

MRL 2025 Shared Task

Sewoong Lee, Ikhyun Cho, Jinu Lee
University of Illinois Urbana-Champaign
{samuel27, ihcho2, jinulee2}@illinois.edu

Korean. Korea has developed a variety of traditional physical activities, many of which were recently introduced to the world through the drama series *Squid Game*. To win these games, there are multiple strategies that can be explained in terms of physical reasoning. However, when simply relying only on the names of the games, it is also possible to generate incorrect answers that may sound plausible but are in fact completely unrealistic. Using this approach, we created 45 examples of physical reasoning tasks based on traditional Korean games.

We also included a number of questions about Korean food. Using ingredients that are commonly used in Korean and recipes that are widely considered common sense among native Korean speakers, the dataset allows us to evaluate whether an LLM can provide correct answers to questions about everyday culinary practices. Since such food-related knowledge is rarely referenced in English or other languages, we constructed 31 questions specifically focused on Korean cuisine.

Last, because of South Korea's high rate of mandatory military service, native speakers of Korean are very familiar with simple forms of physical reasoning related to the military and firearms. For example, whether one should hold their breath or keep breathing while aiming a rifle. We constructed the remaining 23 problems around such reasoning scenarios, drawing from the context of the Korean military as well as everyday experiences in schools and daily life.