Psychology 205: Research Methods in Psychology
Field studies and randomized designs
The example of epidemiology and honors workshops

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

May, 2021

Outline

The effect of preschool

1. A team of educational psychologists examined the effect of early reading in preschool upon later academic attainment.
2. They randomly selected 20 preschools in Evanston and gave a measure of reading skill to 200 children (mean = 5.0, sd =1.0).
3. They followed the progress of the top 50 students (mean score = 6) for a year.
4. At the end of the year they compared their sample students to the mean and found the group was no different from the average. They concluded that preschool hurt these students.
5. Explain this effect

1. A team of educational psychologists examined the effect of early reading in preschool upon later academic attainment.
2. They randomly selected 20 preschools in Evanston and gave a measure of reading skill to 200 children (mean = 5.0, sd =1.0).
3. They followed the progress of the bottom 50 students (mean score = 4) for a year.
4. At the end of the year they compared their sample students to the mean and found the group was no different from the average.
5. They concluded that preschool helped these students.
6. Explain this effect.
7. How does this relate to the previous problem?

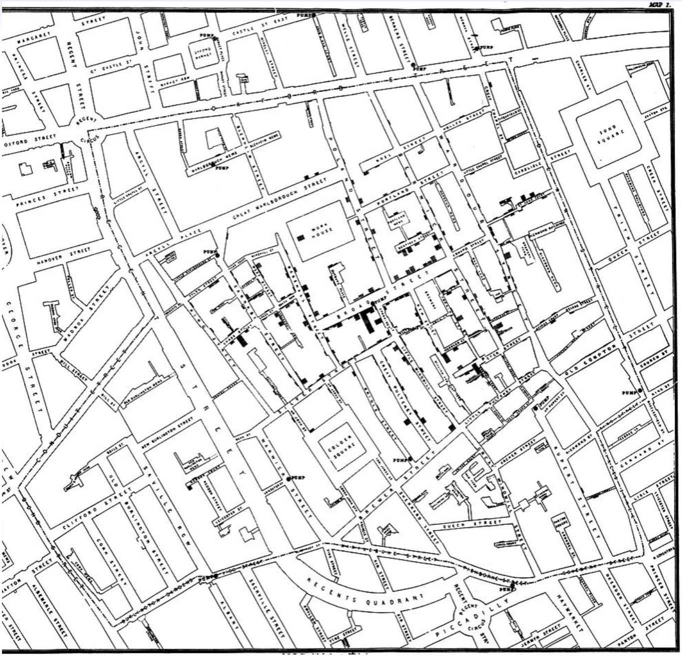
Epidemiology

1. "The branch of medicine that deals with the incidence, distribution, and possible control of diseases and other factors relating to health"
2. Typical study is correlational: Higher levels of variable X are associated with more of disease Y

Epidemiology

Classic case of John Snow and tests of the "Miasma" hypothesis

1. Cholera epidemics would hit London in the summer.
2. Particularly a problem on the South side of the river where the air was very bad.
3. The hypothesis was that cholera was spread by "miasma" or bad air.
4. "On 31 August 1854, after several other outbreaks had occurred elsewhere in the city, a major outbreak of cholera reached Soho. John Snow, the physician who eventually linked the outbreak to contaminated water, later called it "the most terrible outbreak of cholera which ever occurred in this kingdom."[2]
5. "Over the next three days, 127 people on or near Broad Street died. In the next week, three quarters of the residents had fled the area. By 10 September, 500 people had died and the mortality rate was 12.8 percent in some parts of the city. By the end of the outbreak, 616 people had died." (Wikipedia)


MAP 1.
OXFORD STREET
PRINCES STREET
HANOVER STREET
GREAT MARLBOROUGH STREET
WORK HOUSE
BROAD STREET
GOLDEN SQUARE
SOHO SQUARE
REGENT STREET
REGENTS QUADRANT
PICCADILLY
CLIFFORD STREET
OLD COMPTON STREET
QUEEN STREET
PUMP

