

# Ch. 19: Controversy in Testing

## Controversy in Testing

- Historical viewpoints / Gender and IQ
- Race, Ethnicity, Genetics
- IQ testing and Ethnicity
- Eugenics & Immigration Law
- Test Bias
- Test Fairness and the Law
- Test Selection Philosophy
- *The Bell Curve / Critiques*
- The Flynn Effect
- Twin and Adoption Studies

Psychology 402 - Fall 2018 - Dr. Michael Diehr

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Gender and IQ

- In the 1800s, commonly accepted that men were intellectually superior to women
- Darwin, Descent of Man (1871) "The chief distinction in the intellectual powers of the two sexes is shewn by man's attaining to a higher eminence, in whatever he takes up, than can woman - whether requiring deep thought, reason, or imagination, or merely the use of the senses and hands"
- Book was edited by Darwin's daughter Henrietta and wife Emma.
- Darwin was in other ways socially liberal

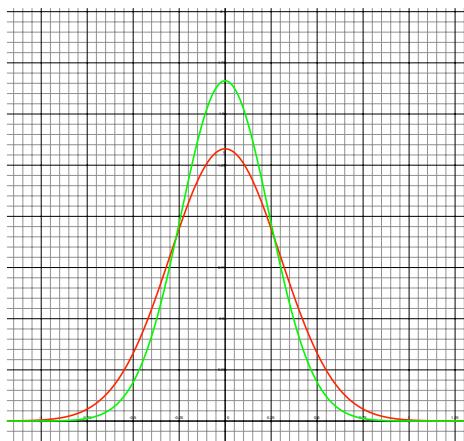
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Gender and IQ

- Modern view: men and women on average have equal IQ scores.
- Differences? Yes but very small (under 3 IQ points if any)
- Other findings:
  - Men's IQ more variable (higher variance)
  - Males better at stereotypical "male" tasks (visuospatial skills) whereas women better at "female" tasks (language). Why?

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Men vs. Women IQ



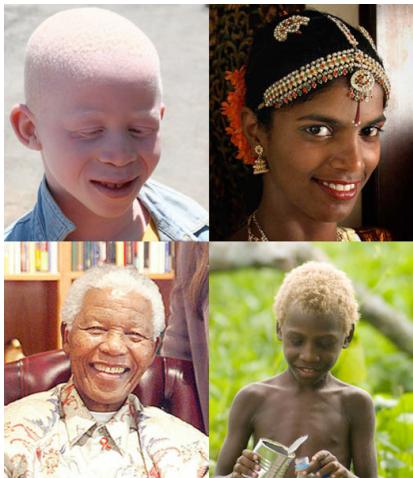
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Nonconcordant traits

- Naive view : Race = Genetics or heritage
- Biochemical view : traits & genes are spread out among groups. Group difference occur, but often the differences have fuzzy edges
- Non concordance : visible traits (skin color, eye shape, hair texture, etc.) don't go together
- Visible vs. Invisible differences: differences on the genetic level often don't track what is seen in surface differences (such as skin color)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Nonconcordant traits



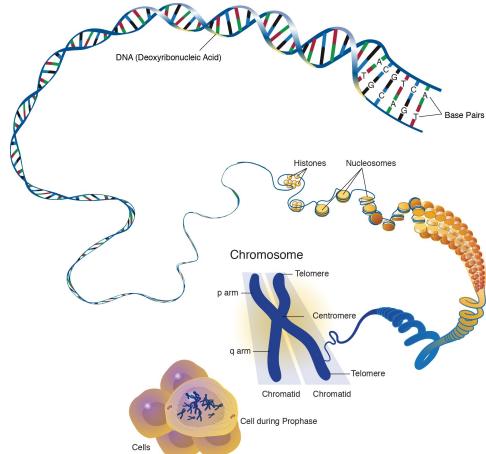
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Pre-DNA views

- Gold, Silver, Brass, Iron -- Plato
- “There is a physical difference between the white and black races which I believe will forever forbid the two races living together on terms of social and political equality.” -- Abraham Lincoln

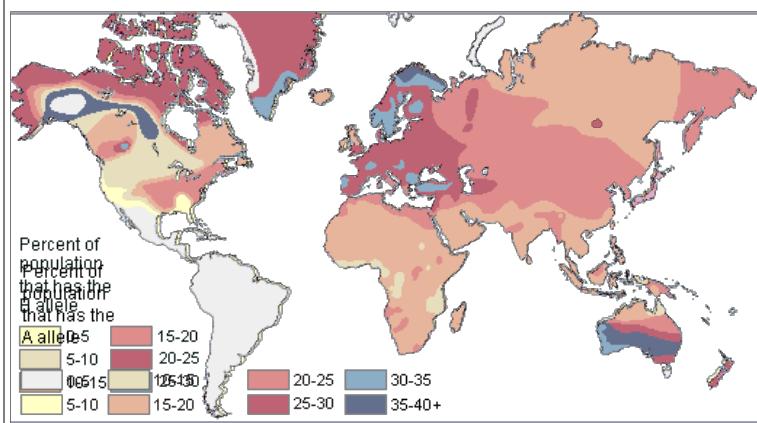
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Genetics : DNA



