Measurement and scaling

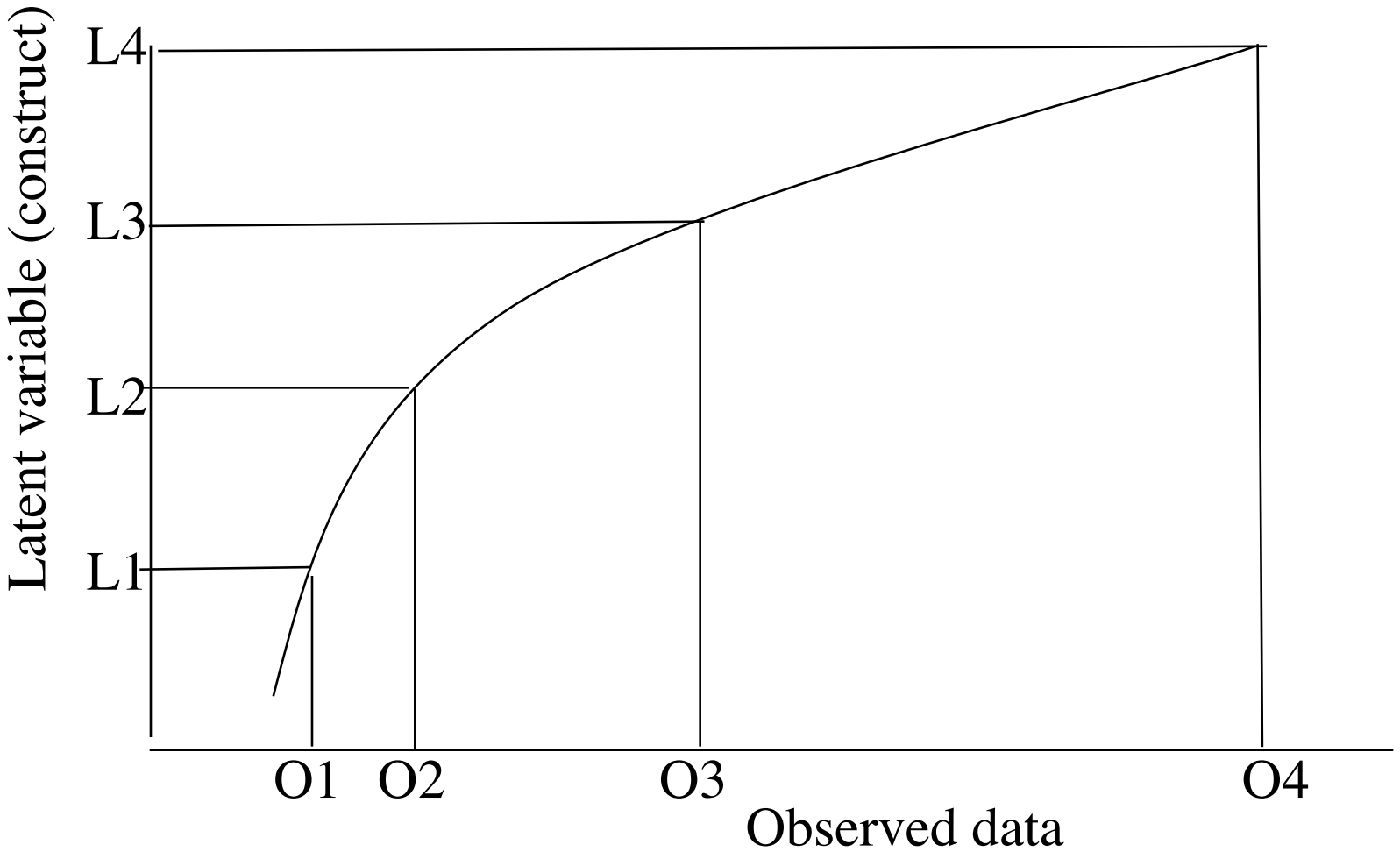
Inferring latent values from observed values

Types of Scales: Inferences from observed variables to Latent variables

- Nominal
- Ordinal
- Interval
- Ratio

- Categories
- Ranks (x > y)
- Differences
  - X-Y > W-V
- Equal intervals with a zero point =>
  - X/Y > W/V

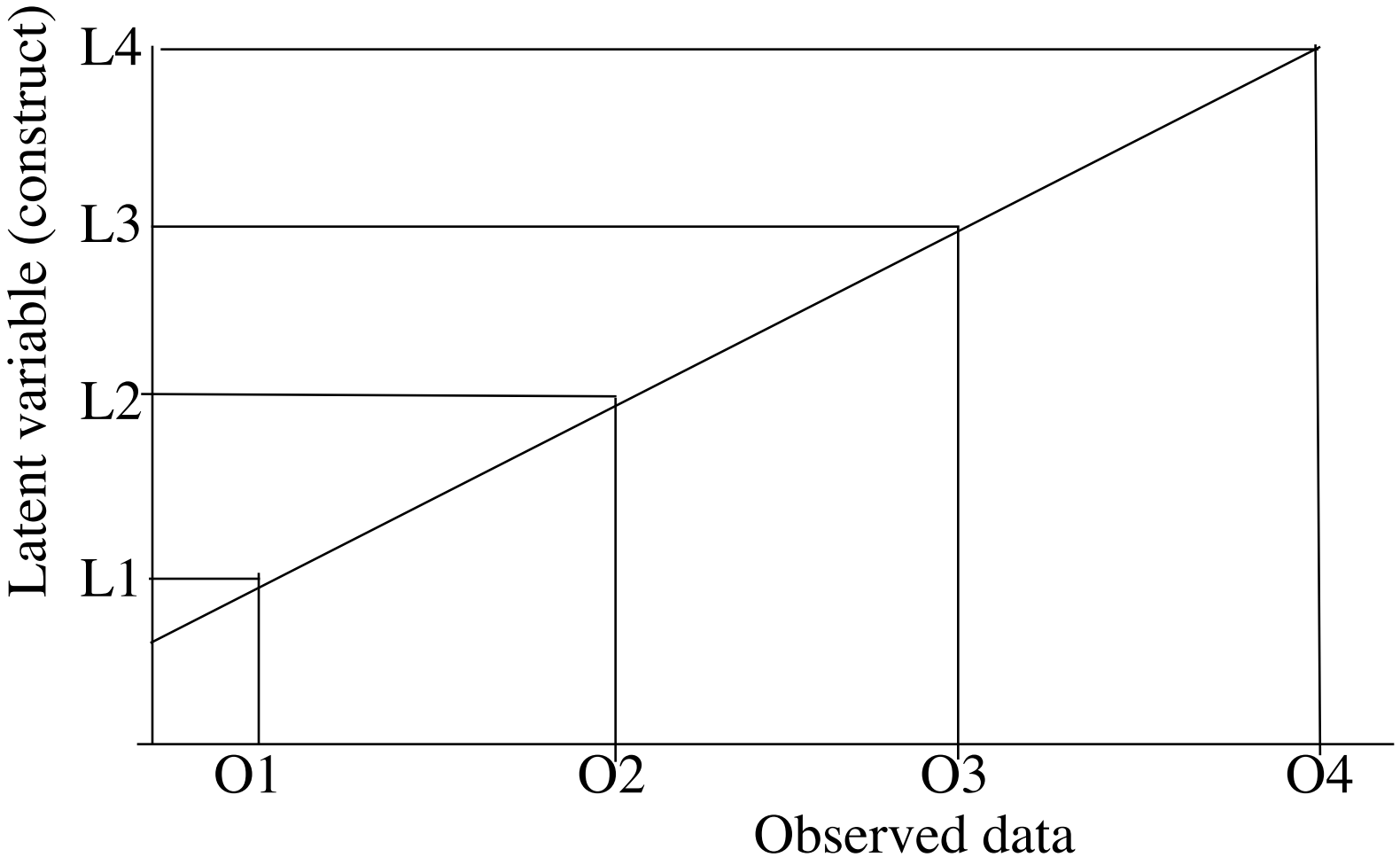
Mappings and inferences
Ordinal scales Ranks (x > y)