Snow and the Broadstreet pump

1. "On proceeding to the spot, I found that nearly all the deaths had taken place within a short distance of the [Broad Street] pump. There were only ten deaths in houses situated decidedly nearer to another street-pump. In five of these cases the families of the deceased persons informed me that they always sent to the pump in Broad Street, as they preferred the water to that of the pumps which were nearer. In three other cases, the deceased were children who went to school near the pump in Broad Street..."
2. "With regard to the deaths occurring in the locality belonging to the pump, there were 61 instances in which I was informed that the deceased persons used to drink the pump water from Broad Street, either constantly or occasionally..."
3. "The result of the inquiry, then, is, that there has been no particular outbreak or prevalence of cholera in this part of London except among the persons who were in the habit of drinking the water of the above-mentioned pump well."

from Wikipedia 1854 Broad Street Cholera outbreak

Lead levels and IQ

1. Intelligence tests scores have increased dramatically over the past 100 years. This has come to known as the "Flynn effect" for the person who first reported this remarkable change (Flynn, 1984, 1987).
2. In a recent paper Factor-Litvak & Krebs (2019) made use of Scottish Mental Survey of 1932 and 1947 to examine one explanation of the effect based upon the change in lead exposure for the two cohorts.
3. The Scottish Mental Survey gave a short paper and pencil Intelligence test to every 11 year old school child in 1932 and then in 1947 and was a major source of what has come to be called "cognitive epidemiology" (Deary & Batty, 2007).
4. "Based on an interpretation of the evidence, and on the knowledge that lead can adversely affect a child's mental development, a theory is proposed to explain the rise in Scottish test scores. Following a fall in environmental lead levels, and in conjunction with higher intakes of dietary calcium, the theory states that the 1947 Cohort experienced lower lead burdens than did children in the 1932 Cohort (see note 9). As a result, children in the 1947 Cohort had an intelligence advantage over children in the earlier cohort and were able to perform better on intelligence tests."
5. Can you think of alternative explanations for this increase in IQ scores between 1932 and 1947?

(Factor-Litvak & Krebs, 2019)

Lead levels and IQ – observational with measurement

1. Dunedin Multidisciplinary Health and Development study is an ongoing prospective cohort of 1972-1973 birth cohort from Dunedin, New Zealand (Moffitt, Caspi, Rutter & Silva, 2001).
2. Analyzed Childhood levels as exposure ascertained as blood levels measured at age11 years.
3. High blood lead levels were observed among children from all socioeconomic status levels in this cohort.
4. Of 1037 original participants,1007 were alive at age 38years,of whom 565(56%)had been lead tested at age 11 years (54% male; 93% white)

(Reuben, Caspi, Belsky, Broadbent, Harrington, Sugden, Houts, Ramrakha, Poulton & Moffitt, 2017)

Adult change in IQ and SES by childhood blood levels

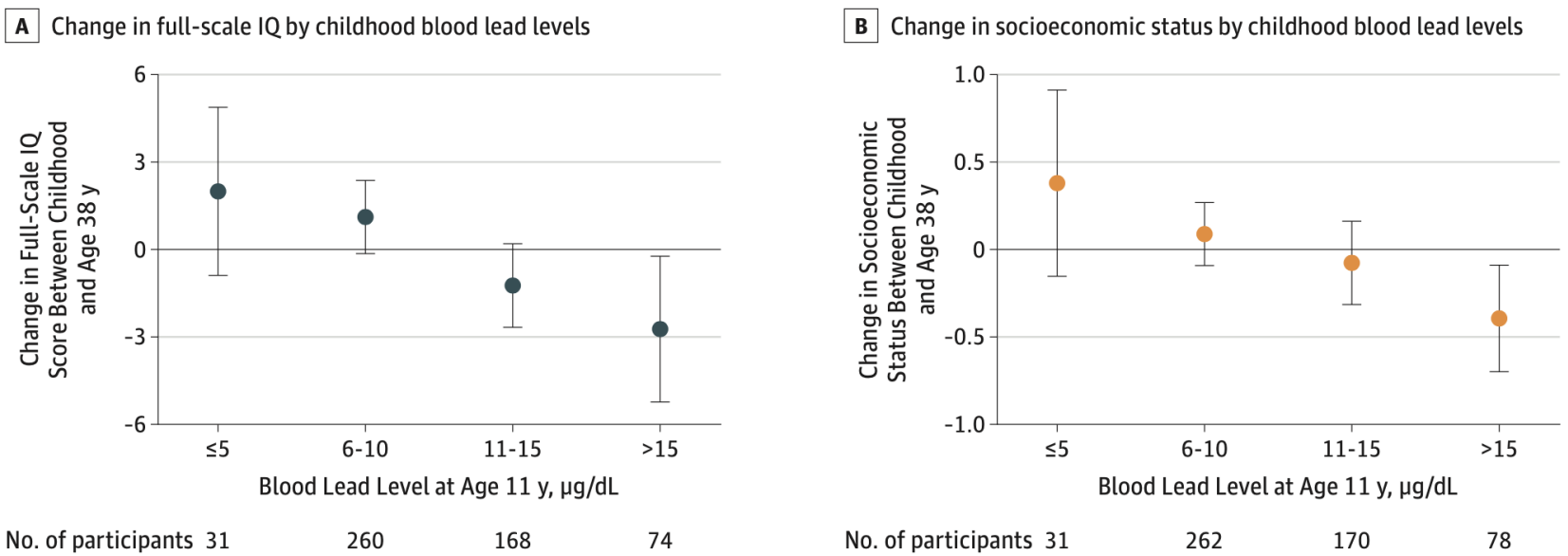
Table 3. Association Between Childhood Blood Lead Levels at Age 38 Years and Adult IQ and Adult Socioeconomic Status[a]

