

Background

Rising Academies, an educational network based in Ghana, has created Rori, an AI-powered math tutor available on WhatsApp. Rori can be used at home or in schools free of charge. The Rori curriculum has one or more micro-lessons for each skill in the math Global Proficiency Framework (GPF), with over 500 micro-lessons to date. Each micro-lesson includes a brief student-friendly explanation of the skill and ten scaffolded practice questions. Many of these questions require open-ended responses, which was a decision taken for pedagogical reasons. Students are expected to write their answers into WhatsApp using the mobile keyboard. If students answer a question incorrectly, they are first shown a hint to help them solve the question and if their second attempt is unsuccessful, they are shown a worked solution. When students finish a micro-lesson, they are encouraged to continue with the next, which incrementally increases in difficulty. Rori will suggest students move either backwards or forwards in the curriculum if they find a lesson too difficult or easy. For more context you can [watch this 2-minute video](#).

Rori's curriculum is built upon the comprehensive and evidence-based GPF. The framework was developed to create uniform global standards for reading and mathematics across the world and was created by USAID by using inputs from experts representing organizations such as the World Bank, the Bill and Melinda Gates Foundation, the UK's Foreign, Commonwealth, and Development Office, the UNESCO Institute for Statistics, and many more. The GPF represents a global standard for the competencies required for learners at different stages. It covers grades 1 to 9, aligns with national standards globally, and the standards are linked across grade levels. The math framework has five domains: "Numbers and operations", "Measurement", "Geometry", "Statistics and probability", and "Algebra". Each domain is split into constructs, then subconstructs, and then in specific skills that a student in each grade should be able to demonstrate. For example, the domain "Numbers and operations" has a topic "Integers and Exponents" that has skills such as "Add and subtract" and "Multiply and divide". For a more detailed description of the structure of the curriculum see [here](#).

Structure

Each response in our dataset was scored by a pre-existing, rules-based classification model, native to Rori, which classifies answer attempts as "correct", "wrong" or "other". The latter was typically returned when a student entered something besides an answer attempt, such as a voice note or a sticker. These classifications were then manually reviewed by humans, and changed where necessary, meaning

the dataset also has a ground truth score for each student answer. The dataset is comprised of students' answers to math questions from Rori lessons from grade levels 6 to 9 in the domains "Algebra" and "Number and operations". Each student answer is paired with the corresponding question, the expected response, a ground-truth correct/incorrect score, the specific learning standard evaluated by the question, the time the student answered, and a UID number that can be used to link student responses across the dataset.

Summary Information		Example attributes of single entry	
Total Answers	53,031	lesson	G9.N5.2.1.1
Correct Answers	34,668	question_number	2
Incorrect Answers	15,278	question_text	$3^2 + 3^1 = ___$
Other Answers	3,085	expected_answer	12
Unique Students	2,508	student_response	$=6+6$ $=12$
Grade Levels Covered	6-9	model_grade	wrong
Domains Covered	Algebra, Numbers and Operations	human_grade	correct
Number of Lessons	151	time	1/9/24 7:57
Number of Skills	35	user_id	17

Figure 1 *Structure of dataset*

The dataset also includes matched but anonymized demographic data on the 2,508 users, such as when they first started using Rori, their country code, self-reported age, and number of messages they sent and active days on Rori. At-home users tend to come from Nigeria, Ghana and South Africa and are mostly between the ages of 10 and 30 and could be using their own or their family members' phones. You can access the AMMORE dataset and data dictionary [here](#).

Potential Uses of the AMMORE Rising Dataset

The dataset's structure enables various potential analyses. For example (a) investigating students' skill mastery across micro-lessons, (b) analyzing the relative difficulty of specific questions or micro-lessons

across students, or (c) exploring how the classification model's judgments compare to those of human raters.

Expanding on the first example, while there are many ways to evaluate student mastery at the micro-lesson level, for simplicity, we define mastery as an 80% correct answer rate for questions from a micro-lesson. As discussed above, a micro-lesson is a set of 10 questions of the same difficulty level focusing on a specific learning standard. We consider the responses labelled “correct” or “wrong” and discard those labelled “other” to compute the percentage of micro-lessons that students “mastered”. Using this threshold, we can determine that students “mastered” 48% of micro-lessons. To further this analysis, one could combine or “roll up” micro-lesson mastery into skill-level mastery. The dataset includes 151 different micro-lessons covering 35 different skills. For instance, if we posit that a student must master at least 75% of the micro-lessons contained within a skill to have mastered that skill, we can determine how many of the 2,508 students in the dataset have mastered each skill. With this example, 1,133 of the 2,508 students in the data set (45%) would have mastered a skill.

Also, because the same student practices skills at different grade levels, it is possible to compare student age to the grade-level of the topics they are practicing. Using the same mastery thresholds as above, we can determine that amongst the 11% of students who master at least two skills (273 students), 28% of them (76 students) master skills at multiple grade levels. One can also estimate whether students are performing at “grade-level”. Our dataset’s lessons span grades 6 to 9, with 38% of all answers at level 9, 29% at level 6, then 20% and 13% at levels 7 and 8 respectively.

Yet another approach could be to use this dataset to test different analytics approaches, such as Bayesian Knowledge Tracing (BKT), which we explore in experiment 2, or other mastery prediction models. The rich data available, including question-level responses and progression through micro-lessons over time, makes this dataset particularly suitable for such analyses. These are just a few potential uses for this novel dataset. The combination of detailed student responses, demographic information, and curriculum structure provides a unique opportunity for researchers to explore various aspects of learning analytics, from individual student progress to broader trends in mathematical skill development across grade levels.