The language of this dataset is Uyghur, which is a Turkic language primarily spoken by the Uyghur people in the Xinjiang Uyghur Autonomous Region of northwest China, as well as in Central Asia and diaspora communities worldwide. Uyghur is written mainly in an Arabic-derived script, though Latin and Cyrillic alphabets have also been used historically. This dataset is written in Uyghur Arabic Script.

This dataset consists of 132 multiple-choice questions, each offering two possible answers (0 or 1). The correct responses are distributed as follows: 68 questions have option 0 as the correct answer (51.6%), while 64 questions have option 1 as the correct answer (48.4%), resulting in a relatively balanced distribution. Each question and answer contains between 1 and 2 sentences, ranging from 1 to 40 words. On average, questions contain 7.7 words, and answers contain 6.0 words. All questions and answers have been proofread by five native speakers and verified using a Uyghur word-based spell checker [1].

The questions were developed manually based on the contents collected from authentic Uyghur literary materials [2-5], including cultural and traditional texts, proverbs and sayings, folklore collections, and instructional manuals. These materials were sourced from online book repositories such as uygur.com and uyghukitap.com. Approximately 65% of the questions focus on general physical reasoning and common sense, while the remaining questions address physical reasoning and common sense within the context of Uyghur culture, traditions, proverbs, and folklore.

To collect the data, a native speaker first scans a selected book to identify passages related to physical common sense, reasoning, culture, and traditions. The concluding parts of the text are used to formulate the questions, while the introductory and descriptive sections provide the basis for the correct answers. Incorrect answers are then generated by modifying the correct ones using several strategies: if the answer is related to a material, a key noun is altered (e.g., changing "wood" to "iron"); if a metric such as size or time is central to the answer, its scale or value is adjusted; and verbs are frequently modified since they are typically positioned at the end of sentences in Uyghur, as in other Turkic languages, allowing them to be easily altered without changing the overall sentence structure.

[1] Uyghur Spell Checker. 2025. *LibreOffice Extensions*. Available at: https://extensions.libreoffice.org
[2] Abdukérim Raxman, Reweydulla Hemdulla, and Shérip Xushtar. 1996. *Uyghur örp-adetiliri*. Shinjang Yash-Ösmürler Neshriyati.
[3] Memetjan Hesen, translator. 2012. *21-esirdiki yüz minglighan néme üchün: Turmush sawatliri*. Milletler Neshriyati.
[4] Muhemmet Rehim.1990. *Uyghur xelq maqal-temsilliri*. Shinjang Xelq Neshriyati.
[5] Erkin Kérimi. 2007. *Uyghur qol hönerwenchilik risalisi series*. Shinjang Xelq Neshriyati.