Center for Advanced Studies in
Measurement and Assessment

CASMA Research Report

Number 48

Similarities Between Equated Equivalents Using Presmoothing and Postsmoothing

Hyung Jin Kim[†]
Robert L. Brennan
Won-Chan Lee

August 2016

[†]Hyung Jin Kim is Associate Research Scientist, Center for Advanced Studies in Measurement and Assessment (CASMA), College of Education, University of Iowa (email: hyungjin-kim@uiowa.edu). Robert L. Brennan is E. F. Lindquist Chair in Measurement and Testing and Founding Director, Center for Advanced Studies in Measurement and Assessment (CASMA), College of Education, University of Iowa (email: robert-brennan@uiowa.edu). Won-Chan Lee is Director, Center for Advanced Studies in Measurement and Assessment (CASMA), College of Education, University of Iowa (email: won-chan-lee@uiowa.edu).

Center for Advanced Studies in
        Measurement and Assessment (CASMA)
College of Education
University of Iowa
Iowa City, IA 52242
Tel: 319-335-5439
Web: https://education.uiowa.edu/casma

Contents

Abstract

Presmoothing and postsmoothing improve equating by reducing sampling error. However, little research has been conducted about similarities in equated-equivalents between presmoothing and postsmoothing. This simulation study considers log-linear presmoothing and cubic-spline postsmoothing. The study mainly focuses on how equated equivalents differ between presmoothing and postsmoothing for different smoothing degrees, and examines which presmoothing degrees give relatively similar equated equivalents to specific postsmoothing degrees in terms of absolute differences in rounded equated scale scores and summary variables for rounded equated scale scores. Results are also discussed concerning which postsmoothing parameter gives the most similar equating results to a presmoothing method with "optimum" degrees of presmoothing based on the Akaike information criterion. For postsmoothing, linear interpolation is usually used to obtain equivalents at the ends of the score scale, where frequencies are relatively low. By contrast, it is unconventional to use linear interpolation with presmoothing. Consequently, it is not too surprising that at the ends of the score scale, differences between equivalents for presmoothing and postsmoothing tend to be quite large. A unique feature of this study is that, for most analyses, linear interpolation at the ends of the score scale was used for both presmoothing and postsmoothing in order to provide a more "fair" comparison of the two methods. On balance, in terms of the summary statistics investigated, it appears that cubic-spline postsmoothing appears preferable to log-linear presmoothing, especially when linear interpolation is not used with presmoothing.

1 Introduction

Since different forms of the same test (say Form X and Form Y) are never perfectly parallel, equating should be performed to adjust for form differences in difficulty. However, when equipercentile relationships are estimated using sample percentiles and percentile ranks, random sampling error can be too large to ignore (Kolen & Brennan, 2014). Since it is typically assumed that smoothness is a characteristic of the population distribution of scores (X for Form X and Y for Form Y) as well as the equipercentile relationship itself, irregularity in score distributions and equipercentile relationships suggests the presence of sampling error. In order to diminish sampling error, smoothing is typically applied to the observed frequency distributions and/or the equipercentile relationships.

1.1 Presmoothing and Postsmoothing

There are two types of smoothing: presmoothing and postsmoothing. The main differences between them are when smoothing is performed and the objects that are smoothed. For presmoothing, score distributions, say $f(x)$ for X and $g(y)$ for Y, are smoothed prior to equating, whereas for postsmoothing, equipercentile equivalents are smoothed after equating. In this study, attention is restricted to a polynomial log-linear model for presmoothing and a cubic-spline method for postsmoothing (see Kolen & Brennan, 2014, for basic equations). Each smoothing method has its own degree(s) of smoothing to indicate how much smoothness is applied relative to the observed frequency distributions or the equipercentile relationships.

When using the polynomial log-linear model, it is important to consider the choice of a lower-order polynomial of degree C (also known as a smoothing parameter/degree). In this study, when a particular value of C was chosen, that choice was based on the Akaike information criterion (AIC) (Akaike, 1981). The AIC is computed based on the fit and the number of parameters in the model, $AIC = \chi_C^2 + 2(C+1)$. The C value with the smallest AIC is taken as the optimum degree. A smoothing degree is separately determined for Y and X. Note that, for the log-linear presmoothing method, a smaller C implies a higher degree of smoothing.