Ordinal Scales

- Any monotonic transformation will preserve order
- Inferences from observed to latent variable are restricted to rank orders
- Statistics: Medians, Quartiles, Percentiles

Mappings and inferences interval scales allow for a comparison of distances X-Y > W-V

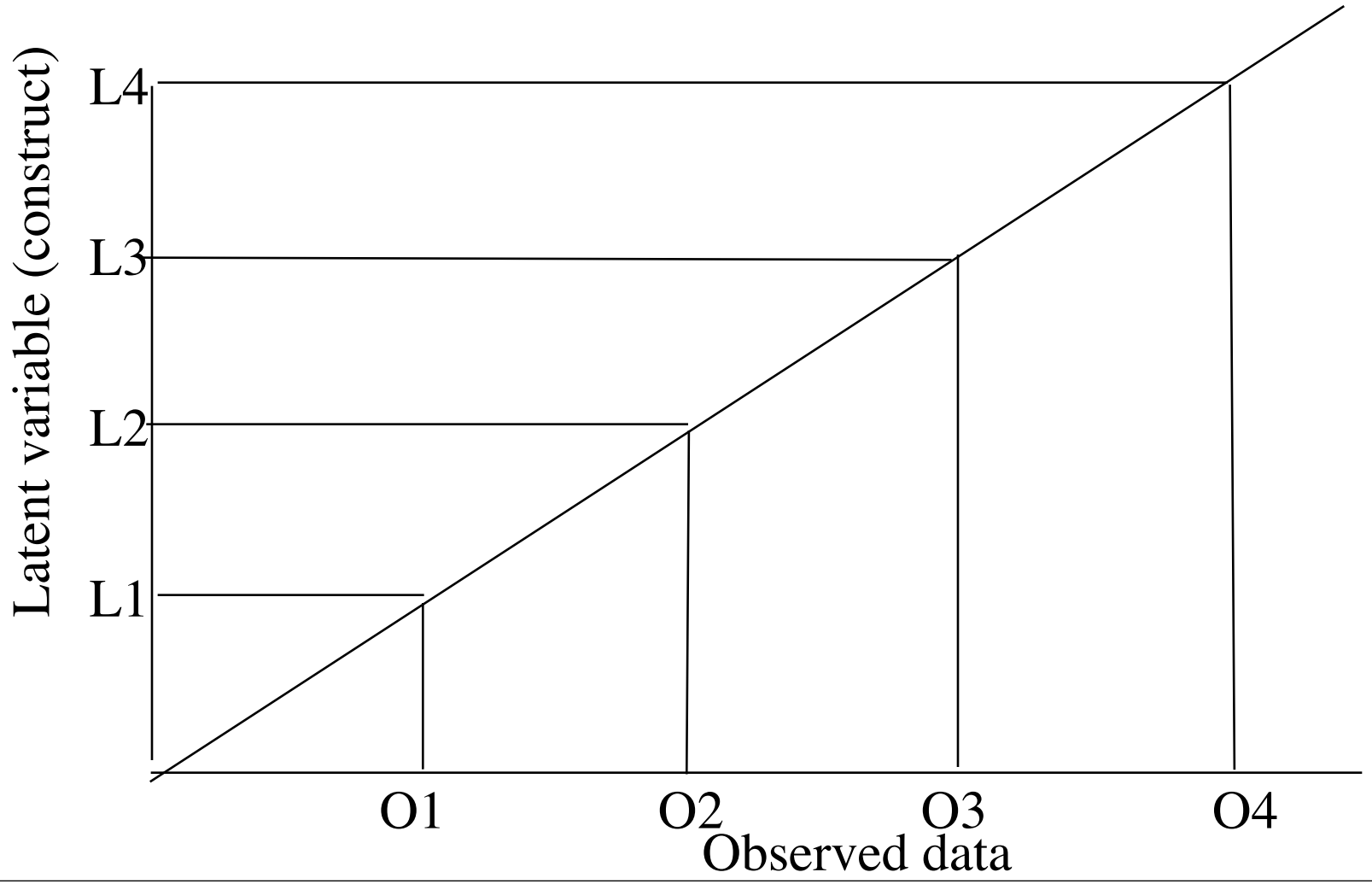


Interval Scales

- Possible to infer the magnitude of differences between points on the latent variable given differences on the observed variable
  X is as much greater than Y as Z is from W
- Linear transformations preserve interval information
- Allowable statistics: Means, Variances

