

## DataDirector™ Test Statistics Glossary

### Terminology

Median score:	the center of a distribution of scores; to find the median score put all test scores in order from lowest to highest and identify the middle score
Mean score:	equal to the total divided by the number of tests; to find the mean score add all test scores together and divide by the number of tests
Percentile:	indication of how data values are spread over the interval from smallest value to largest value
Inter quartile range:	the midspread; a measure of statistical dispersion being equal to the difference between the 3 <sup>rd</sup> and 1 <sup>st</sup> quartiles
Variance:	a measure of how spread out a distribution is; the dispersion of a set of data points around their mean value
Standard deviation:	a measure of how spread out a distribution is; a measure of dispersion or variation in a distribution; equal to the square root of the arithmetic mean of the squares of the deviations from the arithmetic mean
Confidence interval:	a statistical range with a specified probability that a given parameter lies within the range
Kuder Richardson Formula 20 (KR-20)	<ul style="list-style-type: none"> <li>• 1<sup>st</sup> published in 1937</li> <li>• a measure of internal consistency reliability for measures with dichotomous choices</li> <li>• values range from 0.00 to 1.00</li> <li>• high values indicate the exam is likely to correlate with alternate forms (a desirable characteristic)</li> <li>• impacted by exam difficulty, spread in scores, and length of exam</li> </ul>
Coefficient Cronbach Alpha	<ul style="list-style-type: none"> <li>• an extension of the KR-20</li> <li>• 1<sup>st</sup> published in 1951</li> <li>• commonly used measure of the internal consistency reliability of a psychometric instrument</li> <li>• measures how well a set of variables or items measures a single, uni-dimensional latent construct</li> </ul>
P value:	probability with a value ranging from 0 to 1; probability if the test statistic really were distributed as it would be under the null hypothesis, of observing a test statistic as extreme or more extreme than the one actually observed
Point biserial:	<ul style="list-style-type: none"> <li>• quantifies the relationship between two variables</li> <li>• ranges from -1.00 to +1.00</li> <li>• positive biserials tell us that those scoring higher on the exam were more likely to answer the test item correctly</li> <li>• negative biserials tell us that high scorers on the exam answered the test item incorrectly more frequently than low scorers</li> <li>• negative biserials can indicate that an item was keyed incorrectly, that an item was poorly constructed or misleading, or that content of the item was inadequately taught</li> </ul>