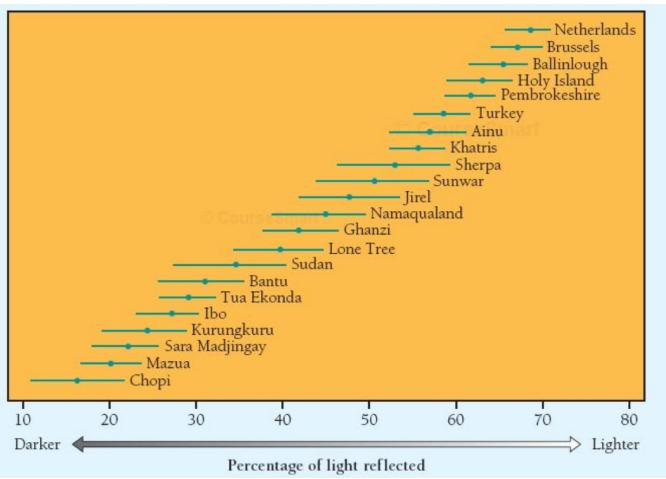
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Worldwide Blood Type



Psychology 402 - Fall 2018 - Dr. Michael Diehr

## “Black” vs. “White”



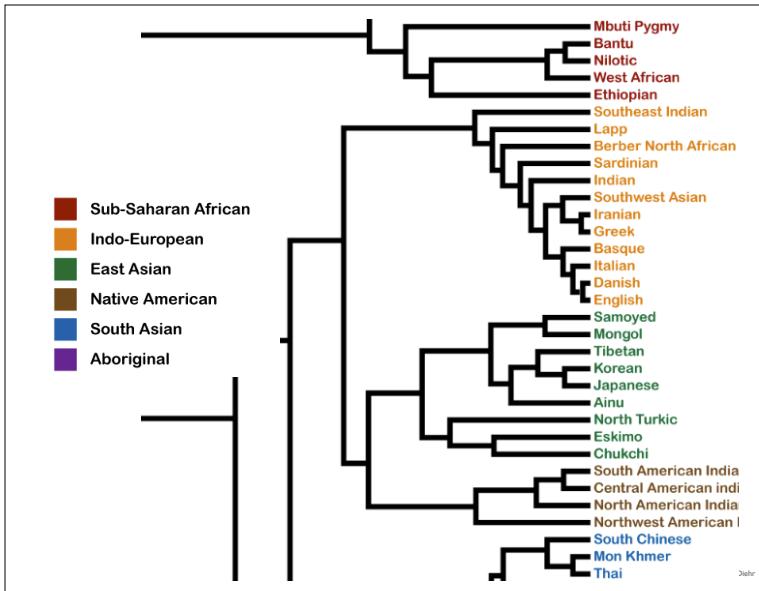
## Visible differences?

Indigenous  
Australian  
Melanesia  
African  
European

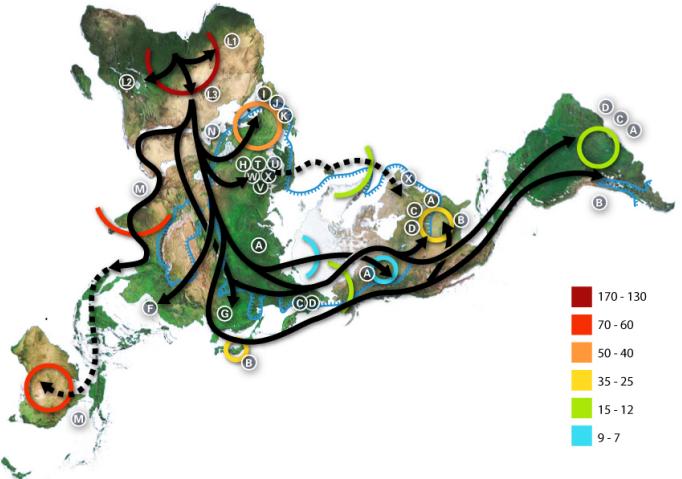
Australian and  
Africans are  
most genetically  
different



