PSP METRICS

The Frick Building
437 Grant St Ste 1900
Pittsburgh, PA 15219

Phone: 412.261.1333
Fax: 412.261.5014

motivator 360

The Paired-Comparison Process
vs.
The Likert Scale
For 360 Measurement

Making judgments about people is a difficult task, and various efforts have been made to achieve greater accuracy in the judgments. The most popular attempt has been the adoption of Likert-type (typically five point) scales for rating people on certain dimensions or competencies. In order to achieve any degree of accuracy from these Likert scaling procedures, considerable training of all raters is required, and there is the threat that the process will degrade as new – and untrained – raters become involved.

Training alone however does not solve the problems with the use of Likerttype scales. Other problems include:

- Most raters don't use either the highest or the lowest ratings.

- Most raters place the majority of their ratings in the middle because it is the "safe" rating. This tendency is known as "average rater error."

- **Each behavior is typically rated only once; in the Paired-Comparison Process, the average 360 rates a behavior multiple times.**

In an attempt to improve the accuracy of Likert-type rating scales, additional scale values are often added so that a five point scale becomes, for example, a ten point scale, the idea being that the more points on a scale, the more accurate it must be. However, the result is less accuracy because additional scale values make it even more difficult for raters to distinguish real differences. Adding more scale values simply introduces more variability (error) into the process.

The central problem with most rating systems is that they require judgments to be made in an absolute sense, i.e., they refer to ideals. But, **most real world judgments** are not made this way and – consciously or not – **usually involve a comparison of alternatives.** The Paired-Comparison Process makes use of this fact by asking raters to make comparisons rather than absolute judgments.

If a group of raters was asked to estimate the length of a piece of rope, there would be substantial disagreement among the judgments of the group members. If, on the other hand, the raters were shown two pieces of rope and asked which one is longer, a high degree of rater agreement would be obtained. Similarly, **the Paired-Comparison Process asks raters to compare a ratee's behavior with another behavior of that same ratee, thereby achieving a higher degree of rater agreement than other procedures that use non-comparative judgments.** The Paired-Comparison Process also requires more judgments or ratings before a final rating is reached. These additional judgments, made by comparing a single behavioral competency to other behavioral competencies, result in greater objectivity and accuracy in the final rating.

The Paired-Comparison Process also **prevents "gaming" the 360**. In a Likerttype rating scale, the ratings are transparent, i.e., the rater can readily "see" what rating he/she is giving the ratee. This enables the rater to adjust his/her ratings based on their own biases about the individual. If the rater wishes to "send a message" to the ratee, either positive or negative, he/she can make their ratings artificially high or low.

The Paired-Comparison Process, however, asks the rater to judge whether the ratee is higher or lower on two behaviors, e.g., is John stronger in giving praise for a job well done or delegating tasks to the right person? To further strengthen the reliability of the rating, the rater is asked to judge the ratee's performance in praising and delegating multiple times, each time against a different behavior. This makes it virtually impossible for a rater to "game" the ratings and further structures an objective end result.

Overall, **the Paired-Comparison Process structures a more objective and effective 360 that ensures greater accuracy and reliability for all concerned**.