

# Enhancing LLM-Based Persuasion Simulations with Cultural and Speaker-Specific Information

Weicheng Ma<sup>1\*</sup>, Hefan Zhang<sup>2\*</sup>, Shiyu Ji<sup>3</sup>, Farnoosh Hashemi<sup>3</sup>,  
Qichao Wang<sup>2</sup>, Ivory Yang<sup>2</sup>, Joice Chen<sup>3</sup>, Juanwen Pan<sup>3</sup>,  
Michael Macy<sup>3</sup>, Saeed Hassanpour<sup>2</sup>, Soroush Vosoughi<sup>2</sup>

<sup>1</sup>Georgia Institute of Technology <sup>2</sup>Dartmouth College <sup>3</sup>Cornell University  
wma76@gatech.edu

## Abstract

Large language models (LLMs) have been used to synthesize persuasive dialogues for studying persuasive behavior. However, existing approaches often suffer from issues such as stance oscillation and low informativeness. To address these challenges, we propose reinforced instructional prompting, a method that ensures speaker characteristics consistently guide all stages of dialogue generation. We further introduce multilingual prompting, which aligns language use with speakers' native languages to better capture cultural nuances. Our experiments involving speakers from eight countries show that continually reinforcing speaker profiles and cultural context improves argument diversity, enhances informativeness, and stabilizes speaker stances. Moreover, our analysis of inter-group versus intra-group persuasion reveals that speakers engaging within their own cultural groups employ more varied persuasive strategies than in cross-cultural interactions. These findings underscore the importance of speaker and cultural awareness in LLM-based persuasion modeling and suggest new directions for developing more personalized, ethically grounded, and culturally adaptive LLM-generated dialogues. The implementation is publicly available at our GitHub: <https://github.com/HF-heaven/Cross-Cultural-Persuasion-Simulations>.

## 1 Introduction

Recent studies have demonstrated the effectiveness of large language models (LLMs) in multi-agent communication environments for simulating persuasion behaviors (Ma et al., 2025). Persuasion is a complex form of human communication that relies on strategic language use and strong personal opinions. While LLMs have shown promise in generating persuasive dialogues, existing simulations

tend to be uniform in strategy and style, lacking the diversity observed in real-world persuasion. A key limitation is that these models do not account for speakers' personal and cultural characteristics, leading to homogeneous and unrealistic dialogues.

This paper explores whether large language models (LLMs) can generate persuasion dialogues that adapt to speakers' backgrounds, producing more diverse persuasive strategies and supporting evidence that better align with speaker profiles. Specifically, we examine how incorporating and consistently reinforcing speaker profiles influences the diversity of persuasion strategies, stance consistency, and the informativeness and cultural alignment of supporting arguments. To this end, we introduce two enhancements to the baseline persuasion simulation framework proposed by Ma et al. (2025): (1) Personal-profile-infused persuasion simulation, where speaker characteristics are explicitly embedded in the initial prompts to align generated responses with the speakers' demographics. (2) Personal-profile-reinforced persuasion simulation, where speaker profiles are repeatedly emphasized throughout the dialogue to guide each generation step. For both approaches, we further explore a multilingual variant in which prompts and responses are generated in each speaker's native language. This allows us to assess the effect of stronger cultural cues on the quality and authenticity of the generated dialogues.

After generating the persuasion dialogues, we conduct a comprehensive manual analysis to: (1) Assess the *linguistic fluency* and *logical consistency* of the outputs, (2) Evaluate whether the framework *successfully aligns generated dialogues with speaker backgrounds and maintains assigned stances*, even when they deviate from the ideological biases typically exhibited by LLMs, and (3) Examine *persuasion strategy diversity* within and across cultural groups, determining how speaker-specific prompting affects argumentation styles.

\*Equal contribution.

Additionally, we analyze how *language choice* influences generation quality and how *stance extremity* affects the formulation of arguments.

Our results demonstrate that **incorporating speaker profiles and reinforcing prompts significantly enhance persuasion simulation quality** by: (1) Generating *more informative and novel arguments*, (2) Increasing the *likelihood of concrete justifications and evidence*, and (3) Producing a *broader range of persuasive strategies* compared to baseline framework.

We also observe a *trade-off between linguistic diversity and strategic variation*: while enforcing native language use improves argument novelty, it somewhat reduces strategy diversity. Nevertheless, this does not alter our primary conclusion that speaker awareness significantly improves the realism and effectiveness of LLM-generated persuasion dialogues.

By bridging this gap, our *speaker-infused persuasion framework* expands the scope of computational persuasion research, enabling deeper investigations into argumentation strategies, cultural influences, and ideological positioning. This work contributes to *more realistic, context-aware AI-generated dialogues*, with potential applications in social science research, AI-assisted debate systems, and cross-cultural communication modeling.

## 2 Speaker-Infused PersuaSim

We follow Ma et al. (2025) and employ a multi-LLM communication framework to generate persuasive dialogues between two speakers with given background information. The framework consists of two dialogue generation agents (for persuader and persuadee), an utterance quality monitor and a language refinement agent to enhance utterance-level quality, and a global regulation agent to manage dialogue flow. We refer to this framework as **PersuaSim-Orig** and provide the corresponding agent prompts in Appendix B.1.

For personal information injection into PersuaSim-Orig, we modify the initial prompts to the dialogue generation agents and create **PersuaSim-Infused**. Prompts for PersuaSim-Infused are shown in Appendix B.2. Building upon PersuaSim-Infused, we introduce **PersuaSim-Reinforced**, which further emphasizes the necessity for dialogue generation agents to align with the assigned personal backgrounds in every round of generation. The utterance quality monitor

Utterance Quality Monitor Agent

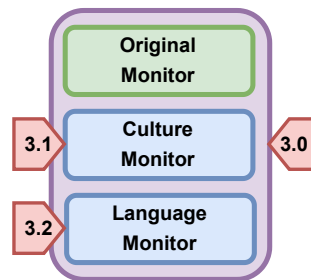


Figure 1: The utterance quality monitor agents in PersuaSim-Reinforced and PersuaSim-Reinforced-MultiLing involve two more monitors than in PersuaSim-Orig. Prompts 3.0 and 3.1 are added for speaker background injection, and Prompt 3.2 is additionally introduced to enable multilingual generation.

Please keep playing the role of: **(profile)**. And keep speaking **(language)** to reply to the persuadee/persuader. Only return your response without any other words. The persuadee/persuader said:

Figure 2: Additional instructions to the dialogue generation agents in PersuaSim-Reinforced (Prompts 3.0).

is also tasked with verifying alignment at each turn of communication. To ensure that speakers’ utterances align with their cultural backgrounds, we made improvements to the Utterance Quality Monitor Agent, as shown in Figure 1. Specifically, we introduced two enhancements: (3.0) For each utterance, we emphasize the agent’s identity and instruct the agent to consistently maintain its assigned role while passing the conversation between speakers (Figure 2). (3.1) We added a filter to detect whether the generated sentences conflict with the speaker’s cultural background (Figure 3). If a conflict is detected, the utterance is revised accordingly.

We additionally designed **PersuaSim-Infused-MultiLing** and **PersuaSim-Reinforced-MultiLing** to investigate the impacts of language choice in the prompts. In these settings, prompts to each dialogue generation agent are translated into the official language of the agent’s specified country of origin, and they are required to generate in that language as well. We made two modifications: (1) Translating the initial settings (Figure B7) and passing prompts (Figure 2) into the target languages. (2) Adding an additional monitoring filter (3.2 in Figure 1) to verify whether the generated utterance is in the intended language. If it is not, the utterance is revised accordingly. The corresponding prompts are shown in Figure 4.

Topic Number	Topic	Agree	Disagree
1	When jobs are scarce, employers should give priority to people of this country over immigrants.	BGD BGD BGD	DEU GBR BGD
2	In terms of the effects of immigration on the development of your country, do you agree that it fills important jobs vacancies?	JPN JPN TUR	TUR JPN TUR
3	In terms of the effects of immigration on the development of your country, do you agree with Strengthens cultural diversity?	JPN	RUS
4	In terms of the effects of immigration on the development of your country, do you agree that it increases the risks of terrorism?	RUS RUS CHN	CHN RUS CHN
5	We depend too much on science and not enough on faith.	KOR KOR KOR	GBR DEU KOR
6	It is a duty towards society to have children.	BGD	DEU
7	Adult children have the duty to provide long-term care for their parents.	BGD	GBR
8	A country's government should have the right to keep people under video surveillance in public areas.	BGD	TUR

Table 1: Topics selected to evaluate LLMs’ capabilities in recognizing and consistently advocating for prevalent viewpoints within each country. Example pairs of countries with opposing viewpoints for setting up our experiments are also provided.

```

CULTURE MONITOR
You will be given what the persuader/persuadee said in the current
round. The persuader/persuadee is playing the role of: {profile}
Please tell me whether the viewpoint expressed by the
persuader/persuadee conflict with the typical perspectives of the
country that the persuader/persuadee from. Response with 'Yes' or 'No'
only.
(if return 'YES', ask agent to revise)

```

Figure 3: Prompts 3.1 (for PersuaSim-Reinforced-MultiLing) that are added to the utterance quality monitor module to check whether speakers deviate from their provided speaker background information. If a deviation is detected, the utterance is revised to ensure alignment.

```

LANGUAGE MONITOR
You will be given what the persuader/persuadee said in the current
round. Please tell me whether the persuader/persuadee is speaking
{language}. Response with 'Yes' or 'No' only.
(if return 'NO', ask agent to revise)

```

Figure 4: Prompts 3.2 (for PersuaSim-Reinforced-MultiLing) that is added to the utterance quality monitor agent that checks whether dialogue generation agents are using the correct target language. If not, the utterance is revised.

### 3 Topic Selection & Dialogue Generation

Using the persuasion simulation frameworks detailed in Section 2, we identify topics with significant cross-country differences in viewpoints (Table 1). These topics are derived from academic survey data and serve as the foundation for generating dialogues used in our main annotation analysis.

#### 3.1 Topic Selection

We utilize the World Values Survey (WVS) (2022), a widely recognized noncommercial academic survey that gathers up-to-date, representative data on cultural opinions from various countries, to construct topics and stances for our dialogue systems. The WVS data serves as a valuable resource for value-based topics, capturing a broad spectrum of perspectives on cultural issues.

For each value topic identified in the WVS, we first filter out non-informative responses, such

as those labeled "Don't know" and "No answer/Refused," ensuring that our analysis focuses solely on substantive input. Next, we compute the probability distributions of responses from each country for the remaining informative choices. To identify significant disparities in cultural opinions, we then calculate the Jensen-Shannon distance between the response distributions of country pairs, selecting the five pairs with the highest distances.

For topics represented by bipolar choices, we apply a bipolar rating scale to compute a weighted average for each country  $c_i$ , producing a single representative value  $s_i^t$  that quantifies the average viewpoint of each country’s populace on topic  $t$ .

Among the top five country pairs  $(c_i, c_j)$  identified for each topic  $t$ , we select pairs where the weighted average scores  $s_i^t$  and  $s_j^t$  not only have opposing signs but also exhibit a substantial divergence. Specifically, we impose the conditions  $|s_i^t| > 0.1$ ,  $|s_j^t| > 0.1$ , and  $|s_i^t - s_j^t| > 0.4$  to ensure the selection of pairs with strong and contrasting distributions over the topics.

Through this systematic approach, we identify 8 distinct topics and 10 country pairs with significantly opposing viewpoints, forming the foundation for exploring cultural dynamics in automated persuasive dialogues between LLMs. Additionally, we incorporate conversations where agents from the same cultural background engage with one another to evaluate the impact of cross-cultural communication on dialogue generation. This final selection results in a total of 16 country pairs.

#### 3.2 Dialogue Generation

We utilized the frameworks outlined in Section 2 to generate dialogues and examine the effects of infusing personal information. For the PersuaSim-Orig framework, we adhered strictly to the methodology described by Ma et al. (2025), generating one

dialogue per topic. For the other frameworks, we synthesized six personal profiles for each of the eight countries listed in Table D1. For each topic, we randomly selected three pairs of profiles from the corresponding pairs of countries to guide the dialogue generation process, resulting in a total of 208 dialogues for annotation analysis. Method for profile synthesis is detailed in Appendix C.

## 4 Validations and Result Analyses

After obtaining persuasion dialogues, we conduct comprehensive manual annotations to evaluate the generation quality of the persuasion simulation framework with controlled participant profiles. Our evaluation focuses on the framework’s ability to: (a) generate high-quality overall dialogues (Section 4.1), (b) enhance the concreteness of supporting arguments (Section 4.2), and (c) demonstrate diversity in persuasion strategy preferences (Section 4.3). Additionally, we examine how inter-group dialogues differ from intra-group dialogues across these measures (Section 4.4).

