Dr. Hinton-Bayre has correctly noted a discrepancy here between the Jacobson and Temkin procedures. His comparison of prediction bounds both confirms this and suggests too that the differences are minor. There is a large-sample explanation: suppose \( \begin{pmatrix} x \\ x' \end{pmatrix} \) denotes a vector of before and after readings on a random variable, following a shift model with fixed variance: its expectation is \( \begin{pmatrix} \mu \\ \mu + \delta \end{pmatrix} \), and its covariance matrix is \( \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \). This is a general formulation and a consequence of the standard random coefficient model (with or without practice effects) that implicitly underlies the change-score problem. Normality need not even be assumed at this point. \( \rho \) is the correlation coefficient between \( x \) and \( x' \).

The difference \( x' - x \) then has mean \( \delta \), and its standard deviation is readily shown to be \( \sqrt{2\sigma^2(1-\rho)} \). The question is how to estimate this from paired-sample data. The Temkin approach is to ignore the form of the expression in the parameters \( \rho \) and \( \sigma^2 \), and just empirically take the sample variance of the observed differences. The Jacobson approach observes the form and plugs in estimates of \( \sigma^2 \) and \( \rho \) by \( r_{xx'} \) (Hinton-Bayre’s symbol—it is just a sample correlation), and \( \sigma^2 \) (strangely perhaps) by \( s^2 \), the sample variance of the initial scores. I say strangely, because there is no reason not to pool the variation in the before and after readings, and get a more efficient (i.e., less wasteful) estimator which would be denoted by \( s^2_{\text{pooled}} = \frac{1}{2}(S^2 + S^2_1) \) with a smaller (indeed, essentially minimal) standard error for estimating \( \sigma^2 \).

Now it can be algebraically verified that the Temkin proposal is exactly equivalent to Hinton-Bayre’s formula with the term \( S^2_1 \) replaced by \( S^2_{\text{pooled}} \). Both are consistent, and for large samples they will be close. To this extent the Temkin and Jacobson procedures are close, but to the extent that \( S^2_{\text{pooled}} \) is better than \( S^1 \), the Temkin procedure is better.