	Adjusted by Sex		Fully Adjusted[b]	
	Regression β Coefficient (95% CI)	*P* Value	Regression β Coefficient (95% CI)	*P* Value
WAIS-IV Adult IQ[c]				
Full-scale IQ[d]	−1.97 (−3.34 to −0.59)	.005	−1.61 (−2.48 to −0.74)	<.001
Verbal Comprehension IQ[e]	−1.39 (−3.01 to 0.23)	.09	−1.01 (−2.18 to 0.16)	.09
Perceptual Reasoning IQ[e]	−2.36 (−3.69 to −1.03)	<.001	−2.07 (−3.14 to −1.01)	<.001
Working Memory IQ[e]	−1.52 (−2.95 to −0.08)	.04	−1.26 (−2.38 to −0.14)	.03
Processing Speed IQ[e]	−0.91 (−2.19 to 0.37)	.16	−0.70 (−1.85 to 0.45)	.23
Adult socioeconomic status[d,f]	−1.94 (−3.50 to −0.37)	.02	−1.79 (−3.17 to −0.40)	.01

(Reuben et al., 2017)

Adult change in IQ and SES by childhood blood levels

Figure 3. Association of Childhood Blood Lead Level With Cognitive Decline and Downward Socioeconomic Mobility Into Adulthood



(Reuben et al., 2017)

The benefits and problems of field studies

1. Two recent studies of the effect of diet on longevity have shown that:
   - People who have low fat diet live longer (Nurses Health Study)
   - but there is no effect of a low fat diet on your health (Womens' Health Initiative).
2. Both studies were very well powered (many participants) and carefully done.
3. How can this be?

(Partially adapted from Gary Taubes: "Do we really know what makes us healthy". New York Times Magazine, September 16, 2007.)

Hormone Replacement Therapy

1. Nurses study (observational)
   - HRT is good
2. Women's Health Initiative (random assignment)
   - HRT is slightly bad

Nurses Health Study

1. Observational study of nurses
2. Positive effect of estrogen on heart disease
3. But also observed reduction in death by homicide, suicide, and accidents

The bias of healthy users

1. 'People who faithfully engage in activities that are good for them – taking a drug as prescribed, ... or eating what they believe is a healthy diet – are fundamentally different from those who don't.
2. Nurses who took HRT were thinner, fewer risk factors for heart disease, more educated, wealthier, exercise more, more health conscious.

