

Ch. 19: Controversy in Testing

1555

Psychology 402 - Spring 2025 - Dr. Michael Diehr

Copyright © 2025 Michael Diehr
All Rights Reserved
For use only by students enrolled
in my sections of Psyc 402
through the end of the semester.
May not be posted, shared or uploaded
online without permission.

1556

Psychology 402 - Fall 2024 - Dr. Michael Diehr

Ch. 19 - Part 1

1557

Psychology 402 - Spring 2025 - Dr. Michael Diehr

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

1564

Psychology 402 - Spring 2025 - 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

1565

Psychology 402 - Spring 2025 - 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 slightly more variable (higher variance)
 - Males better at stereotypical “male” tasks (visuospatial skills) whereas women better at “female” tasks (language). Why?

1566

Psychology 402 - Spring 2025 - Dr. Michael Diehr

Race vs. Ethnicity

- Race - genetic heritage
- Ethnic group -- population whose members identify with each other
- “National, religious, geographic, linguistic and cultural groups do not necessarily coincide with racial groups: and the cultural traits of such groups have no demonstrated genetic connection with racial traits. Because serious errors of this kind are habitually committed when the term “race” is used in popular parlance, it would be better when speaking of human races to drop the term “race” altogether and speak of ‘ethnic groups’.”

1568

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Pre-DNA views

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

1569

Psychology 402 - Spring 2025 - Dr. Michael Dohr

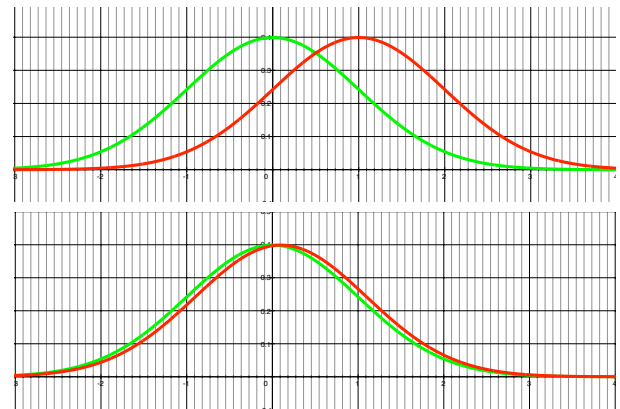
Nonconcordant traits

- Naive view: Ethnicity → Race → DNA
- Biochemical view
 - Traits & genes show population group differences, but not always, boundaries can be fuzzy
- Non concordance: phenotypes don't correlate with geography.
- Visible vs. Invisible differences: differences on the genetic level often don't track what is seen in surface differences
- Example: Skin color vs. Blood Type

1575

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Between Group vs. Within Group Variance



1580

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Post-DNA views

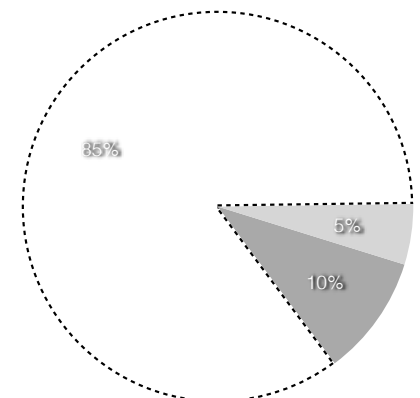
- Variance
 - variation between individuals
 - aka variation *within groups*
 - variation *between groups*
- Variance
 - variation between individuals : 3mbp / person
 - variation within groups : 85%
 - variation between groups: 15%
 - about 5% - within “races”
 - about 10% - between “races”

1581

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Variance: Genetic Variation

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



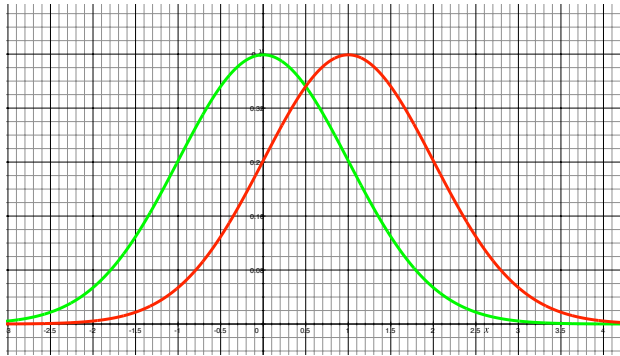
For example:

