

# Information Inequality: the Gender of Knowledge

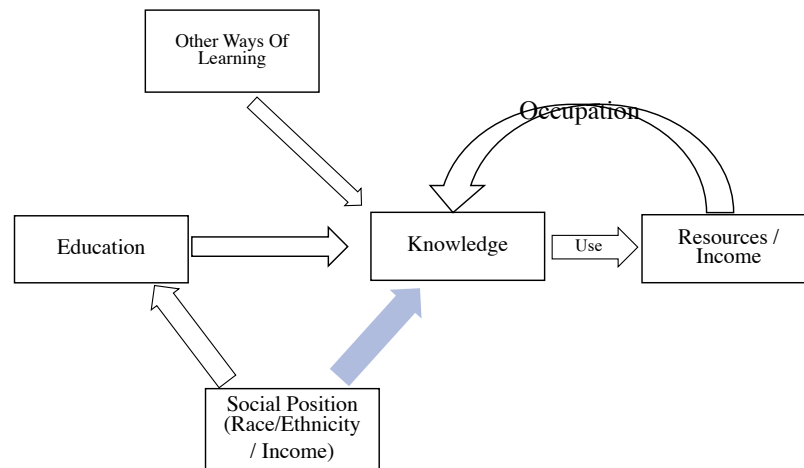
Molly King | Presentation Notes

- I argue that a concept I am calling knowledge inequality is important as both cause and outcome of social inequality.

– knowledge inequality = inequality in the possession and/or usage of information

## ***Why care about knowledge inequality?***

1. Differences in knowledge capacity itself are, by definition, a dimension of ‘inequality’;
  - (a) Here I want to note that I do NOT mean to equate knowledge or information capacity with intelligence. Knowledge of facts is only one aspect of knowledge, after all, and IQ measures something more akin to analytic abilities.
2. Knowledge is a potential cause of later inequality in outcomes and access to resources.
  - (a) Information discrepancies enable differential access to resources and institutional positions, thereby causing later inequality as well.
3. Differences in the amount of knowledge people have are influenced by unequal social positions in our society; and
  - (a) In this research I focus on the idea that social status causes knowledge inequality.
  - (b) Specifically, I focus on the impact of income, race, and ethnicity on knowledge inequality.



## Methods

- It is surprising how little we know about the knowledge stock of Americans.
- So I wanted to perform a wide scan analysis of knowledge inequality, looking at who has and does not have knowledge in different domains, and how those inequalities might compare to each other.
- So I set out to answer the research question:

*Research Question: How does the status gap in knowledge vary by domain?*

- So I wanted to ask, How does the status gap in knowledge vary by domain?
- And here I mean status in the sense of status characteristics, like gender, class, and race.
- People have studied knowledge gaps in many of these domains before. What is unique to my study is gathering these domains all together in one comparative framework, allowing us to look at the structured acquisition of knowledge.
- My data include 48 nationally representative data sets from between the years 2005 and 2015, each including at least one knowledge question.
- With such limited knowledge questions in any one of these surveys, one of the things I think is interesting to consider is: what or whose knowledge is selected for inclusion. Knowledge is linked to power, and which knowledge items are included in surveys reflects what knowledge is valued in our society.
- Anyway, I compiled these data from places like main public opinion survey repositories, the General Social Survey and Pew Research Center.
- A question was included if it asked respondents about factual knowledge - a question with a generally agreed-upon answer
- These are true/false or multiple-choice questions that asked things like:
  - — “True or false: A laser is a concentrated soundwave. The answer is false - lasers are concentrated light waves.”
  - — “Who is the vice president?”
- For each question, I mark for each individual whether they got the question correct or incorrect.
- I curated these data and categorized them by domain.
- This resulted in 16 topical domains, which we will see soon in the results.

## Model

- For each question, I then use logistic regression to predict the probability that an individual will get the question correct.
- I regressed the outcome of correct answer on the independent variables income, gender, race and ethnicity categories, and controls education categories, and age and age squared.

| Factors          | Outcome      |
|------------------|--------------|
| Gender           | Probability  |
| Income           | that you get |
| Race / Ethnicity | the question |
| Age + $age^2$    | correct      |
| (Education)      |              |

## Results

Turning over the page, we will now move on to discussing the results.

**How likely are the wealthy to get each answer correct compared to men?**



Figure 1: Income correlates with mean knowledge advantage in 13/14 domains

- So here we see that each open circle represents an individual question. I have taken about 450 questions and divided them into these 16 domains across the x-axis.
- The yellow square is the mean knowledge level within each domain, and the confidence interval represents 95% simulated certainty around that mean for each domain.
- If there were no difference in the average difference between incomes, we would see the confidence interval bar cross the 0 line.
- Here I tested whether income had a significant effect on knowledge within each entire domain. For each domain, the simulated mean and confidence intervals allow us to see whether there is a significant difference of income on knowledge and the direction of that difference.
- Here we see this model using income to predict the likelihood that respondents get the factual knowledge question correct, controlling for all other factors. This is essentially the income elasticity of knowledge.
- While there is no average gender difference in only 1 of 14 domains, men have greater average knowledge in 9 domains.
- As a reminder, this is after controlling for education - making this a conservative test for the effect of income (since in effect we are controlling twice for class).
- So in this slide, I tested whether the effect of income on the average knowledge in the whole domain was significant. For each domain, we decide whether there is a significant effect of income, and then see that that effect largely favors those with higher incomes having more knowledge.