For cubic-spline postsmoothing, there is no current statistical criterion for choosing the degree of smoothing, S. (Note that, a larger S implies a higher degree of smoothing.) Ideally, the choice of S should result in a smooth function that does not depart too much from the unsmoothed equivalents (e.g., located within one standard error bands). Visual inspection is an important tool in choosing S, and an important question is how much smoothness is smooth enough to represent the population distribution.

1.2 Research Objectives

Previous studies show that smoothing improves the performance of equating (Kolen, 1984; Livingston & Feryok, 1987; Hanson, Zeng, & Colton, 1994; Kim, 2014). When Kim (2014) compared cubic spline smoothing and log-linear presmoothing methods under the common-item nonequivalent groups design, cubic spline postsmoothing tended to produce smaller random error, whereas log-linear presmoothing tended to perform better in terms of bias and root mean square error. However, little research has been conducted

to ascertain similarities and differences in equating relationships between different degrees of presmoothing and postsmoothing. This study focuses mainly on differences in equated equivalents between presmoothing and postsmoothing for different degrees of smoothing under the random groups design.[1] The study also examines which presmoothing degrees give relatively similar equated equivalents to a specific postsmoothing degree. Finally, this study considers the optimum Cs for X and Y based on the AIC and selects a postsmoothing parameter S which gives the most similar equating results.

2   Methodology

This section describes the methodology of the simulation study, including study factors, data generation, and evaluation of results.

2.1   Study Factors

This study considered three conditions for test length: 15 items, 40 items, and 60 items. The sample sizes used for (new group, old group) were $(1,000, 1,000)$, $(3,000, 3,000)$, and $(6,000, 6,000)$. The study also considered two conditions for differences in difficulty between new and old forms: .05 and .10. Under the assumption that difficulty parameters follow a normal distribution, Table 1 provides distributions for difficulty parameters for the new and old forms for each targeted form difference in difficulty.

For postsmoothing, often linear interpolation is used to obtain equivalents for extreme percentile ranks, say below 0.5% and above 99.5%. However, there is no current literature that studies linear interpolation with presmoothing. For this study, it was judged that both presmoothing and postsmoothing should employ the same procedure for dealing with scores with relatively small frequencies. Otherwise, comparisons of results would be suspect, or, at best, ambiguous. Therefore, this study applied linear interpolation for both postsmoothing and presmoothing outside the same percentile rank range from 0.5% to 99.5%.

For presmoothing, seven log-linear smoothing degrees from 2 to 8 were considered for total scores Y and X, respectively referred to as C_Y and C_X, hereafter. Thus, for log-linear presmoothing degrees, there were $7 \times 7 = 49$ pairs for the combination of C_Y and C_X (i.e., $C_Y{:}C_X$) ranging from 2:2 to 8:8. For cubic-spline postsmoothing, nine smoothing degrees (S) were considered: 0.01, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.75, and 1.00. Equating was performed using the equipercentile equating method under the random groups design.

2.2   Simulation

In order to construct simulated tests, the 3 parameter logistic (PL) item response theory (IRT) model was used. Discrimination parameters, a, were generated from the lognormal distribution with the mean 0 and the standard deviation 0.5. Guessing parameters, c, were generated from the beta distribution with parameters of 5 and 17. For both the

---

[1]Note that Antal, Kim, and Lee (2011, April) considered a comparison of presmoothing and postsmoothing under the nonequivalent groups with anchor test design.

new and old groups, ability parameters were generated from the normal distribution with the mean 0 and the standard deviation 1.

For each study condition (test length, sample size, and form difference in difficulty), item responses were generated using the 3PL model for the new group taking Form X and the old group taking Form Y. Scoring was conducted based on number-correct scores.

All analyses were conducted using rounded scale scores (RSS), which were obtained as follows. First, raw scores (number-correct scores) were transformed using the arcsine transformation (see Kolen & Brennan, 2014, p. 405). A linear transformation was then applied and rounding was performed such that the resulting RSS ranged from 10 to 50 with an increment of one. The net result is an RSS ranging from 10 to 50 for each raw score, with a raw-to-RSS conversion table that is monotonically non-decreasing. As a result, for a test length of 40 items, the number of possible distinct raw-score points is the same as the number of possible distinct RSS. However, for a test length of 60 items, the number of possible distinct raw-score points is larger than that of RSS; and, for a test length of 15 items, the number of possible distinct raw-score points is smaller than that of RSS.