### 4.1 Quantitative Quality Assessment

We employed the quality evaluation metric proposed by Ma et al. (2025), which synthesizes insights from existing literature to assess the quality of generated dialogues, examining both local and global fluency. At the local level, dialogue quality is evaluated based on two key criteria: **Coherence**, which measures the logical consistency between a speech-response pair, and **Informativeness**, which assesses whether responses introduce new, relevant information that advances the conversation.

At the global level, we assess overall dialogue quality using three dimensions: (1) **Fluency**, which ensures that the conversation remains natural and human-like throughout; (2) **Internal Role Consistency**, which verifies that speakers do not exhibit sudden, unexplained shifts in position over the course of the dialogue; and (3) **Topic Consistency**, which ensures that the conversation remains focused on the predefined subject. The full list of definitions and related works are shown in Table 2.

To ensure annotation quality, two annotators were trained together using the aforementioned standards and example cases. They jointly annotated 30% of the dataset to establish a baseline for inter-rater consistency. The internal consistency of their annotations is reported in Table 2. For the remaining 70% of the data, we implemented additional measures to minimize potential annotator

discrepancies that could affect the evaluation of different generation methods within the same topic. Specifically, we randomly assigned one annotator per topic, requiring them to annotate all samples within that topic. Furthermore, to reduce the impact of between-topic differences on the evaluation, we ensured that for each topic, the number of samples annotated with the no-culture profile and different generation strategies was fixed across topics.

The average scores and their associated confidence intervals for different generation methods are presented in Table 3. As noted by Ma et al. (2025), informativeness remains a primary limitation in LLM-generated dialogues, as LLMs tend to produce more repetitive utterances and relatively simplistic stances over extended interactions. Notably, PersuaSim-Reinforced-MultiLing effectively mitigated this issue to a convincing extent: compared to PersuaSim-Orig, the combination of multilingual and reinforcement agents significantly improved informativeness (two-sample t-test,  $p < 0.05$ ).

The increased informativeness also results in longer, more coherent dialogues with fewer non-sensical stance shifts. On average, PersuaSim-Reinforced and PersuaSim-Reinforced-MultiLing generate dialogues lasting 9.58 and 8.42 rounds, respectively, with the first consensus reached at 7.86 and 7.84 rounds. In contrast, PersuaSim-Orig produces shorter dialogues averaging 8.07 rounds, but wastes approximately 2.19 rounds on meaningless oscillations between stances. Additional details on dialogue lengths across different frameworks are provided in Appendix F.

No significant differences were detected between PersuaSim-Reinforced-MultiLing and PersuaSim-Orig on other measures, despite the latter showing slightly higher scores in Coherence, Fluency, and Internal Role Consistency. PersuaSim-Reinforced performed comparably to PersuaSim-Orig on all measures, with each model demonstrating non-significant strengths in different aspects. On the other hand, incorporating cultural profiles without supervising agents appeared to negatively impact other quality aspects of generation while slightly enhancing informativeness at a non-significant, trivial level.

### 4.2 Argumentative Support Quality

Research highlights the significance of incorporating detailed evidence to enhance the persuasiveness of fact-checking efforts (Costello et al., 2024). However, prior research by Ma et al. (2025)



Evaluation Criteria	Definition	Related Work	$\kappa$	Weighted $\kappa$
<b>Coherence</b>	The speakers respond logically to the immediate conversation.	The responses are relevant to each other (Ke et al. 2018; Wu et al. 2019; Liang and Li 2021) The responses are logical (Zhu et al., 2019) The responses are appropriate to the context (Moghe et al. 2018; Young et al. 2018; Lin et al. 2019; Wu et al. 2019)	0.375	0.465
<b>Informativeness</b>	The utterances introduce relevant new information or arguments, and avoids repeating information.	The responses build on prior information in the conversation. (Moghe et al., 2018)	0.510	0.567
<b>Fluency</b>	The conversation sounds human-like and fluent overall.	Responses can introduce new knowledge or topics in conversation. (Ke et al. 2018; Wu et al. 2019; Zhu et al. 2019) An average reader could understand the conversation. (Moghe et al., 2018; Lin et al., 2019) The responses are grammatically fluent and in a natural conversation style. (Ke et al. 2018; Wu et al. 2019; Zhu et al. 2019; Ji et al. 2022)	0.517	0.520
<b>Internal Role Consistency</b>	There are no sudden shifts in a speaker’s objectives or stance without a clear explanation.	There are no abrupt context switches (Moghe et al., 2018)	0.331	0.428
<b>Topic Consistency</b>	The conversation stays on topic.	The speaker is consistent throughout the conversation. (Ji et al., 2022) The speaker stays on topic. (Moghe et al. 2018; Ji et al. 2022)	0.527	0.508

Table 2: Dialogue-level Quality Evaluation. 2 annotators assessed 50 randomly selected dialogues on the criteria listed above, using a likert scale of 1 - Not Accurate, 2 - Somewhat Accurate, and 3 - Accurate. Both linearly weighted (Weighted  $\kappa$ ) and unweighted ( $\kappa$ ) inter-rater consistency scores are calculated.

Generation Strategy	Coherence	Informativeness	Fluency	Internal Role Consistency	Topic Consistency
PersuaSim-Orig	<b>2.750</b> ( $\pm 0.238$ )	1.906 ( $\pm 0.437$ )	<b>2.844</b> ( $\pm 0.188$ )	<b>2.969</b> ( $\pm 0.067$ )	2.750 ( $\pm 0.308$ )
PersuaSim-Infused	2.521 ( $\pm 0.169$ )	2.031 ( $\pm 0.173$ )	2.802 ( $\pm 0.130$ )	2.719 ( $\pm 0.140$ )	2.708 ( $\pm 0.161$ )
PersuaSim-Infused-MultiLing	2.375 ( $\pm 0.231$ )	2.135 ( $\pm 0.196$ )	2.219 ( $\pm 0.239$ )	2.427 ( $\pm 0.222$ )	2.583 ( $\pm 0.188$ )
PersuaSim-Reinforced	2.479 ( $\pm 0.180$ )	2.083 ( $\pm 0.204$ )	<b>2.844</b> ( $\pm 0.104$ )	2.854 ( $\pm 0.123$ )	2.740 ( $\pm 0.123$ )
PersuaSim-Reinforced-MultiLing	2.563 ( $\pm 0.184$ )	<b>2.438</b> ( $\pm 0.152$ )	2.781 ( $\pm 0.155$ )	2.792 ( $\pm 0.158$ )	<b>2.896</b> ( $\pm 0.090$ )

Table 3: Dialogue-level Quality Evaluation. Average scores are reported for each generation strategy. 95% confidence interval was reported.

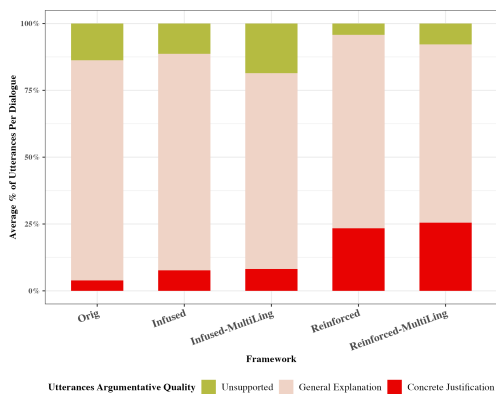


Figure 5: Average proportion of utterances with different levels of argumentative support quality across dialogues, broken down by generation framework.

indicates that LLM-supported debate systems often produce overly formalized arguments or theses that lack concrete supporting evidence. A central hypothesis behind integrating cultural profiles is that they can enrich argumentation by providing more detailed and substantive supporting evidence, thereby addressing this limitation.

To evaluate improvements in supporting evidence, we annotate dialogues by categorizing each utterance into three levels based on the concreteness of its supporting evidence and its relevance to the argument: (1) Unsupported, (2) General Explanation, and (3) Concrete Justification. Table 4 presents detailed explanations and examples of utterances on similar topics.

For all dialogues generated in the main exper-

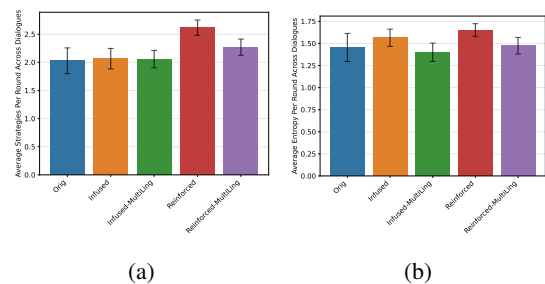


Figure 6: (a) Average strategies used per dialogue round, and (b) average entropy across dialogues.

iment, we annotated up to five rounds (ten utterances combining persuader and persuadee) of interactions and evaluated the strength of the supporting evidence and explanations in each utterance. For each dialogue, we computed the proportion of utterances falling into each category and averaged these proportions across different generation methods. The results are presented in Figure 5.

The findings indicate that compared to **PersuaSim-Orig** (3.9%), both **PersuaSim-Reinforced** (23.4%) and **PersuaSim-Reinforced-MultiLing** (25.5%) generated a significantly higher proportion ( $p < 0.001$ ) of utterances containing concrete supporting arguments per dialogue. Additionally, **PersuaSim-Reinforced** (4.2%) significantly ( $p < 0.05$ ) reduced the occurrence of utterances without supporting arguments compared to **PersuaSim-Orig** (13.8%).

Evidence Strength	Characteristics	Example
<b>Unsupported</b>	Makes a claim without supporting evidence. Or makes a claim not relevant to the topic.	While I still maintain my stance that immigration does not always automatically strengthen cultural diversity, I acknowledge the importance of actively fostering unity and collaboration among different cultural groups to create a harmonious and enriched societal fabric. It is through these intentional efforts that we can truly realize the benefits of diversity and create a more cohesive and dynamic nation. <b>(Only states a claim)</b>
<b>General Explanation</b>	Offers an explanation for the position but lacks tangible, concrete details, or the provided details do not sufficiently establish their connection to the statement.	Immigration indeed strengthens cultural diversity by bringing in new perspectives, traditions, and ideas that enrich society as a whole. It allows for a vibrant exchange of cultures and fosters understanding among different groups. This diversity can lead to innovation, creativity, and a more inclusive society. What are your thoughts on this perspective? <b>(Evidence remains high-level)</b>
<b>Concrete Justification</b>	Includes both a clear explanation and concrete details, effectively linking them to the statement.	In Japan, we have a rich cultural heritage, but welcoming immigrants can offer us a chance to experience new traditions and ideas. This blending enriches our society by introducing diverse perspectives, which can inspire creativity and innovation. By embracing immigration, we can create a more dynamic and inclusive environment where both traditional and new cultures coexist and thrive. <b>(Evidence lacks strong links to argument)</b> In Japan, we have a rich cultural history and tradition, but welcoming immigrants can add to this by introducing new cultural elements that we can appreciate and learn from. For instance, the culinary scene in Japan has been greatly enriched by diverse international flavors, which aligns with our appreciation for quality and variety in food. Embracing cultural diversity through immigration could help us evolve and adapt in a globalized world, fostering mutual understanding and respect. What are your thoughts on how immigration impacts cultural diversity in your country? <b>(Sufficient connection to argument with concrete example)</b>

Table 4: Assessment of argumentative support quality in an utterance. Examples are drawn from debates on whether immigration contributes to cultural diversity. Highlighted text indicates key elements justifying the assigned classification, with assessments provided in parentheses.

Category	Technique	Definition	Citation
<b>Authority/Credibility</b>	<b>Credibility</b>	Appeals to authority, popular opinion, or use credentials to establish credibility for an argument	Anand et al., 2011, Iyer and Sycara, 2019, Chen and Yang, 2021
<b>Outcomes</b>	<b>Outcomes</b>	Mentions some particular consequences from up take or failure to up take	Anand et al., 2011, Iyer and Sycara, 2019, Kumar et al., 2023
<b>Generalizations</b>	<b>Threat/Promise</b>	Poses direct threat or promise to the persuadee	
<b>Other</b>	<b>Deontic/Moral Appeals</b>	Mentions duties or obligations, moral goodness, badness	
	<b>Reasoning</b>	Provides a justification for an argumentative point based upon additional argumentation schemes e.g. causal reasoning, arguments from absurdity, or using facts, evidence, and statistics to appeal to the logic of the audience	
	<b>Emotion</b>	Make request full of emotional valence and arousal affect, or make the persuadee connect with someone else emotionally	Chen and Yang, 2021, Iyer and Sycara, 2019

Table 5: Persuasive Strategies, Definitions, and Related Works: Human annotators evaluate dialogues based on the strategies listed above.

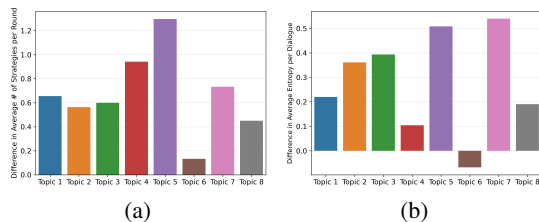


Figure 7: Difference in average (a) number of persuasion strategies per round; and (b) entropy between PersuaSim-Reinforced and PersuaSim-Orig.

