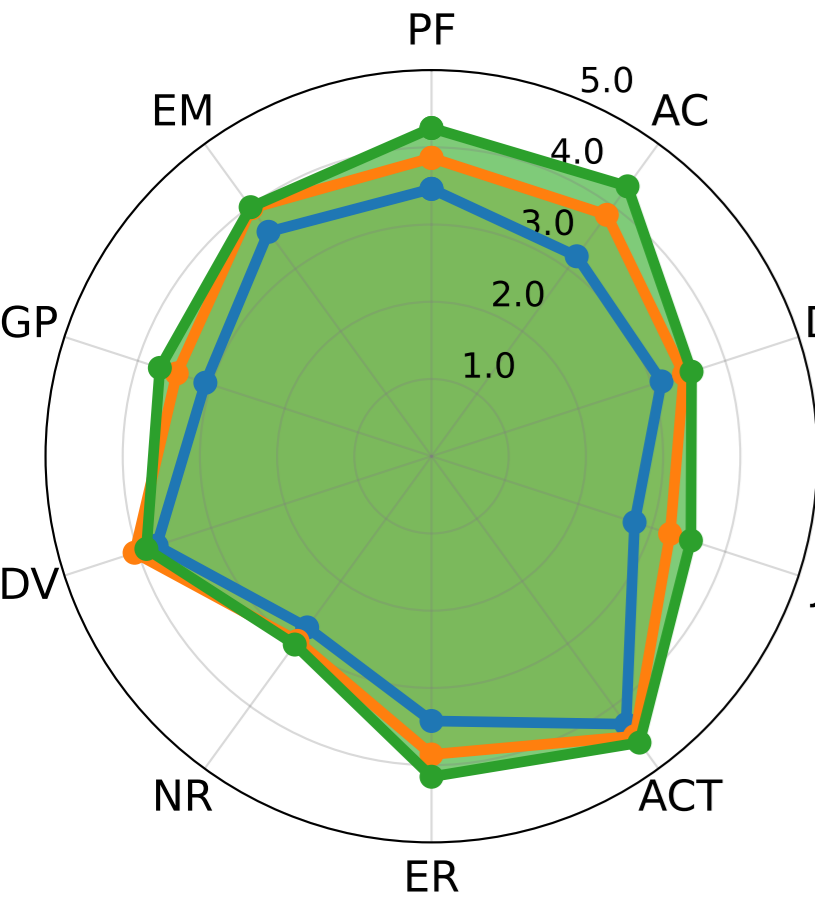
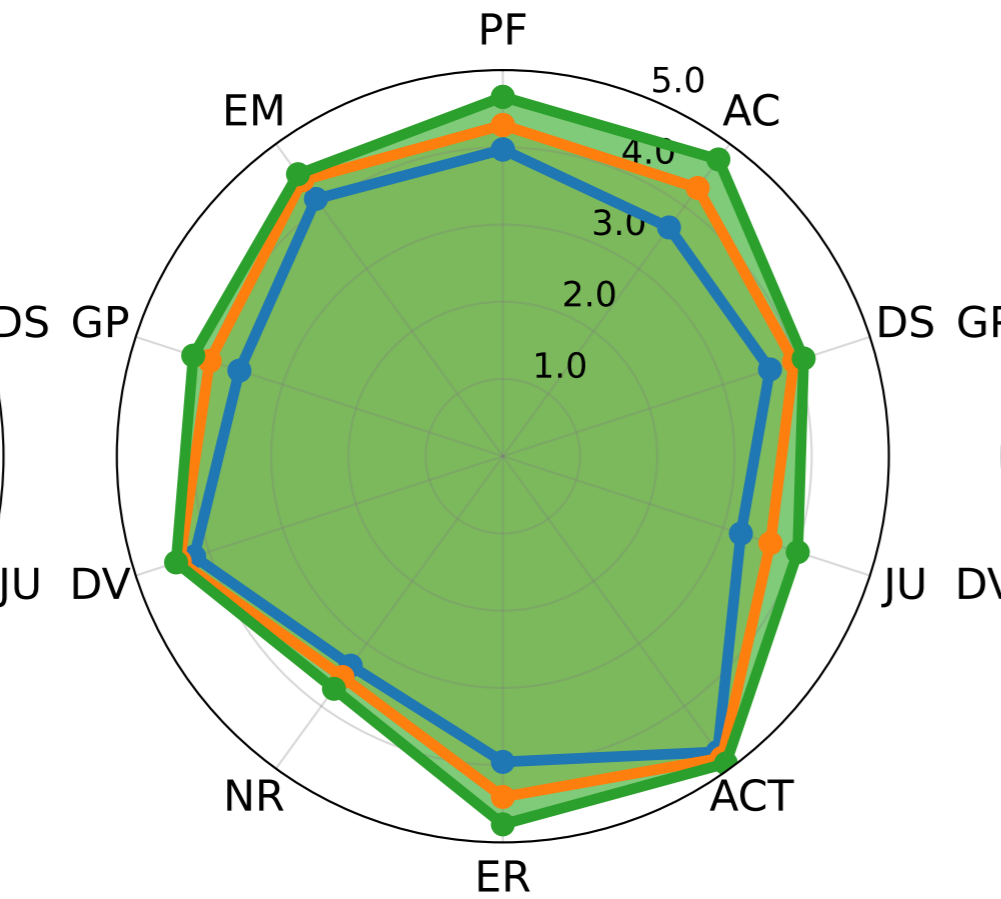