Most literature compares unrounded equivalents only. However, in practice, RSS are the scores actually reported to examinees, and are also frequently used for making important decisions such as college admission and professional licensure. Since presmoothing and postsmoothing perform smoothing using different methods at different stages of equating, this might lead to differences in smoothness in unrounded scale scores, which can lead to marked differences in RSS (including differences in gaps and many-to-one conversions). Therefore, it is worthwhile to compare similarities and differences in RSS between presmoothing and postsmoothing methods.

All possible equating procedures were performed on each simulated data set. All possible equating procedures included all crossed factors of smoothing methods and degrees of smoothing. For each study condition, one hundred replications were conducted.

2.3   Evaluation

The equating results for presmoothing and postsmoothing are compared here with respect to absolute differences in RSS and summary numbers. Moreover, for a presmoothing method with the optimum C_Y and C_X selected based on the AIC, the study examines which postsmoothing degree yields the smallest absolute differences.

2.3.1   Absolute Differences in RSS

For each pair of nine postsmoothing degrees and 49 presmoothing degrees, weighted absolute differences in RSS (i.e., WAD) were computed with weights proportional to relative frequencies for the new group:

$$\text{WAD} = \frac{\sum_{j=1}^{100} \sum_i w_{ij} \times |\hat{e}_{Y_S}(x_i) - \hat{e}_{Y_C}(x_i)|}{100}, \tag{1}$$

where i is the ith score point; j is the jth replication; $\hat{e}_{Y_S}(x_i)$ refers to Form Y equivalent RSS of Form X at score x_i using a postsmoothing method with a degree S; $\hat{e}_{Y_C}(x_i)$

refers to Form Y equivalent RSS of Form X at score x_i using a presmoothing method with degrees C_Y and C_X (abbreviated as C here to avoid double subscripts); and, w_{ij} is a relative frequency for score point x_i for the new group at the jth replication. For unweighted absolute differences in RSS (i.e., UAD), w_{ij} in Equation 1 is replaced with $1/n$ where n refers to the number of raw-score points.

2.3.2 Summary Variables for RSS

Equating results in RSS are also summarized in terms of four summary variables: Many-To-One (MTO), Score-Gaps (SG), RS-to-lprss (RTL), and RS-to-hprss (RTH). MTO refers to the number of times for which more than one raw score is converted to the same RSS; SG refers to the number of gaps in the raw-to-RSS conversion table; RTL refers to the number of raw scores converted to the lowest possible RSS; and, RTH refers to the number of raw scores converted to the highest possible RSS. For SG, this study counts the number of times score gaps are greater than or equal to 2. For MTO, this study counts cases where the number of raw scores converged to the same rounded scale score is greater than or equal to 3.

In real operational settings, lower values for MTO and SG are generally preferable. Also, in most equating contexts, policy constraints require that RTL and RTH should be at least one.

2.3.3 For AIC-selected Presmoothing Method

For each study condition, the presmoothing parameters for Y and X are selected based on the AIC and are hereafter referred to as $C_{Y_{AIC}}$ and $C_{X_{AIC}}$, respectively. It is then determined which postsmoothing degree gives the smallest WAD compared to the AIC-selected presmoothing method. Note that, unlike the first two evaluations, this evaluation against AIC-selected presmoothing method was conducted using both unrounded and rounded scale scores. For each presmoothing method with a $C_{Y_{AIC}}$:$C_{X_{AIC}}$ degree, this study examines a frequency distribution of how many times each S is selected as a degree with the smallest WAD. Similarly, the study also considers the smallest UAD for a presmoothing method selected using the AIC.

3 Results

This section consists of three subsections. The first section gives results using the WAD for all pairs of presmoothing and postsmoothing methods. The UAD yielded virtually the same patterns as the WAS, so there results are omitted. The second section compares presmoothing and postsmoothing methods in terms of the four summary variables. The last section describes which postsmoothing degree gives the most similar equating results to a presmoothing method with $C_{Y_{AIC}}$ and $C_{X_{AIC}}$.

3.1 Weighted Absolute Differences in RSS

The first section shows the overall WAD between all 49 presmoothing methods and all nine postsmoothing methods. The following sections discuss observations for the WAD

in more detail depending on different test lengths, group differences in ability, and sample sizes.