Psychology 402 - Fall 2018 - Dr. Michael Diehr



## Prehistorical Migration

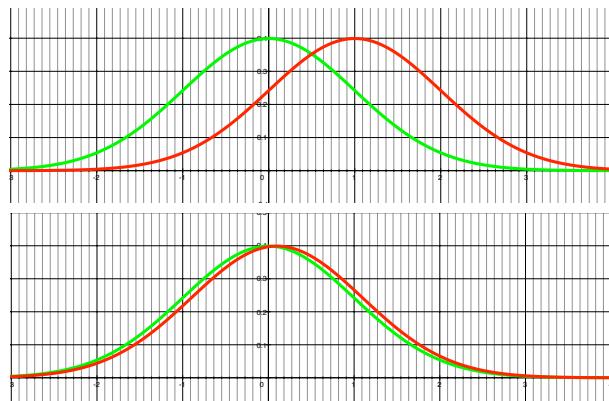


## Post-DNA views

- Variance
  - variation between individuals
    - aka variation *within groups*
    - variation *between groups*
- Variance
  - variation between individuals : 3mbp / person
  - variation within groups : 85%
  - variation between groups: 15%
    - about 5% - within “races”
    - about 10% - between races

Psychology 402 - Fall 2018 - Dr. Michael Diehr

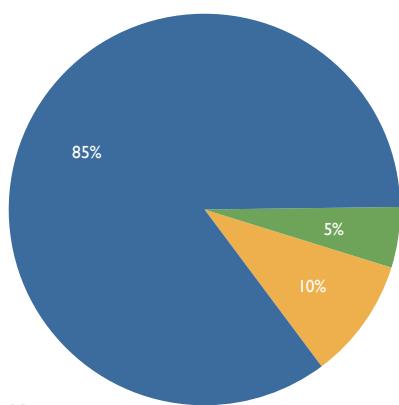
## Between Group vs. Within Group Variance



Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Variance: Genetic Variation

- Within local populations
- Within “race”
- Between “race”



For example:

- 85% within Japanese
- 5% between Japanese & Korean
- 10% between Asian and Caucasian

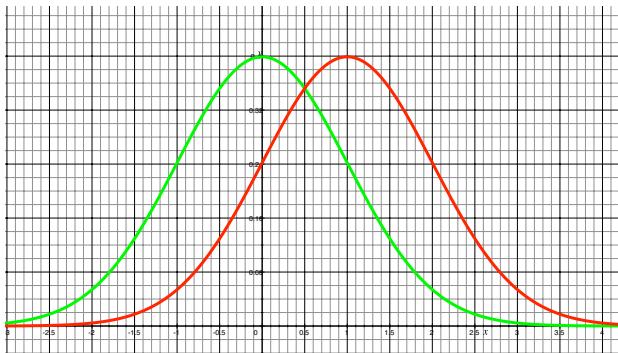
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Ethnicity and IQ

- Asian-American students perform better on Mathematics, but IQ test scores about average (same as majority / White Americans)
- Hispanic Americans & Native Americans -- do well on Performance & Spatial tests, less well on Verbal tests. Overall performance somewhere between White & Black
- African Americans were thought to score about 1 SD below the mean (e.g. 85). Controversial, and difference has been shrinking (13 points below for young children, 10 point for older children, 9 or less in more recent studies)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# Normal Curve: 1SD difference



Psychology 402 - Fall 2018 - Dr. Michael Diehr

# Ethnicity and IQ -- 2

- Why might different ethnic groups score differently?
- Environment?
  - wealth, school, language, culture, values, attitude, trust, nutrition, tutoring...
- Genetics?
  - neurophysiological issues
  - genetics interacting with environment
- Test Bias?
  - Achievement vs. IQ test? (AA score on NAEP = 1.1SD below mean in 1978, only .65 by 1990)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# Eugenics

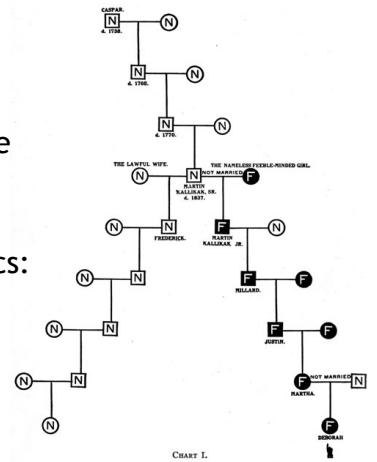
- Social philosophy advocating improvements of human hereditary traits through active intervention.
- Long history ("The best men must have intercourse with the best women as frequently as possible, and the opposite is true of the very inferior" Plato, Republic)
- Francis Galton (1860s) : First scientific formulation. (Note: Galton was Darwin's Cousin)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# Henry H. Goddard & the Feeble Minded Kallikak Family

## Henry H. Goddard & the Feeble Minded Kallikak Family

- Intelligence as Mendelian gene
- Single gene for IQ
- Dominant / Recessive
- Terminology: moron, imbecile, idiot
- Proponent of Eugenics:
  - institutionalization
  - sterilization
  - immigration restrictions