Mappings and inferences

ratio scales allow ratio comparisons $X/Y > W/V$



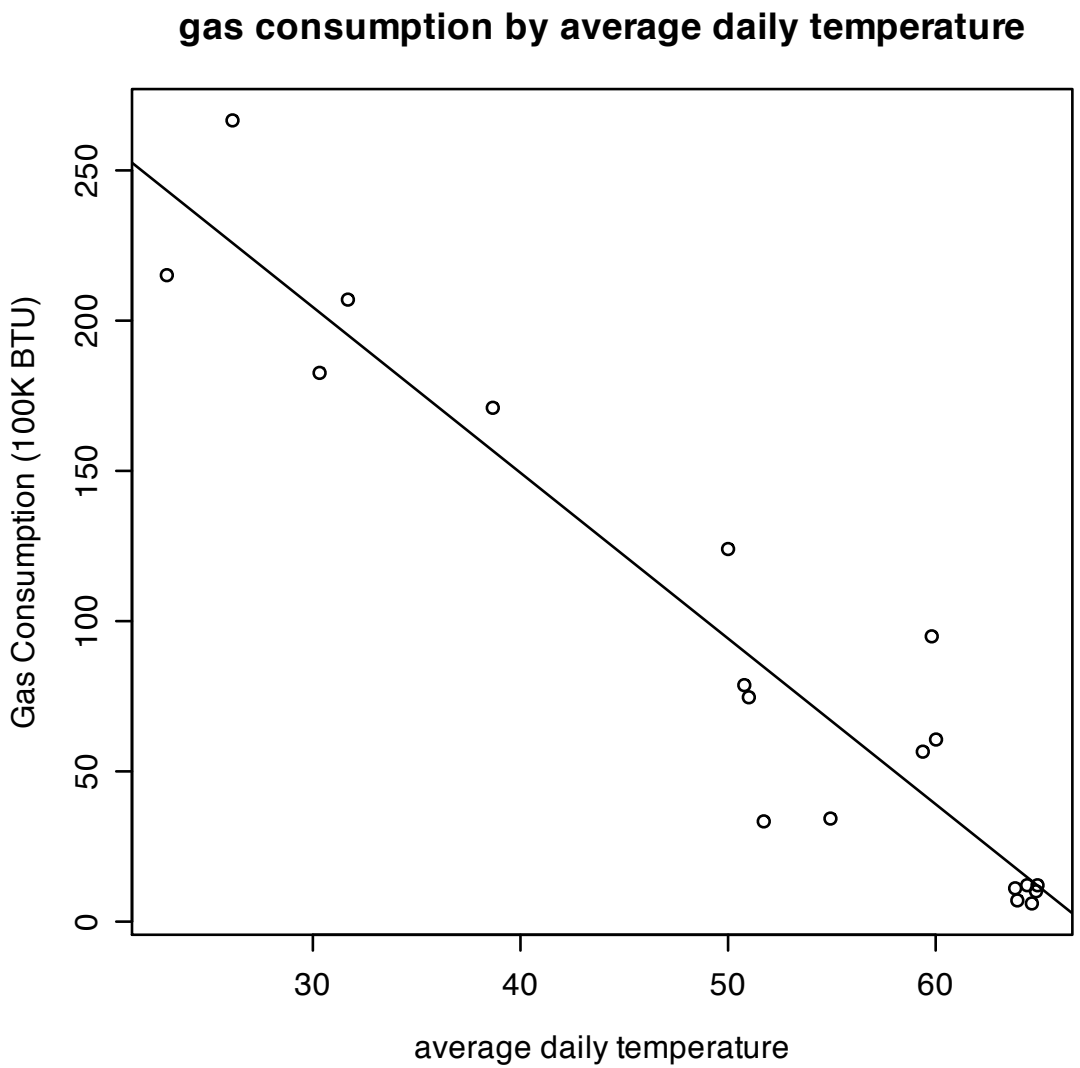
Ratio Scales

- Interval scales with a zero point
- Possible to compare ratios of magnitudes (X is twice as long as Y)

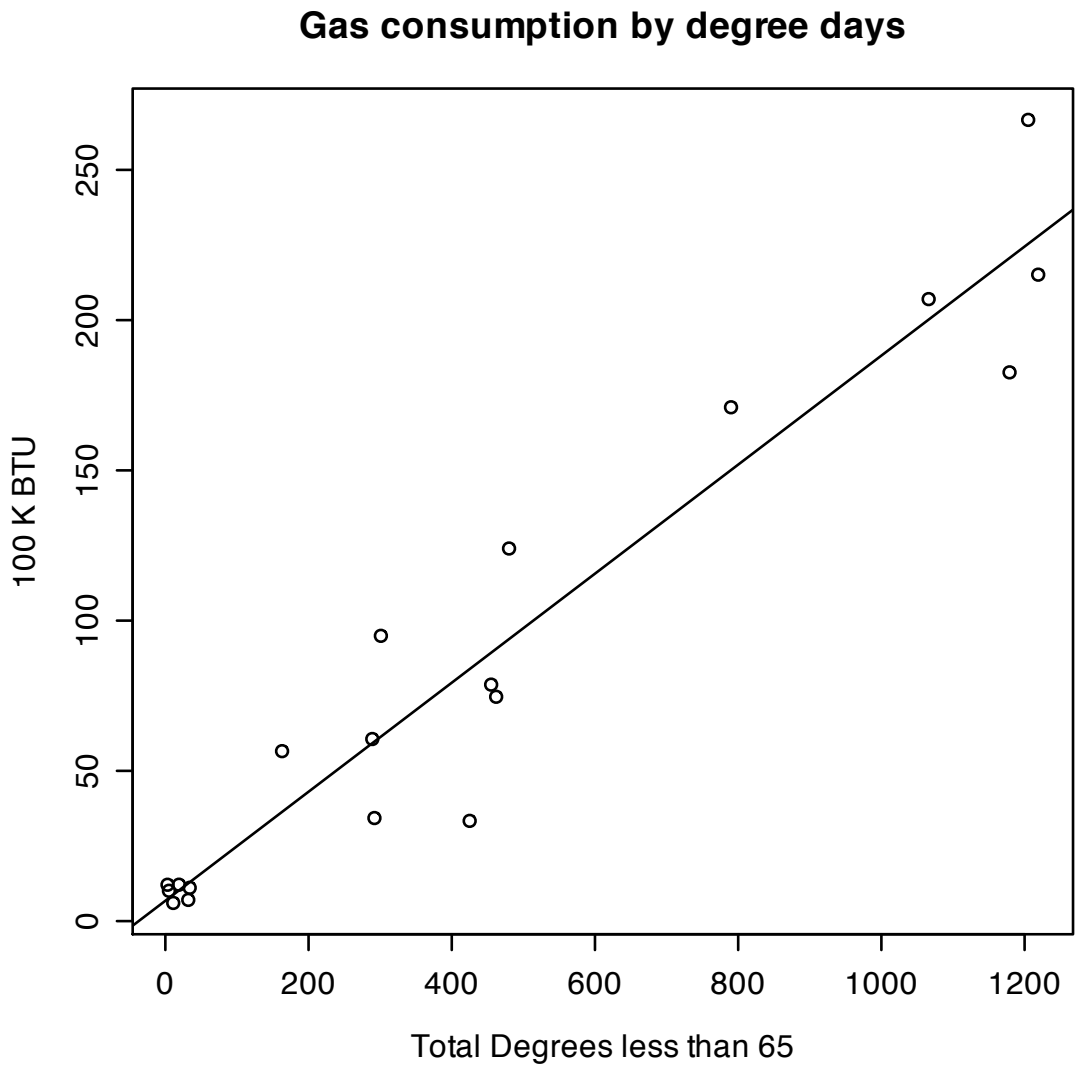
The search for appropriate scale

- Is today colder than yesterday? (ranks)
- Is the amount that today is colder than yesterday more than the amount that yesterday was colder than the day before? (intervals)
  - 50 F - 39 F < 68 F - 50 F
  - 10 C - 4 C < 20 C - 10 C
  - 283K - 277K < 293K -283K
- How much colder is today than yesterday?
  - (Degree days as measure of energy use)
  - K as measure of molecular energy

Gas consumption by average temperature



Gas consumption by degree days is practically a ratio scale



Latent and Observed Scores
The problem of scale

Much of our research is concerned with making inferences about latent (unobservable) scores based upon observed measures. Typically, the relationship between observed and latent scores is monotonic, but not necessarily (and probably rarely) linear. This leads to many problems of inference. The following examples are abstracted from real studies. The names have been changed to protect the guilty.

