2.6: The Role of Chance

“The human understanding supposes a greater degree of order and equality in things than it really finds; and although many things in nature be sui generis and most irregular, will yet invest parallels and conjugates and relatives where no such thing exists.” — Francis Bacon

Focus:
Distinguishing between signal and noise

Main learning points:

1. Gilovich: Random data is usually misperceived - tend to see patterns in random noise
   - Clustering illusion – seeing clusters as structure — e.g. “hot hands”
     - cause: representative heuristic - like goes with like
   - Causality illusion – attributing correlation to some cause
     - cause: confabulation - need to invoke an agency of some kind
   - Regression fallacy – expecting extremes to stay extreme
     - cause: prediction should resemble the predictor
       - if an event is an extreme, the subsequent event should stay extreme, rather than have (on average) the mean value.

2. Abelson: Can test systematically whether chance is the responsible factor. Learning to do this correctly requires careful thinking. People are bad at thinking about statistics:
   a) Statistical tests always require a comparison — often with results of chance
      - explanation is a difference that makes a difference
   b) Look at whether chance can explain given data – null hypothesis
      - check on quality of evidence
      - note: language is misleading – not really accept/reject
        - better to say that test supports / discredits the hypothesis
        - note: implicit weighting of relative likelihoods – not an absolute
      - also, need to keep context in mind
        - was the data specially selected?
        - were the categories specially selected?

Remedy: once categories are found, test using new, independent data