3.1.1   Overall Results

Table 2 presents the overall WADs between all 49 presmoothing and nine postsmoothing methods. The overall WADs were aggregated and averaged over all study factors. For each C_Y, a light-gray cell indicates that a presmoothing method with a corresponding C_X gave the smallest WAD compared to a postsmoothing method with a corresponding degree S. For example, for $S = .10$, the presmoothing method with $C_Y = 3$ gave the smallest WAD when $C_X = 3$, but the presmoothing method with $C_Y = 5$ gave the smallest WAD when $C_X = 6$. Among all combinations of C_Y and C_X, a dark-gray cell refers to the combination with the smallest WAD compared to each postsmoothing method with degree S. For example, for the postsmoothing degree $S = .20$, the presmoothing method with $C_Y = C_X = 5$ gave the smallest WAD. Additionally, boldface numbers refer to postsmoothing degrees that gave the smallest WADs compared to the presmoothing methods with $C_Y = C_X = 4$ and $C_Y = C_X = 6$. $C_Y = C_X = 4$ and $C_Y = C_X = 6$ are considered because they are the degrees that are commonly used in operational equating.

Based on Table 2, for each $C_Y \geq 4$, the C_X with the smallest WAD tends to decrease as S increases. For example, for $C_Y = 6$, $C_X = 8$ gave the smallest WAD to $S = .01$; $C_X = 7$ had the smallest WAD to $S = .05$ and $.10$; $C_X = 5$ gave the smallest to $S = .20$, $.30$, and $.40$; and, $C_X = 4$ had the smallest value to S greater than or equal to $.50$. It appears that, for each S, equivalents between presmoothing and postsmoothing methods become more similar as C_Y and C_X increase. However, the WADs generally do not differ much as long as $C_Y, C_X \geq 4$ for the same S. For $C_Y \leq 3$, C_X should be the same as C_Y to give the smallest WADs for all S; i.e., for $C_Y = 2$, $C_X = 2$ gave the smallest WAD to all S degrees, and for $C_Y = 3$, $C_X = 3$ gave the smallest to all S degrees. For any S, when $C_Y, C_X \leq 3$ and $C_Y \neq C_X$, the WADs are large relative to the other pairs of presmoothing and postsmoothing methods. It can be assumed that, for $C_Y = 2$ or 3, performing equating with $C_X \neq C_Y$ resulted in different equated equivalents.

Among all combinations of C_Y and C_X, $C_Y = C_X = 8$ gave the smallest WADs compared to $S = .01$ and $.05$; $C_Y = 6$ and $C_X = 7$ gave the smallest value compared to $S = .10$; and, $C_Y = C_X = 5$ gave the smallest WAD compared to $S = .20$; $C_Y = C_X = 3$ gave the smallest values compared to $S = .30$ to $.75$. For $S = 1.00$, $C_Y = C_X = 2$ gave the smallest WAD. For all Ss, the presmoothing methods with the smallest WAD tended to have the same smoothing degree for Y and X (i.e., $C_Y = C_X$). Additionally, Table 2 shows that, based on the boldface numbers, $S = .10$ gave the smallest WADs for the presmoothing methods with $C_Y = C_X = 4$ and with $C_Y = C_X = 6$.

3.1.2   Effect of Test Length

The WADs for all 49 presmoothing and nine postsmoothing methods are presented in Tables 3, 4, and 5 for the test lengths of 15 items, 40 items, and 60 items, respectively. Patterns of results are similar to those found based on the overall WAD from Table 2. For $C_Y \geq 4$, C_X with the smallest WAD decreases as S increases. For $C_Y = 2$ (or 3), $C_X = 2$ (or 3) gave the smallest WAD for all S degrees. Among all combinations of

C_Y and C_X, presmoothing methods giving the smallest WADs tended to have the same degree for Y and X.

Based on the boldface numbers, the S degrees that gave the smallest WADs for the presmoothing methods with $C_Y = C_X = 4$ and with $C_Y = C_X = 6$ depended on the test length: S tended to increase for a longer test. For the test lengths of 15 items, 40 items, and 60 items, for $C_Y = C_X = 4$, $S = .10, .10$, and .20 gave the smallest WADs, respectively. For $C_Y = C_X = 6$, $S = .05, .10$, and .10 gave the smallest WAD, respectively.