Effect of teaching upon performance

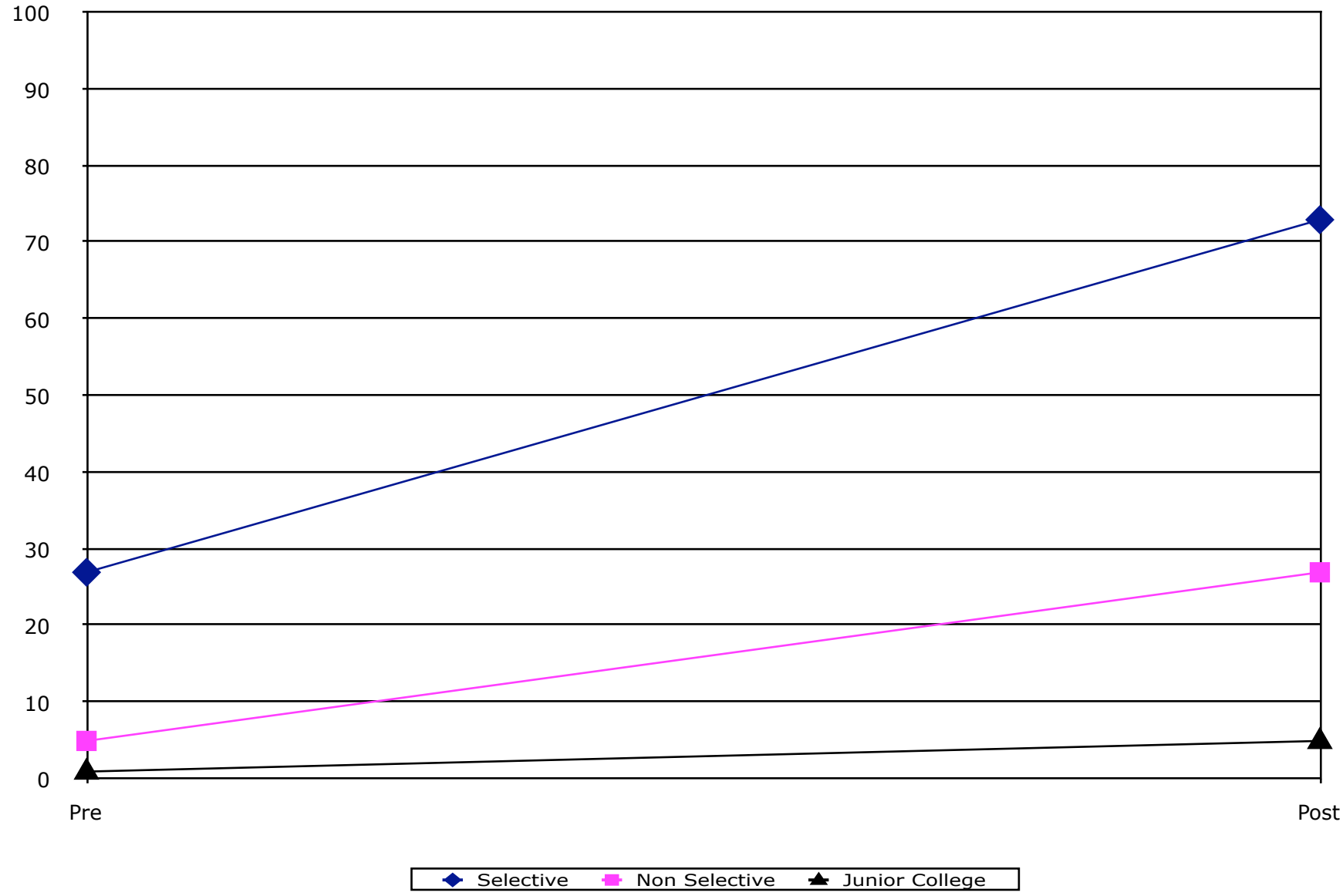
- A leading research team in motivational and educational psychology was interested in the effect that different teaching techniques at various colleges and universities have upon their students. They were particularly interested in the effect upon writing performance of attending a very selective university, a less selective university, or a two year junior college. A writing test was given to the entering students at three institutions in the Boston area. After one year, a similar writing test was given again. Although there was some attrition from each sample, the researchers report data only for those who finished one year. The pre and post test scores as well as the change scores were as shown below:

Effect of teaching upon performance

	Pretest	Posttest	Change
Junior College	1	5	4
Non-selective university	5	27	22
Selective university	27	73	45

From these data, the researchers concluded that the quality of teaching at the very selective university was much better and that the students there learned a great deal more. They proposed to study the techniques used there in order to apply them to the other institutions.

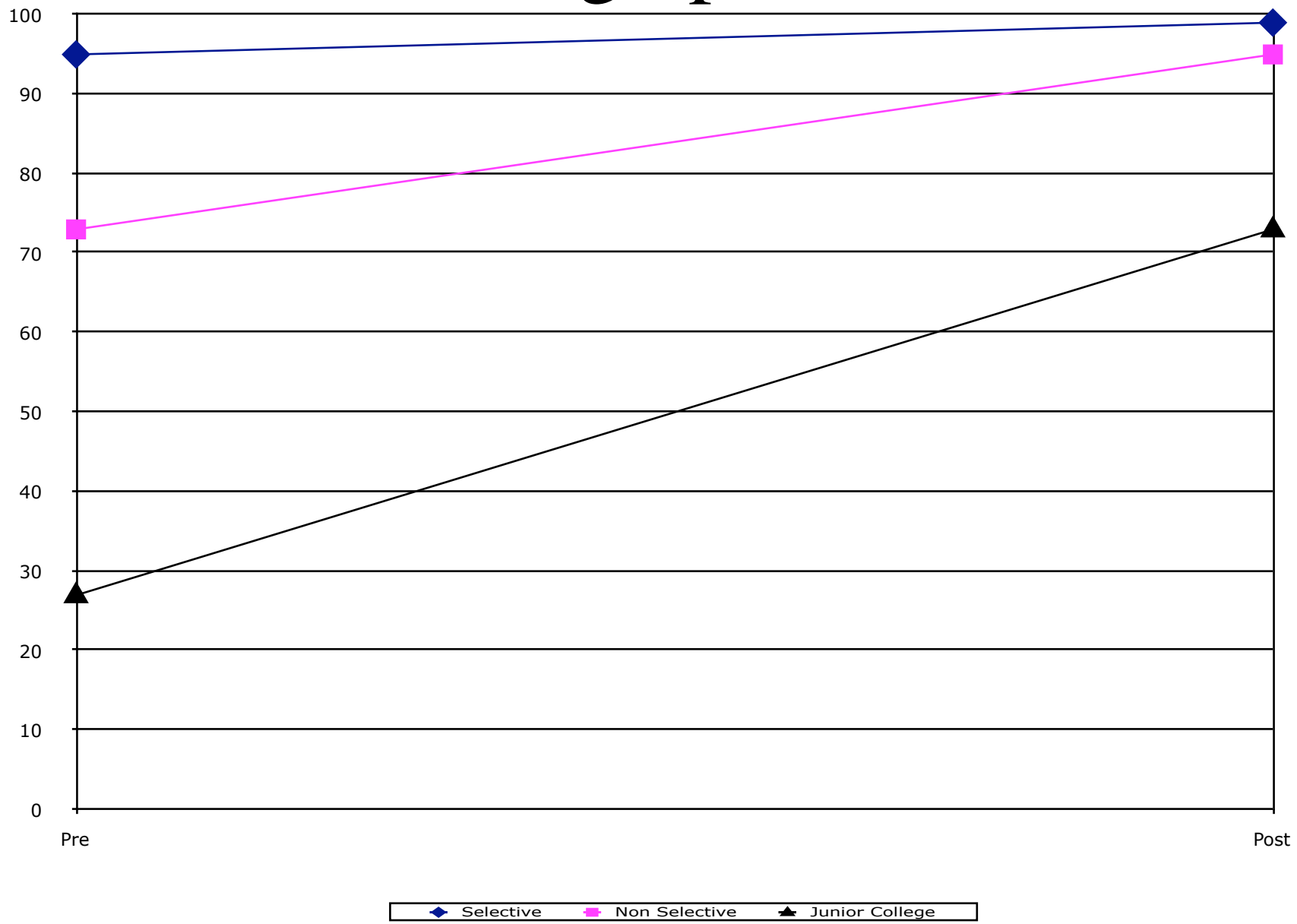
Effect of Teaching upon Performance?