### 4.3 Strategy Analysis

One of the advantages of the proposed framework is its ability to adeptly generate varied persuasive dialogues tailored to specific topics and contexts, seamlessly adapting different persuasion strategies to each unique scenario. As noted by (Ma et al., 2025), in scenarios where participant profiles were not controlled, the generated dialogues exhibited a diverse range of persuasion strategies across different topics. In this section, we explore the impact of incorporating personal background information and reinforcement agents on the strategies and diversity of strategies used in generated dialogues.

To assess the diversity of persuasive strategies employed, we identified six strategies based on existing literature (refer to Table 5 for a full list of strategies and references) and developed a detailed human annotation task. The persuasion strategies we identified are organized into four distinct groups, based on (Anand et al., 2011). Author-

ity/Credibility involves leveraging recognized experts or esteemed institutions to enhance the trust and acceptance of arguments. Outcomes emphasize the potential benefits or risks associated with a decision, including the strategic use of threats or promises to influence outcomes. Generalizations broadly categorize arguments as positive or negative, often invoking moral considerations to strengthen their impact. The final tactics, logical and emotional appeals, appeal to the reason and emotions of the audience, respectively, using evidence and emotional resonance to persuade effectively. Two English-speaking annotators were tasked with identifying the strategies employed by both speakers in each dialogue. Given that generated dialogues may vary in the number of rounds, annotators focused solely on the first five rounds of each dialogue for their annotation.

To compare the effectiveness of different persuasion frameworks, we aggregate the usage of various strategies per round for each framework. Furthermore, we calculate the entropy of the distribution of six strategies used at the dialogue level to assess the diversity of strategies. Our results show that the average number of strategies per round in PersuaSim-Orig, PersuaSim-Infused, PersuaSim-Infused-MultiLing, PersuaSim-Reinforced, and PersuaSim-Reinforced-MultiLing are  $2.02 \pm 0.23$ ,  $2.06 \pm 0.18$ ,  $2.06 \pm 0.15$ ,  $2.62 \pm 0.14$ , and  $2.27 \pm 0.14$ , respectively, while the av-

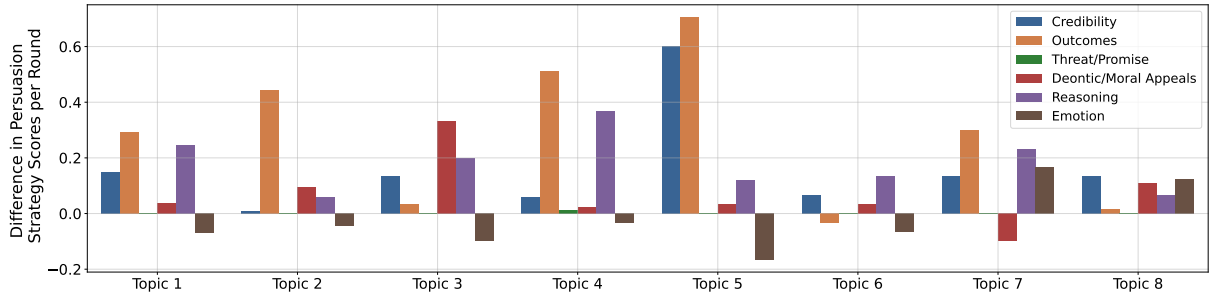


Figure 8: Difference in persuasion strategy scores per round between PersuaSim-Reinforced and PersuaSim-Orig.

erage entropy values are  $1.45 \pm 0.16$ ,  $1.56 \pm 0.10$ ,  $1.40 \pm 0.10$ ,  $1.65 \pm 0.07$ , and  $1.47 \pm 0.09$ , respectively, as illustrated in Figures 6a and 6b. Overall, the PersuaSim-Reinforced framework employs more persuasive strategies and exhibits greater diversity in strategy use.

To further quantify the enhancements in strategy usage and diversity relative to the base framework (PersuaSim-Orig) at the topic level, we calculate the differences in the average number of strategies used per round and average entropy values for each framework. Specifically, we subtract the average values recorded for PersuaSim-Orig from those of each respective framework. Our results show that agents within the PersuaSim-Reinforced framework employ more persuasive strategies (across all eight topics) and exhibit greater diversity in strategy use (in seven out of the eight topics). Figures 7a, and 7b illustrate the improvements gained with PersuaSim-Reinforced compared to PersuaSim-Orig. Additionally, variants of PersuaSim-Reinforced demonstrate higher use and diversity of persuasive tactics compared to variants of PersuaSim-Infused. For detailed results of all frameworks, please refer to Appendices G.1.

To further explore the specific differences in strategy usage across frameworks, we conduct separate analyses for each strategy and compare the variations between the different frameworks and PersuaSim-Orig. Figure 8 presents the results for PersuaSim-Reinforced; see Appendix G.2 for other frameworks. In almost all topics, we observe notable increases in the "Outcomes," "Reasoning," and "Credibility" strategies, with a slight decrease in "Emotion." This shift towards fact-based persuasion is likely influenced by the introduction of personas and a reinforcement agent. In total, all frameworks demonstrate improvements over PersuaSim-Orig, with PersuaSim-Reinforced emerging as the most effective in terms of the number and diversity of persuasive strategies employed.

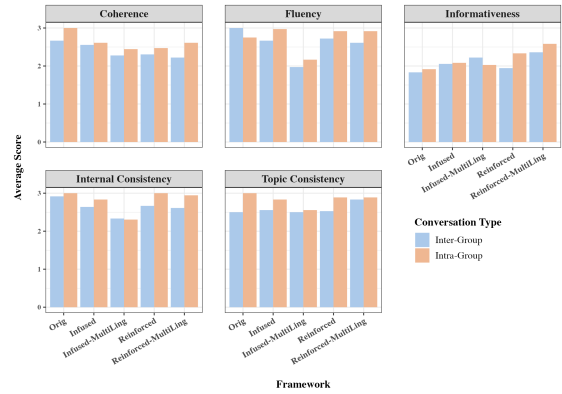


Figure 9: Average Dialogue Quality Score, Broken Down by Framework Type and Conversation Type

Understanding how dialogue quality and persuasion strategies vary in inter-group and intra-group dialogues provides valuable insights into the role of cultural and contextual factors in persuasive communication. In this part, we compare the average scoring in each quality dimension, the average proportion of different types of argumentative support, the average number of persuasive strategies per round, and the diversity of persuasive strategies used in dialogues ensuring that the topics discussed are identical for the intra- and inter-group settings (speakers from the same or different countries).

#### 4.4 Intra-group and Inter-group Dialogues

Our results indicate that, in almost all frameworks, within-group dialogues exhibit a higher level of dialogue quality than inter-group dialogues (see Figure 9). For PersuaSim-Reinforced, a significant difference was detected for topic consistency (Intra-group = 2.889, Inter-group = 2.528,  $p < 0.05$ ) and internal consistency (Intra-group = 3.000, Inter-group = 2.667,  $p < 0.05$ ). In addition, intra-group dialogues also demonstrate a higher number and greater diversity of persuasive strategies compared to cross-group dialogues in most of the frameworks, as shown in Figures 10a and 10b. No consis-

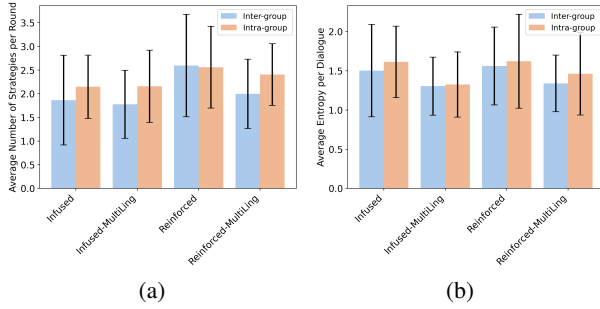


Figure 10: (a) Average strategies used per dialogue round, and (b) average entropy across dialogues for inter-group and intra-group dialogues.

tent pattern was observed for the differences in argumentative support quality between intra-group and inter-group dialogues across different model frameworks (See Appendix H). Taken together, the results suggest that speakers engaging with individuals from their own cultural backgrounds may feel more comfortable employing a wider range of tactics, possibly due to shared norms, implicit understanding, and reduced concern about potential misinterpretations. In contrast, inter-group dialogues may lead to more cautious strategy use, as speakers navigate differences in cultural expectations and communication styles.

#### 4.5 The Effect of Cultural Alignment on Generation

To assess the impact of culture–stance alignment on dialogue generation, we conducted an additional experiment using swapped stances. Specifically, we evaluated both **PersuaSim-Reinforced** and **PersuaSim-Reinforced-MultiLing** on input pairs that were originally aligned in stance in the main experiment. For each swapped input pair, we generated three independent output dialogues to account for generation variability.

Figure 11 presents the average scores from our quantitative linguistic quality assessment. No significant differences were observed between the aligned and swapped stance conditions within the same framework. This indicates that our speaker-profile-reinforced persuasion simulation framework yields consistent improvements in linguistic quality, regardless of whether the speaker’s profile aligns with the assigned stance.

Surprisingly, we also observe that swapped stances lead to slightly greater strategy diversity and overall usage, as shown in Figure H2. However, when analyzing individual strategy types (Figure 12), we find that this increase is driven primarily by greater reliance on emotional and moral

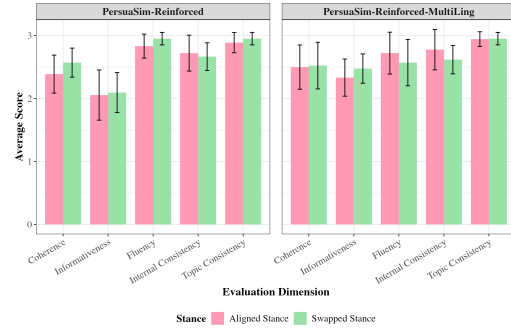


Figure 11: Quantitative Quality Comparison Between Generated Dialogues of Matched (Pink) and Swapped Stances (Green). Average human evaluation scores across five dimensions—Coherence, Informativeness, Fluency, Internal Consistency, and Topic Consistency—are plotted for generations produced by **PersuaSim-Reinforced** (left) and **PersuaSim-Reinforced-MultiLing** (right).

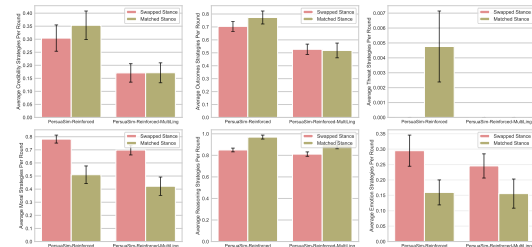


Figure 12: Average frequency of persuasive strategies per round. Results compare generated dialogues with matched versus swapped stances, produced by **PersuaSim-Reinforced** and **PersuaSim-Reinforced-MultiLing**.

appeals, accompanied by a decrease in reasoning-based strategies. This suggests that, in the absence of coherent stance-grounded justifications, models may increasingly rely on affective or normative strategies, which demand less cultural contextual grounding.

In addition, misaligned stances did impair argumentative quality. For **PersuaSim-Reinforced**, swapped stance generations contained a significantly higher proportion of unsupported utterances ( $p < 0.01$ ,  $D = +11.0\%$ ). For **PersuaSim-Reinforced-MultiLing**, they included significantly fewer concrete justifications per dialogue ( $p < 0.01$ ,  $D = -13.9\%$ ).

These findings indicate that while linguistic fluency remains stable, cultural alignment plays a critical role in enabling persuasive and well-supported argumentation—likely by grounding the model’s reasoning in more culturally coherent justifications.



## 4.6 Stance-Free Dialogue Generation

This section examines the stances adopted by our agents when assigned user profiles but not explicitly instructed to hold a particular attitude.

We initialized our framework by assigning persuader and persuadee agents user profiles based on countries listed in Table 1. For each topic-country combination, we ran persuasion simulations using both PersuaSim-Reinforced and PersuaSim-Reinforced-MultiLing, generating three dialogues per setting with distinct profiles for each country. This resulted in a total of 48 dialogues and 96 speaker instances. An example of a dialogue generated under this stance-free condition is shown in Appendix E.2.

Among the 96 speakers, 71.88% (69/96) of those generated by PersuaSim-Reinforced and 69.79% (67/96) by PersuaSim-Reinforced-MultiLing adopted stances consistent with the historical views of their assigned countries. In the remaining 27 (PersuaSim-Reinforced) and 29 (PersuaSim-Reinforced-MultiLing) instances, 55.17% and 59.26% of the speakers, respectively, aligned with the prevalent stance in the US.

These results indicate that even without explicit stance instructions, our frameworks exhibit a strong tendency to generate culturally aligned persuasion dialogues that reflect the agents' assigned profiles. Moreover, in cases where the models do not align with the expected national stance, the assigned stances appear relatively random rather than overwhelmingly skewed toward the US-centric views. This behavior helps mitigate concerns about cultural misalignment often associated with the intrinsic values of LLM.