Tables 6, 7, and 8 present the WADs between nine postsmoothing methods and seven presmoothing methods for $C_Y = C_X$ for the test lengths of 15 items, 40 items, and 60 items, respectively. For each postsmoothing degree, a shaded cell refers to a presmoothing degree with the smallest WAD; and, for each presmoothing degree, a boldface number refers to a postsmoothing degree with the smallest WAD. For example, for the test length of 15 items, $C_Y = C_X = 4$ gave the smallest WAD for $S = .20$; and, among nine postsmoothing degrees, $S = .10$ gave the smallest WAD for $C_Y = C_X = 4$.

Results show that similarities between postsmoothing and presmoothing depend on test length. For each postsmoothing degree, the presmoothing degree with the smallest WAD tends to increase as test length increases. For example, for $S = .20$, $C_Y = C_X = 4$, 5, and 6 gave the smallest WADs for 15 items, 40 items, and 60 items, respectively. Note that the pattern for $S \geq .30$ is not as evident as for $S \leq .20$. For each presmoothing degree, the postsmoothing degree with the smallest WAD also tends to increase as test length increases. For example, for $C_Y = C_X = 4$, $S = .10, .10$, and .20 gave the smallest WAD for test lengths of 15 items, 40 items, and 60 items, respectively. Note that the pattern of increasing S for each presmoothing degree is not as evident as that for increasing $C_Y = C_X$ for each postsmoothing degree. Considering that, for the test lengths of 40 and 60 items, 2 and 3 were never selected as the optimum degree based on the AIC (discussed in the section 3.3.2), the S with the smallest WAD tended to be less than or equal to .20. In other words, as long as $C_Y, C_X \geq 4$, $S \geq .30$ did not give the smallest WAD.

Based on Tables 9 and 10, it can also be observed that, for $C_Y, C_X \geq 4$, differences in RSS between presmoothing and postsmoothing methods tend to decrease for a longer test; i.e., equating results of presmoothing become more similar to those of postsmoothing as test length increases. According to Table 9, as test length increases from 15 to 40, for each S, the decreases in WAD seem to be almost constant for all $C_Y, C_X \geq 4$; however, for each combination of C_Y and C_X, the decreases in WAD seem to be larger for a higher S. In contrast, for either C_Y or C_X equal to 2 or 3 and $C_Y \neq C_X$, the WADs tend to increase for a longer test; i.e., as a test length increased, presmoothing resulted in equated RSS values that were less similar to those using postsmoothing.

Similar patterns can be seen as test length increases from 40 to 60 based on Table 10; however, the patterns are not as evident as those shown on Table 9 due to the relatively small changes. In other words, the degree of similarity between presmoothing and postsmoothing is similar for test lengths 40 and 60.

3.1.3   Effect of Form Differences in Difficulty

Results are similar for various levels of form differences in difficulty; i.e., the WAD between presmoothing and postsmoothing methods does not depend on how forms differ in their difficulty. General results are similar to those observed based on Table 2. Since the WAD values are very similar to those shown in Table 2, separate tables are not provided here.

For each $C_Y \geq 4$, the C_X with the smallest WAD tends to decrease as S increases. However, for $C_Y = 2$, $C_X = 2$ gave the smallest WAD for all S degrees. Similarly, for $C_Y = 3$, $C_X = 3$ gave the smallest WAD for all S degrees. Additionally, for each S, presmoothing methods with the smallest WAD tended to have the same smoothing degree for Y and X.

3.1.4   Effect of Sample Sizes

For different sample sizes, similar results are observed. The resulting patterns are also similar to those found using the overall WAD. Therefore, results are not separately presented in this paper. One thing to be noted is that the WAD tended to decrease as sample sizes increased; i.e., the degree of similarity between presmoothing and postsmoothing increased with larger sample sizes. When sample sizes increased from 3,000 to 6,000, the decreases in WAD were small relative to those when sample sizes increased from 1,000 to 3,000.

3.2   Summary Variables for RSS

This section is divided into two subsections. The first subsection explains overall differences in summary variables between presmoothing and postsmoothing methods. The second subsection describes how presmoothing and postsmoothing methods give different values for summary variables for different test lengths.

In order to compare presmoothing and postsmoothing methods in terms of summary variables, the test length of 15 items was excluded. For a test length of 15 items, 16 distinct raw scores are scaled onto 41 distinct RSS. Thus, it is highly likely that MTO and SG will be zero in the raw-to-RSS conversion tables. Therefore, for the analyses using summary variables, only the test lengths of 40 and 60 items are considered.

