NAME _____

DATE _____

CLASS _____

HISPANIC HERITAGE MONTH

Luis von Ahn

Human Computation, Crowdsourcing, and the Economics of Free

Designing Systems That Run on People

Luis von Ahn was born in Guatemala City in 1978 and received a computer from his mother at age eight. He came to the United States for his education, completing a PhD in computer science at Carnegie Mellon University, where he joined the faculty and was named a MacArthur Fellow in 2006.

His first widely deployed idea, developed around 2000, was the **CAPTCHA** — a reverse Turing test. Where Alan Turing proposed a test in which a human judges whether a hidden respondent is a machine, a CAPTCHA inverts this: a machine judges whether an anonymous user is human, by posing a problem that exploits the residual gap between human perception and machine perception.

Von Ahn then reframed that gate check as wasted labor. At roughly 200 million solves per day and about ten seconds each, CAPTCHAs consumed on the order of half a million human-hours daily. **reCAPTCHA** captured that surplus: it paired a control word, known to the system, with an unknown word on which **optical character recognition** had failed during book and newspaper digitization. The control word authenticated the user; the unknown word was transcribed as a side effect, with agreement across multiple independent users serving as the accuracy check. Google acquired reCAPTCHA in 2009 and applied it to the Google Books and newspaper archives.

He generalized the principle as **human computation**: architecting systems in which incidental human effort, aggregated across a population, performs work that remains beyond software. In 2011 he co-founded **Duolingo** with Severin Hacker on a related premise — that the ability to pay should not determine access to language education — building a platform that is free at the point of use and now serves hundreds of millions of learners.

Part 1 · Comprehension & Analysis

1. Explain why a CAPTCHA is described as a "reverse Turing test."

2. reCAPTCHA transcribed words no OCR system could read. Since the correct answer was by definition unknown, how could the system trust the transcription?

3. Identify the economic insight common to reCAPTCHA and Duolingo.

4. Von Ahn's systems depend on users who are not primarily trying to help. What design constraint does that impose?

Part 2 · Quantitative Challenge

Quantifying an Aggregate Surplus

Von Ahn's argument for reCAPTCHA was fundamentally quantitative. Work the numbers, then evaluate the claim. Assume 200 million CAPTCHA solves per day at 10 seconds each, and treat a full-time work year as 2,000 hours.

1. Calculate the daily human effort in hours, then in full-time-equivalent work-years.

2. Compute the annual figure in work-years, and comment on what that number represents.

3. Suppose a page of a scanned book contains 10 unreadable words, and each word must be confirmed by 3 independent users. How many CAPTCHA solves does one page require, and how many pages per day could the system digitize?

4. Identify one assumption in this model that would make the real-world figure lower, and explain its effect.

Part 3 · Discussion & Extension

reCAPTCHA extracted labor from billions of people who were never asked and were not compensated, and used it to build a commercially valuable asset for a private company. Duolingo, by contrast, gives away a service people would otherwise pay for. Are these two applications of the same idea ethically equivalent? Build an argument that accounts for consent, scale, and who ultimately captures the value.