Another research team in motivational and educational psychology was interested in the effect that different teaching techniques at various colleges and universities have upon their students. They were particularly interested in the effect upon mathematics performance of attending a very selective university, a less selective university, or a two year junior college. A math test was given to the entering students at three institutions in the Boston area. After one year, a similar math test was given again. Although there was some attrition from each sample, the researchers report data only for those who finished one year. The pre and post test scores as well as the change scores were:

	Pretest	Posttest	Change
Junior College	27	73	45
Non-selective university	73	95	22
Selective university	95	99	4

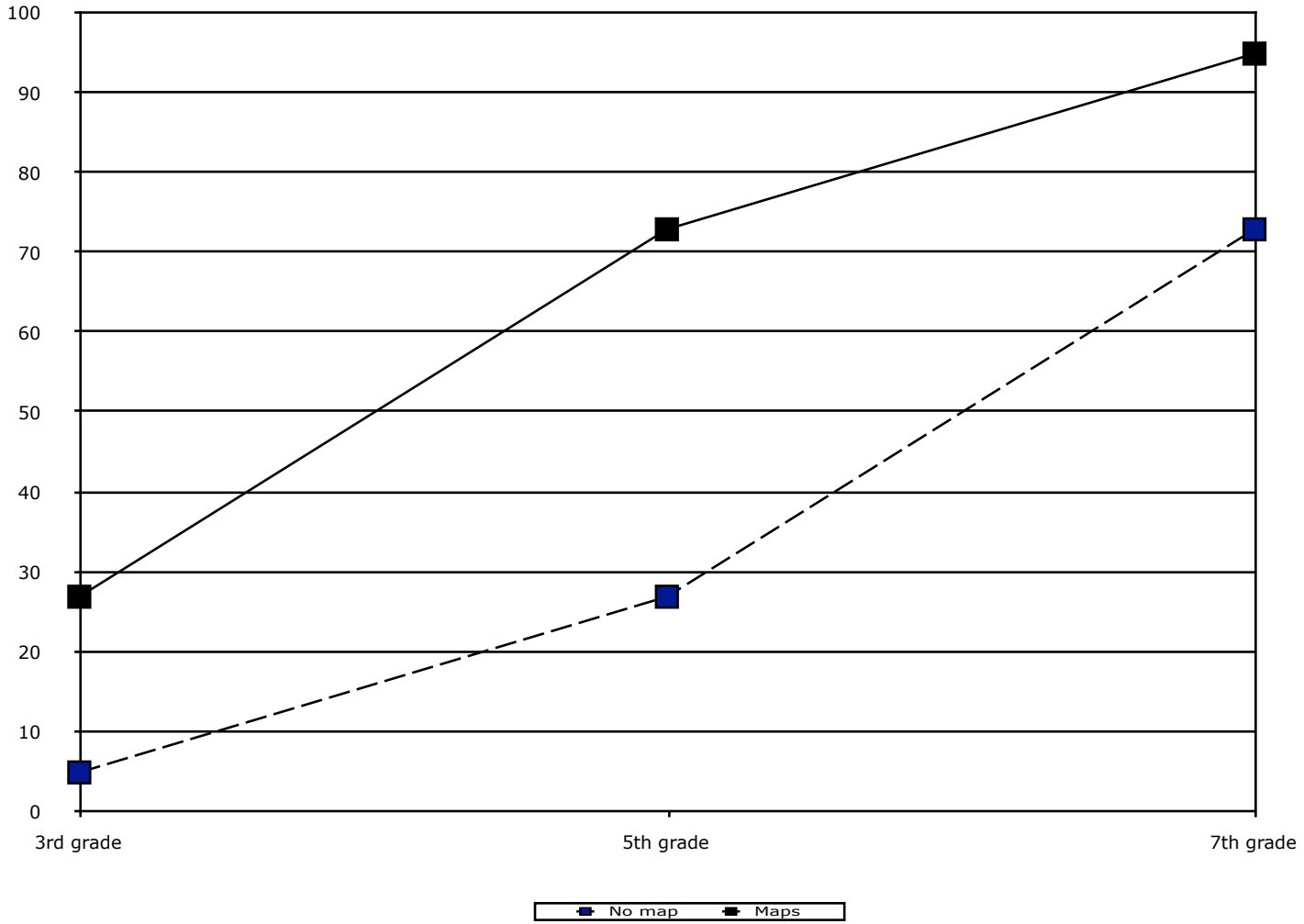
Effect of Teaching upon Performance?



A leading cognitive developmentalist believed that there is a critical stage for learning spatial representations using maps. Children younger than this stage are not helped by maps, nor are children older than this stage. He randomly assigned 3rd, 5th, and 7th grade students into two conditions (nested within grade), control and map use. Performance was measured on a task of spatial recall (children were shown toys at particular locations in a set of rooms and then asked to find them again later. Half the children were shown a map of the rooms before doing the task.