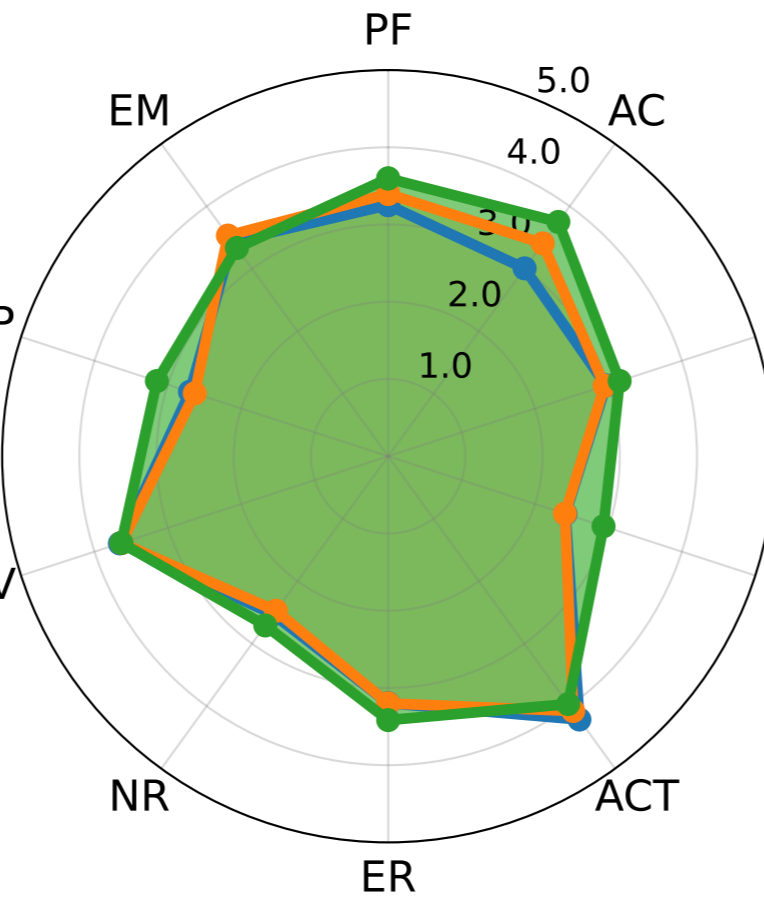
Responder: GPT-4.1-mini
(Evaluator: GPT-4.1)