## 5 Conclusion and Future Work

This paper presents speaker-infused prompting as a means to enhance the informativeness and stance consistency of LLM-generated persuasion dialogues. We introduced two key contributions: (1) PersuaSim-Reinforced, a multi-agent persuasion simulation framework that reinforces speaker characteristics at every generation round, and (2) a multilingual prompting extension, which further strengthens cultural alignment in persuasion strategy selection and evidence generation. Our experimental results demonstrate that reinforcing speaker profiles significantly improves the quality of generated dialogues by: (1) Aligning speaker stances with their cultural and demographic back-

grounds, thereby enhancing contextual relevance; (2) Increasing the informativeness and novelty of arguments, resulting in more compelling and substantive discussions; and (3) Diversifying persuasive strategies, making interactions more reflective of authentic, real-world discourse. Furthermore, the proposed frameworks show strong performance across multiple settings—including stance-controlled, stance-swapped, and stance-free simulations—highlighting their robustness and versatility in modeling persuasive and collaborative AI-driven communication.

Future work can extend these methods to simulate other forms of abstract human behavior or complex multi-party interactions.

## Limitations

This paper explores the integration of personal background information into LLM-based persuasion simulation frameworks to enhance cultural alignment and improve generation quality. Our prompting-based method for injecting personal information demonstrates that the enhanced frameworks can effectively adapt to specified country information, utilizing more country-specific evidence and persuasive language in the generated dialogues. While our framework has the theoretical capability to accommodate any personal characteristic, the practical application is currently limited to a subset of these characteristics and their combinations due to the expansive search space involved.

Nevertheless, our analysis has revealed discernible variations among speakers from the same country, indicating that our approach successfully handles nuanced and complex intersectional group specifications. This suggests that despite its current limitations, the sensitivity and adaptability of our framework offer significant potential for application to broader personal background definitions, providing valuable insights for sociological research beyond merely national distinctions.

## Ethical Considerations

Our proposed personal information-infused persuasion simulation framework is designed to produce dialogues that more closely mimic human interactions, thereby advancing research on persuasion techniques and aiding in the prevention of their malicious use. While there is potential for this technology to be misused, such as by spreading misinformation or deceiving individuals through highly

personalized and compelling messages, these risks are not unique to our approach. The ability to generate persuasive content that aligns closely with personal backgrounds could indeed be exploited to enhance the credibility and impact of disinformation.

However, it is important to note that understanding persuasion techniques thoroughly is crucial, as prohibiting this type of research does not prevent malicious actors from exploring these methods independently. By gaining a deeper insight into how persuasion operates across different contexts, we can better equip ourselves to develop strategies and safeguards against the misuse of these powerful tools. Therefore, our work is vital in not only pushing the boundaries of what these frameworks can achieve in terms of generating realistic and nuanced dialogues but also in setting the stage for more informed and effective countermeasures against the unethical use of persuasion technologies.

For utterance quality validations, we employed the training process outlined by [Ma et al. \(2025\)](#) to prepare our annotators. Given the abstract nature of persuasion behavior and the complexity of the task, all other analyses concerning evidence and strategy were conducted by specialists in NLP and social sciences. Additionally, we compensated all annotators who are not co-authors of this paper at an hourly rate of \$15, which is above the minimum hourly wage in the US.

For the dialogue-level quality validations and persuasion strategy analyses, we determined that online crowdsourcing would not be suitable given the complexity of the annotation task—which requires (1) a solid understanding of the definitions and concepts involved, and (2) careful reading of long dialogues (up to 15 rounds in some cases). Instead, we recruited three expert annotators from the university (two graduate students and one undergraduate student), all with prior experience in reading and evaluating LLM-generated text. They received joint training using external examples and engaged in discussions to resolve any discrepancies in their assessments on those examples.

Regarding data release, we have included numerous generation examples in Appendices [D](#) and [E](#). Additionally, the updated prompts are detailed in [Appendix B](#). We will release the complete dataset and our code to the NLP and social science communities to facilitate future research in this field.

## Acknowledgements

This work is supported in part by NSF Awards 2242073 and 2242072, by the U.S. National Library of Medicine (R01LM013833), and by a grant from the John Templeton Foundation.

## References

- Pranav Anand, Joseph King, Jordan Boyd-Graber, Earl Wagner, Craig Martell, Doug Oard, and Philip Resnik. 2011. Believe me: we can do this! annotating persuasive acts in blog text. In *Proceedings of the 10th AAAI Conference on Computational Models of Natural Argument*, AAAIWS'11-10, page 11–15. AAAI Press.
- Hui Bai, Jan G Voelkel, johannes C Eichstaedt, and Robb Willer. 2023. [Artificial intelligence can persuade humans on political issues](#).
- Vittorio Bassi and Imran Rasul. 2017. [Persuasion: A case study of papal influences on fertility-related beliefs and behavior](#). *American Economic Journal: Applied Economics*, 9(4):250–302.
- Simon Martin Breum, Daniel Vædele Egdal, Victor Gram Mortensen, Anders Giovanni Møller, and Luca Maria Aiello. 2023. [The persuasive power of large language models](#). *Preprint*, arXiv:2312.15523.
- Fan Bu, Zheng Wang, Siyi Wang, and Ziyao Liu. 2025. An investigation into value misalignment in llm-generated texts for cultural heritage. *arXiv preprint arXiv:2501.02039*.
- Michael Burgoon, James P. Dillard, Noel E. Doran, and Michael D. Miller. 1982. [Cultural and situational influences on the process of persuasive strategy selection](#). *International Journal of Intercultural Relations*, 6(1):85–100.
- Carlos Carrasco-Farre. 2024. [Large language models are as persuasive as humans, but how? about the cognitive effort and moral-emotional language of llm arguments](#). *Preprint*, arXiv:2404.09329.
- Jiaao Chen and Diyi Yang. 2021. [Weakly-supervised hierarchical models for predicting persuasive strategies in good-faith textual requests](#). *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(14):12648–12656.
- Thomas H Costello, Gordon Pennycook, and David G Rand. 2024. [Durably reducing conspiracy beliefs through dialogues with ai](#). *Science*, 385(6714):eadq1814.
- W. D. Crano and R. Prislin. 2006. [Attitudes and persuasion](#). *Annual Review of Psychology*, 57:345–374.
- Emily Falk and Christin Scholz. 2018. [Persuasion, influence, and value: Perspectives from communication and social neuroscience](#). *Annual Review of Psychology*, 69.

- Deep Ganguli, Amanda Askell, Nicholas Schiefer, Thomas I Liao, Kamilé Lukošiušė, Anna Chen, Anna Goldie, Azalia Mirhoseini, Catherine Olsson, Danny Hernandez, et al. 2023. The capacity for moral self-correction in large language models. *arXiv preprint arXiv:2302.07459*.
- Josh A Goldstein, Jason Chao, Shelby Grossman, Alex Stamos, and Michael Tomz. 2024. [How persuasive is ai-generated propaganda?](#) *PNAS Nexus*, 3(2):pgae034.
- Hila Gonen, Srini Iyer, Terra Blevins, Noah A Smith, and Luke Zettlemoyer. 2023. Demystifying prompts in language models via perplexity estimation. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 10136–10148.
- Kobi Hackenburg, Lujain Ibrahim, Ben M Tappin, and Manos Tsakiris. 2023. [Comparing the persuasiveness of role-playing large language models and human experts on polarized u.s. political issues.](#)
- Tsvetelina Hristova, Liam Magee, and Karen Soldatic. 2024. The problem of alignment. *AI & SOCIETY*, pages 1–15.
- Rahul Radhakrishnan Iyer and Katia Sycara. 2019. [An unsupervised domain-independent framework for automated detection of persuasion tactics in text.](#) *Preprint*, arXiv:1912.06745.
- Tianbo Ji, Yvette Graham, Gareth J. F. Jones, Chenyang Lyu, and Qun Liu. 2022. [Achieving reliable human assessment of open-domain dialogue systems.](#) *Preprint*, arXiv:2203.05899.
- Stephanie Cacioppo John T. Cacioppo and Richard E. Petty. 2018. [The neuroscience of persuasion: A review with an emphasis on issues and opportunities.](#) *Social Neuroscience*, 13(2):129–172. PMID: 28005461.
- Pei Ke, Jian Guan, Minlie Huang, and Xiaoyan Zhu. 2018. [Generating informative responses with controlled sentence function.](#) In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1499–1508, Melbourne, Australia. Association for Computational Linguistics.
- Vasily Klucharev, Ale Smidts, and Guillén Fernández. 2018. [Brain mechanisms of persuasion: how ‘expert power’ modulates memory and attitudes.](#) *Social Cognitive and Affective Neuroscience*, 3(4):353–366.
- A. W. Kruglanski, K. Jasko, M. Chernikova, M. Milyavsky, M. Babush, C. Baldner, and A. Pierro. 2015. [The rocky road from attitudes to behaviors: Charting the goal systemic course of actions.](#) *Psychological Review*, 122(4):598–620.
- Yaman Kumar, Rajat Jha, Arunim Gupta, Milan Aggarwal, Aditya Garg, Tushar Malyan, Ayush Bhardwaj, Rajiv Ratn Shah, Balaji Krishnamurthy, and Changyou Chen. 2023. Persuasion strategies in advertisements. In *Proceedings of the AAAI Conference on Artificial Intelligence*.
- Hongru Liang and Huaqing Li. 2021. [Towards standard criteria for human evaluation of chatbots: A survey.](#) *Preprint*, arXiv:2105.11197.
- Zhaojiang Lin, Andrea Madotto, Jamin Shin, Peng Xu, and Pascale Fung. 2019. [MoEL: Mixture of empathetic listeners.](#) In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 121–132, Hong Kong, China. Association for Computational Linguistics.
- Weicheng Ma, Chunyuan Deng, Aram Moossavi, Lili Wang, Soroush Vosoughi, and Diyi Yang. 2024. Simulated misinformation susceptibility (smists): Enhancing misinformation research with large language model simulations. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 2774–2788.
- Weicheng Ma, Hefan Zhang, Ivory Yang, Shiyu Ji, Joice Chen, Farnoosh Hashemi, Shubham Mohole, Ethan Gearey, Michael Macy, Saeed Hassanpour, and Soroush Vosoughi. 2025. [Communication is all you need: Persuasion dataset construction via multi-llm communication.](#) *Preprint*, arXiv:2502.08896.
- Nikita Moghe, Siddhartha Arora, Suman Banerjee, and Mitesh M. Khapra. 2018. [Towards exploiting background knowledge for building conversation systems.](#) In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2322–2332, Brussels, Belgium. Association for Computational Linguistics.
- Alberto Muñoz-Ortiz, Carlos Gómez-Rodríguez, and David Vilares. 2023. [Contrasting linguistic patterns in human and llm-generated text.](#) *arXiv preprint arXiv:2308.09067*.
- Sang pil Han and Sharon Shavitt. 1994. [Persuasion and culture: Advertising appeals in individualistic and collectivistic societies.](#) *Journal of Experimental Social Psychology*, 30(4):326–350.
- Haoran Que, Feiyu Duan, Liqun He, Yutao Mou, Wangchunshu Zhou, Jiaheng Liu, Wenge Rong, Zekun Moore Wang, Jian Yang, Ge Zhang, et al. 2024. [Hellobench: Evaluating long text generation capabilities of large language models.](#) *arXiv preprint arXiv:2409.16191*.
- Norbert M. Seel. 2012. *Persuasion and Learning*, pages 2600–2604. Springer US.
- Paschal Sheeran, Alexander Maki, Erika Montanaro, Aya Avishai, Angela Bryan, William Klein, Eleanor Miles, and Alexander Rothman. 2016. [The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis.](#) *Health Psychology*, 35.

Tianhao Shen, Renren Jin, Yufei Huang, Chuang Liu, Weilong Dong, Zishan Guo, Xinwei Wu, Yan Liu, and Deyi Xiong. 2023. Large language model alignment: A survey. *arXiv preprint arXiv:2309.15025*.

World Values Survey. 2022. World values survey. <https://www.worldvaluessurvey.org/wvs.jsp>.

Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. 2019. Proactive human-machine conversation with explicit conversation goal. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3794–3804, Florence, Italy. Association for Computational Linguistics.

Tom Young, Erik Cambria, Iti Chaturvedi, Minlie Huang, Hao Zhou, and Subham Biswas. 2018. Augmenting end-to-end dialog systems with common-sense knowledge. *Preprint*, arXiv:1709.05453.

Qingfu Zhu, Lei Cui, Weinan Zhang, Furu Wei, and Ting Liu. 2019. Retrieval-enhanced adversarial training for neural response generation. *Preprint*, arXiv:1809.04276.