Psychology 402 - Fall 2018 - Dr. Michael Diehr

# Henry H. Goddard & the Feeble Minded

Child Genotype	Mental Age	IQ range	Terminology	
GG	17+	100+	"normal"	
Gg	13-16	70-100	"dull"	
gg	8 to 12	51-70	moron, "high-grade defective"	"Feeble Minded"
gg	3 to 7	26-50	imbecile	
gg	< 3 years	0-25	idiot	

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# Goddard at Ellis Island

- "We picked out one young man whom we suspected was defective, and through the interpreter, proceeded to give him the test [Binet's IQ test]. The boy tested 8 [Mental Age] on the Binet scale. The interpreter said 'I could not have done that when I came to this country' and seemed to think the test unfair. We convinced him that the boy was defective."
- Encouraged by this "success" he did more testing

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Goddard's IQ Testing - 1917

- Goddard hired two women\* to “identify feeble-minded on sight”
- Tested 35 Jews, 22 Hungarians, 50 Italians, and 45 Russians.
- Results:
  - 83% of Jews, 80% of Hungarians, 79% of Italians, and 87% of Russians were Feeble Minded

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Goddard on Politics

- “Democracy means that the people rule by selecting wisest, most intelligent, and most human to tell them what to do to be happy. Thus, democracy is a method for arriving at a truly benevolent aristocracy” (Gould, p. 191)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Eugenics

- Was a formal academic discipline in many colleges in early 1900s
- Notable supporters : Alexander Graham Bell, the Rockefeller Foundation
- Was adopted by the Nazis in the mid 1930s as a scientific basis for racism, segregation, human experimentation, forced sterilization, euthanasia and ultimately genocide.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Eugenics

“This person suffering from hereditary defects costs the community 60,000 Reichsmark during his lifetime. Fellow Germans, that is your money, too.” (propaganda poster for newspaper advocating compulsory euthanasia program circa 1938. The patient looks to have cerebral palsy, a non-hereditary disorder)



Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Eugenics in the USA

- Discriminatory “eugenics” policies were adopted in many states
- Compulsory sterilization (1907-1963), over 64000 people. A report of this program’s success in the USA was cited by Nazi scientists in the Nuremberg trials
- Marriage licenses denied to those with genetic disorders
- Immigration controls...

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## USA Immigration History

- Few laws/enforcement in 1600s, 1700s, 1800s
- 1865 Civil War / Emancipation of Slaves
- 1882 Chinese Exclusion Act
- 1917 Immigration Act
  - banned “illiterates, feeble-minded” and many other Asians (see Goddard’s theories)
- 1924 Immigration Act
  - Restricted Southern & Eastern Europeans, banned Africans, Arabs, Asians...
- Immigration and Nationality Act of 1965
- Immigration Reform and Control Act of 1986

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Modern Conceptions

- Many principles of eugenics were scientifically wrong : For example, you can't eradicate single-gene heterozygous recessive traits via phenotypic selection alone.
- Many genetic "disorders" also have a benefit -- e.g. sickle-cell anemia protects against malaria
- Some genetic-based programs are underway currently, based primarily on genetic screening.
- Question about what will happen as genetic tests get cheaper and more available.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Test Bias

- Content Validity
- Criterion Validity

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Test Bias - Content Validity?

- Perhaps score differences between ethnic groups are simply due to test item differences? Different cultures/ethnic groups are exposed to different information growing up?
  - Examples: "Petrol", "Opera", "Shilling" "Bourbon" > "Tequila"
  - Some item differences are clear
- However, large-scale testing hasn't shown big differences.
- Quay (1971): gave Stanford Binet in African-American dialect. Result: about 1 point increase.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Test Bias - Content Validity 2

- Clarizo (1979) - AA children can understand mainstream dialect. (Reverse, however is not necessarily true)
- Flaugher (1978) - experts judged "fairness" of items on IQ test and removed unfair items (16%). Result: test scores did not change
- Zores & Williams (1980) - There is a bias in the race, gender, ethnicity of people & situations portrayed in IQ tests.
- More research needed -- but little evidence that test bias can explain score differences.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Test Bias - info you don't have?