(Taubs, 2007)

The bias of compliance

1. People who comply with their doctors' orders are healthier than those who don't
2. Effects are even true for placebo takers (Gallagher, Viscoli & Horwitz, 1993)!
3. General regression to the mean of placebo effects (Davis, 2002)

Doctors' prescribing effect

1. People who are eager to take particular drugs are probably different than those who are not

Randomized field trials as an alternative

1. Observational studies have all kinds of biases, what about doing random assignment?
2. How to do it?
3. The example of the Women's Health Initiative

Women's Health Initiative

1. A continuing longitudinal study, started in 1992.
2. Original study question was the effect of (low fat) diet and Hormone Replacement Therapy on pre and post-menopausal women.
3. Older (pre and post menopausal) women
4. Randomized field trial
5. HRT vs placebo
6. Reduced fat versus normal
7. Calcium supplements versus placebo

Women's Health Initiative

1. The Women's Health Initiative (WHI) is a long-term national health study funded by the National Heart, Lung, and Blood Institute, or NHLBI. The original WHI study began in the early 1990s and concluded in 2005. Since 2005, the WHI has continued as Extension Studies, which are annual collections of health updates and outcomes in active participants. The second Extension Study enrolled 93,500 women in 2010 and follow-up of these women continues through another extension that will conclude in 2020. NHLBI intends to fund the next Extension Study for the follow-up of WHI participants through 2027.
2. As with the original WHI study, the main areas of research are cardiovascular disease, cancers, and osteoporotic fractures. While WHI continues to focus on strategies to prevent the major causes of death, disability, and frailty in older women, the breadth and richness of the WHI data allow for the exploration and investigation of many more research questions on women's health and aging.
3. To learn more about the original WHI study that began in the early 1990's, including specific details about the three clinical trials and the observational study, visit the WHI program page on the NHLBI website.

Women's Health Initiative

1. The WHI is one of the largest women's health projects ever launched in the United States, having enrolled more than 161,000 women at 40 clinical centers.
2. The WHI randomized controlled clinical trial enrolled more than 68,000 postmenopausal women between the ages of 50 and 79.
3. The WHI observational study tracked the medical histories and health habits of more than 93,000 women, providing information to complement the clinical trial.
4. The WHI found that hormone replacement therapy did not prevent heart disease in postmenopausal women as once thought.
5. The scientific knowledge gained from the WHI hormone trials has helped save an estimated $35.2 billion in direct medical costs in the United States.

WHI, partial results

1. By the end of the first year, the low-fat diet group reduced average total fat intakes to 24 percent of calories from fat, but did not meet the study's goal of 20 percent. At year six, the low-fat diet group was consuming 29 percent of calories from fat. The comparison group averaged 35 percent of calories from fat at year one and 37 percent at year six. Women in both groups started at 35-38 percent of calories from fat. The low fat diet group also increased their consumption of vegetables, fruits, and grains.
2. Women were aged 50-79 at trial enrollment in 1993-98 and were followed for an average of 8.1 years. The study diet focused on reducing total fat, and unlike diets used to reduce heart disease risk, did not differentiate between "good fats" found in fish, nuts, and vegetable oils, and "bad" fats like saturated fat and trans fat found in processed foods, meats, and some dairy products. The study design reflected a widely believed but untested theory that reduction of total fat would reduce risks of breast or colorectal cancers. For heart disease, it was anticipated that reduction in total fat would be accompanied by a reduction in saturated fats, which are known to contribute to heart disease risk.

Participation bias

1. Who participates in a random study?
2. Who complies with instructions?
3. Effect of assignment versus effect of actual treatment
4. Must do the analysis on all subjects (effect as assigned)

WHI results

1. Stopped HRT trials after slightly greater risk of heart attack.
2. Effect of dietary modification was minimal unless one looked just at the compliant subjects (but see above).
3. Study has continued with random assignment to various dietary supplement conditions (with placebo controls) to examine the effect on general healthy aging.

The crises in science education

1. STEM majors are decreasing
   - Science, Technology, Engineering, Math
2. Particularly, women and minorities are not enrolling in or not continuing in STEM courses
3. Why is this happening?

Alternative explanations for STEM differences

1. Ability
2. Interests
3. Discrimination
4. Stereotype threat

Honors Workshops in STEM courses

1. Evidence from calculus classes that study groups help performance.
2. Treisman (1992) at UCB found that white and Asian males used study groups, females and African-American students did not.
3. Interpreted differences in test performance as motivational effect.