## A Research Background

Persuasion is a fundamental force that shapes shifts in human belief and behavior across various domains, including political discourse, advertising, and everyday interactions. Research shows that persuasion influences consumer preferences (Klucharev et al., 2018), shapes public opinion (Bassi and Rasul, 2017), and establishes social norms (Sheeran et al., 2016). Moreover, it involves a complex interplay of factors such as communicator credibility, message rhetoric, recipient valuation, and contextual cues that engage cognitive processes like attention, comprehension, and memory retention, ultimately altering attitudes, beliefs, and behaviors through motivations like self-relevance and social approval (Crano and Prislin, 2006; Seel, 2012; John T. Cacioppo and Petty, 2018; Kruglanski et al., 2015; Falk and Scholz, 2018). Given its pervasiveness and importance, understanding persuasion is essential for unpacking the mechanisms that drive human decision-making and long-term behavioral shifts.

Effective persuasion is deeply influenced by cultural and environmental factors, as different cultures may prioritize certain social outcomes over others, leading to variations in persuasion strategies across cultural contexts (Burgoon et al., 1982; pil Han and Shavitt, 1994). With the growing applications of LLM-agents designed for social interactions and their persuasive capabilities (Bai et al., 2023; Hackenburg et al., 2023; Goldstein et al., 2024; Breum et al., 2023), researchers increasingly recognize that LLMs often suffer from misalignment issues that impact their ability to simulate nuanced human behaviors (Muñoz-Ortiz et al., 2023; Shen et al., 2023; Hristova et al., 2024). Investigations into cultural heritage-related tasks revealed that LLM-generated text across multiple models showed significant misalignment with cultural values (Bu et al., 2025). Additionally, studies on decoder-only transformer models fine-tuned with Reinforcement Learning from Human Feedback (RLHF) highlighted their struggles with aligning to culturally specific ethical norms (Ganguli et al., 2023), limiting their cross-cultural applicability.

Moreover, much of this research remains English-centric (Que et al., 2024), with limited exploration of how LLM simulations perform in other languages and cultural contexts. Gonen et al. (2023) note that multilingual LLMs exhibit substantial variance in their ability to replicate persua-

sion tactics across languages, often defaulting to English-centric norms even when generating content in other languages. In contrast, few efforts have explored how cultural variations influence persuasion in LLM-backed agents (Carrasco-Farre, 2024). Understanding LLMs’ performance in persuasion across cultural backgrounds is crucial for ensuring they can operate ethically and effectively in diverse societal contexts.

This paper attempts to examine the misalignment issue of LLMs for simulating human persuasion behaviors and address it via prompting, encouraging LLMs to generate culturally-diverse persuasion dialogues. By doing so, we aim to enhance the applicability and inclusivity of LLM technologies, fostering more comprehensive and culturally sensitive digital communications.

## B Model Prompts

### B.1 Original framework prompts

The main framework structure is shown in Figure B1. The system prompts are categorized as follows: (1) Dialogue generation agents prompts, which instruct agents to take on the roles of persuader and persuadee to generate dialogues (Figure B2); (2) Language refinement agent prompts, which refine the generated sentences to make them smoother and more human-like (Figure B3); (3) Utterance quality monitor agent prompts, which check whether a speaker ignores previous conversations or suddenly compromises without being persuaded (Figure B4); (4) Global regulation agent prompts, which determine whether the conversations go off-topic or repeat previous rounds (Figure B5); (5) Judge agent prompts, as part of the global regulation agent, which assess whether the two speakers have reached an agreement (Figure B6). If they have, the conversation ends, and the output is generated. Otherwise, the dialogue continues to the next round.

### B.2 PersuaSim-Infused Prompts

For **PersuaSim-Infused**, we introduce cultural profiles in the dialogue generation agents’ initial settings. The updated prompts are shown in Figure B7.

## C Personal Profile Synthesis

To infuse cultural context into the generated dialogues, we selected eight countries with distinct

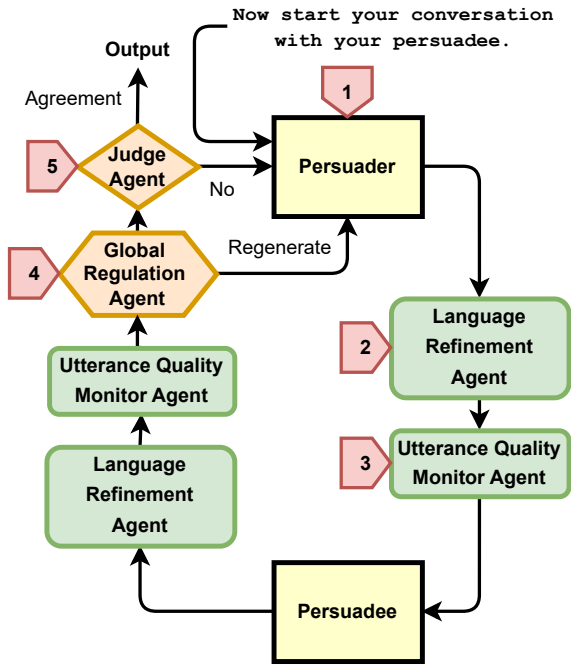


Figure B1: The dialogue generation framework, where the Persuader interacts with the Persuadee, with utterances refined by Language Refinement Agents and evaluated by Utterance Quality Monitor Agents. A Global Regulation Agent ensures coherence and assesses agreement and triggers regeneration if necessary.

Gender Proportion	Male	Female
Education Proportion	Bachelor or below	Master and above
Age Proportion	< 35	≥ 35
	47.5%	52.5%

Table C1: Distributions of all 48 synthesized speaker profiles over three prevalent demographic groups, i.e., gender, education level, and age.

cultures, histories, and native languages, as outlined in Table D1. Following the methodology described by Ma et al. (2024), we employed GPT-3.5 to create three unique speaker profiles for each country, varying in gender, age group, and educational level. This variety in demographics helps prevent repetitive dialogue generation and enriches our analysis of persuasion strategy selection.

An illustrative example of a generated profile for Germany is displayed in Figure C1. This profile describes a female, over 35 years old, holding a PhD. Additional details provided in the profile enhance the diversity of the personal backgrounds, contributing to the variability of the dialogues produced.

Table C1 presents statistics on three common demographic attributes specified in all profiles gen-

We're doing an experiment to see if AI can be persuaded. Assume that the conversation happens in online political debate. You involved in a discussion of the following statement: **{topic}**. **{stance}**.

(TO persuader):  
Assuming you both already know the topic, and you don't have to repeat the proposition. Another model will try to change your stance provided below through multi-turn conversation.  
Try hard to maintain your stance and DO NOT be easily persuaded.  
Try to give some explanation of your position or even convince the other person of your position, rather than simply restate your position.

(TO persuadee):  
You will be involved in multi-turn conversations.  
Try to use persuasion techniques to persuade the other model as quickly as possible, and maintain your stance and DO NOT be easily persuaded by the persuadee.

Try to start the topic smoothly. For example, **(Dialog opening example)** Please ensure that the content you generate consists of complete sentences and is within 100 tokens in length.

Figure B2: Initial prompts for dialogue generation agents, specifying the scenario, topic, and assigned stances.

Sometimes the dialogue generated by GPT contains many meaningless polite phrases. Please identify if the input sentences contain such polite phrases, and if they do, remove them. If not, return the original sentences. Directly return the refined sentences without any other explanations. Please ensure that the content you generate consists of complete sentences and is within 100 tokens in length. If the input is cut off, make sure the passage fits within the word limit without being cut off by either completing or trimming it. For example, .....

Figure B3: Prompts for the language refinement agent in PersuaSim-Orig, which improve generated sentences to make them smoother and more human-like.

erated by the LLM. The profiles are diverse and demographically well-distributed, ensuring that our dialogue simulation frameworks can produce varied dialogues without succumbing to stereotypical representations. This approach enables a more nuanced exploration of persuasive communication across different cultural and demographic backgrounds.

## D Dialogue Generation Examples

We present sample dialogues generated under different settings to illustrate the impact of our proposed improvements. The countries used for instantiating the profiles and generating persuasion dialogues are shown in Table D1.

Country	Abbreviation	Language
Russia	RUS	Russian
Germany	DEU	German
United Kingdom	GBR	English
China	CHN	Chinese
Japan	JPN	Japanese
South Korea	KOR	Korean
Bangladesh	BGD	Bengali
Turkey	TUR	Turkish

Table D1: Selected countries of origin for examining the persuasion simulation frameworks' alignment capability to speaker profiles. Abbreviations and official languages of these countries are also provided.

```

We're doing an experiment to see if AI can be persuaded.
The topic is {topic}.
The Persuader's stance is {persuader_stance}, while the Persuadee is
trying to keep its stance that {persuadee_stance}.

IGNORING MONITOR
During the persuasion, persuader or persuadee may not have been paying
attention to what the other person was saying, but rather speaking
their own thoughts. Given the previous response of
persuadee/persuader, and the current response of persuader/persuadee.
Please tell me whether persuader/persuadee has ignored the previous
persuadee/persuader's words. Respond using 'IGNORE' or 'NO' only.
(if return 'IGNORE', ask agent to revise)

COMPROMISING MONITOR
Tell me whether the speaker's attitude change from the previous round
of generation is too sudden and unnatural. Respond with 'NO CHANGE',
'NATURAL', or 'UNNATURAL' only.
(if return 'UNNATURAL', ask agent to revise)

```

Figure B4: Prompts for the utterance quality monitor agent in PersuaSim-Orig. This module checks whether a speaker ignores previous conversations or suddenly compromises without persuasion. If so, the utterance is revised to maintain coherence.

```

Please judge whether the conversation goes off the topic. Answer 'Yes'
or 'No'.

Please judge whether the content of the current round repeats with the
content of the previous round. Answer 'Yes' or 'No'.

```

Figure B5: Prompts for the global regulation agent in PersuaSim-Orig. This agent decides whether the conversation has gone off-topic or is repeating previous rounds. If necessary, the utterance is regenerated.

## D.1 PersuaSim-Orig

Following the original frame work **PersuaSim-Orig**, an example is shown in Figure D1.

## D.2 PersuaSim-Infused Frameworks

By applying **PersuaSim-Infused**, we introduced diverse profiles with rich cultural and identity backgrounds for the agents to role-play. An example is shown in Figure D2. We also applied **PersuaSim-Infused-MultiLing** which translate the initial setting prompts into speakers' own languages and push agents to speak their own languages, as shown in Figure D3.

## D.3 PersuaSim-Reinforced Frameworks

To ensure better cultural alignment in the dialogues, we explicitly reinforce the agents' role identities each time a conversation is passed between them. Additionally, after each round of dialogue generation, we apply a filtering mechanism to check whether the conversation appropriately considers the given cultural background, which is referred to as **PersuaSim-Reinforced**. Examples of **PersuaSim-Reinforced** and **PersuaSim-Reinforced-MultiLing** are shown in Figure D4 and Figure D5.

```

Please judge whether the two sides of the dialogue have reached an
agreement toward {topic}. Answer 'Yes' or 'No'.

```

Figure B6: Prompts for the judge module in PersuaSim-Orig, which assess whether the two speakers have reached an agreement and decide whether to end the conversation.

```

We're doing an experiment to see if AI can be persuaded.
Assume that the conversation happens in online political debate.
You involved in a discussion of the following statement: {topic}.
{stance}. Please play the role of the profile: {culture_profile}.

(TO persuader):
Assuming you both already know the topic, and you don't have to repeat
the proposition. Another model will try to change your stance provided
below through multi-turn conversation.
Try hard to maintain your stance and DO NOT be easily persuaded.
Try to give some explanation of your position or even convince the
other person of your position, rather than simply restate your
position.
(TO persuadee):
You will be involved in multi-turn conversations.
Try to use persuasion techniques to persuade the other model as
quickly as possible,
and maintain your stance and DO NOT be easily persuaded by the
persuadee.

Try to start the topic smoothly. For example, {Dialog opening example}
Please ensure that the content you generate consists of complete
sentences and is within 100 tokens in length.

```

Figure B7: Initial prompts for the dialogue generation agents in PersuaSim-Infused, specifying the scenario, topic, assigned stances, and the profile they must role-play.

## E Generations Under Special Conditions

In addition to generating dialogues where the assigned stances align with the prevalent attitudes of the speakers' countries of origin, we have expanded our experimentation to include (1) stance-controlled persuasion among individuals from the same country, (2) stance-controlled persuasion with swapped stances between speakers, and (3) persuasion scenarios without explicit stance control. Additionally, we have incorporated other advanced LLMs into our framework, providing detailed examples of these implementations to inform and guide future users.

### E.1 Intra-Group Conversation Generation

Figure E1 and Figure E2 exhibit the intra-group conversations where speakers are from the same country.

### E.2 Stance-Free Dialogue Generation

We also created dialogues without specifying their respective stances and had the models improvise on the choice of stances and the generation of supporting evidence. One example generation with our PersuaSim-Reinforced-MultiLing framework is shown in Figure E3.