- Mainstream conception:
  - "Biased test" = test with information that I don't know. It's unfair.
- Reality:
  - Many people know information outside their immediate day-to-day culture.
  - Amount of this info is probably correlated with knowledge, IQ (and perhaps Intelligence?)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

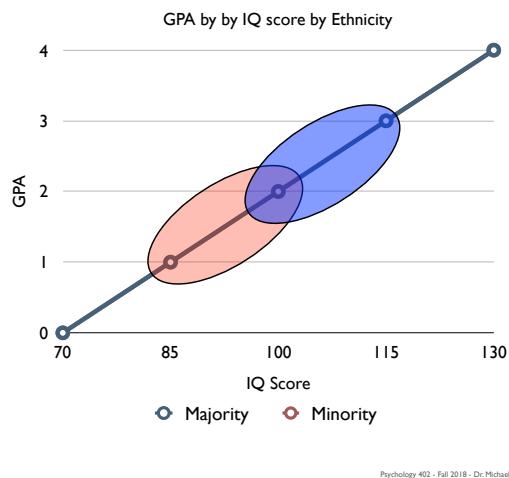
## Test Bias - Criterion Validity

- Generally, criterion-validity is considered more important than content validity.
- How well does an IQ test predict later academic success? Is this prediction the same across ethnic groups?
- Three basic scenarios:
  - Regression line is the same
  - Same slope, different intercept
  - Different slopes

Psychology 402 - Fall 2018 - Dr. Michael Diehr

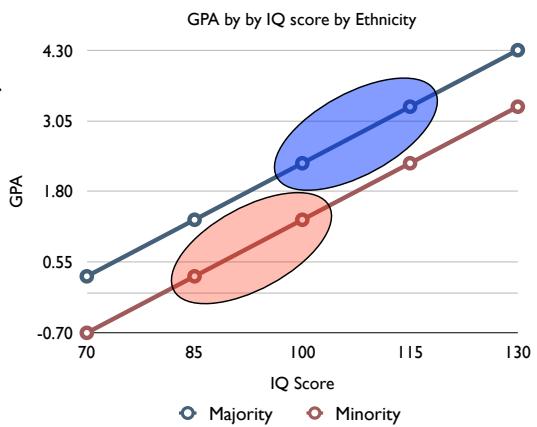
## Same regression line

Prediction is equally accurate for both races but one race generally scores higher



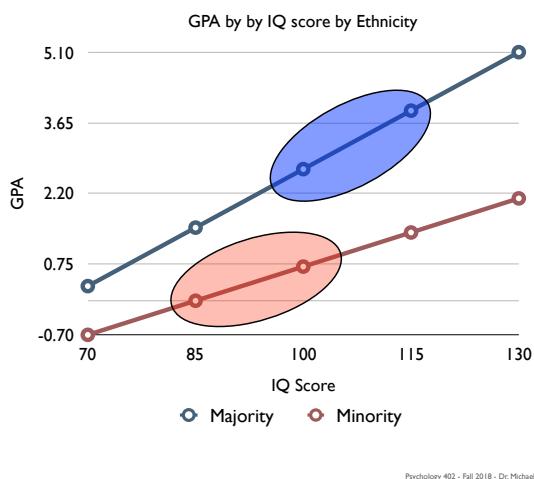
## Parallel regression lines

Prediction is accurate for both races if we use two regression lines with same slope different intercept



## Different regression lines

Clear example of bias -- predictions for the two ethnicities have different validity



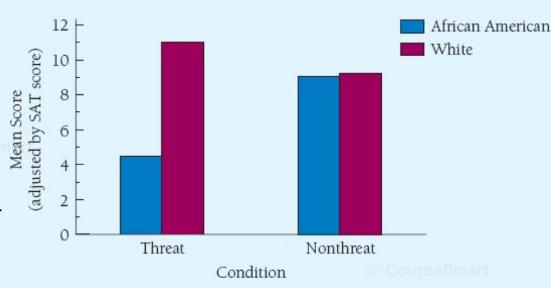
## Test Bias

- The 2nd example (parallel regression lines with different intercepts) is what the data seems to suggest, at least with the SAT test
- If you use a single regression line, that line over-predicts the academic performance of minority students while under-predicting that of majority students - Cleary (1968), Jensen (1984)
- Similar findings for different tests (IQ) in variety of ethnic groups, and in other countries.

## Stereotyping & Test Performance

AA students perform worse under "threat" condition

Conclusion: self-defeating cognition?



**FIGURE 19.4** Effects of stereotyping upon test performance. When told they were taking a test of intellectual abilities, white students scored significantly higher than African American students. However, some students were randomly assigned to take the same test but under conditions where there was no threat. Without a threat present, white and African American students performed equivalently.  
(Adapted from Steele, 1997, p. 621)

## Test Fairness and the Law

- 1964 Civil Rights Act
  - Created EEOC
- EEOC Guidelines
  - 1970, 1978
  - Uniform Guidelines on Employee Selection Procedures
- Adverse Impact : minority applications rejected at higher rate than non-minority
- May be acceptable if test is shown to be Valid

## Test Selection Philosophy

- Unqualified Individualism
  - high scores overall are selected
- Quota System
  - high scores within each group are selected
- Qualified Individualism
  - high scores overall are combined with other information to improve differential prediction

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## The Bell Curve (HM94)