Study group effect on motivation

1. Student by him/her self
   - I don't know how to do problem 6
   - I must be stupid
2. Student in study group
   - I know how to do problem 5,
   - you know how to do problem 6,
   - let's teach each other
   - I am not stupid, the material is hard!

Biology Honors workshops at NU

1. The problem: students tend to do badly and drop out of core course for Pre-Meds.
2. An quasi-experimental study of biology workshops.
3. Students asked if interested in participating
   - Volunteers more interested in biology
   - Volunteers more anxious
4. Among those willing to participate, random assignment to honors study groups or not
5. Workshop students did better, more likely to complete the course than those who volunteered but did not participate

Born, Revelle & Pinto (2002)

Peer led biology workshops: an experimental study

1. This 2-year quasi-experiment evaluated the effect of peer-led workshop groups on performance of minority and majority undergraduate biology students. The workshop intervention used was modeled after a program pioneered by Treisman (1992).
2. Majority volunteers randomly assigned to workshops (n = 61) performed significantly better than those assigned to the control group ($n = 60, p < 0.05$) without spending more time studying.
3. Workshop minority students (n = 25) showed a pattern of increasing exam performance in comparison to historic control minority students (n = 21), who showed a decreasing pattern ($p < 0.05$).

(Born et al., 2002)

Born et al. (continued)

1. Volunteers (n = 121) initially reported that biology was more interesting and more important to their futures than did nonvolunteers ($n = 435, p < 0.05$).
2. Volunteers also reported higher levels of anxiety related to class performance ($p < 0.05$).
3. The relationship of anxiety to performance was moderated by volunteer status.
4. Performance of volunteers was negatively associated with self-reported anxiety ($r = -0.41, p < 0.01$).
5. Performance of nonvolunteers was unrelated to self-reported anxiety ($r = -0.02$).

(Born et al., 2002)

Born et al. (continued)

1. Results suggest elevated anxiety related to class performance may increase willingness to participate in activities such as workshop interventions.
2. In addition, students who volunteer for interventions such as workshops may be at increased risk of performance decrements associated with anxiety.
3. Even so, workshop programs appear to be an effective way to promote excellence among both majority and minority students who volunteer to participate, despite the increased risk of underperformance associated with higher levels of anxiety.

(Born et al., 2002)