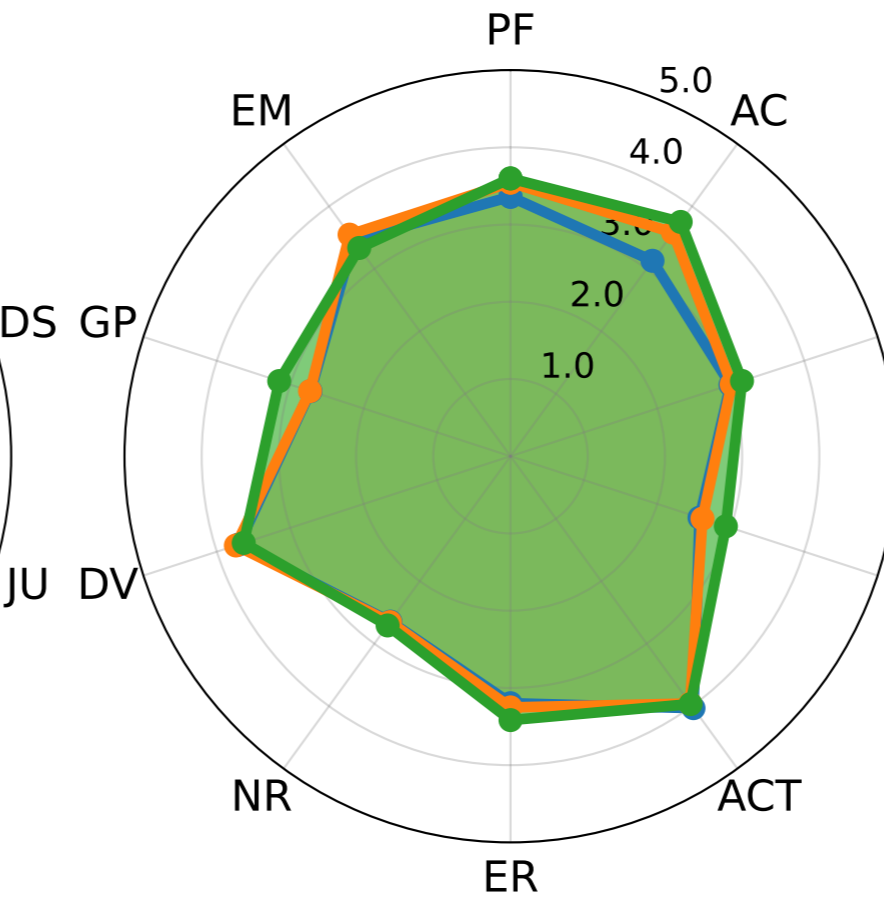
Responder: GPT-4.1
(Evaluator: GPT-4.1)



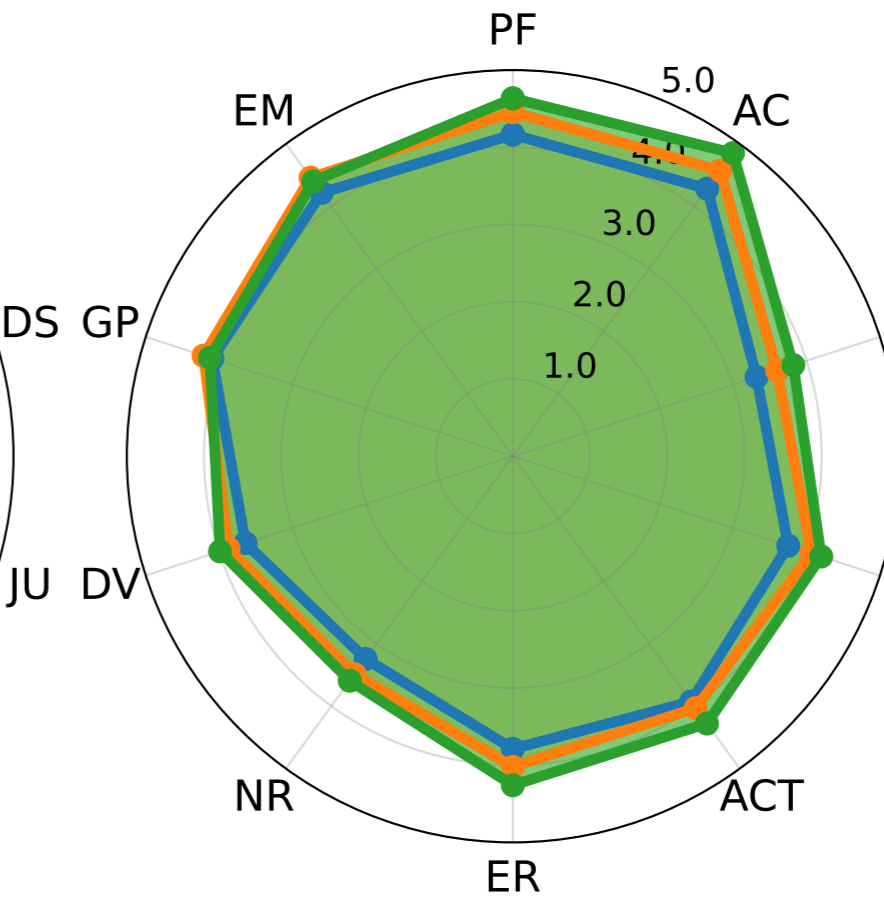
Responder: GPT-4o-mini
(Evaluator: GPT-4.1)



