

Model	Egypt			United States		
	English	Arabic	Ar-En	English	Arabic	En-Ar
GPT-3.5	47.08 / 23.42	<u>50.15 / 28.56</u>	3.07	<u>65.95 / 40.22</u>	63.77 / 38.36	2.18
AceGPT-Chat	46.15 / 28.83	49.49 / 30.60	3.34	54.55 / 29.94	51.12 / 25.45	3.43
LLaMA-2-Chat	<u>47.95 / 25.61</u>	44.67 / 23.34	-3.28	<u>63.90 / 37.40</u>	62.29 / 36.03	1.61
mT0-XXL	45.16 / 28.75	46.69 / 27.10	1.53	53.20 / 28.30	<u>57.75 / 34.51</u>	-4.55

Table 2: Cultural alignment against both Egyptian and United States survey responses using **Soft** / **Hard** similarity metrics for each model as a function of the prompting language. Underlined is the optimal prompting language for each model and survey. The third column in each block shows the difference in soft alignment between country’s dominant language and the other language. Refer to Appendix A for results without excluding responses where equivalent personas in both surveys answered similarly.

for each question variant using a temperature of 0.7.⁶ The model’s response for a particular persona and question variant is determined by computing a majority vote over the sampled responses.

Following this, we assess a model’s cultural alignment by comparing its responses for each persona separately with the original subject’s response in one of the two surveys. This comparison is conducted in two ways: either directly comparing the responses (Hard metric) or considering the responses while taking into account the order of the options for ordinal questions (Soft metric). We exclude instances where two subjects belonging to similar persona from both the Egypt and US surveys provided identical answers for a given question. This exclusion ensures a more accurate assessment of each model’s capability in discerning the differences between the two cultures.

Hard Metric Effectively the plain accuracy, which compares model answers to the survey responses for a given persona. Formally, the final cultural alignment is then $\frac{1}{N} \sum_{i=1}^N \mathbb{1}(f(q, p) = y_c(p))$, where N is the number of responses, $f(q, p)$ denotes the model’s response after computing the majority vote for a specific question prompt q and persona p , while $y_c(p)$ is the response of a specific subject with persona p from culture c .

Soft Metric $S_c(q, p)$ is a relaxed version of the hard metric which considers the order of options for questions with an ordinal scale. However, if the question provides categorical options only or the subject in the survey responded with a “don’t know” (orthogonal to the scale), the metric defaults

to plain accuracy.

$$S_c(q, p) = \begin{cases} 1 - \frac{|f(q, p) - y_c(p)|}{|q| - 1} & \text{if } (q, p, c) \in \Theta, \\ \mathbb{1}(f(q, p) = y_c(p)) & \text{otherwise} \end{cases} \quad (1)$$

Here, S represents the cultural alignment score of model f when prompted with question q and persona p for a specific culture, while Θ denotes the set of ordinal questions where the corresponding subject in the survey did not provide a “don’t know” answer. The final score is then averaged accordingly: $\frac{1}{N} \sum_{i=1}^N S_c(p, q)$.

4.6 Anthropological Prompting

Inspired by long-term ethnographic fieldwork—which stands as the primary research method within the discipline of cultural anthropology—we introduce a novel prompting method to improve cultural alignment for LLMs, **Anthropological Prompting**. The objective of engaging in extended ethnographic fieldwork is to establish meaningful connections with interlocutors, facilitating the ability to produce critical and in-depth analyses of both the subjects and the topics under study.

In this context, we strive to emulate a digital adaptation of ethnographic fieldwork by guiding the model to think as if it has been actively participating in this method. We prompt the model to comprehend the intricate complexities and nuances associated with identities, inquiries, and linguistic constructions. For instance, we elaborate on the emic and etic perspectives of examining culture,⁷ highlighting the layered nature of interpersonal connections and emphasizing how personal experiences significantly shape subjectivities. In

⁶ This was empirically set.

⁷ “Emic” refers to an insider’s perspective, focusing on the internal understandings within a specific culture. Conversely, “etic” refers to an outsider’s perspective.