Exam performance during Quarter 1

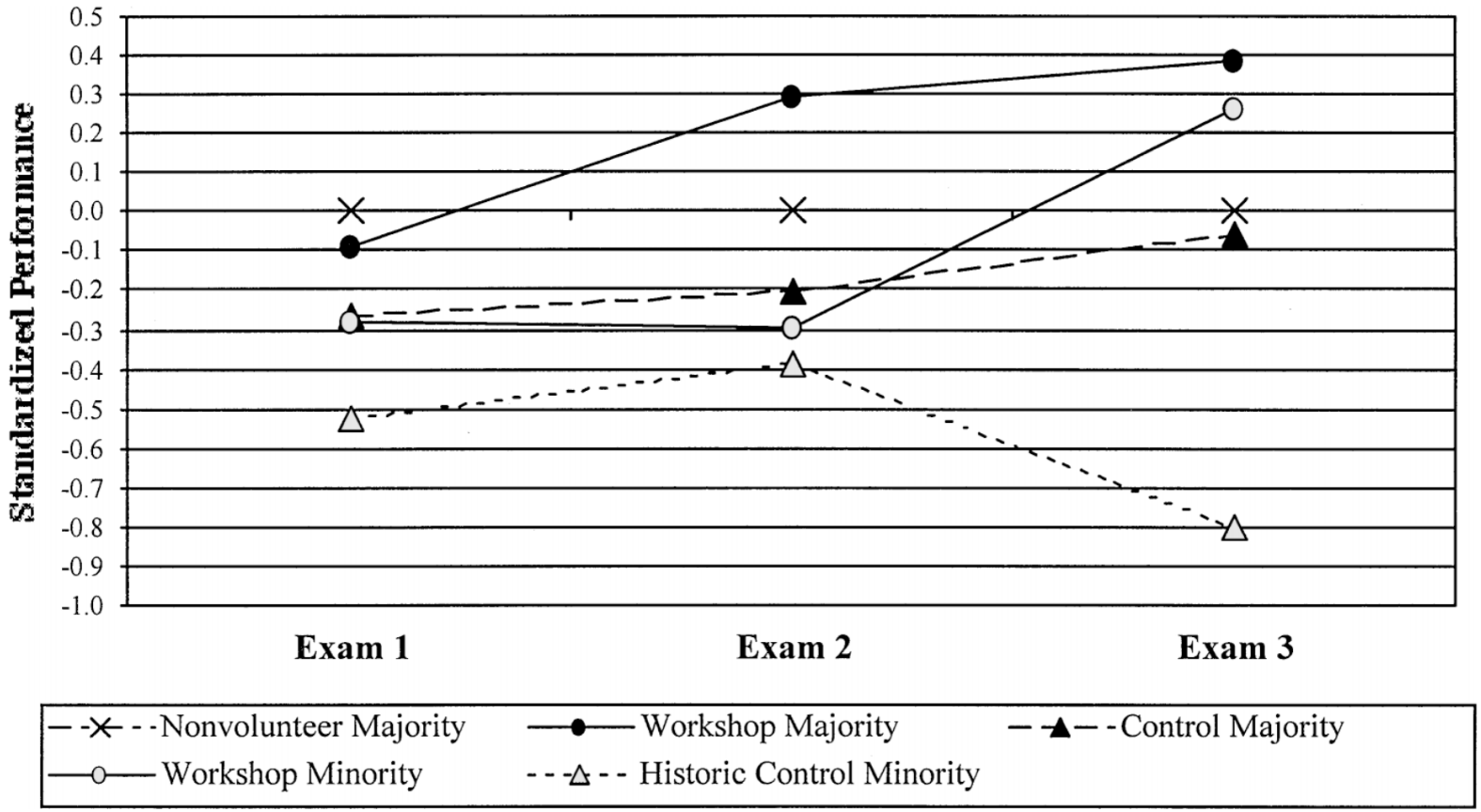


Fig. 1. Standardized Quarter 1 exam performance as a function of group, controlling for prior cumulative grade point average.

(Born et al., 2002)

