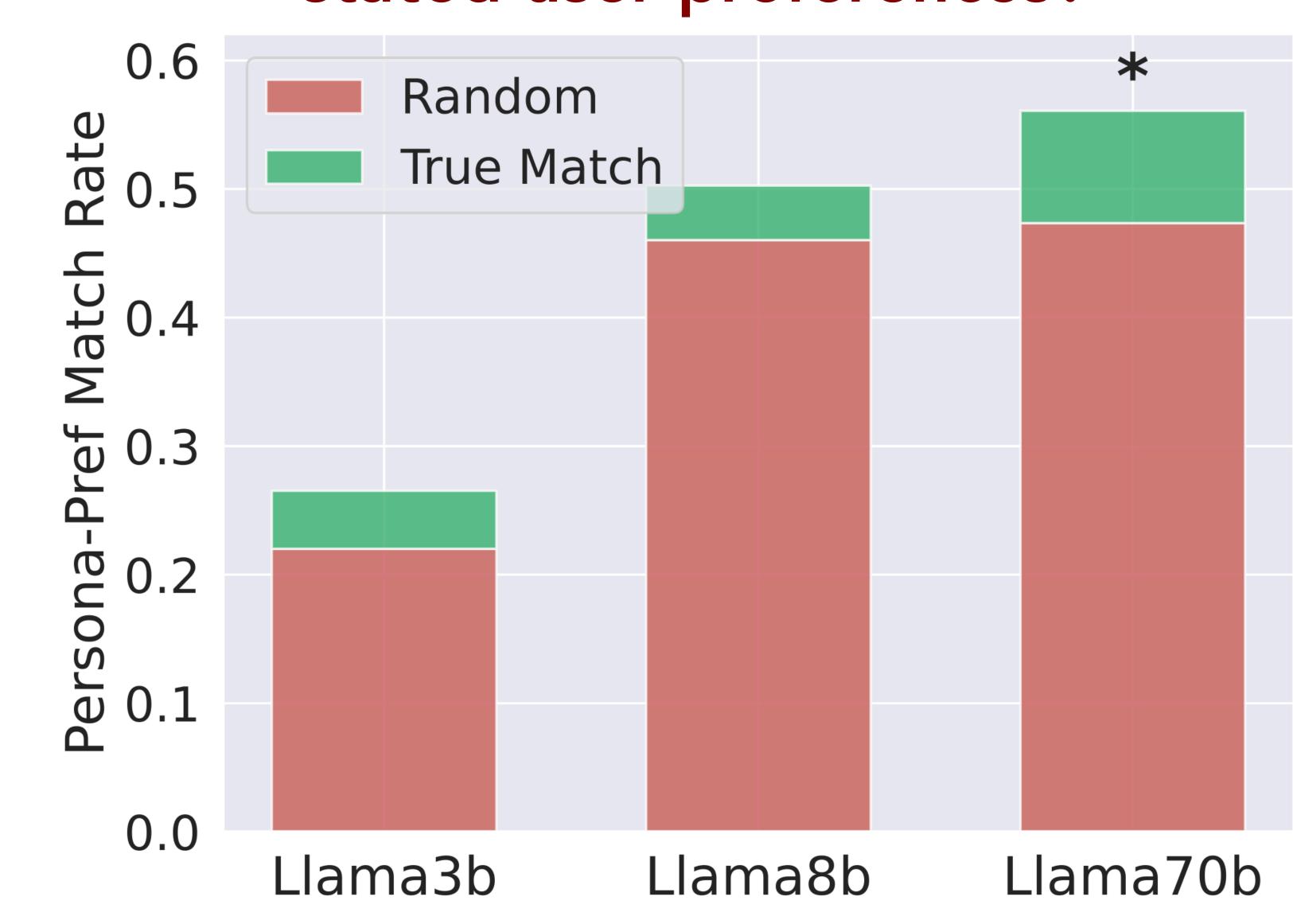
Does SynthesizeMe improve with more data? y = 53.22 + 0.78x (p = 2.38e-11)Average Accuracy No Personalization 65.0 <u>©</u> 62.5 **⋖** 60.0 55.0 52.5

Number of Context Preferences Yes! About 0.8% improvement per interaction

Does SynthesizeMe match stated user preferences?