- Controversial book by R Herrnstein\* and C. Murray (uses research by Arthur Jensen)
  - “g” exists and is measurable by IQ scores
  - Social stratification (difference between rich and poor) increasing, due to IQ differences
  - IQ predicts “success” (poverty, crime, etc.) better than many other measures
  - IQ differences between ethnic groups are the cause of social/economic differences
  - Recommends policy changes: ending welfare, etc.
  - Argued Griggs v. Duke Power was wrong
- \* died before publication

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Social correlates of IQ

Measure	IQ		
	<75	90 to 110	> 125
Married by age 30	72%	81%	67%
Unemployed > 1 month/year	12%	7%	2%
Lives in Poverty	30%	6%	2%
Chronic welfare recipient	31%	8%	0%

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Criticisms of The Bell Curve

- “The authors seem to show the evidence and leave the implications for the reader to figure out; discussing scientific work on intelligence, they never quite say that intelligence is all important and tied to one's genes, yet they signal that this is their belief and that readers ought to embrace the same conclusions.”
- - Howard Gardner

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Gould's Criticisms of The Bell Curve

- The Bell Curve's claim requires 4 logical arguments:
  - Intelligence can be expressed as IQ
  - Ordinal - people can be ranked by worth
  - Genetic
  - Immutable
- “most of the premises are false”  
(Gould, p. 368)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Criticisms of The Bell Curve

- “I believe this book is a fraud, that its authors must have known it was a fraud when they were writing it, and that Charles Murray must still know it's a fraud as he goes around defending it. [...] After careful reading, I cannot believe its authors were not acutely aware of [...] how they were distorting the material they did include.”
- - M Nunley

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# IQ : Genes vs. Environment

- What evidence do we have?

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# The Flynn Effect

- IQ tests are re-normed over time
- Using today's IQ tests (mean = 100) the mean in 1932 would have been 80.

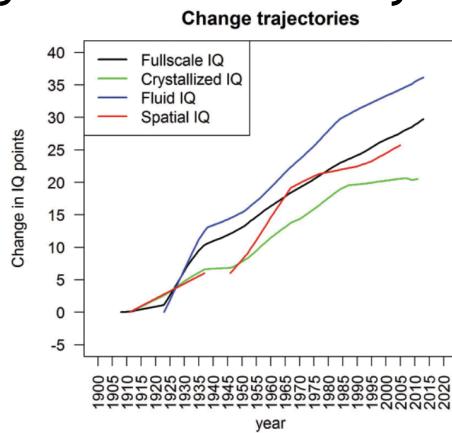
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## The Flynn Effect

- If IQ (g) is primarily influenced by genetics not environment, then IQ must be stable
- Yet, looking over time, we see that
  - IQ test scores on the whole are rising about 3 pts/decade
  - The gap between ethnic groups is declining
- The changes are too fast to be genetic
- Thus, something else is happening.
- Environmental causes? complexity of experience, nutrition, healthcare, parents literacy, family structure...

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Flynn Effect - 100 years



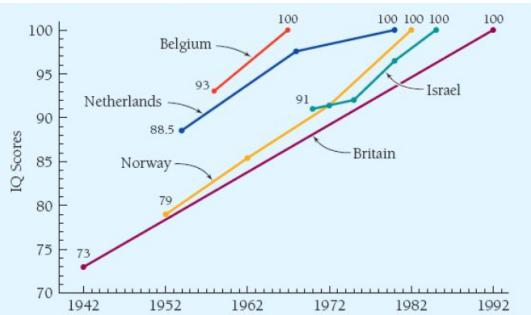
**Fig. 1.** Domain-specific IQ gain trajectories for 1909–2013. Changes are based on weighted average annual IQ changes in all available data.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Flynn Effect by Country

**FIGURE 19.2**

Gains in average IQ over time in five countries.



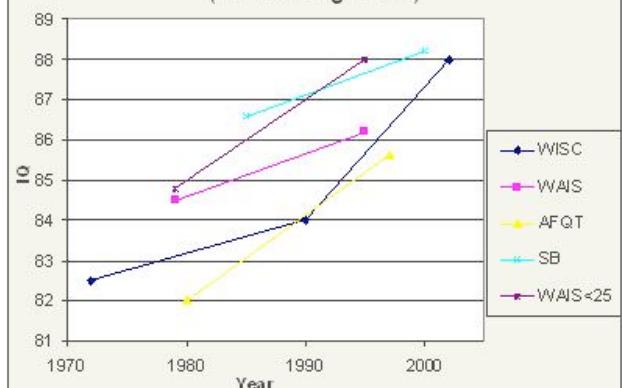
Note: Every nation is normed on its own samples. Therefore, although nations can be roughly compared in terms of different rates of IQ gain, they cannot be compared in terms of IQ scores. That is, the fact that the mean IQ of one nation appears higher than another at a given time is purely an artifact.

