If there is any single statistic that we practitioners should know so deeply that it is integral to our view of the world, it is the standard deviation. This statistic is a beautiful one for evidence-based practice and, well, just beautiful in general. The reason for its beauty is that it depicts variation among individuals.

Our practice is with individuals. We design our intervention plans around the individual. When we use research findings to inform practice, we need to know how individuals vary in their responses during evaluation procedures, and to know how they vary in their responses to intervention. The standard deviation, as well as other measures of variation such as the range of scores, provides us direct and useful information about individuality and uniqueness against the backdrop of general human tendencies. Although the standard deviation is commonly reported in research reports, my experience is that readers often overlook it and fail to recognize its importance. The standard deviation can validate our humane and individualized approach to service and can transform that service. Before I describe how this statistic can do these things, let me refresh your knowledge about what it is.

You can find the mathematical equation for the standard deviation in any introductory statistics or quantitative research methods textbook (e.g., Portney & Watkins, 1999; Rosenthal & Rosnow, 1991). The look of this equation is perhaps the primary reason that people do not immediately see the beauty of the statistic so it is not shown here. Instead I describe its logic as it is calculated by hand in its most simple and intuitively understandable form.

In essence, the standard deviation is an average of difference scores. The first in the four steps of the calculation of the standard deviation is to calculate a difference score for each individual in the group: the individual’s score on a test or assessment minus the group’s mean score. This calculation describes how far the score of the single individual deviates (or differs or varies) from the group as a whole. The group’s mean score is a measure of central tendency in that it is the central score around which each individual’s score differs to some degree. Other measures of central tendency are the median (middle score) and mode (most frequent score) but these are not used in the calculation of the standard deviation. The second step in the calculation of the standard deviation is to square each individual’s difference score. This step makes all of these scores positive in sign. A difference score that is positive, because the individual’s original score was higher than the mean of the group, remains positive when squared, while a difference score that is negative, because of an original score lower than the mean, becomes positive. The reason for making all difference scores positive becomes clear in the third step of the calculation. At this third step, the variance is calculated. It is calculated by adding together the individuals’ squared difference scores and then dividing by the number of individuals.1 If we had not squared the difference scores in the second step of the calculation, the sum of the difference scores would be equal or nearly equal to 0. It so happens that the scores for many client attributes are distributed in the clinical population symmetrically around the group mean: half of the scores fall below the mean, and half above. The positive and negative sign difference scores would cancel each other out when added together and we would have no useful measure of the average deviation of individuals’ scores from the mean. The fourth and final step is to take the square root of the variance. This final step creates the standard deviation, which is on a numerical scale that is the same as the individual’s original score on the test or assessment. A standard deviation of 5 tells us that on the average, individuals’ scores differ from the group mean by 5 points. If the group mean is 25, and the standard deviation is 5, then individual scores tend to fall between 20 and 30.

The standard deviation becomes a useful tool when we understand its relationship to what is called the normal distrib-

---

1This is the variance for an entire population of individuals. In studies that do not consist of an entire population of individuals (i.e., most studies), an adjustment is made to the calculation of the variance. Instead of dividing the sum of the squared difference scores by N (the number of participants), the sum is divided by N minus 1. In studies in which there are a very large number of participants, the population variance and the adjusted variance, which is called the sample or unbiased estimate of the variance, are nearly equivalent.
tions. Figure 1 shows that client scores follow a normal distribution when their frequency follows a bell shaped curve. Scores with values close to the mean are the most frequent. Scores that are farther away from the mean, on either side of it, are less frequent. And as discussed in the previous paragraph, half of the scores fall below the mean, while half fall above it. The standard deviation helps to quantify the frequency of individuals whose scores fall a certain distance from the mean. Take for example our group mean of 25 with a standard deviation of 5. The majority of the individuals, about 68%, have scores that differ from the mean by 1 standard deviation or less. Their scores range from 20 to 30 points. About 27% of the entire group of individuals score between 1 and 2 standard deviations from the mean: between 15 and 20 or between 30 and 35 points. Only 5% of the individuals score 2 or more standard deviations from the mean, with scores of less than 15 or more than 35 points. These percentages—68%, 27%, and 5%—describe patterns of scoring for any normal distribution of scores, regardless of the value of the group mean or the value of the standard deviation. Therefore, when a group mean is 100 and the standard deviation is 25, it will still be the case that 68% of the individuals have scores that differ from the mean by 1 standard deviation or less. Since many types of scores are normally distributed, we can use 68%, 27%, and 5% to validly describe patterns of scoring among populations of clients on many types of tests and assessment tools.