	No map	Maps
3rd grade	5	27
5th grade	27	73
7th grade	73	95

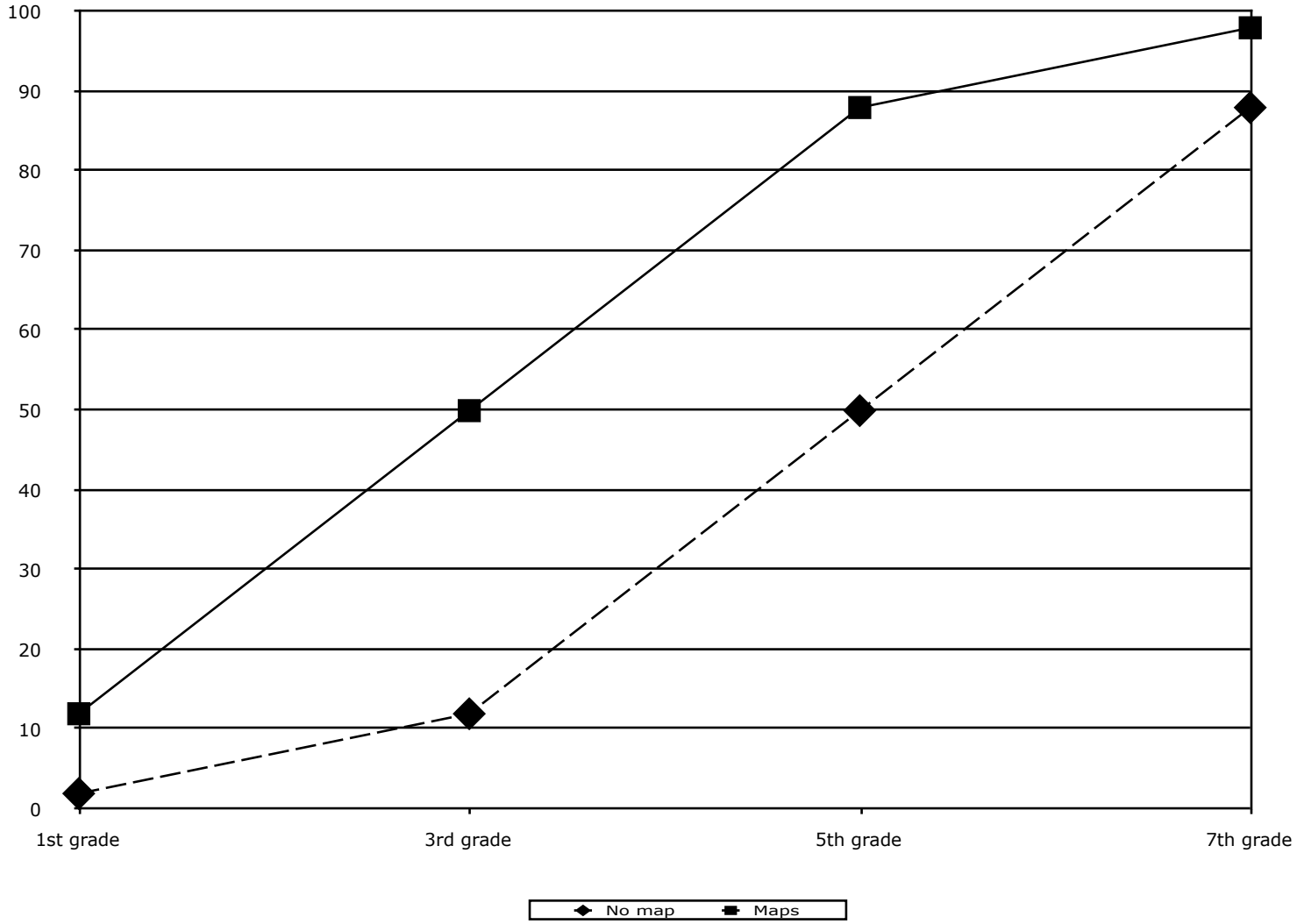
Spatial reasoning facilitated by maps at a critical age



Another cognitive developmentalist believed that there is a critical stage but that it appears earlier than previously thought. Children younger than this stage are not helped by maps, nor are children older than this stage. He randomly assigned 1st, 3rd, 5th, and 7th grade students into two conditions (nested within grade), control and map use. Performance was measured on a task of spatial recall (children were shown toys at particular locations in a set of rooms and then asked to find them again later. Half the children were shown a map of the rooms before doing the task.

	No map	Maps
1st grade	2	12
3rd grade	12	50
5th grade	50	88
7th grade	88	98

Spatial Reasoning is facilitated by map use at a critical age

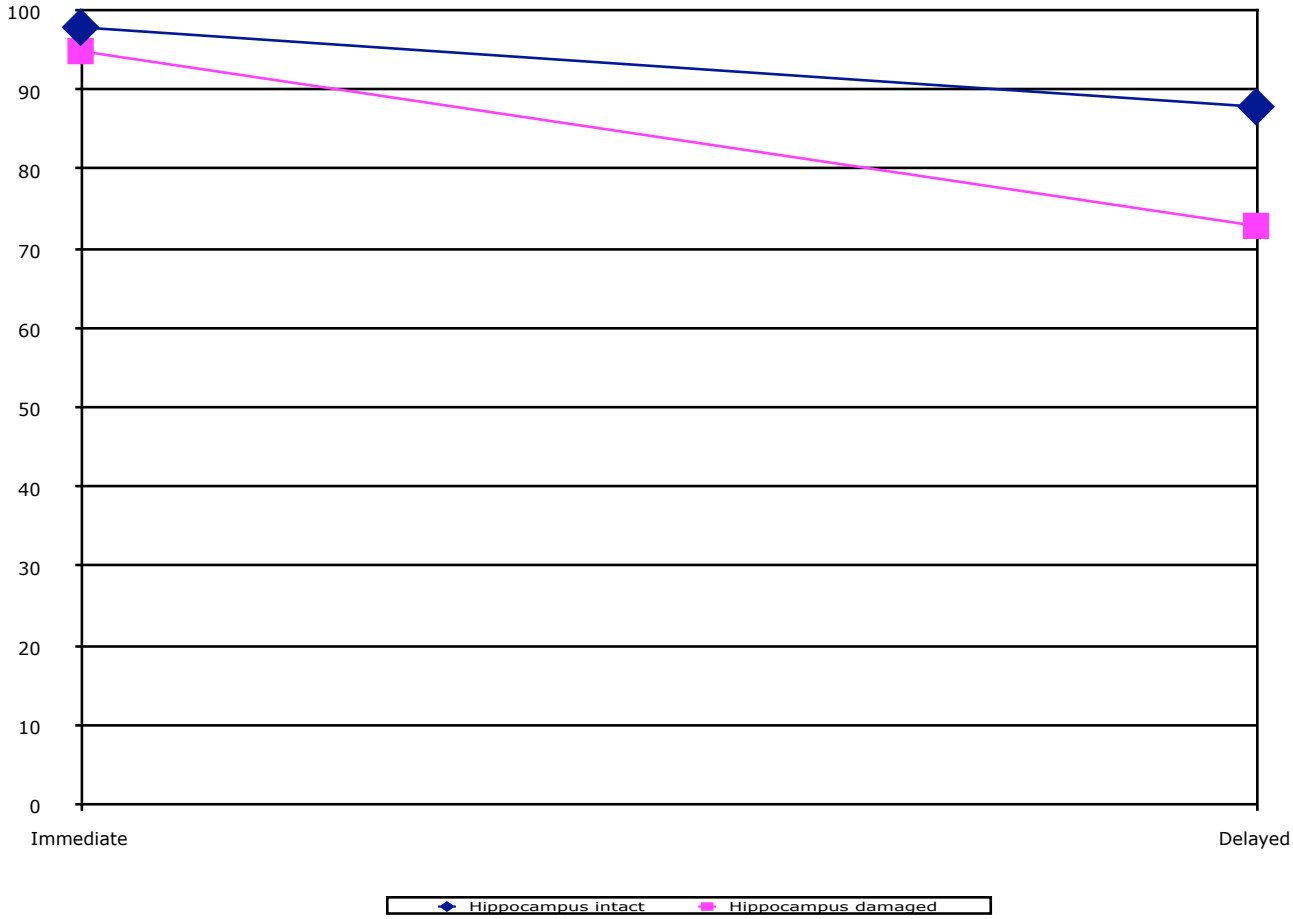


Cognitive-neuro psychologists believe that damage to the hippocampus affects long term but not immediate memory. As a test of this hypothesis, an experiment is done in which subjects with and without hippocampal damage are given an immediate and a delayed memory task. The results are impressive:

	Immediate	Delayed
Hippocampus intact	98	88
Hippocampus damaged	95	73

From these results the investigator concludes that there are much larger deficits for the hippocampal damaged subjects on the delayed rather than the immediate task. The investigator believes these results confirm his hypothesis. Comment on the appropriateness of this conclusion.