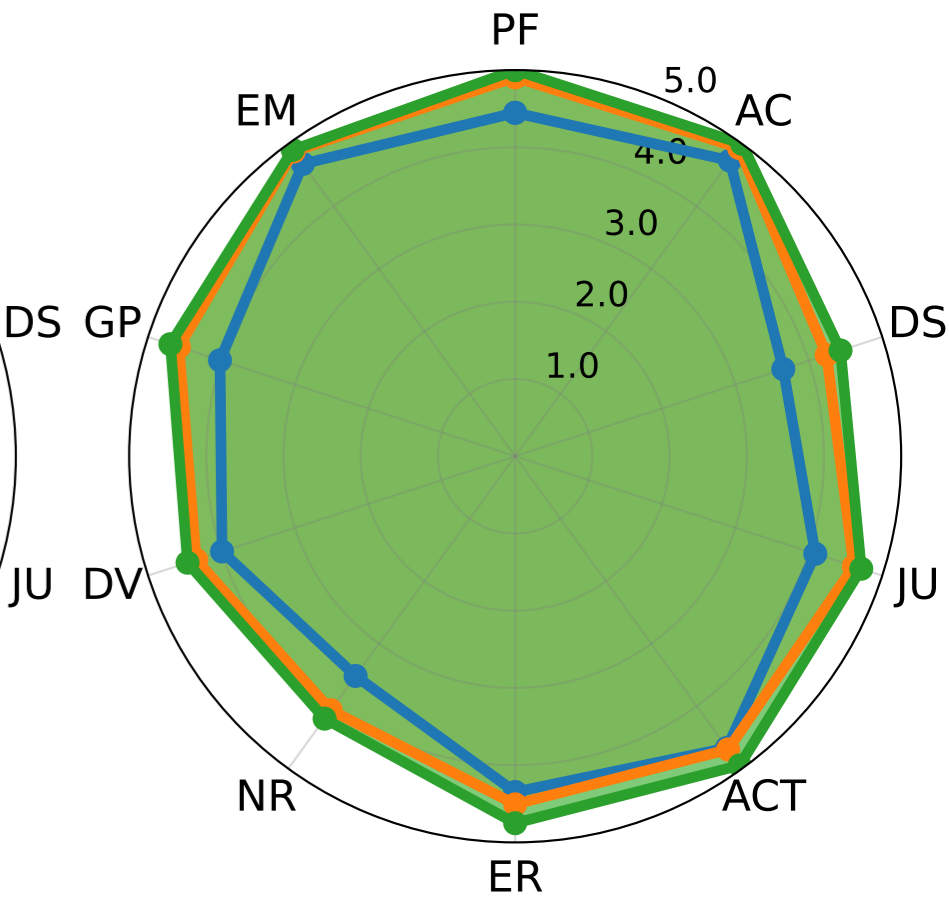
Responder: GPT-4o
(Evaluator: GPT-4.1)



Responder: GPT-4o
(Evaluator: Gemini-2.5-flash)



Responder: GPT-4.1
(Evaluator: Gemini-2.5-flash)



Metrics: Personalization Fit | Attribute Coverage | Depth / Specificity | Justification | Actionability
Effort Reduction | Novelty & Relevance | Diversity of Suggestions | Goal Progress Alignment | Engagement/Motivation Potential