A Correct-Response Rate Is Not a Property of the Question but a Relation Between Question and Group

— The same question (discrimination **1.7**, difficulty **0**) is answered correctly by **0.240743** of a group of mean ability **−1** and by **0.759257** of a group of mean **+1**, a ratio of **3.153803** — what does not depend on the group is not the correct-response rate but the difficulty — [Paper 579]

Yuuki YAMAGISHI (山岸 勇輝)
Yamagishi Complexity Torus Lab, Japan
September 2026

— From thinking to syncing. —

Abstract

The difficulty of an exam question is often stated by its correct-response rate, as in "a hard question answered by 30 per cent".What this paper shows is that a correct-response rate is not a property of the question alone but a relation to the group taking it, and that for two questions of different discrimination even the order of difficulty reverses from group to group.No new theorem or law is claimed. First, the same question's correct-response rate differs 3.153803 times between groups——a question of discrimination 1.7 and difficulty 0 is answered by 0.240743 of a group of mean ability −1, 0.500000 of mean 0 and 0.759257 of mean +1 (Section 2). **Second, the construction was checked two ways——when the mean equals the difficulty the rate is exactly 0.5, the rates at ±1 sum exactly to 1, and an independent approximation gives 0.238035 (a difference of 0.002708)** (Section 2). **Third, and this is the core.For a highly discriminating question A and a weakly discriminating question B, B is easier for groups of low ability and A for groups of high ability——the order reverses at a group mean of 0.327653, where both rates meet at 0.538597** (Section 3). **Fourth, the gap is large in low groups——in a group of mean −1.5, A gives 0.101094 and B 0.329708; in a group of mean +1.5, A gives 0.836491 and B 0.670292** (Section 3). **Fifth, what does not depend on the group is the difficulty, not the correct-response rate——questions of equal discrimination keep their order in every group** (Section 4). **Sixth, the separator is whether the two questions discriminate equally——if not, their curves cross somewhere, and where the group sits decides the order** (Section 5). **Keywords**: item response theory, proportion correct, item difficulty, logistic model, test theory.

Scope of this paper (scope note): **No new theorem or law is claimed——the two-parameter logistic model of item response theory is standard.No guessing parameter is included——answering correctly by guessing among options is the problem Paper 553 dealt with, and it is not included here.Ability is normally distributed——the group's ability is normal with standard deviation 1, and only its mean is moved; a different shape would change the numbers.The item parameters are representative——the first question has discrimination 1.7 and difficulty 0; the two compared have (2.0, 0.2) and (0.5, 0).No parameters are estimated from data——this paper looks only at the form of the model, not at numbers from real exams.No exam is designed——which questions to use is not stated.Relation to earlier papers**: **Paper 553 showed that in a triangle test guessing is right one time in three, so a correct-response rate of 0.5 means a discriminating fraction of 0.25——the same question of what a rate means; there it changes with the format of the test, here with the group taking it.Paper 470 showed that a test ninety-nine per cent accurate can be wrong nine times in ten when prevalence is low——the shared shape is that a test's measure changes with the group taking it; this paper adds the reversal of order between two questions.Paper 300 showed that sharing a root is decidable——"the rate changes with the group" and "the order reverses" share one root.**What is added is **confirming by symmetry and by an independent approximation that one question's rate changes 3.153803 times with the group mean**, **finding that the order of two questions of different discrimination reverses at a group mean of 0.327653**, and **showing that equal discrimination prevents the reversal, and placing the separator there**.

1. Notation and the Question

Setting: **the probability that a person of ability θ answers correctly is written with the two-parameter logistic**

$$P(\theta) = \frac{1}{1 + e^{-a(\theta - b)}},$$

where a is the discrimination and b the difficulty (the ability at which exactly half answer correctly).If a group's ability is normal with mean μ and standard deviation 1, the group's correct-response rate is

$$p(\mu) = \int P(\theta)\, \phi(\theta - \mu)\, d\theta.$$

Observation 1.1

the difficulty of a question is often stated by its correct-response rate."A hard question answered by 30 per cent" sounds as though the difficulty belonged to the question itself——yet the rate may be a number that cannot be fixed without saying who answered.

Question: **is a correct-response rate a property of the question?Does which of two questions is harder stay the same whoever takes them?**

2. One Question, Different Groups

Computation 2.1: a question of discrimination 1.7 and difficulty 0 taken by five groups of different mean ability.

group mean ability	correct-response rate
-1.0	$\mathbf{0.240743}$
-0.5	$\mathbf{0.361958}$
0.0	$\mathbf{0.500000}$
$+0.5$	$\mathbf{0.638042}$
$+1.0$	$\mathbf{0.759257}$

> **Observation 2.2**
>
> one question's correct-response rate differs $\mathbf{3.153803}$ times between groups.A "hard question" solved by one in four in a group of mean -1 is solved by three in four in a group of mean $+1$——the question has not changed at all.

> **Observation 2.3**
>
> the construction was checked two ways.When the mean equals the difficulty the rate is exactly 0.5, and the rates at means ± 1 sum exactly to 1 (both by symmetry, to 10^{-10})——and the independent approximation for a logistic averaged over a normal, $\operatorname{logistic}(a\mu/\sqrt{1+\pi a^2/8})$, gives $\mathbf{0.238035}$ at mean -1, differing from the numerical integral by 0.002708.