Performance over three quarters

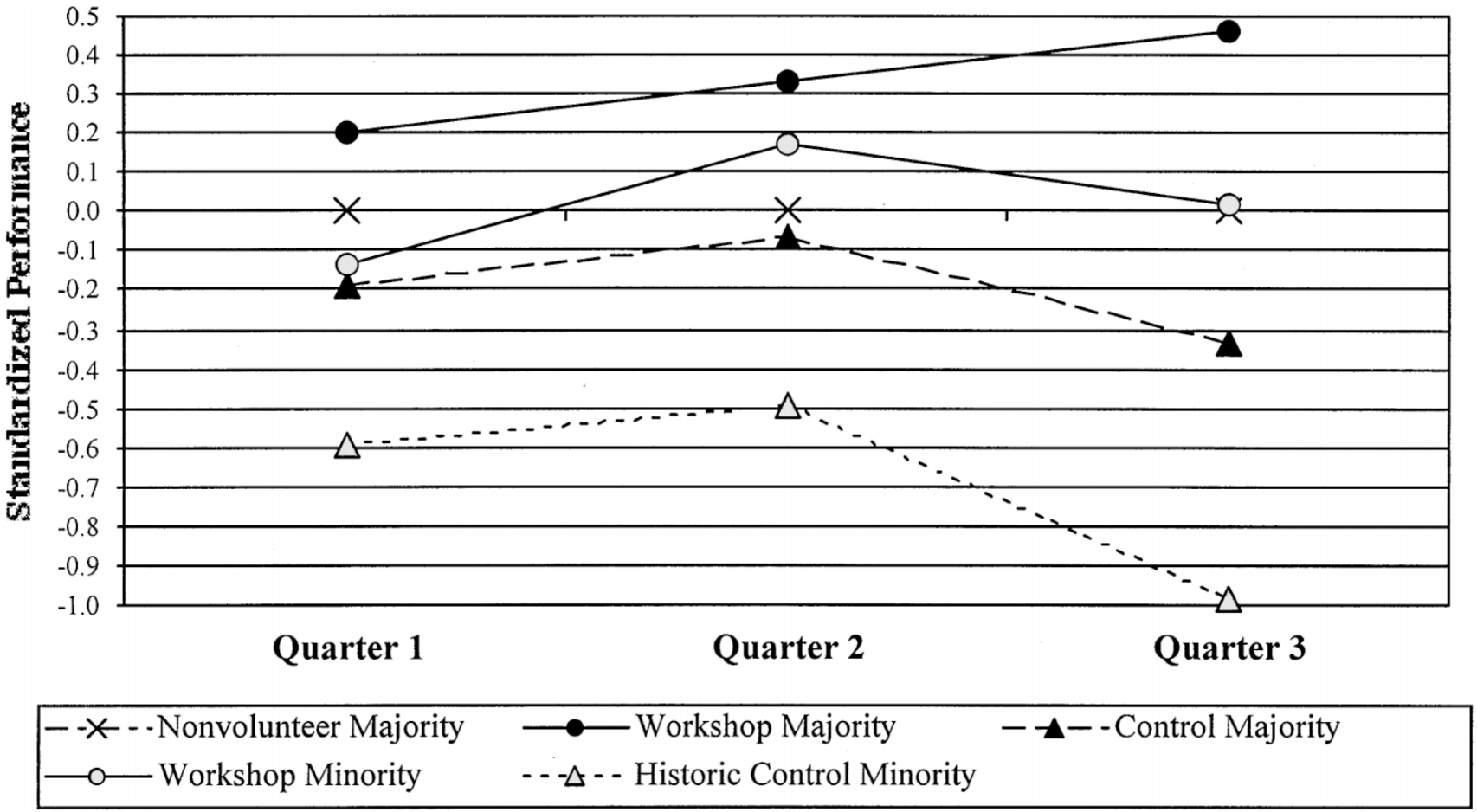


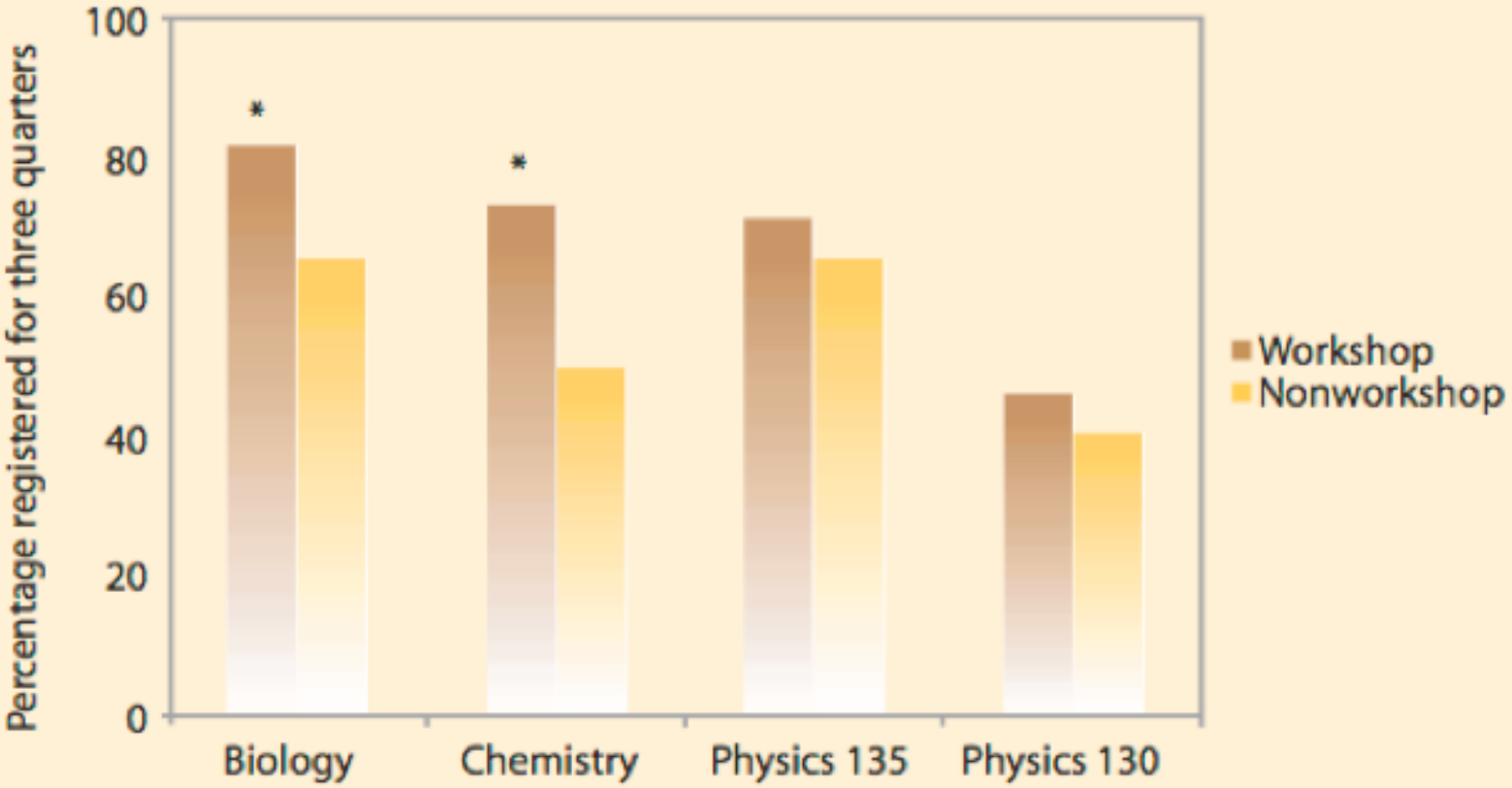
Fig. 2. Standardized performance in each quarter as a function of group, controlling for prior cumulative grade point average.