- 85% within Japanese
- +5% between Japanese vs. Korean
- +10% between Asian vs. Caucasian

1582

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Normal Curve: 1SD difference



1584

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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)

1585

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Eugenics

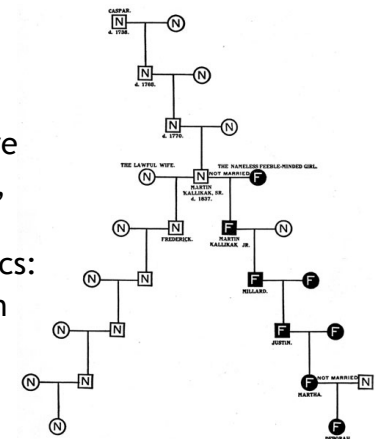
- Social philosophy advocating improvements of human hereditary traits through active intervention.
- Long history ("The best men must have intercourse with the best women as frequently as possible, and the opposite is true of the very inferior" Plato, Republic)
- Francis Galton (1860s) : First scientific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

1586

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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



1587

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Eugenics

- A formal academic discipline in many USA 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.

1594

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Eugenics in the USA

- Discriminatory eugenics policies were adopted in many states
- Compulsory sterilization (1907-1963), over 64000 people. 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...

1596

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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

1600

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Modern Conceptions

- Eugenics - many theories are scientifically wrong : e.g. impossible to eradicate single-gene heterozygous recessive traits via phenotypic selection alone.
- Genetic “disorders” also have a benefits -- e.g. sickle-cell trait protects against malaria
- Genetic factors have been over-estimated in some research studies...

1601

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Ch. 19 - Part 2

1605

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Possible causes of IQ score differences across population groups...

1611

- Test Bias?
 - the test is unfair to certain groups
- Environmental factors
 - wealth, school, language, culture, values, attitude, trust, nutrition, tutoring...
- Genetics?
 - actual biological brain differences
 - genetics interacting with environment (GxE interactions)

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Test Bias

- Content Validity
- Criterion Validity

1612

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Test Bias - Content Validity?

- Test differences between groups due to test item differences? Different cultures have different knowledge?
 - 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.

1613

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Test Bias - Content Validity 2

- Clarizo (1979) - minority children can understand majority dialect. (but not necessarily vice-versa)
- Flaughner (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.

1614

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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?)

1615

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Test Bias - Criterion Validity

- Criterion-validity considered more important than content validity.
- Does IQ score predict later academic success?
- Is linear regression prediction same for all ethnic groups?
- Three scenarios:
 - Regression line is the same
 - Same slope, different intercept
 - Different slopes

1616

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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.

1620

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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

1624

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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

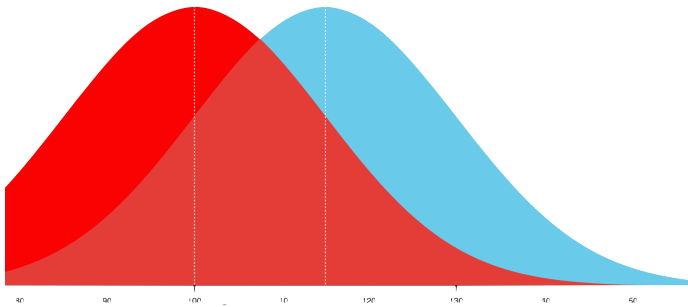
1625

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Selection Philosophy

1627

- Unqualified Individualism (Meritocracy)
- Quota System
- Qualified Individualism



Psychology 402 - Spring 2025 - Dr. Michael Dohr

Affirmative Action

- 1961 President Kennedy - government contractors must “take affirmative action to ensure that applicants are employed, and that employees are treated during employment, without regard to their race, creed, color, or national origin”
- Later cases interpreted to mean race-based quotas might be required
- In USA, laws vary by state
- In 2023..2025, various Supreme Court cases under review

1636

Psychology 402 - Spring 2025 - Dr. Michael Dohr

SAT adds ‘Adversity’ Score

- May 2019
- Measures 15 facets in 3 factors:
 - neighborhood environment
 - crime & poverty, housing cost...
 - family environment
 - parent’s education, single parent...
 - high school environment
 - AP classes, free lunch (poverty)...
- Single Score from 1 to 100