To sum up this brief “refresher course,” the standard deviation documents both the degree and the pattern of variation among individuals against the backdrop of general human tendencies. The value or magnitude of the standard deviation tells us the degree to which individuals vary from the group mean on a particular attribute as measured by a test or assessment procedure. And the relationship of the standard deviation to the normal distribution gives us a pattern to that variation in that it relates the value of a score to its frequency in the overall population. Values close to the group mean are frequent whereas values far away from the mean are rare. When practitioners use evidence about degree and pattern of variation in individuals they are bringing what Stephen J. Gould (1996) calls a variation perspective into their clinical reasoning and performance. When Gould was diagnosed in 1982 with abdominal mesothelioma, the literature about the disease delivered what he called a “brutal message” (p. 47), a prognosis of death at a median of 8 months post-diagnosis and no known cure. The most typical interpretation of this message would be “I will most probably be dead in eight months.” Yet his reading of the literature was anything but typical in that he was a paleontologist whose theories on biodiversity and variation had changed the face of evolutionary theory. He knew that the median is a measure of central tendency and does not describe the variation of mortality outcomes. Therefore he focused on measures of variation. In his words:

Central tendency is an abstraction, variation the reality. . . I am not a measure of central tendency, either mean or median. I am one single human being with mesothelioma, and I want a best assessment of my own chances—for I have personal decisions to make, and my business cannot be dictated by abstract averages. I need to place myself in the most probable region of the variation based upon particulars of my own case; I must not simply assume that my personal fate will correspond to some measure of central tendency. (pp. 48–49)

With his variation perspective, Gould attempted to position himself on the low mortality end of the distribution of results, in part by trying what he thought to be a theoretically sound yet experimental treatment. He lived and worked as a prolific scholar, researcher, and teacher until 2002 when he died, 2 decades after his initial diagnosis.

Gould’s message is a vivid one because it is about survival, yet the message is the same for the optimization of quality of life outcomes. A variation perspective requires an atypical way of reading the literature, an explicit use of variation evidence in clinical reasoning and collaboration with clients, and an eye to the patterns of variation in one’s own clinical practice (Glassiou et al., 1998; Mant, 1999; Rosser, 1999; Sackett, Strauss, Richardson, Rosenberg, & Haynes, 2000; Tickle-Degnen, 2001). The possibilities for how to implement the variation perspective are endless and certainly our literature on client-centered practice provides much of the guidance. However, here I focus on specific recommendations for using standard deviation evidence in the clinic:

1. Find the standard deviation in tables and text of the results sections in quantitative research studies about client quality of life and occupation, the validity of assessment tools, and the effectiveness of intervention. Use the standard deviation as it relates to the normal distribution (Figure 1: A normal distribution of scores) to determine which score values were the most frequent ones found among research partici-
pants in a study. These values would be the ones that fall within 1 standard deviation from the mean. Also determine which score values were found to be least frequent, that is the values that fell greater than 2 standard deviations from the mean. Compare these percentages from the study to your estimate of the percentages of your clients with different scores. Are these percentages similar or dissimilar? Even if you do not score clients’ attributes, such as upper-extremity strength, speed of donning clothing, or satisfaction with therapy, there may be a perceivable systematic variation among clients. Many clients may tend to exhibit a moderate level of an attribute, while fewer may be extremely low or high on that attribute. Alternatively, because of referral sources or the type of clinical practice, clients may tend to fall on the extreme ends of the distribution compared to published findings. For example, more of your clients may have low upper extremity strength compared to published findings. For example, for more of your clients may have low upper extremity strength compared to what would be expected from the published results about similar clinical populations. The point of this exercise is to become sensitized to degrees and patterns of variation among clients during actual clinical practice.