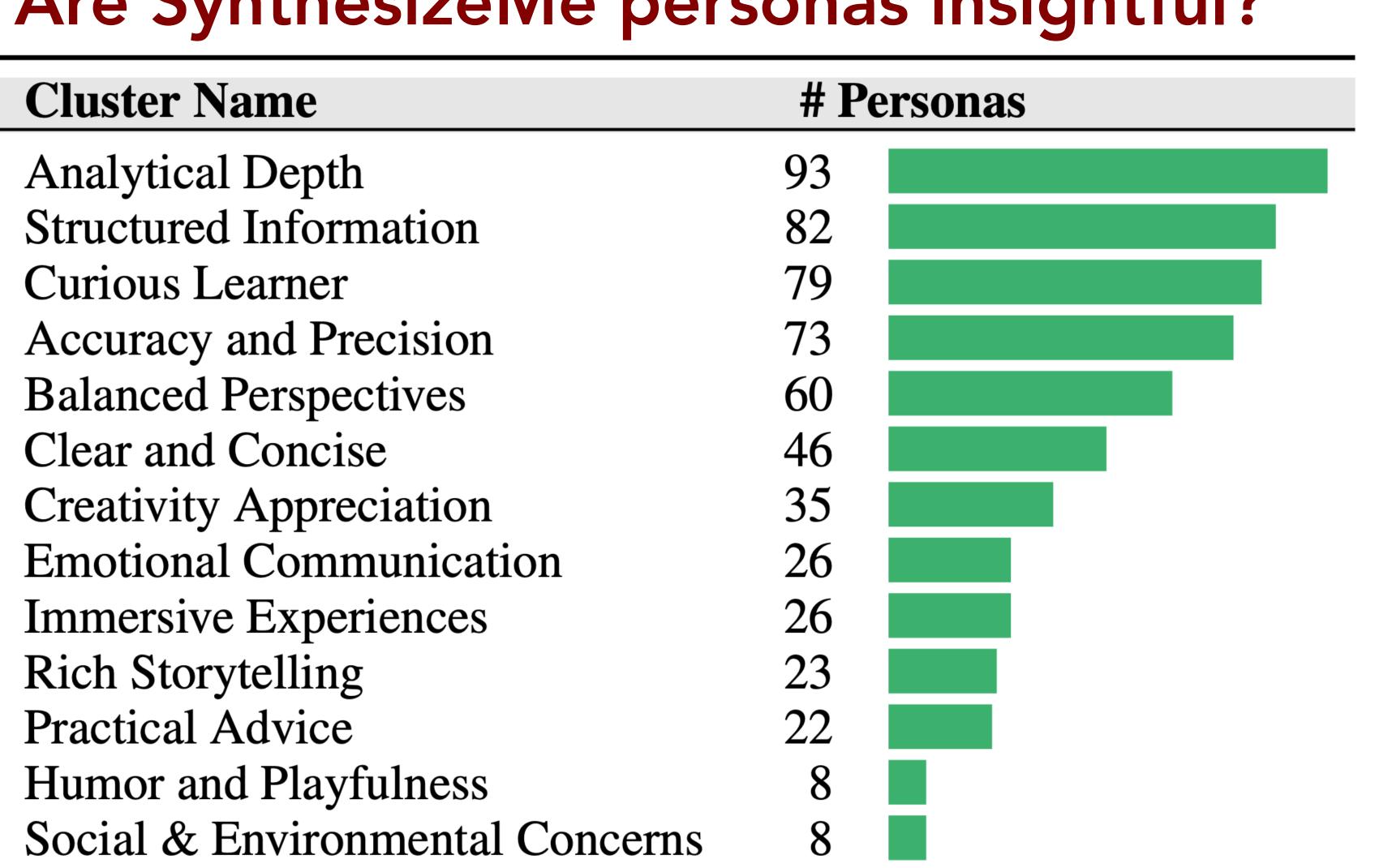
Llama 3.18B

Llama 3.3 70B



Yes! SynthesizeMe personas match with the stated preferences of the users they aim to personify at a statistically signficant rate over a random baseline (p<0.01)

