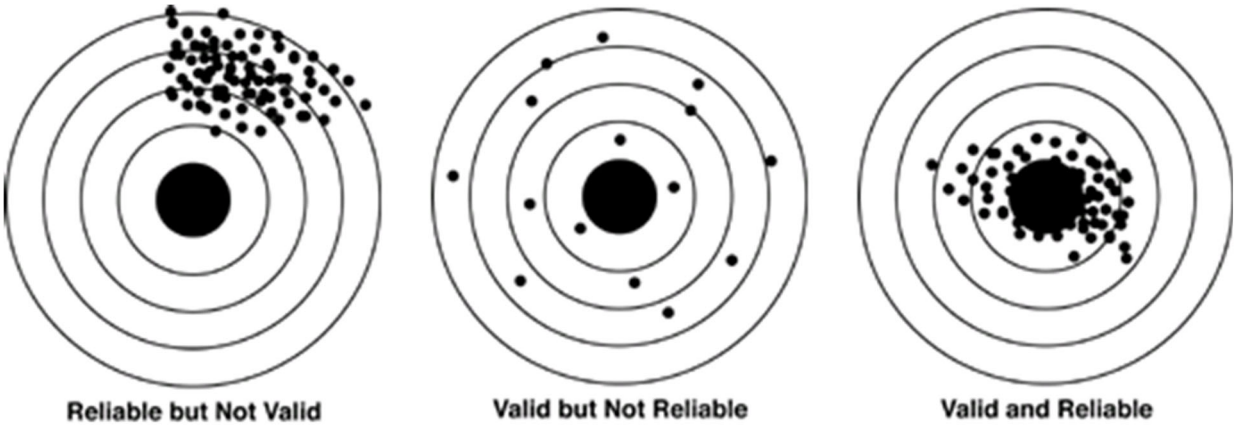
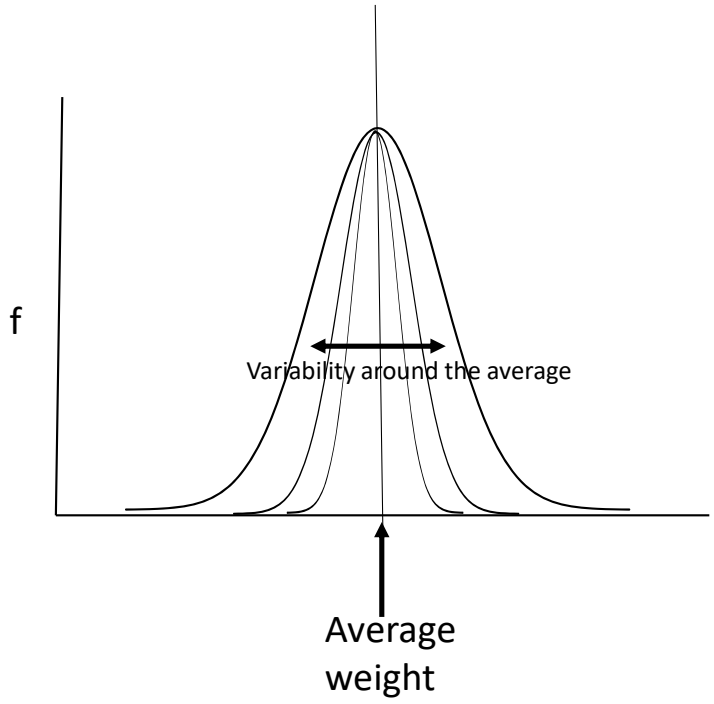
Reliability vs. Validity

Target analogy for reliability and validity. Imagine a dartboard. When the throws are highly consistent over time, they are reliability. When they are more scattered, they are less reliable. You could potentially have reliable throws but they tend to miss on average, so they would be not valid (or perhaps biased in this case)



Effect of Measurement Error on Variance



Remember that random measurement error sometimes leads to overestimates and sometimes leads to underestimates. On average the estimate will be accurate.

Spearman-Brown Prophecy Formula

Measures with more items have higher reliability. The Spearman-Brown formula estimates the reliability if the same average inter-item correlation was obtained but there were more items in the test.

$$R_{xx-revised} = \frac{nR_{xx-original}}{1+(n-1)R_{xx-original}}$$

n is the factor by which the size is increased

If length is increased from a 10-item test is increased to 20 items (with the same average inter-item correlation), n = 2, because the length is increased by a factor of 2

Example:
If length is increased from a 10-item test is increased to 20 items (with the same average inter-item correlation), n = 2, because the length is increased by a factor of 2. Assume the original reliability $R_{xx\text{-original}}$ is .6.

$$R_{xx-revised} = \frac{nR_{xx-original}}{1+(n-1)R_{xx-original}}$$

$$= \frac{2(.6)}{1+(2-1).6} = \frac{1.2}{1.6} = .75$$

The reliability coefficient is not biased if it has fewer items, but with just a few items, it may be more difficult to obtain an “acceptable” level of reliability. It makes sense that more repeated measurements (more items) increases the precision of a measure.