3. The Order Reverses

Computation 3.1: a highly discriminating question A ($a = 2.0$, $b = 0.2$) and a weakly discriminating question B ($a = 0.5$, $b = 0$), taken by the same five groups.

group mean	question A	question B	easier
-1.5	**0.101094**	**0.329708**	B
-0.5	**0.297934**	**0.441248**	B
0.0	**0.439671**	**0.500000**	B
$+0.5$	**0.590045**	**0.558752**	A
$+1.5$	**0.836491**	**0.670292**	A

Computation 3.2: **the group mean at which the two rates are equal is $\mathbf{0.327653}$, where both are $\mathbf{0.538597}$ (to 10^{-8}).**

> **Observation 3.3**
>
> **this is the core.For groups of low ability the weakly discriminating B is easier; for groups of high ability the highly discriminating A is easier——which of the same two questions is harder cannot be answered without naming the group. Ranking questions by correct-response rate reflects where the group taking them sits.**

> **Observation 3.4**
>
> **for a single person, the two curves cross at $\theta^* = 4/15$.The group boundary of 0.327653 departs from it because a group's ability is spread out——over that spread the steep and the shallow curve pick up different weights.**

4. What Does Not Depend on the Group

> **Observation 4.1**
>
> **what does not depend on the group is the difficulty b.b is "the ability at which exactly half answer correctly" and belongs to the question whoever takes it——the correct-response rate is a relational quantity set by question and group together; the difficulty is a quantity of the question.**

Observation 4.2

questions of equal discrimination keep their order.With equal discrimination, the curve of the easier question lies above at every ability, so averaging over any group preserves the order——a reversal happens only when discriminations differ and the curves cross.

5. The Separator

Observation 5.1

the separator is whether the two questions discriminate equally.If they do, the order of their rates is the same in every group——if not, their curves cross somewhere, and whether a group sits to the left or right of it decides the order.

Observation 5.2

the question of Paper 553, with a different reason.There, guessing was right one time in three in a triangle test, so a rate of 0.5 meant a discriminating fraction of 0.25——what a rate means changed with the format of the test there, and changes with the group here.

Observation 5.3

the shape of Paper 470.There, a test ninety-nine per cent accurate was wrong nine times in ten at low prevalence——a test's measure is a relation to the group it is applied to. This paper adds that the order of two questions can reverse.

Observation 5.4

one root, under Paper 300.“The rate changes with the group” and “the order reverses” come from one root, the way a question's curve answers the distribution of ability in a group——the second appears when the slope of that answer, the discrimination, differs from question to question.

Observation 5.5

what is not said.No exam is designed, and which question to use is not stated——what is shown is that stating difficulty by correct-response rate silently carries who took the test.

6. Roll Call and Closing

1. **one question from 0.240743 to 0.759257** (Computation 2.1): **changing with the group mean.**
2. **a ratio of 3.153803** (Observation 2.2): **the question has not changed.**
3. **exactly 0.500000 when the mean equals the difficulty** (Observation 2.3): **a check by symmetry.**
4. **0.238035 by an independent approximation** (Observation 2.3): **a difference of 0.002708.**
5. **the order of A and B reverses** (Computation 3.1): **the core.**
6. **boundary group mean 0.327653** (Computation 3.2): **rates meet at 0.538597.**
7. **in a low group, A 0.101094 and B 0.329708** (Computation 3.1): **the shallow question is easier.**
8. **what does not depend on the group is the difficulty** (Observation 4.1): **the rate is relational.**
9. **equal discrimination, no reversal** (Observation 4.2): **the curves do not cross.**
10. **the separator is equal discrimination** (Observation 5.1): **if they cross, the group's place decides.**

Closing: **The difficulty of an exam question is often stated by its correct-response rate.Yet a correct-response rate is not a property of the question alone but a relation to the group taking it——a question of discrimination 1.7 and difficulty 0 is answered by 0.240743 of a group of mean -1 and 0.759257 of a group of mean $+1$, a ratio of 3.153803. When the mean equals the difficulty the rate is exactly 0.5, the rates at ± 1 sum exactly to 1, and an independent approximation returns a close 0.238035.For two questions of different discrimination, even which is harder reverses——for low groups the weakly discriminating B is easier (at mean -1.5, A 0.101094, B 0.329708), for high groups the highly discriminating A (at $+1.5$, A 0.836491, B 0.670292). The order reverses at a group mean of 0.327653, where both rates meet at 0.538597, away from the single-person crossing at $4/15$ because a group is spread out. What does not depend on the group is the difficulty, not the rate, and questions of equal discrimination never reverse.The separator is whether the two questions discriminate equally——if not, their curves cross, and whether the group sits left or right of that point decides the order. Under Paper 300, the dependence on the group and the reversal share one root.Placed among the earlier papers——like Paper 553 this asks what a rate means; there the format of the test changed it, here the group does. Like Paper 470, a test's measure is a relation to the group it is applied to.To be honest——only the two-parameter logistic was used, with no guessing parameter; the group's ability was normal with standard deviation 1 and the item parameters representative, and no real exam data**

were used. And no exam is designed: what is shown is that stating difficulty by correct-response rate silently carries who took the test.

References

1. Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Sage.
2. Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Lawrence Erlbaum.
3. Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research.
4. Yamagishi, Y. (2026). *The Three Hundredth Census: Whether Two Things Share a Root Is Decidable* [Paper 300]. Zenodo. 10.5281/zenodo.22171914
5. Yamagishi, Y. (2026). *A Test That Is Ninety-Nine Per Cent Accurate Can Be Wrong Nine Times in Ten* [Paper 470]. Zenodo. 10.5281/zenodo.22733434
6. Yamagishi, Y. (2026). *Half of Them Answered Correctly, and One in Four Could Tell* [Paper 553]. Zenodo. 10.5281/zenodo.23014198

— From thinking to syncing. —

Yuuki YAMAGISHI (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)