2. Follow the literature on a particular outcome measure that is relevant to practice. Is the standard deviation in Study 1 similar to that in Studies 2 and 3? Or do the studies together represent a range of possible variation outcomes? Reading literature in this cumulative and synthesizing manner will hone skills in detecting and understanding variation.

3. Once you have gained mastery over the basics of perceiving systematic variation, look for complex patterns of variation. Do males and females have different degrees of variation in their responses on a particular test or to a particular type of intervention? Children and adults? Individuals in a secure versus insecure home environment? Information about complex patterns of variation can be found in a single study’s report of standard deviation findings for blocks of research participants or the different conditions of an experiment (Oxman & Guyatt, 1992). Information can also be found by comparing standard deviation findings across more than one study. For example, did the studies that were about children demonstrate the same degree and pattern of variation as the studies about adults?

4. Use a variation perspective in clinical reasoning. Based upon published standard deviation findings and personal observation of variation among clients, what are reasonable expectations about the attributes and responses of new clients? What are the likely central tendencies and the variations? We all know that individuals are variable. What statistical information does is guide our perception of systematic variability. We form open expectations about the odds, proportions, and probabilities of seeing certain levels of attributes in our clinical population. These expectations are open in that they guide rather than dictate clinical reasoning. For example, we may expect younger clients to have a more variable response to a particular intervention compared to older clients. To form assessment and intervention strategies by weighing alternatives with an informed uncertainty is fundamental to a variation perspective.

5. Communicate from a variation perspective. The key sign that practitioners have a variation perspective is the content of their language when informing colleagues and clients about research findings and when making assessment and intervention recommendations. An overfocus on central tendency findings related to assessment procedures yields statements like “your child is going to develop these skills over time” or “not develop these skills over time.” For intervention recommendation, this overfocus shows up in statements like “studies have shown that intervention X works better than intervention Y, therefore, I recommend you do X” or simply “intervention X works and Y does not.” The only time such certainty is warranted in our language is if studies have shown that an assessment procedure has perfect reliability and validity (unlikely) or that everyone in intervention X did better than everyone in Y (also unlikely).

Central tendency is important because it is about majority responses. It is very sensible to make practice decisions that are likely to result in the maximal best outcome for our group of clients as a whole. The reality of the situation, however, is that we do not know if an individual client will be in the minority or the majority. Our language should reflect our uncertainty while providing useful guidance for decision making. A variation statement would be something like “in studies of children like yours, children who received scores similar to your child’s score were likely to have difficulty developing skills X, Y, and Z. Some of the children, though a minority, were able to develop these skills given enough time and help.”

6. My last recommendation is to follow Gould’s lesson (1996). First, locate where an individual client falls on the “probable region of the variation based upon particulars of [his or her] own case” (p. 49). Consider the responses of research participants who are similar to this client in gender, age, or any number of possible physical, cognitive, emotional and social attributes that may be reported in the findings. The result is better advice and effective planning. Second, if the client’s profile does not put him or her in a probable region that is positive, successful, or beneficial, work with the client to help reposition him or her to a more favorable region. Use theory, research, and experience to make a reasoned evidence-based intervention plan that will optimize this particular client’s potential to achieve beneficial outcomes.

Sensitizing yourself to variation results in the literature and variation patterns in your own practice can validate a humane and individualized approach to service and transform that service. Noticing and using standard deviation results is the start. There is a reason that this statistic is one of the first discussed in a basic statistics textbook or course. It is fundamental to every parametric statistic and calculation procedure discussed in the remainder of the book or course: standard errors, confidence intervals, z- and t-tests, analysis of variance, correlation and regression. Once the standard deviation is understood, all else follows. That’s the beauty of it.

References


---

**CORRECTIONS**

To Suzanne Peloquin’s article, “Reclaiming the Vision of Reaching for Heart as Well as Hands,” (September/October 2002, Volume 56[5], pp. 517–526). Founder George Edward Barton was erroneously identified as a nurse. Although Barton published in nursing journals, took nursing courses, and hoped for the name *occupational nursing*, by training he was an architect.