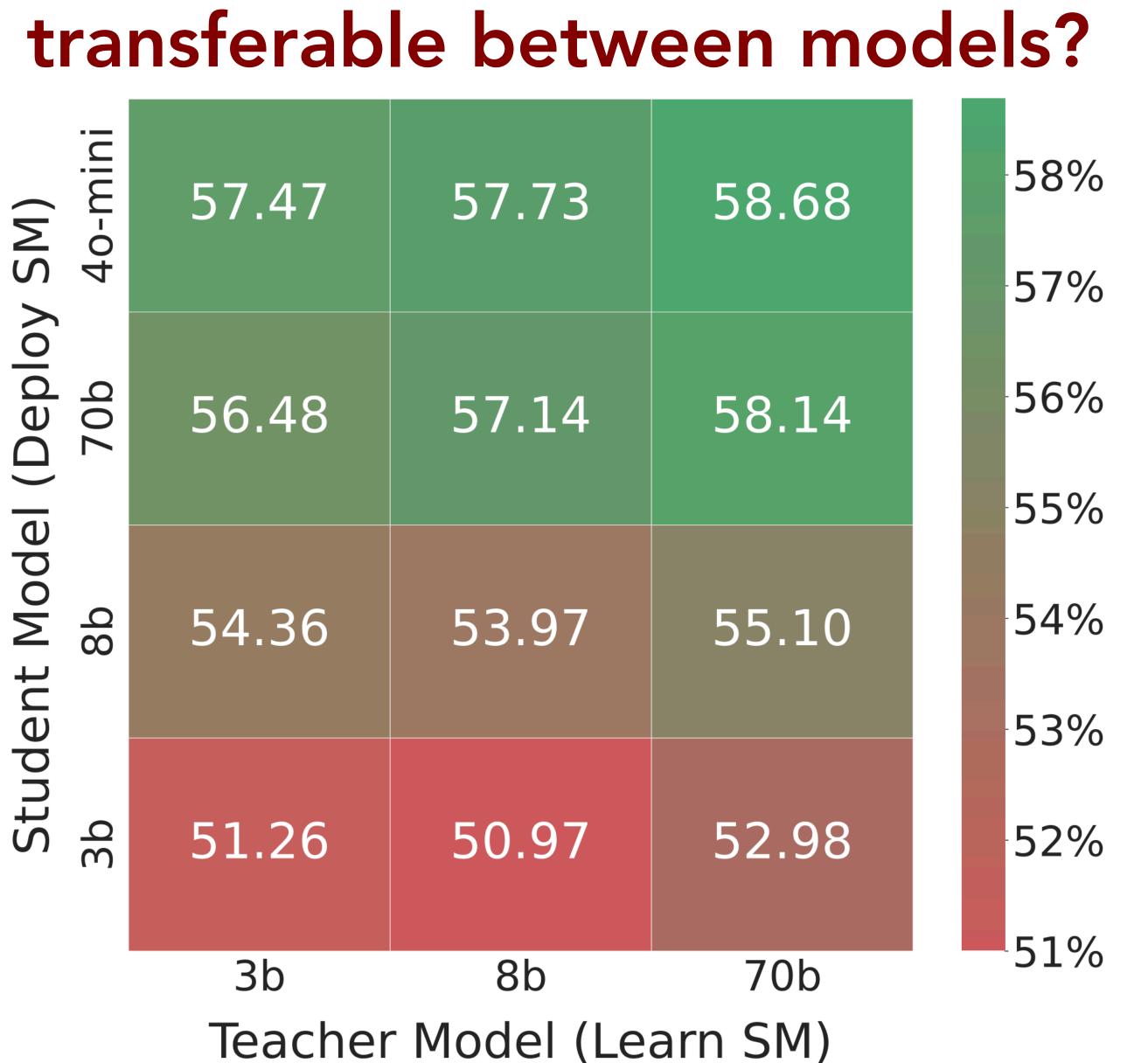
Are SynthesizeMe personas insightful?



insights about the users that are interacting with our platform!

Insights can be specific (e.g. 8 users have environmental concerns) or broad (e.g. most users like structured responses)

Are SynthesizeMe prompts



Yes! You can learn prompts with one model & deploy with another!









Yes! We can cluster personas and find