Understanding Psychological Test Scores

Accurate measurement of human behavior and mental processes involves the use of assessment tools with established validity (the test measures what it was actually designed to measure) and reliability (the test measures the trait or ability in a consistent manner and without excessive error).

Such tests are administered in standardized fashion, referring to the fact that they are administered in precisely in accordance with established instructions and proper testing conditions. Once the test has been devised and constructed, it is typically given to a large sample of normal and healthy individuals often across a broad range of ages, genders, and educational levels. This normative process allows the examiner to establish what "average" is for an individual of the examinee's age. Whenever possible, consideration is also given to gender, race, and educational achievement when these factors are found to also impact test scores. The goal is for the best "apples to apples" comparison. This process allows for establishing a range of normalcy around which to compare the examinee's performance on that particular measure.

There are many possible ways to depict an individual's score relative to the normative sample. Such data could be represented in the form of 1) **standard scores**, 2) **scaled scores**, 3) **T-scores**, or 4) **percentile scores**. Each reporting method allows for comparison relative to the normative sample.

The method of reporting selected by an individual psychologist may vary based on training and personal discretion. In most cases, scores are depicted in the form of either **standard scores** or **scaled scores**. Standard scores reflect a traditional method of descriptive reporting most frequently observed in areas of intellectual assessment. For example, a standard score of 100 represents precisely average performance, indicating that 50% of the normative sample scored better than that individual, and 50% scored lower. However, very few individuals obtain such a precisely average score.

It is generally accepted in psychological practice that normal range of scores would fall between 90 and 110, and any scores within this range could be considered to be within normal limits for that particular individual. Standard scores between ranges of 80 to 89 would fall in a *low average* range of performance, and scores between 70 and 79 in the range of *borderline deficiency*. Only 2% of the normative sample scores below a standard score of 70, marking this as the commonly accepted cutoff point for the identification of deficient performance. Conversely, scores between 110 and 120 represent a high average range of performance with scores above 120 representing superior performance, with only 2% of the population achieving standard scores at or above 130. Scores within this range with represent unusually high level of a particular trait or ability.

Scaled scores offer a somewhat different descriptive framework. A scaled score of 10 represents average (50th percentile) performance with a normal range between 8 and 12. Scaled scores of 6 or 7 represent low average abilities, with scaled scores of 4 and 5 representing a borderline deficient range. Scaled scores below 4 represent deficient performance.

T Scores are yet another method of depicting a person's score. In this context, a T Score of 50 is precisely average, with a normal range of roughly 40 to 60. T Scores below 30 (2nd percentile) or above 70 (98th percentile) represent unusually low or high performance. Many common personality tests use T Scores to represent the findings.

Finally, individual performance may also be represented in the form of a **percentile** score. As noted above, a score in the 50th percentile represents precisely average performance with scores between the 25th and 75th percentile representing the normal range. Scores below the 2nd percentile and above the 98th percentile are very unusual, and thus are traditionally seen as "abnormal."

There is no particular reason to select one scoring method over another other than preference of the test developer. Scoring systems are interchangeable. A T score of 50 means the same as a Standard Score of 100, which means the same as a Scaled Score of 10. All are equivalent to a percentile score of 50. Regardless, these scoring and reporting methods allow for relatively precise quantification of behavior and mental processes, and aid in diagnosis by determining the extent to which a deficiency or abnormality of behavior might be present.