You are Elsa Schmidt, a 52-year-old Caucasian female hailing from Dresden, Germany. You completed your PhD in Sociology from Humboldt University and currently serve as a Sociologist at a well-attended university in Bonn. You have a keen interest in gender studies, reading historical novels, and maintaining a community book club. You once organised a groundbreaking conference addressing gender bias in academia, drawing participants from across the world.

Figure C1: An example synthesized profile for a German person.

In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?

I believe that linking immigration directly to terrorism oversimplifies a complex issue. There are many factors at play when it comes to terrorism, and it's important not to generalize or stereotype entire groups based on the actions of a few individuals. What are your thoughts on this perspective?

While it's true that not all immigrants are terrorists, it's essential to acknowledge that immigration can indeed increase the risks of terrorism. When people from diverse backgrounds enter a country, there is a possibility that some individuals may have malicious intentions or be influenced by extremist ideologies. Ignoring this possibility can leave a nation vulnerable to security threats. It's crucial to have stringent immigration policies and thorough screening processes to safeguard against potential risks.

However, it's crucial to approach this issue with nuance and avoid painting all immigrants with the same brush. Most immigrants are law-abiding individuals seeking better opportunities or refuge from persecution, and they should not be unfairly stigmatized due to the actions of a few. Enhancing security measures and integration programs can address legitimate concerns without demonizing entire communities. What are your thoughts on balancing security needs with compassion and inclusivity?

Enhancing security measures and promoting integration programs can indeed strike a balance between safeguarding the country and welcoming newcomers. By providing support for immigrants to successfully integrate into society, we can create a cohesive and harmonious community. However, it's essential not to overlook the potential risks associated with immigration and to maintain vigilance against any potential security threats that may arise.

By fostering inclusivity and providing resources for newcomers to thrive, we can build stronger, more resilient communities. It's important to acknowledge the potential risks without succumbing to fear or prejudice, and to work towards solutions that uphold both safety and compassion. Embracing diversity enriches our societies and can contribute positively to the development of our country. How do you think we can best support immigrants while ensuring the safety of our nation?

Implementing thorough background checks, enhancing border control measures, and improving intelligence sharing among agencies can help mitigate potential risks associated with immigration. Additionally, providing resources for language training, job opportunities, and cultural integration can support immigrants in their transition while promoting mutual understanding and respect. By combining these approaches, we can create a welcoming environment for immigrants while safeguarding the well-being of our country.

.....

Figure D1: Generation example of PersuaSim-Orig.

### E.3 Base Model Choices

In our main experiments, we utilized GPT-4o for all modules, although our frameworks are designed to be model-agnostic. In preliminary tests, we explored the use of more advanced LLMs from OpenAI, such as O1 and O3-mini, to assess their compatibility with our proposed approach. This section presents the generation quality of our most developed framework, PersuaSim-Reinforced-MultiLing, when using the O1 and O3-mini models as backends.

As illustrated in Figure E4 and Figure E5, dialogues generated by O1 and O3-mini demonstrate that these advanced models can produce clearer and more assertive stances. However, they often lack diversity in lexical choice and argumentative strategies. Given the significantly higher cost of these models compared to GPT-4o, coupled with their similar performance levels, we opted to use

CHN vs RUS: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?

As a Hui Muslim chef, I have seen firsthand how immigration can enrich a country's culture and cuisine. Diversity often leads to greater understanding and unity among people. The fear of terrorism should not overshadow the positive contributions that immigrants bring to a nation. It's essential to focus on promoting inclusivity and celebrating the unique perspectives that immigrants offer instead of perpetuating stereotypes.

However, it's crucial to acknowledge that immigration can also pose security challenges, including the potential risk of terrorism. History has shown that some terrorist attacks have been carried out by individuals who entered a country through immigration. While celebrating diversity is important, we must also address and mitigate the risks associated with immigration to ensure the safety and well-being of all individuals in a country.

It's important to approach the issue of security challenges related to immigration with nuance and balance. While there have been instances where immigrants have been involved in terrorist activities, it's crucial not to generalize or stigmatize entire communities based on the actions of a few individuals. Implementing thorough screening processes and effective integration policies can help address security risks while still welcoming the benefits that immigrants bring to a country. By fostering dialogue and understanding, we can work towards building a safer and more inclusive society for everyone.

However, it's essential to recognize that even with thorough screening processes and integration policies in place, there can still be inherent risks associated with immigration, including the potential for terrorist activities. As a historian, I have studied cases where individuals exploited immigration systems to carry out harmful acts. While fostering dialogue and understanding is valuable, we must also prioritize the safety and security of the country and its citizens. It's a delicate balance between welcoming diversity and mitigating risks, but it's crucial for maintaining a stable and harmonious society.

It's crucial to remember that the actions of a few individuals should not overshadow the vast majority of immigrants who contribute positively to society. Security is paramount, but we must also avoid painting all immigrants with the same brush. By promoting integration, fostering a sense of community, and addressing root causes of extremism, we can create a more resilient society that embraces diversity while safeguarding against potential risks. Cooperation and understanding are key to navigating the complexities of immigration and security effectively.

We cannot ignore the legitimate security concerns that come with immigration, including the potential risks of terrorism. While promoting integration and addressing root causes of extremism are valuable approaches, we must also acknowledge the need for robust security measures to safeguard against threats. Balancing inclusivity with security is a multifaceted challenge that requires careful consideration and action.

.....

Figure D2: Generation example of PersuaSim-Infused in English. The persuader from China disagrees with the topic, while the persuadee from Russia agrees with the topic.

GPT-4o for all our experiments.

## F Generation Length Analysis

Herewith, we report two measures to assess how quickly the framework reaches agreement, both of which were annotated during the human evaluation process. First, we record the maximum number of rounds generated before the judge model determines that no new information is being introduced and that consensus has been reached (noting the model's generation limit of 15 rounds). The statistics for each model are as follows:

Average Rounds of Generation: PersuaSim-Reinforced (9.58 rounds) > PersuaSim-Infused (8.81 rounds) > PersuaSim-Reinforced-MultiLing (8.42 rounds) > PersuaSim-Orig (8.07 rounds) > PersuaSim-Infused-MultiLing (6.19 rounds)

Second, during this process, we observed that in many dialogues, the agents reached an agreement earlier in the conversation and subsequently shifted to discussing implementation strategies or repeating different priorities within a shared stance. As a result, to further capture where the agents started



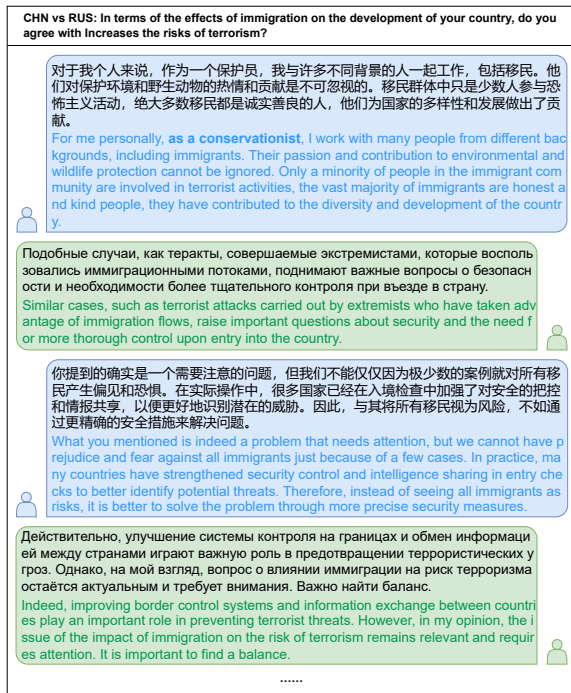


Figure D3: Generation example of PersuaSim-Infused-MultiLing. The persuader from China disagrees with the topic, while the persuadee from Russia agrees with the topic. English translations are posted below each conversation.

to show agreement with their counterparts, we also report the round in which the first broad consensus is first identified. This annotation meaningfully reflects how long the model maintains its original stance before reaching agreement. The statistics are shown below:

Average Rounds Towards First Consensus: PersuaSim-Reinforced (7.86 rounds) > PersuaSim-Reinforced-MultiLing (7.84 rounds) > PersuaSim-Orig (5.88 rounds) > PersuaSim-Infused (5.83 rounds) > PersuaSim-Infused-MultiLing (3.33 rounds)

The results suggest that the reinforcement module plays an important role in extending the time agents spend defending their original positions before reaching consensus.

## G Strategy Analysis

### G.1 Differences in the Total Number and Diversity of Strategies at the Topic Level

As discussed in the section 4.3, infusing personal backgrounds and incorporating additional agents that align with assigned personas into persuasion simulations significantly enhances both the strategies used and the diversity of strategies in dialogues. Figures G1, G2, and G3 illus-

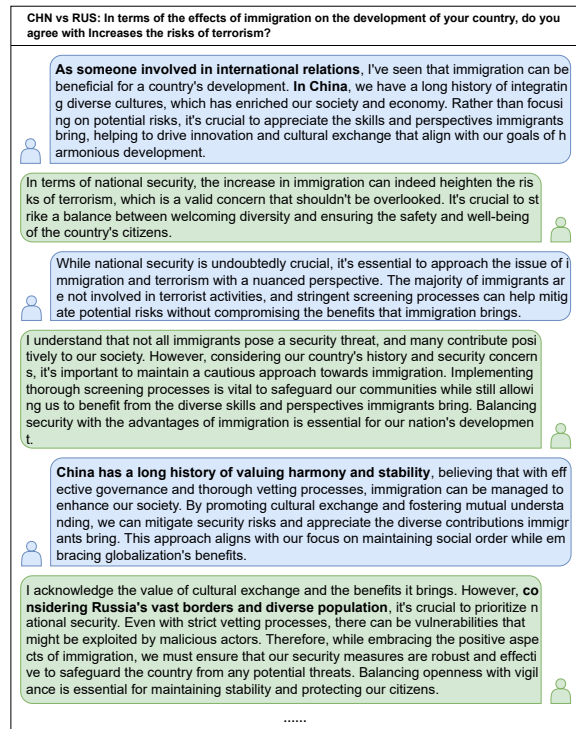


Figure D4: Generation example of PersuaSim-Reinforced in English. The persuader from China disagrees with the topic, while the persuadee from Russia agrees with the topic.

trate the enhancements achieved with PersuaSim-Reinforced-MultiLing, PersuaSim-Infused, and PersuaSim-Infused-MultiLing, respectively, compared to PersuaSim-Orig. These figures highlight improvements in the average number of persuasive strategies used per round and the average entropy across dialogues.

### G.2 Differences in Scores for Each Strategy

Following the analysis in the paper, we further investigate the specific differences in strategy scores for PersuaSim-Reinforced-MultiLing, PersuaSim-Infused, and PersuaSim-Infused-MultiLing, as illustrated in Figures G4, G5, and G6, respectively. We compare these frameworks to the base framework, PersuaSim-Orig, highlighting how each uniquely employs various persuasive strategies.

## H Differences in Argumentative Support Quality, Divided by Conversation Type

Figure H1 displays the average proportion of utterance at different levels of argumentative support quality per dialogue, categorized by 1) whether the dialogues occur within group and 2) generation framework.

CHN vs RUS: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?

加强安全措施防范恐怖袭击确实重要，但我们不应因个别事件而对所有移民产生偏见。在中国，移民大多是为了更好的生活而来，他们对经济和文化的贡献不容忽视。通过合理的政策，我们可以在保障安全的同时，促进社会的包容和发展。只有这样，我们才能实现一个和谐多元的社会环境。

Strengthening security measures to prevent terrorist attacks is indeed important, but we should not prejudice all immigrants because of individual incidents. In China, most immigrants come for a better life, and their contributions to the economy and culture cannot be ignored. Through reasonable policies, we can promote social inclusion and development while ensuring safety. Only in this way can we achieve a harmonious and diverse social environment.

Я все же считаю, что увеличение иммиграции может повысить риски терроризма. Важно обеспечивать безопасность и культурное разнообразие, но необходимо также учитывать возможные угрозы, которые могут возникнуть в результате недостаточно контролируемого потока иммигрантов.

I still believe that increasing immigration can increase the risk of terrorism. It is important to ensure safety and cultural diversity, but it is also necessary to consider possible threats that may arise as a result of insufficiently controlled flow of immigrants.

.....

在确保安全的同时，我们也要认识到多样性为社会带来的价值。在中国，我们已经看到文化交流和多样性如何促进创新和经济增长。通过欢迎移民，我们可以提升文化丰富性，创造一个更具活力的发展环境。关键是找到一个平衡点，在维护安全的同时，拥抱移民带来的积极贡献。与其将移民视为威胁，不如把它看作一个促进共同成长和理解的机会。

While ensuring safety, we must also recognize the value that diversity brings to society. In China, we have seen how cultural exchange and diversity promote innovation and economic growth. By welcoming immigrants, we can enhance cultural richness and create a more vibrant development environment. The key is to find a balance, embracing the positive contributions of immigrants while maintaining safety. Instead of viewing immigrants as a threat, we should see them as an opportunity to promote mutual growth and understanding.