(Born et al., 2002)

Swarat, S. and Drane, D. and Smith, H.D. and Light, G. and Pinto, L., 2004

FIGURE 2

2001–2002 all student retention over three quarters by workshop participation and discipline.



Retention is defined as the proportion of students who complete all three quarters of the course sequence

*Significant chi square, $P < 0.05$

Important problems can be examined by the use of randomized field trials

1. Observational studies are useful, but are very susceptible to problems of subject participation.
2. Marriage leads to happiness, but only amongst those who stay married.
3. Healthy diets prolong life, but perhaps for those who do other healthy activities.
4. People who volunteer for studies are different than those who do not.

Born, W. K., Revelle, W., & Pinto, L. H. (2002). Improving biology performance with workshop groups. *Journal of Science Education and Technology*, *11*(4), 347–365.

Davis, C. E. (2002). Regression to the mean or placebo effect? In H. A. Guess (Ed.), *The science of the placebo: Toward an interdisciplinary research agenda* chapter 7, (pp. 158–166). London: BMJ Books.

Deary, I. J. & Batty, G. D. (2007). Cognitive epidemiology. *British Medical Journal*, *61*(5), 378–384.

Factor-Litvak, P. R. & Krebs, C. (2019). Lead poisoning and intelligence: A search for cause and effect in the scottish mental surveys. *Journal of Environmental and Public Health*, *2019*, 8980604.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*(1), 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*(2), 171–191.

Gallagher, E. J., Viscoli, C. M., & Horwitz, R. I. (1993). The relationship of treatment adherence to the risk of death after myocardial infarction in women. *JAMA*, *270*(6), 742–744.

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behavior: Conduct disorder, delinquency, and violence in the Dunedin longitudinal study*. Cambridge, England: Cambridge University Press.

Reuben, A., Caspi, A., Belsky, D. W., Broadbent, J., Harrington, H., Sugden, K., Houts, R. M., Ramrakha, S., Poulton, R., & Moffitt, T. E. (2017). Association of Childhood Blood Lead Levels With Cognitive Function and Socioeconomic Status at Age 38 Years and With IQ Change and Socioeconomic Mobility Between Childhood and Adulthood. *JAMA*, *317*(12), 1244–1251.

Taubs, G. (2007). Do we really know what makes us healthy".