
Male and female intelligence. Males and females tend to exhibit different types of intelligence. Although there is substantial variability between individuals of the same gender, males on average tend to perform better at navigational and spatial tasks, and females tend to perform better at verbal fluency and memory tasks. This is not a controversial conclusion. Some researchers, however, ask whether males and females differ, on average, in their *overall* intelligence, and that *is* controversial for the following reason: any single measure must rely on premises about the types of intelligence that are important. Even if the researchers don't make a subjective judgement about a type of intelligence being tested, they are constrained by the available tools for measuring intelligence. Mathematical knowledge is easy to test, for example, but wisdom, creativity, practical knowledge, and social skill are not.

Display 2.19 shows a partial listing of a data set with several intelligence test scores for roughly 6,000 American men and 6,000 American women between the ages of 16 and 24 in 1981. The column AFQT shows the scores on the Armed Forces Qualifying Test. This test is designed for evaluating the suitability of military recruits, but it has also been used as a general intelligence test. The AFQT score is actually a combination of scores from four component tests: Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Mathematical Knowledge. The data set represented in Display 2.19 includes each individual's score on these components. (The overall AFQT score reported here, called AFQT89, is based on a non-trivial combination of the component scores; it's not simply the sum.)

Treating the males and females as samples from populations, analyze the data to see whether male and female distributions of AFQT scores differ and by how much they differ. Then do the same for each of the four components. What conclusion can you draw from all five of these comparisons about male and female differences and about the suitability of AFQT as an overall measure of intelligence? Write a statistical report that includes graphical displays, a summary of statistical conclusions (like those in the case studies of Section 2.1), and a section of details upon which the conclusions were based (such as a listing of the computer output showing the results of two-sample *t*-tests and confidence intervals).

Notes: Although these are random samples of American men and women between the ages of 16 and 24, they are not simple random samples. The data come from the National Longitudinal Study of Youth (NLSY), which uses variable probability sampling (see Section 1.5.4). To estimate the means of the larger populations, it is inappropriate to treat them as simple random samples; more advanced techniques are needed. For the purpose of *comparing* male and female distributions, however, treating them as simple random samples will not produce misleading conclusions. The data come from the National Longitudinal Survey of Youth, U.S. Bureau of Labor Statistics, <http://www.bls.gov/nls/home.htm> (8 May 2008). The AFQT scores were computed from the raw component test scores using the formula for AFQT89.

Display 2.19

First 20 rows from a data set that contains the Armed Forces Qualification Test (AFQT) intelligence test scores and component test scores in arithmetic reasoning, word knowledge, paragraph comprehension, and mathematical knowledge for 5,927 U.S. women and 5,951 U.S. men in 1981.

sex	afqt	arith	word	parag	math
male	122	34	22	32	38
female	117	36	24	26	37
male	122	35	25	26	37
male	122	36	22	20	44
female	116	31	25	29	35
male	121	39	24	20	38
female	116	32	20	29	40
male	116	39	24	26	33
female	109	35	21	29	30
male	107	32	20	23	33
female	115	30	24	29	35
male	121	36	24	29	35
male	117	35	25	23	38
female	102	30	20	26	30
female	121	32	28	26	35
female	115	32	24	32	33
female	112	30	21	29	38
female	109	30	21	29	35
female	116	34	24	26	38
male	112	31	24	20	37

afqt is the AFQT intelligence test score, which is a weighted average of scaled versions of component test scores in arithmetic reasoning (*arith*), word knowledge (*word*), paragraph comprehension (*parag*), and mathematical knowledge (*math*)