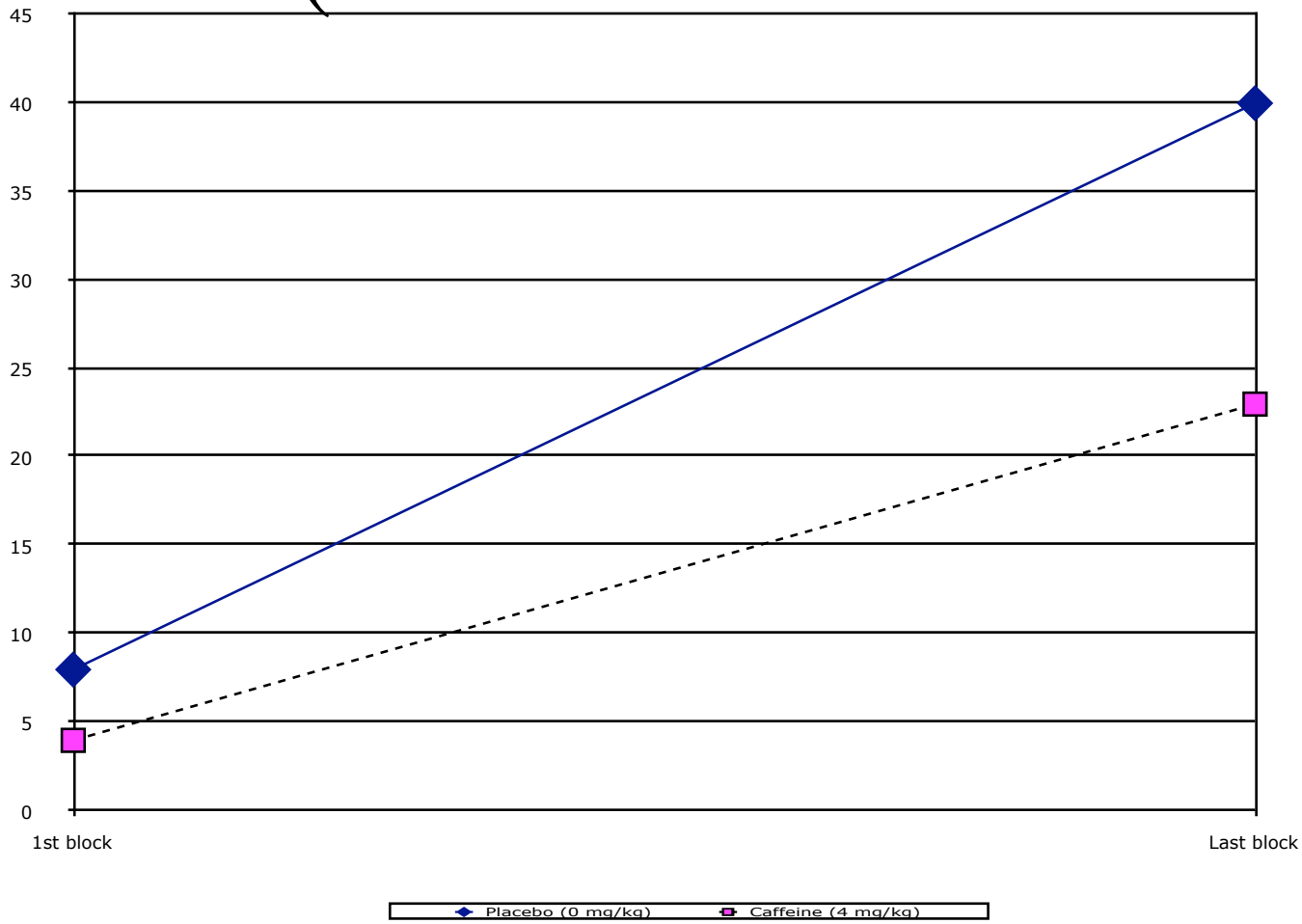
Memory = f(hippocampal damange * temporal delay)



An investigator believes that caffeine facilitates attentional tasks such that require vigilance. Subjects are randomly assigned to conditions and receive either 0 or 4mg/kg caffeine and then do a vigilance task. Errors are recorded during the first 5 minutes and the last 5 minutes of the 60 minute task. The number of errors increases as the task progresses but this difference is not significant for the caffeine condition and is for the placebo condition.

	1st block	Last block
Placebo (0 mg/kg)	8	40
Caffeine (4 mg/kg)	4	23

Errors=f(caffeine * time on task)



Arousal is a fundamental concept in many psychological theories. It is thought to reflect basic levels of alertness and preparedness. Typical indices of arousal are measures of the amount of palmer sweating. This may be indexed by the amount of electricity that is conducted by the fingertips. Alternatively, it may be indexed (negatively) by the amount of skin resistance of the finger tips. The Galvanic Skin Response (GSR) reflects moment to moment changes, SC and SR reflect longer term, basal levels.

High skin conductance (low skin resistance) is thought to reflect high arousal.

Anxiety is thought to be related to arousal. The following data were collected by two different experimenters. One collected Resistance data, one conductance data.