3.2.1   Overall Results

Table 11 presents the minimum, the mode, and the maximum of SG, RTL, RTH, and MTO for all presmoothing and postsmoothing methods. Not much difference is noted between these summary variables. For all methods, the minimum and the mode of SG tend to be 0 and 1, respectively. The maximum tends to be either 3 or 4. With respect to RTL, the minimum, the mode, and the maximum are 1 for almost all methods; for RTH, the minimum and the mode are 0 and 1, respectively, for all methods. The maximum of RTH tends to be 1 for all Ss and $C_Y \geq 4$. However, for $C_Y \leq 3$, the maximum of RTH tends to be greater than for postsmoothing. With respect to MTO, both the minimum and the mode are 0 for all methods. The maximum of MTO tends to be similar for presmoothing and postsmoothing. Most differences between presmoothing

and postsmoothing tend to occur for $C_Y \leq 3$. However, considering that any combination of $C_{Y_{AIC}}$ and $C_{X_{AIC}}$ for the test lengths of 40 and 60 items did not include 2 or 3 (discussed in the section 3.3.2), it can be concluded that there is not much difference observed in summary variables between presmoothing and postsmoothing methods.

Similar results were observed regardless of sample sizes and form differences in difficulty. Therefore, results for different sample sizes different form differences in difficulty are not separately presented here.

3.2.2 Effect of Test Length

Based on Tables 12 and 13, it can be observed that, while RTL and RTH do not change much, MTO seems to increase and SGs seem to decrease as test length increases from 40 to 60 items. As shown in Table 13, for the test length of 60 items, the maximum MTO was between seven and nine for all equating methods, whereas the maximum was at most four for the test length of 40 items. The maximum mode was 5 for the test length of 60 items, whereas the mode was all 0 across all methods for the test length of 40 items.

This phenomenon can be explained simply: since, for a test length of 40 items, 41 distinct raw-scores were scaled onto 41 distinct RSS, it is highly unlikely that MTO would occur, especially when MTO refers to a case where the number of raw scores converted to the same RSS is greater than or equal to 3. However, for a test length of 60 items, the number of distinct raw-scores is greater than the number of distinct RSS by 20; given this scenario, some sets of consecutive numbers would be assigned to the same RSS. Therefore, for a test length of 60 items, it can be expected that MTO would occur more frequently than for a test length of 40 items.

For similar reasons, it can also be expected that the number of SG decreases as test length increases. For a test length of 60 items, a larger number of distinct raw-scores than the number of distinct RSS reduces the chance that two consecutive raw-scores will be converted to two RSS with a gap. However, for a test length of 40 items, the number of distinct RSS is equal to the number of distinct raw-scores, and the chance that an SG occurs is higher than for a longer test.

3.3 For AIC-selected Presmoothing Method

This section consists of two subsections. The first section presents the overall frequency distributions of the S parameter giving the most similar equating relationship to presmoothing methods with degrees of $C_{Y_{AIC}}$:$C_{X_{AIC}}$. The second section also presents the frequency distributions of the S parameter for each test length.

The statement and conclusions using UAD are similar to those using WAD. Therefore, results are provided using the WAD only.

3.3.1 Overall Results

Table 14 presents frequencies of each S parameter with the smallest WAD (in both unrounded and rounded scale scores) relative to presmoothing methods for degrees of $C_{Y_{AIC}}$ and $C_{X_{AIC}}$. For each $C_{Y_{AIC}}$:$C_{X_{AIC}}$, a shaded cell refers to a postsmoothing degree that was selected the most frequently for giving the most similar equating results; i.e., a

shaded cell refers to the mode in a frequency distribution of S with the smallest WAD for each $C_{Y_{AIC}}$:$C_{X_{AIC}}$. For example, among study conditions for which the optimum degrees were 4 for both C_Y and C_X based on the AIC, an unrounded scale scores of $S = .10$ gave the smallest WAD the most frequently (205 times). For rounded scale scores, $S = .01$ had the smallest WAD the most frequently (134 times) for $C_{Y_{AIC}}$:$C_{X_{AIC}} = 4$:4.