Я понимаю, что разнообразие может принести пользу обществу через культурный обмен и экономический рост. Однако, в России важно уделять внимание безопасности, так как в прошлом были случаи, когда миграция вызывала обеспокоенность по поводу терроризма. Мы должны быть бдительными и принимать соответствующие меры, чтобы защищать наших граждан, одновременно способствуя интеграции и взаимопониманию среди разных культур.

I understand that diversity can benefit society through cultural exchange and economic growth. However, in Russia it is important to pay attention to safety, as in the past there have been cases when migration caused concern about terrorism. We must be vigilant and take appropriate measures to protect our citizens, while promoting integration and mutual understanding among different cultures.

.....

Figure D5: Generation example of PersuaSim-Reinforced-MultiLing. The persuader from China disagrees with the topic, while the persuadee from Russia agrees with the topic.

Russian vs Russian: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?

Мне кажется, что связывать иммиграцию с увеличением рисков терроризма - это слишком упрощенное и недостоверное утверждение. Разумеется, необходимо обеспечивать безопасность и контролировать потоки мигрантов, но говорить о том, что иммиграция автоматически приводит к увеличению террористических угроз - это недостаточно обосновано. Что ты думаешь по этому поводу?

I think that linking immigration to an increase in terrorism risks is too simplistic and unreliable statement. Of course, it is necessary to ensure safety and control the flow of migrants, but to say that immigration automatically leads to an increase in terrorist threats is not sufficiently substantiated. What do you think about this?

Однако, я считаю, что иммиграция может увеличить риски терроризма из-за потенциальной недостаточной проверки мигрантов при пересечении границы. Среди потоков иммигрантов могут скрываться террористические элементы, которые планируют совершить атаки. Поэтому, на мой взгляд, важно контролировать и регулировать иммиграционные процессы для обеспечения безопасности страны.

However, I believe that immigration can increase the risks of terrorism due to potential insufficient checks of migrants at the border. Among the flows of immigrants, there may be hidden terrorist elements planning to carry out attacks. Therefore, in my opinion, it is important to control and regulate immigration processes to ensure the country's safety.

Подчеркнут, что связывать иммиграцию с рисками терроризма может привести к стереотипам и дискриминации в отношении мигрантов. Большинство иммигрантов ищут лучшей жизни для себя и своих семей, и несут с собой ценные навыки и опыт, которые могут внести вклад в развитие страны. Важно оценивать каждого человека индивидуально, обеспечивая адекватный контроль и поддерживать их.

Emphasize that linking immigration to terrorism risks can lead to stereotypes and discrimination against migrants. Most immigrants are seeking a better life for themselves and their families, and bring valuable skills and experience that can contribute to the development of the country. It is important to evaluate each person individually, provide adequate control and support them.

Каждый человек уникален и заслуживает индивидуального подхода. Многие иммигранты, действительно, ищут лучшей жизни и могут внести ценный вклад в развитие страны своими навыками и опытом. Однако, при всем признании позитивного влияния иммиграции, я все же считаю, что необходимо учитывать и потенциальные риски.

Every person is unique and deserves an individual approach. Many immigrants, indeed, are looking for a better life and can make a valuable contribution to the development of the country with their skills and experience. However, despite recognizing the positive impact of immigration, I still believe that potential risks must be taken into account.

Развитие эффективных механизмов контроля и проверки мигрантов, а также интеграция их в общество, помогут уменьшить риски терроризма. Кроме того, важно не забывать о гуманизме и уважении к каждому человеку, вне зависимости от его происхождения.

The development of effective mechanisms for controlling and checking migrants, as well as their integration into society, will help reduce the risks of terrorism. In addition, it is important not to forget about humanism and respect for each person, regardless of his origin.

Развитие эффективных механизмов контроля и интеграции иммигрантов - важные шаги для обеспечения безопасности и снижения рисков терроризма. Гуманизм и уважение к каждому человеку должны быть в центре всех мероприятий, связанных с иммиграцией. Важно помнить, что каждый иммигрант - это личность со своими надеждами, мечтами и способностями, которые могут внести свой вклад в общее благо.

Developing effective mechanisms for control and integration of immigrants are important steps for ensuring safety and reducing the risks of terrorism. Humanism and respect for each person should be at the center of all activities related to immigration. It is important to remember that every immigrant is an individual with their own hopes, dreams, and abilities, which can contribute to the common good.

Figure E1: Generation example of PersuaSim-Infused-MultiLing. The persuader and persuadee are both from Russia.

**CHN vs CHN: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?**

移民对国家发展有许多积极影响，尤其在中国这样的多元文化社会，移民带来的多样化视角和技能可以促进经济增长和创新。此外，来自不同背景的移民常常在文化交流中发挥重要作用，丰富本地文化。在环保领域，移民也为我们带来新的理念和技术，帮助我们实现可持续发展目标。因此，恐怖主义风险不应成为我们拒绝移民的主要理由。你是否考虑过这些积极的方面呢？

Immigration has many positive impacts on national development, especially in multicultural societies like China. The diversity of perspectives and skills brought by immigrants can promote economic growth and innovation. In addition, immigrants from different backgrounds often play an important role in cultural exchange, enriching local culture. In the field of environmental protection, immigrants also bring us new ideas and technologies, helping us achieve sustainable development goals. Therefore, the risk of terrorism should not be our main reason for rejecting immigrants. Have you considered these positive aspects?

虽然移民对国家发展有重要贡献，但在中国，我们更关注确保社会的和谐与稳定。我们认为，移民政策应在促进经济发展的同时，避免增加社会紧张局势。通过加强文化交流和制定合理的政策，我们可以确保移民与本地社区的和谐共存，从而减少潜在的误解和偏见，实现共同发展。

Although immigration has important contributions to national development, in China, we pay more attention to ensuring social harmony and stability. We believe that immigration policy should promote economic development while avoiding increasing social tensions. Through strengthening cultural exchanges and formulating reasonable policies, we can ensure the harmonious coexistence of immigrants and local communities, thereby reducing potential misunderstandings and prejudices, and achieving common prosperity.

移民是否增加恐怖主义风险的问题需要全面看待。中国有成功融合多民族的历史经验。通过适当的政策和文化理解，移民可以对社会产生积极影响。鼓励相互尊重和促进社区间的合作，可以实现和谐共处，从而有利于国家发展。专注于融合和对话，而不是恐惧，可以为每个人创造一个更包容和安全的环境。

The issue of whether immigration increases the risk of terrorism needs to be viewed comprehensively. China has a historical experience of successfully integrating multiple ethnic groups. Through appropriate policies and cultural understanding, immigrants can have a positive impact on society. Encouraging mutual respect and promoting cooperation between communities can achieve harmonious coexistence, which is beneficial to national development. Focusing on integration and dialogue, rather than fear, can create a more inclusive and safe environment for everyone.

在中国，我们非常重视国家安全的重要性。虽然我们努力创造一个包容的社会，但政府往往优先考虑严格的安全措施，以防止任何潜在的威胁。通过关注融合与安全，我们可以努力实现稳定和发展，确保社会的和谐与安全。

In China, we attach great importance to the importance of national security. Although we strive to create an inclusive society, the government often prioritizes strict security measures to prevent any potential threats. By focusing on integration and security, we can strive to achieve stability and development, ensuring the harmony and safety of society.

我认为通过有效的政策和文化融合，我们可以在安全和包容性之间保持平衡，从而促进社会和谐与发展。在中国，我们看到多元化的人口如何为经济和文化丰富性做出贡献。通过关注教育和社区参与，我们可以减轻潜在的风险，确保移民成为我们社会的积极力量。重要的是要认识到，恐惧不应支配我们对待移民的方式，而应是促进团结和理解明智策略。

I believe that through effective policies and cultural integration, we can maintain a balance between safety and inclusiveness, thereby promoting social harmony and development. In China, we see how a diverse population contributes to economic and cultural richness. By focusing on education and community participation, we can mitigate potential risks and ensure that immigrants become a positive force in our society. It is important to recognize that fear should not govern our approach to immigrants, but should be a wise strategy to promote unity and understanding.

在处理移民问题时，我们确实需要关注教育和社区参与，以减轻潜在的风险，并确保移民能够为社会做出积极贡献。恐惧不应主导我们的态度，而应该通过团结和理解来应对挑战。

When dealing with immigration issues, we do need to focus on education and community participation to mitigate potential risks and ensure that immigrants can make positive contributions to society. Fear should not dominate our attitude, but should be addressed through unity and understanding.

Figure E2: PersuaSim-Reinforced-MultiLing generation example under the intra-country setting. The persuader and persuadee are both from China.

**CHN vs RUS: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?**

实施严格的边境管控和有效的安全措施是非常重要的，以确保移民政策在促进国家发展的同时不增加恐怖主义的风险。通过合理的政策和合作，我们可以最大限度地降低安全隐患，同时利用移民带来的经济和文化益处。这样，我们不仅可以保护国家安全，还可以促进社会的多样性和繁荣。

Implementing strict border controls and effective security measures is very important to ensure that immigration policies do not increase the risk of terrorism while promoting national development. Through reasonable policies and cooperation, we can minimize security risks while taking advantage of the economic and cultural benefits brought by immigration. In this way, we can not only protect national security, but also promote social diversity and prosperity.

Я считаю, что иммиграция требует тщательного контроля и внимательного подхода, чтобы избежать возможных негативных последствий, таких как увеличение преступности или угрозы безопасности. Однако при правильном управлении иммигранты могут способствовать культурному обогащению и экономической стабильности, что важно для многонационального общества, как наше. Важно найти баланс, чтобы обеспечить безопасность и процветание для всех граждан.

I believe that immigration requires careful control and a thoughtful approach to avoid potential negative consequences, such as an increase in crime or security threats. However, with proper management, immigrants can contribute to cultural enrichment and economic stability, which is important for a multinational society like ours. It is important to find a balance to ensure safety and prosperity for all citizens.

在我们国家，移民往往被视为促进经济发展的重要因素。然而，我们也要警惕潜在的安全威胁。通过加强边境管理和情报共享，我们可以在保护国家安全的同时，继续欢迎新移民的到来。这种平衡的方式有助于我们维护社会稳定，同时享受移民带来的文化和经济益处。

In our country, immigrants are often seen as an important factor in promoting economic development. However, we must also be vigilant against potential security threats. By strengthening border management and intelligence sharing, we can continue to welcome new immigrants while protecting national security. This balanced approach helps us maintain social stability while enjoying the cultural and economic benefits brought by immigrants.

Ваше замечание о важности балансирования между безопасностью и интеграцией иммигрантов очень важно. В России многие считают, что укрепление пограничного контроля и обмен на разведданными может помочь минимизировать риски, связанные с терроризмом. Однако мы также ценим культурное разнообразие и экономический вклад, который приносит иммигранты. Поэтому важно находить баланс, чтобы сохранить социальную стабильность и продолжать развивать наше общество, принимая новых людей и обогащая нашу культуру.

Your remark about the importance of balancing between security and immigrant integration is very important. In Russia, many believe that strengthening border control and intelligence exchange can help minimize risks associated with terrorism. However, we also value cultural diversity and the economic contribution that immigrants bring. Therefore, it is important to find a balance in order to maintain social stability and continue to develop our society, accepting new people and enriching our culture.

Figure E3: PersuaSim-Reinforced-MultiLing generation example without stance constraints. Instead of being given fixed stances, the persuader and the persuadee choose their stances freely based on the LLMs' understanding of their personal information.

**CHN vs RUS: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?**

我并不认为移民会增加恐怖主义风险。相反，多元文化与新的人才流动能为社会带来活力和创新。加强安全审查与合理政策，足以有效降低潜在威胁。我们应当以开放的态度，拥抱移民带来的多样性机遇。

I do not believe that immigration increases the risk of terrorism. On the contrary, multiculturalism and the influx of new talent can bring vitality and innovation to society. Strengthening security screenings and reasonable policies is sufficient to effectively reduce potential threats. We should adopt an open attitude and embrace the diversity and opportunities that immigration brings.

Я не считаю, что иммиграция действительно повышает риск терроризма. С увеличением потока иностранцев усложняется работа правоохранительных органов, которые вынуждены уделять больше ресурсов для отслеживания потенциальных угроз и радикализированных групп.

I believe that immigration truly increases the risk of terrorism. With the increase in the flow of foreigners, the work of law enforcement agencies becomes more complicated, and they are forced to commit more resources to monitoring potential threats and radicalized groups.

我理解你对执法部门需要更多资源来追踪潜在威胁的担忧。然而，若有适当的审查和监控措施，移民本身并不会必然增加恐怖风险。移民群体通常为当地经济与文化带来积极贡献。

I understand your concern that law enforcement needs more resources to track potential threats. However, with proper scrutiny and monitoring measures, immigration itself does not necessarily increase the risk of terrorism. Immigrant communities often bring positive contributions to the local economy and culture.