1639

Psychology 402 - Spring 2025 - Dr. Michael Dohr

The Bell Curve (1994)

- Controversial book by Herrnstein* and Murray (uses research by Arthur Jensen) with these claims:
 - “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

1644

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Social correlates of IQ

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

1646

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Gould’s Criticisms of The Bell Curve

- The Bell Curve’s claim requires 4 logical arguments:
 - Intelligence is measured by IQ score
 - Ordinal - people can be ranked by worth
 - Genetic
 - Immutable
- “most of the premises are false” (Gould, 1996, p. 368)

1651

Psychology 402 - Spring 2025 - Dr. Michael Dohr

IQ : Genes vs. Environment

- What evidence do we have?

1652

Psychology 402 - Spring 2025 - Dr. Michael Dohr

The Flynn Effect

- IQ tests are re-normed over time
- Distribution of Raw scores forms normal curve
- Average score is defined as IQ=100
- Using today's IQ tests (mean = 100) the mean in 1910 would have been 70.

1653

Psychology 402 - Spring 2025 - Dr. Michael Dohr

The Flynn Effect

- If IQ (g) is primarily genetic, IQ must be stable over time
- But instead we find:
 - IQ scores rising about 3 points/decade
 - Ethnic group differences shrinking
- This is too fast to be genetic
- Thus, something else is happening.
- Environmental causes? complexity of experience, nutrition, healthcare, parents literacy, family structure...

1655

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Flynn Effect - 100 years

1656

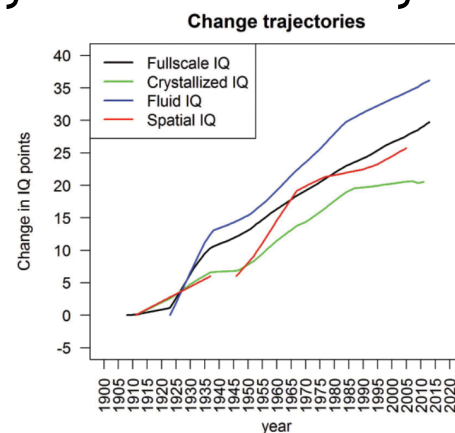
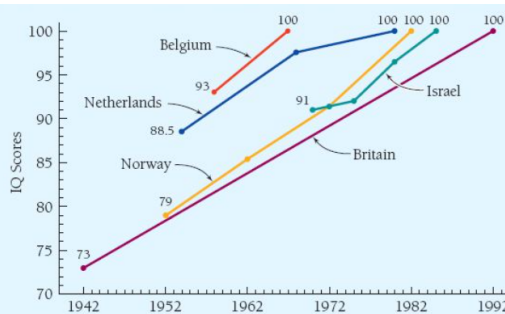


Fig. 1. Domain-specific IQ gain trajectories for 1909–2013. Changes are based on weighted average annual IQ changes in all available data.

102 - Spring 2025 - Dr. Michael Dohr

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.

1657

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Ch. 19 - Part 3

1665

Psychology 402 - Spring 2025 - Dr. Michael Dohr

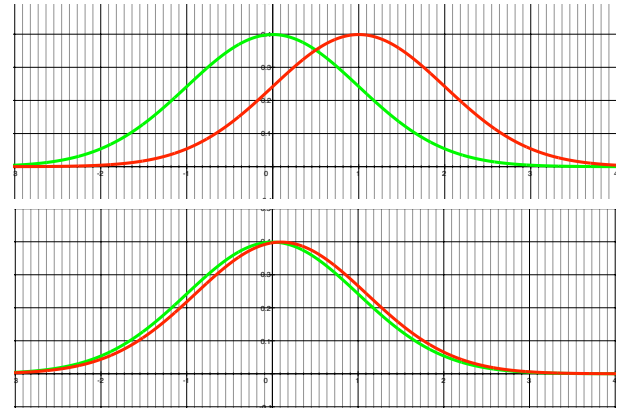
Between vs. within group

- 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?

1671

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Between Group vs. Within Group Variance



1672

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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?