From J. R. Flynn, Searching for justice: The discovery of IQ gains over time. *American Psychologist*, Jan V 54 (n1), 1999, 5–20. Copyright © 1999 American Psychological Association. Reprinted by permission.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Flynn Effect: shrinking racial differences

**Black Scores on Four Tests of Cognitive Ability (white average = 100)**

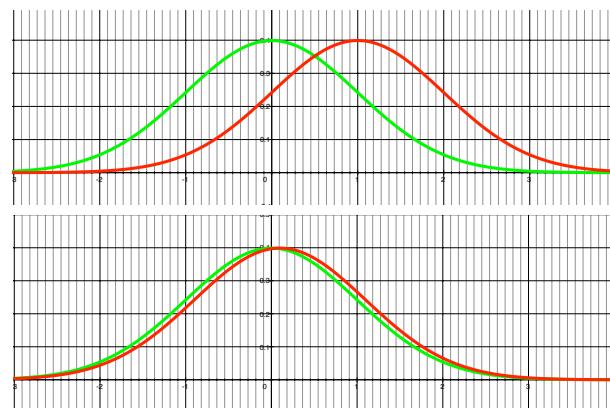


Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Between vs. within group

- Within a racial group, evidence that IQ is partially genetic. Thus it is *heritable*.
- Between racial groups, large differences in IQ scores.
- Therefore, difference between races is genetic.
- Sound logic? Or a fallacy?

## Between Group vs. Within Group Variance

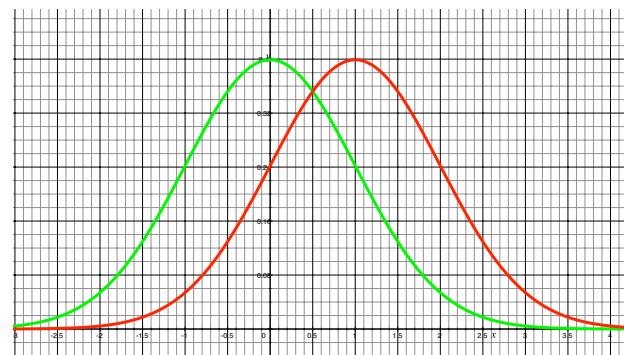


Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Between vs. within group

- Example: height of adult males
  - tall fathers tend to have tall sons, and vice versa.
- Village A : average height 5'6"
- City : average height 5'9"
- Is the between-group difference due to
  - genetics?
  - other factors?
- How to test this theory?

## Normal Curve: 1SD difference



Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Heritability

- Offspring are not identical to their parents -- the amount of variation in children can be expressed as a measure of variance.
- This variance can be partitioned in that which is genetically inherited ( $h^2$ ) version the portion that remains ( $1 - h^2$ ) which is attributed to the environment.
- We can not ethically manipulate these variables in humans, so our research has to be observational, rather than experimental

## Heritability is not always Biological

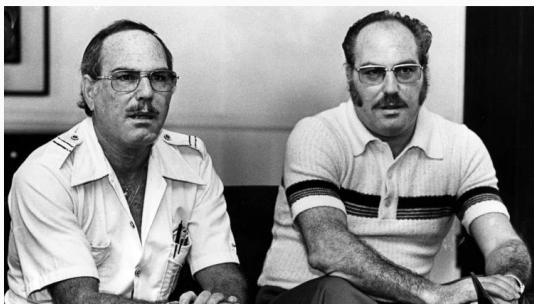
- Vocabulary has a high heritability constant
- Yet vocabulary, which consists of knowledge about word meanings, is clearly 100% environmental -- all words are learned.
- Important to remember that high (statistical) heritability does not prove or imply biological or genetic reality.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Recent Events

Jack Yufe dies at 82; he was raised Jewish, his identical twin as a Nazi



It sounded like a tabloid headline: Identical twins separated after birth. One grew up Jewish, the other a Nazi.

(upon first meeting)  
“We had identical clothes. I got mine in Israel and he got his in Germany. Exactly the same color, with two buttons”

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Twin & Adoption Studies

- Twins are either Monozygotic (MZ) or Dizygotic (DZ). MZ twins have identical DNA, sharing 100% of their genes, while DZ twins are no more related than siblings, sharing 50% of their genes.
- Siblings are sometimes adopted into separate families, thus providing a nice 2x2 quasi-experimental design to look at family vs. genetics
- Actual calculation of  $h^2$  is difficult, but it can be roughly estimated as twice the difference in correlation between MZ and DZ twins.