As shown in Table 14, 4 and 4 were selected the most frequently for the optimum $C_{Y_{AIC}}$:$C_{X_{AIC}}$. In terms of unrounded scale scores, for almost all $C_{Y_{AIC}}$:$C_{X_{AIC}}$, $S = .05$ and .10 gave the most similar equated equivalents. However, in terms of rounded scale scores, for almost all AIC-selected presmoothing degrees, $S = .01$ tends to give the most similar equated equivalents. Based on Table 14, it is evident that using rounded scale scores yields choices of S that tend to be more spread out over all values of S, which is not the case for unrounded scale scores.

Note that, regardless of form differences in difficulty and sample sizes, results were similar to those found based on Table 14, using both unrounded and rounded scale scores. When additional analyses were conducted using the chi-square difference as a criterion, similar results were found. One thing to note is that, based on the chi-square difference criterion, higher degrees were selected as optimum more frequently than they were based on the AIC. However, the result was still the same; namely, 4 and 4 were selected the most frequently for the optimum $C_{Y_{AIC}}$ and $C_{X_{AIC}}$.

3.3.2   Choices Depending on Test Length

Table 15 presents frequencies of the S parameter with the smallest WAD given $C_{Y_{AIC}}$:$C_{X_{AIC}}$. Results are reported in terms of unrounded equated scale scores, separately for different test lengths. For almost all $C_{Y_{AIC}}$:$C_{X_{AIC}}$, for the test length of 15 items, $S = .01$ and $S = .05$ tend to give the most similar unrounded equated scale scores; and, for the test length of 40 items, $S = .05$ and $S = .10$ tend to give the most similar unrounded equated scale scores. For the test length of 60 items, $S = .10$ gave the most similar unrounded equated scale scores for $C_{Y_{AIC}}$:$C_{X_{AIC}}$. This shows that, as test length increases, the S with the smallest WAD also tends to increase. Moreover, for test lengths of 40 and 60 items, 2 and 3 were never selected for either $C_{Y_{AIC}}$ and $C_{X_{AIC}}$. Even for test length of 15 items, 2 was seldom selected for either $C_{Y_{AIC}}$ and $C_{X_{AIC}}$.

Based on Table 15, as test length increases, higher degrees of presmoothing tended to be selected as optimum more frequently. For example, for test length of 15 items, 2 and 3 were selected as optimum for C_Y and C_X, whereas 2 and 3 were never selected as optimum for test length of 40 items: frequencies for 2 and 3 for a test length of 15 items occurred for higher degrees of $C_{Y_{AIC}}$ and $C_{X_{AIC}}$ for test lengths of 40 and 60 items.. Table 15 shows that, as test length increases from 40 to 60, higher degrees tend to have larger frequencies, although 4:4 still has the largest frequencies for both test lengths.

Similarly, Table 16 presents frequencies of a S parameter giving the smallest WAD to presmoothing methods with $C_{Y_{AIC}}$:$C_{X_{AIC}}$ using rounded equated scale scores, separately for different test lengths. Based on Table 16, similar patterns are found using RSS, but with somewhat smaller Ss. For test lengths of 15 and 40 items, $S = .01$ tends to give the most similar rounded equated scale scores to those for presmoothing methods with $C_{Y_{AIC}}$:$C_{X_{AIC}}$. For test length of 60 items, $S = .10$ tended to give the smallest WAD.

Therefore, based on Table 15 and Table 16, it appears that choices of S depend

on test length. That is, for a longer test, a larger S tends to give the most similar unrounded/rounded scale scores to the presmoothing method with $C_{Y_{AIC}}$ and $C_{X_{AIC}}$. Based on Tables 15 and 16, for test lengths of 40 and 60 items, $C_{Y_{AIC}}$ and $C_{X_{AIC}}$ are all greater than or equal to 4.

4 Summary and Discussion

This section consists of two subsections. The first section summarizes the most important findings from the results presented in Section 3. The second section provides discussion and conclusions based on the findings of this study.

4.1 Summary

Based on this study, several important findings are summarized below.

1. For each $C_Y \geq 4$, a smaller C_X tended to give the smallest WAD for a higher S.

   - For $C_Y = C_X \geq 4$, a postsmoothing degree with the smallest WAD tended to be less than or equal to .20.

2. For each S, equated equivalents between presmoothing and postsmoothing methods became more similar as C_Y and C_X increased, especially when $C_Y = C_X$.

