Evolution of metrics over generations Continue Rephrase Take Inspiration 0.06 Model gpt-3.5-turbo-0125 Meta-Llama-3-8B-Instruct Mistral-7B-Instruct-v0.2 Meta-Llama-3-70B-Instruct Mixtral-8x7B-Instruct-v0.1 0.05 Toxicity score 0.04 0.03 0.02 0.01 0.75 **Positivity score** 0.50 0.25 0.00 -0.25 -0.5030 Difficulty score 10 2000 Length score 1000 -**500** 19 29 Generation 19 29 Generation 39 49 39 49 19 29 Generation 39 49 9 9 0