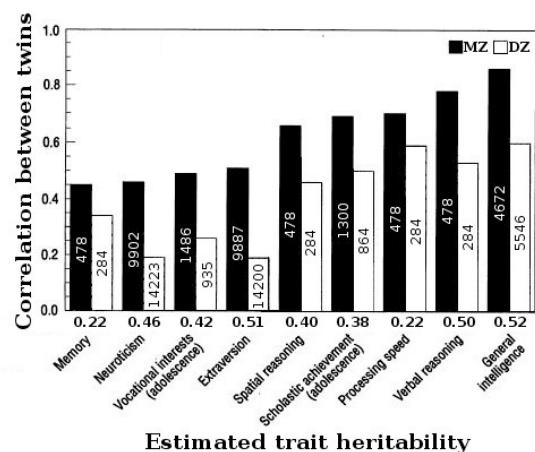
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Twin & Adoption Studies

	MZ Twins	DZ Twins
Reared together	100% genes + environment	50% genes + environment
reared apart	100% genes 0% environment	50% genes + 0% environment

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Heritability from Twin Studies



Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Twin Studies Criticism

- Generalizability : % of women having DZ twins varies with age, may run in families
- Poor model / statistics : In-utero environment : MZ twins adopted and raised apart still share the same environment for 9 months. Delvin et al (1997) estimates this as high as 20% of variance
- Adoptive families are often very similar (middle class, white, etc.) which may lead to underestimation of environmental influence
- GxE interactions -- genetic factors may feedback on environment resulting in overestimate of IQ heritability

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Delvin et al 1997 (DD97)

- Most twin adoption studies ignore the shared environment of twins (both in-utero, and in home prior to adoption)
- They performed a meta analysis of 212 correlations from prior studies
- HM94 had estimated  $h^2$  at 60% to 80%
- DD97 arrives at an estimate of 34% to 48%
- Big difference with large policy implications

Psychology 402 - Fall 2018 - Dr. Michael Diehr

# IQ Score Correlations

Group	R	R <sup>2</sup>
Same person (tested twice)	0.95	90%
Identical twins raised together	0.86	74%
Identical twins raised apart	0.76	58%
Fraternal twins reared together	0.55	30%
Fraternal twins raised apart	0.35	12%
Siblings raised together	0.47	22%
Siblings raised apart	0.24	6%
Unrelated children raised together	0.28	8%

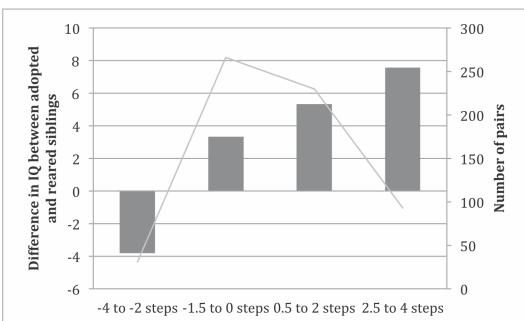
Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Latest Research: Kendler et al. (2015)

- Swedish study of male siblings
- One child raised at home
- One child adopted
- IQ test at age 18
- Measured adopted family Educational level
- Largest study to date (436 pairs)
- Question:
  - How would IQ of adopted siblings vary?

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Kendler et al. (2015)



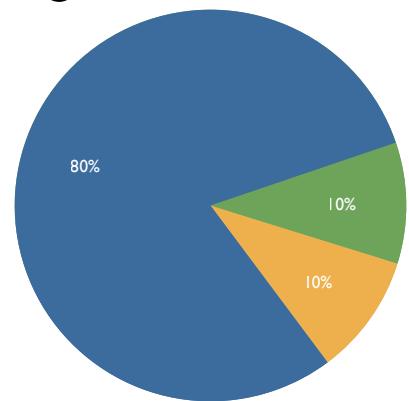
**Fig. 1.** Magnitude of IQ difference (black bars and left y axis) between adopted and nonadopted full-siblings as a function of the difference in educational level between biological and adoptive parents of the adopted siblings (x axis). The four bars represent (from left to right) -4 to -2 steps; -1.5 to 0 steps; 0.5-2 steps; and 2.5-4 steps difference on the education scale. The gray line (right y axis) illustrates the number of pairs in each group.

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Explaining Variance

- Genetics
- Environment
- Unexplained

Heritability / Genetics variance estimates range from 34% to 80%



Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Conclusions

- Gender differences on IQ tests are small (less than 0.2 SD) but still controversial especially at extremes
- Ethnic differences are fairly large (1.0 SD, but decreasing)
- Explanations:
  - Test bias? some but not enough to explain
  - Genetic differences? yes, but recent results suggest this % has been vastly over-estimated.
  - Environmental differences? Yes. Explains recent reductions in differences (e.g. Flynn effect)
  - Latest research : IQ is very mutable through environment (Kendler et al. Sibling study)

Psychology 402 - Fall 2018 - Dr. Michael Diehr

## Conclusions 2

- IQ and Race is a Highly Controversial, Politicized topic
- Gould suggests “scientific racism” exists today
  - conscious or unconscious biases?

Psychology 402 - Fall 2018 - Dr. Michael Diehr