	Resistance	Conductance
Anxious	2, 2	.5, .5
Low anx	1, 5	1, .2

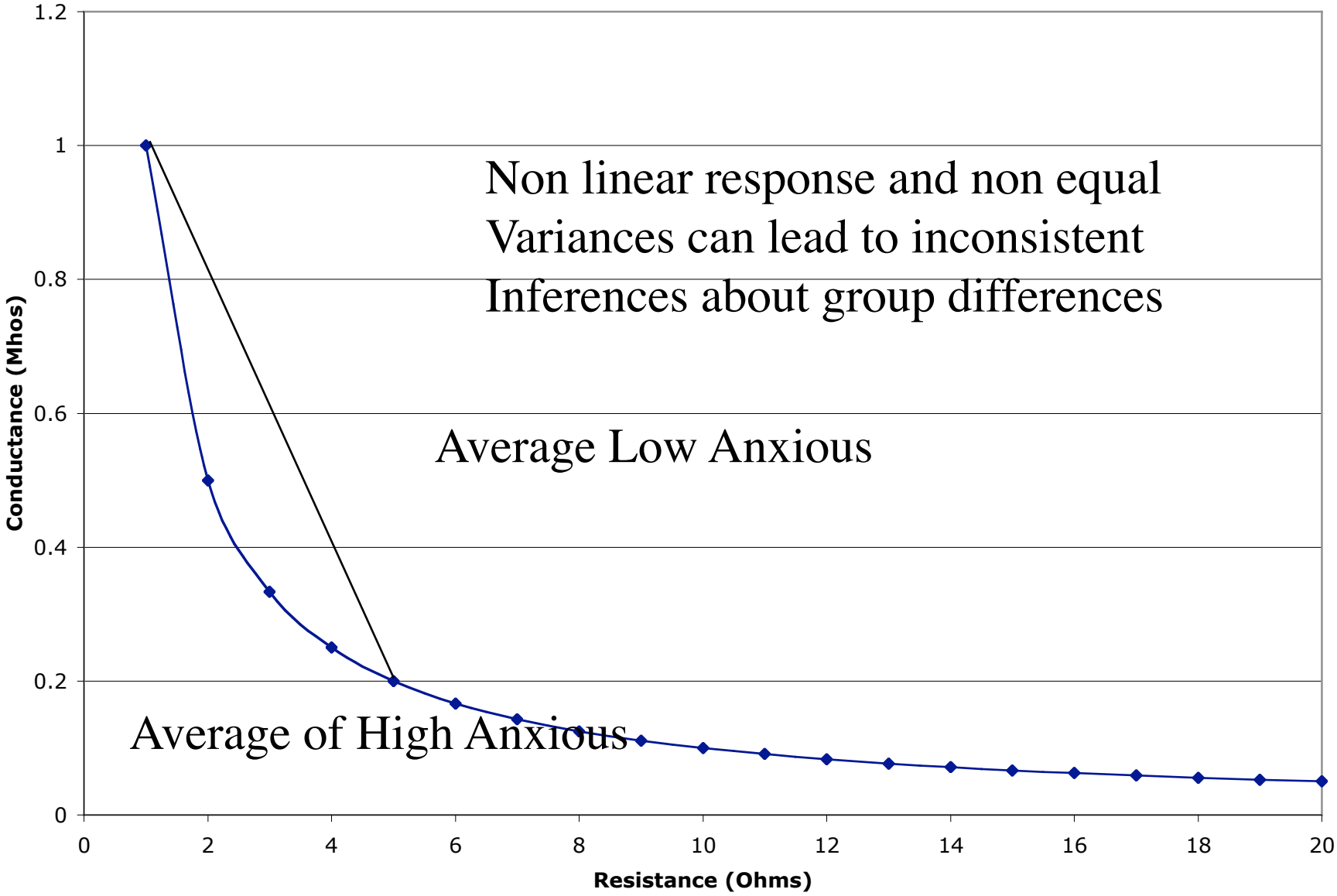
The means were

	Resistance	Conductance
Anxious	2	.5
Low anx	3	.6

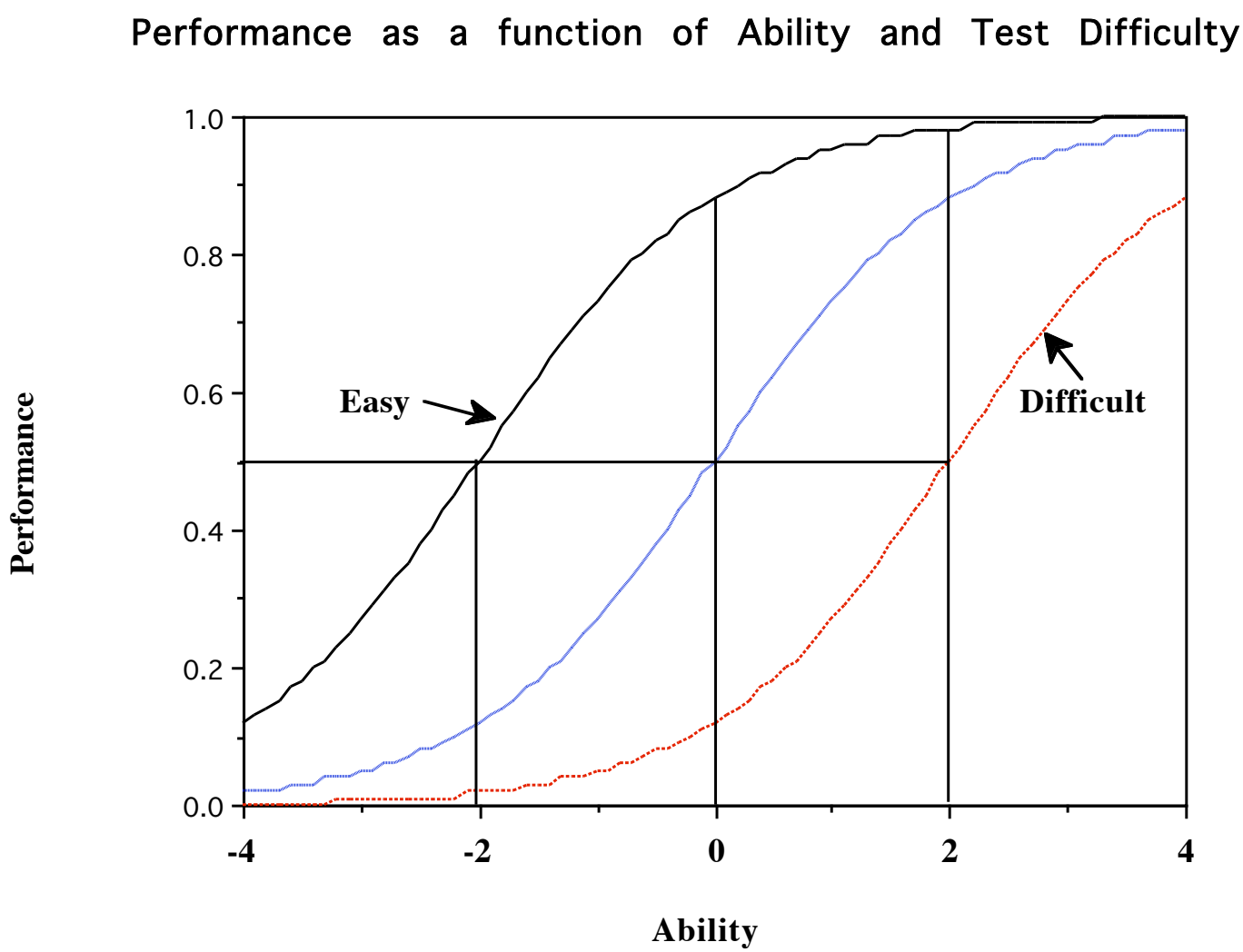
Experimenter 1 concluded that the low anxious had higher resistances, and thus were less aroused. But experimenter 2 noted that the low anxious had higher levels of skin conductance, and were thus more aroused.

How can this be?

Conductance = 1/Resistance



Performance and task difficulty



Performance, ability, and task difficulty

	Difficulty				
	-2	-1	0	1	2
Latent Ability					
-4.00	0.12	0.05	0.02	0.01	0.00
-2.00	0.50	0.27	0.12	0.05	0.02
0.00	0.88	0.73	0.50	0.27	0.12
2.00	0.98	0.95	0.88	0.73	0.50
4.00	1.00	0.99	0.98	0.95	0.88
Change from					
-4 to -2	0.38	0.22	0.10	0.04	0.02
-2 to -0	0.38	0.46	0.38	0.22	0.10
0 to 2	0.10	0.22	0.38	0.46	0.38
2 to 4	0.02	0.04	0.10	0.22	0.38

Performance and Task Difficulty

Note that equal differences along the latent ability dimension result in unequal differences along the observed performance dimension. Compare particularly performance changes resulting from ability changes from -2 to 0 to 2 units.

This is taken from the standard logistic transformation used in Item Response Theory that maps latent ability and latent difficulty into observed scores. IRT attempts to estimate difficulty and ability from the observed patterns of performance.

$$\text{Performance} = 1/(1+\exp^{(\text{difficulty-ability})})$$