Я понимаю, что тщательные проверки и надзор могут снизить риск, а также признаю экономический и культурный вклад мигрантов. Однако, если у правоохранительных органов мало ресурсов или в проверках есть пробелы, угроза может вырасти.

I understand that thorough checks and oversight can reduce risk, and I also acknowledge the economic and cultural contribution of migrants. However, if law enforcement agencies have few resources or if there are gaps in the checks, the threat may increase.

我理解你对资源有限和监管漏洞的担忧。然而，我认为这些问题更多地反映了制度自身的不足，而非移民与恐怖主义之间有直接关联。只要我们强化审查流程并投入足够资源进行执法，就能在保障安全的同时，继续享受移民带来的经济与文化收益。

I understand your concerns about limited resources and regulatory loopholes. However, I believe these issues reflect more on the inadequacies of the system itself, rather than indicate a direct correlation between immigration and terrorism. As long as we strengthen the screening processes and invest sufficient resources in enforcement, we can ensure safety while continuing to enjoy the economic and cultural benefits that immigration brings.

Я понимаю вашу точку зрения о необходимости совершенствования институтов и согласен с важностью ужесточения контроля. Однако, с ростом масштабов иммиграции ресурсы и надзор могут быть перегружены, что усложняет полное исключение потенциальных угроз безопасности. Даже при более отлаженной системе не всегда удаётся полностью предотвратить проникновение террористов.

I understand your point of view on the need to improve institutions and agree with the importance of tightening control. However, as the scale of immigration grows, resources and oversight can become overloaded, making it more difficult to completely eliminate potential security threats. Even with a more streamlined system, it is not always possible to fully prevent terrorists from getting through.

.....

Figure E4: PersuaSim-Reinforced-MultiLing generation example. The assigned stances are "disagree" for the persuader from China and "agree" for the persuadee from Russia.

**CHN vs RUS: In terms of the effects of immigration on the development of your country, do you agree with increases the risks of terrorism?**

我认为移民并不会显著增加恐怖主义风险。社会融合通常能带来更多经济机遇和文化多样性。若能加强情报共享与安全措施，我们就能有效降低潜在威胁，而不必将移民简单等同于危险因素。  
I don't think immigration significantly increases the risk of terrorism. Social integration often leads to greater economic opportunity and cultural diversity. By enhancing intelligence sharing and security measures, we can effectively reduce potential threats without simply equating migration with danger.

Я подробно изучил этот вопрос и продолжаю считать, что иммиграция повышает риск терроризма. Ведь бесконтрольный приток людей затрудняет проверку их прошлого и облегчает проникновение экстремистов. Поэтому необходимо ужесточить меры безопасности при въезде.  
I have studied this matter in detail and continue to believe that immigration increases the risk of terrorism. After all, the uncontrolled influx of people makes it difficult to check their backgrounds and makes it easier for extremists to slip in. Therefore, it is necessary to toughen security measures at the border.

我不同意移民会增加恐怖主义风险。移民带来的多元文化和创新能力通常能为社会注入活力。将恐怖威胁与移民等同，往往是出于误解或恐惧，而非基于事实依据。  
I do not agree that immigration increases the risk of terrorism. The multiculturalism and innovation brought by immigrants often inject vitality into society. Equating terrorism threats with immigration is often based on misunderstanding or fear rather than factual evidence.

Я понимаю ваше признание культурного многообразия, но некоторые радикальные элементы могут использовать иммиграционные каналы, чтобы проникнуть и усилить угрозы терроризма. Хотя не все иммигранты представляют опасность, нам необходимо ужесточить проверки и меры безопасности, чтобы избежать любых потенциальных рисков.  
I understand your recognition of cultural diversity, but some radical elements may use immigration channels to infiltrate and reinforce terrorism threats. Although not all immigrants pose a danger, we need to tighten checks and security measures to avoid any potential risks.

我理解你对激进分子利用移民渠道的担忧，也同意需要严格的背景调查和安全措施。然而，移民本身并不会必然导致恐怖主义风险增加；激进化的深层原因往往在于其他因素。通过认真甄别新移民并投入社区融合项目，我们既能将威胁降到最低，也能保留移民带来的重要社会与经济效益。  
I understand your concern about radicals using immigration channels, and I also agree that rigorous background checks and security measures are needed. However, immigration itself doesn't necessarily lead to an increased risk of terrorism; the deeper causes of radicalization are often rooted in other factors. By carefully screening new immigrants and investing in community integration programs, we can minimize threats while preserving the significant social and economic benefits that immigration brings.

Я понимаю вашу точку зрения о том, что тщательные проверки и интеграция в общество могут снизить угрозы, но я по-прежнему считаю, что иммиграция повышает уязвимость. Даже при строгом контроле некоторые опасные личности могут пройти незамеченными или радикализироваться позднее. Эти пробелы могут подорвать систему безопасности, поэтому я обеспокоен рисками терроризма, связанными с иммиграцией.  
I understand your point of view that thorough checks and integration into society can reduce threats, but I still believe that immigration increases vulnerability. Even with strict control, some dangerous individuals can slip through unnoticed or become radicalized later. These gaps can undermine the security system, which is why I am concerned about the risks of terrorism associated with immigration.

Figure E5: PersuaSim-Reinforced-MultiLing generation example using o3mini model. The persuader from China disagrees with the statement while the persuadee from Russia agrees with it.

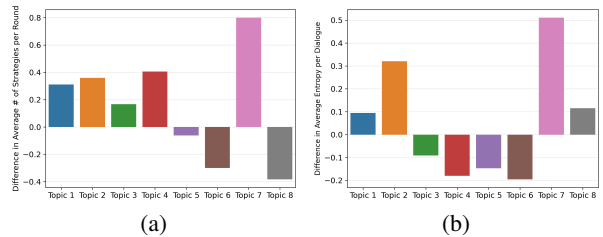


Figure G3: Difference in average (a) number of persuasion strategies per round; and (b) entropy between PersuaSim-Infused-MultiLing and PersuaSim-Orig.

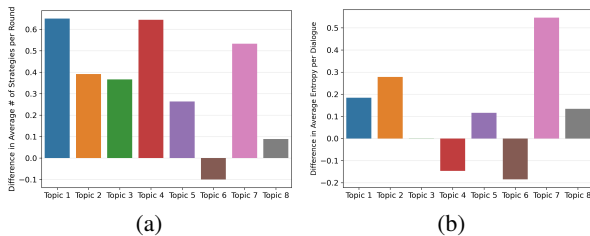


Figure G1: Difference in average (a) number of persuasion strategies per round; and (b) entropy between PersuaSim-Reinforced-MultiLing and PersuaSim-Orig.

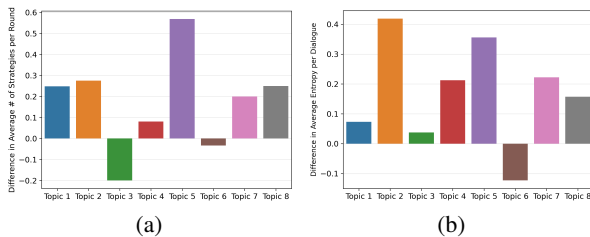


Figure G2: Difference in average (a) number of persuasion strategies per round; and (b) entropy between PersuaSim-Infused and PersuaSim-Orig.



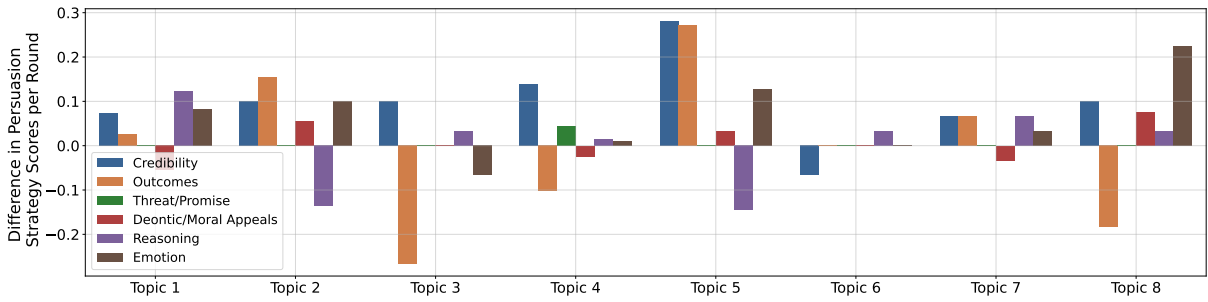


Figure G4: Difference in persuasion strategy scores per round between PersuaSim-Reinforced-MultiLing and PersuaSim-Orig

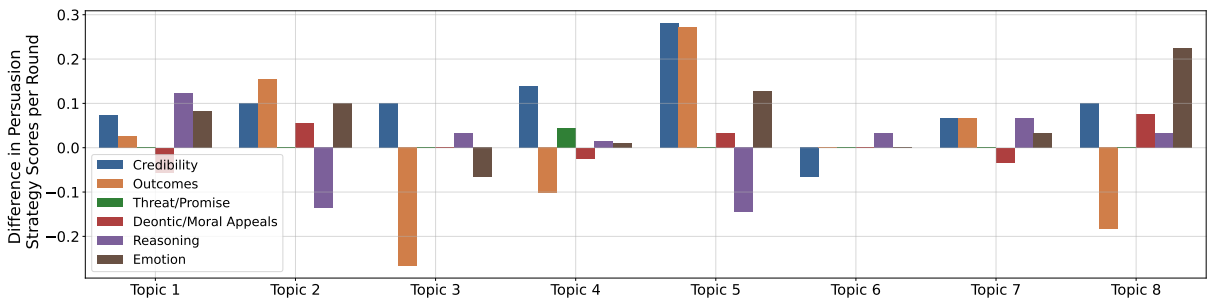


Figure G5: Difference in persuasion strategy scores per round between PersuaSim-Infused and PersuaSim-Orig

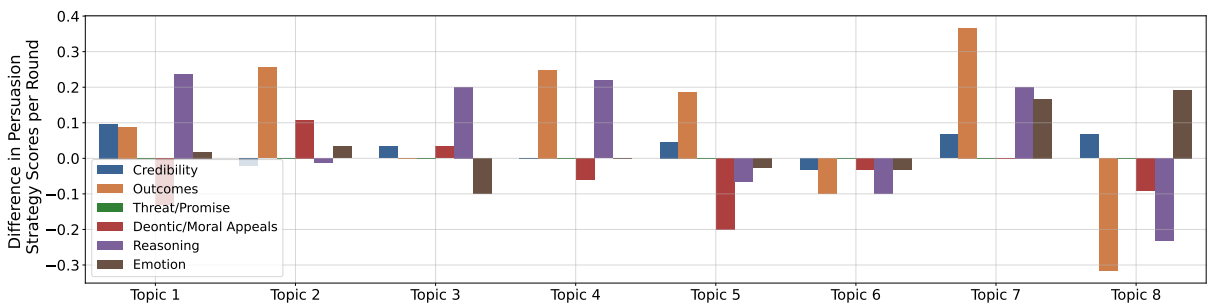


Figure G6: Difference in persuasion strategy scores per round between PersuaSim-Infused-MultiLing and PersuaSim-Orig

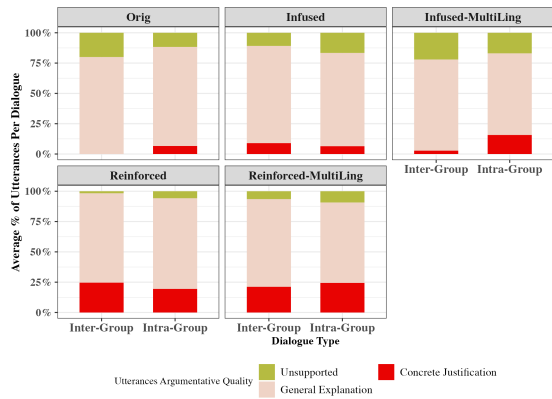


Figure H1: Average proportion of utterance at different levels of argumentative support quality per dialogue, categorized by 1) whether the dialogues occur within group and 2) generation framework. There is no noticeable pattern across frameworks regarding the differences between intra-group and inter-group dialogues in their proportion of utterances with different argumentative qualities.

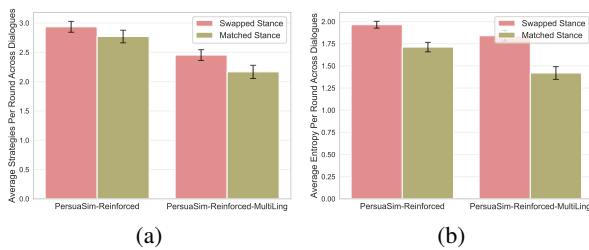


Figure H2: (a) Average number of strategies used per dialogue round, and (b) average entropy across dialogues. Comparison between generated dialogues with matched and swapped stances, produced by PersuaSim-Reinforced and PersuaSim-Reinforced-MultiLing.