## What proportion of questions are the wealthy more likely to answer correctly?

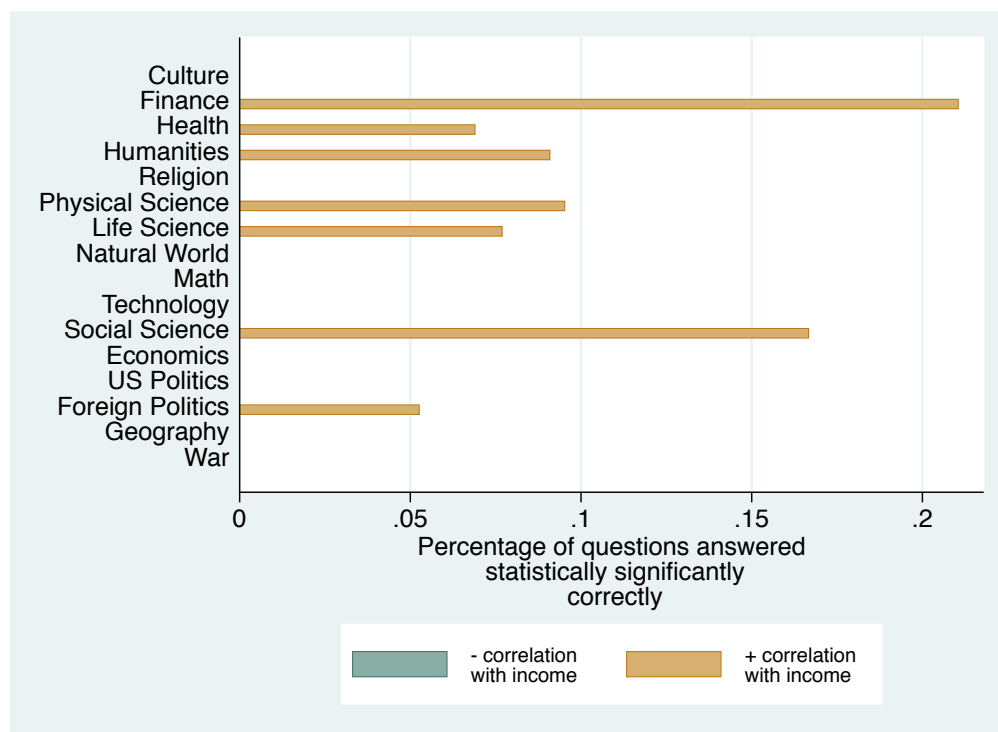


Figure 2: Those with higher incomes answer a greater proportion of questions correctly

- But this first figure doesn't tell us how many of these questions show knowledge differences that are statistically significant.
- So for each of these 450 questions, I tested whether the gender difference in knowledge was significant.
- For example, in the domain of culture, we saw that the individual coefficients clustered fairly evenly between the 0 and 0.5 log odds line.
- However, in this case, we see that none of those coefficients are statistically significant.
- We also see that a higher income predicts answering a greater proportion of questions correctly in the domains of finance, health, humanities, the physical, life, and social sciences, and foreign politics.

What proportion of questions are Blacks and whites more likely to answer correctly?

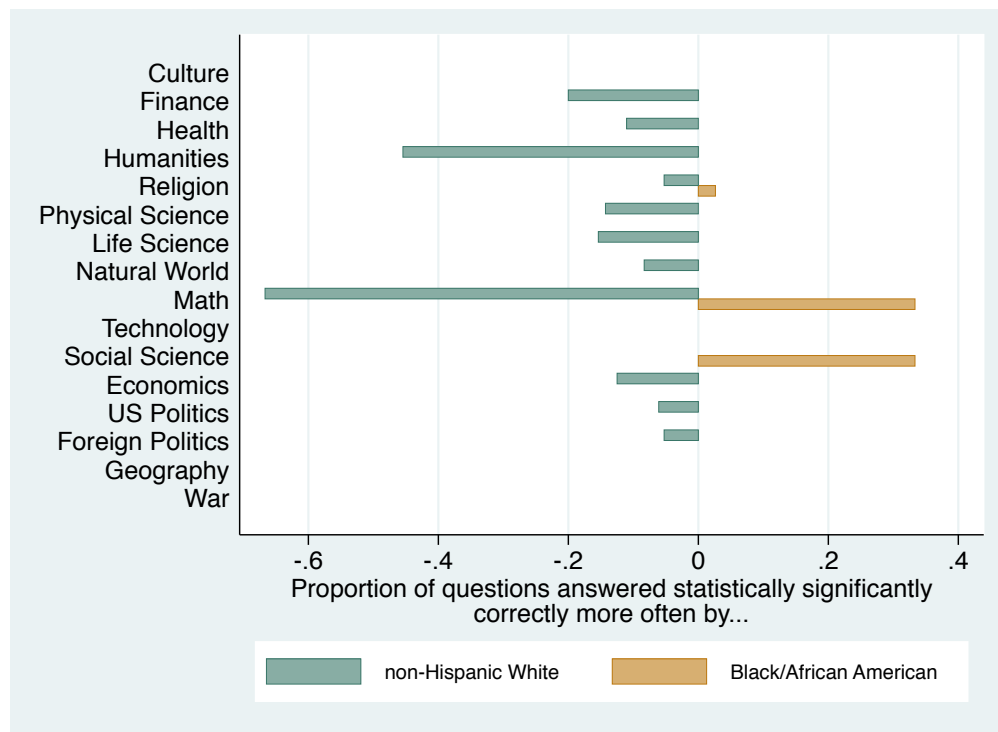


Figure 3: Proportion of questions answered by non-Hispanic Whites vs Blacks

## Relevance

- Again, which knowledge is deemed important to know, and the acquisition of that knowledge, is all shaped by social structures, institutional contexts, social roles, and status positions. Here i've begun to explore inequalities in those knowledge items that are deemed important to measure by groups in power, an important first step to descriptively understanding inequalities in the knowledge landscape.