1673

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Heritability

- Children are similar to but not identical to parents
- Variation in children's traits: a kind of variance
 - h^2 : genetically inherited
 - $1 - h^2$: environmental influences
- Can not ethically alter these variables in humans, so research must be observational, rather than experimental

1675

Psychology 402 - Spring 2025 - Dr. Michael Dohr

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.

1676

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Twin & Adoption Studies

- Monozygotic (MZ) twins: identical DNA (100%)
- Dizygotic (DZ) share 50% DNA (like siblings)
- Twins adopted into same families, or separated to different families
- 2x2 quasi-experimental design
 - family vs. genetics
- h^2 is approximately twice the difference in correlation between MZ and DZ twins.

1678

Psychology 402 - Spring 2025 - Dr. Michael Dohr

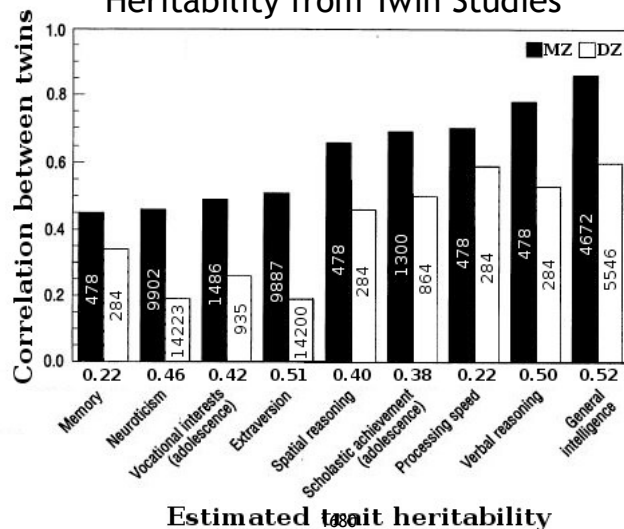
Twin & Adoption Studies

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

1679

Psychology 402 - Spring 2025 - Dr. Michael Doherty

Heritability from Twin Studies



Isabel Doherty

Twin Studies Criticism

1681

Psychology 402 - Spring 2025 - Dr. Michael Doherty

Twin Studies Criticism

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

1683

Psychology 402 - Spring 2025 - Dr. Michael Doherty

GxE interactions

- Biological Sex: genetic
- “Girls are bad at math”
- Pupils, families, friends, teachers have lower expectations
- Girls get poorer math education
- Girls show lower scores on Math tests
- Comparison of Math abilities vs. Sex
 - differences appear to be genetic
 - could really be 100% environmental

1684

Psychology 402 - Spring 2025 - Dr. Michael Doherty

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

1685

Psychology 402 - Spring 2025 - Dr. Michael Doherty

IQ Score Correlations

Group	R	R ²
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%

1686

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Recent 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?

1687

Psychology 402 - Spring 2025 - Dr. Michael Dohr

Kendler et al. (2015)

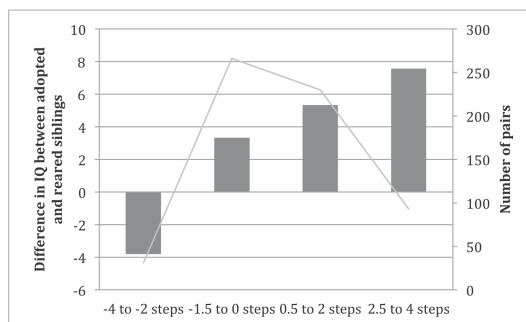


Fig. 1. Magnitude of IQ difference (black bars and left y axis) between adopted and nonadopted full-siblings as a function of the difference in educational level between biological and adoptive parents of the adopted siblings (x axis). The four bars represent (from left to right) -4 to -2 steps; -1.5 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 - Spring 2025 - Dr. Michael Dohr

Conclusions

- Gender differences on IQ tests are small (less than 0.2 SD) but still controversial especially at extremes
- Ethnic differences were large (1.0 SD) but are shrinking
- Explanations:
 - Test bias? Some, but small.
 - Genetic differences? Perhaps, but recent results suggest this % was overestimated.
 - Environmental differences? **Yes.** Explains Flynn effect and shrinking ethnic differences.
- Recent research : IQ is very mutable through environment (Kendler et al., 2015) sibling study

1690

Psychology 402 - Spring 2025 - Dr. Michael Dohr