   - However, for each S, the WADs did not differ much as long as $C_Y, C_X \geq 4$.
   - For each S, when $C_Y, C_X \leq 3$, C_X should be the same as C_Y to yield the smallest WADs. Otherwise, the WADs were large relative to the other pairs of presmoothing and postsmoothing methods.

3. Similarities between postsmoothing and presmoothing depended on test length.

   - For each S, C_Y and C_X with the smallest WAD tended to increase as test length increases.
   - For each pair of C_Y and C_X, the S with the smallest WAD also tended to increase for longer tests.
   - As long as $C_Y, C_X \geq 4$, the postsmoothing degrees yielding the smallest WADs were never larger than 0.30.
   - For each S, the decrease in WAD tended to be almost constant for all $C_Y, C_X \geq 4$ as test length increases.
   - However, for each pair of C_Y and C_X, the decrease in WAD tended to be larger for a higher S.

4. The summary variables were similar for presmoothing and postsmoothing methods when linear interpolation was implemented for both methods.

5. 4 and 4 were selected the most frequently for the optimum presmoothing degrees based on the AIC.

6. Regardless the optimum C_Y and C_X based on the AIC, the S that gave the most similar equated-equivalents tended to be the same.

7. As test length increased, the S with the smallest WAD also tended to increase.

4.2   Discussion

As in postsmoothing, this study implemented linear interpolation when equating was conducted after presmoothing. Although doing so is unconventional, it was done here so that results could provide a "fair" comparison of presmoothing and postsmoothing.

Based on the summary of findings, since the WADs were relatively large for $C_Y, C_X \leq 3$ and $C_Y \neq C_X$, it can be assumed that performing equating with such presmoothing degrees can result in equating relationships that are quite different from the other presmoothing and postsmoothing methods. For each S, although presmoothing methods with $C_Y = C_X$ tended to give the smallest WAD, the WADs did not differ much as long as $C_Y, C_X \geq 4$. It can be assumed that equating relationships based on presmoothing are similar as C_Y and C_X get higher. However, postsmoothing results in very different equating relationships depending on S. Note that, as S gets closer to 1, the smoothed relationship becomes closer to a straight line.

When analyses were conducted without linear interpolation for presmoothing (not reported here), the results using the WAD were very similar to those obtained with linear interpolation. However, the results using the UAD were quite different from those with linear interpolation. This indicates that many (perhaps most) differences between presmoothing and postsmoothing tend to occur at the two ends of the scale where relative frequencies are low. Note that, even when the UAD was considered without linear interpolation, for $C_Y = 2$ or 3, different degrees for X gave relatively large WADs for all S degrees. It can be assumed that, when either C_Y or C_X is 2 or 3 and $C_X \neq C_Y$, performing equating with log-linear presmoothing results in quite different equated equivalents.

In order to compare summary variables (SG, RTL, RTH, and MTO), the test length of 15 items was excluded from the analyses. Since the number of distinct raw scores for a test length of 15 items is smaller than the number of possible distinct RSS, it can be expected that there will be many SG and few MTO. Therefore, this study included test lengths of 40 and 60 items in the analyses. Overall, the results suggest that summary variables are similar for presmoothing and postsmoothing methods. Note that linear interpolation was implemented for both presmoothing and postsmoothing, which ensured that there was at least one RTH and RTL. When additional analyses were conducted without linear interpolation for presmoothing, summary variables for presmoothing were worse than those for postsmoothing. Especially, most presmoothing results did not have RTH and RTL. Since implementing linear interpolation is unconventional for presmoothing, cubic-spline postsmoothing might be preferable to log-linear presmoothing in terms of the summary variables investigated in this study.

For a presmoothing method with $C_{Y_{AIC}}$:$C_{X_{AIC}}$, this study also found a postsmoothing degree S that gave the most similar equated scale scores. Overall, 4 tended to be selected most frequently for both $C_{Y_{AIC}}$ and $C_{X_{AIC}}$. Choices of S with the smallest WAD doesn't seem to depend on which $C_{Y_{AIC}}$ and $C_{X_{AIC}}$ were selected. This also confirms that equating relationships become similar as C_Y and C_X get higher. However, choices

of S with the smallest WAD seem to depend on test length: a higher S tended to give the smallest WAD compared to the AIC-selected presmoothing methods as test length increased.

It is important to note that presmoothing is a stand-alone procedure independent from equating. By contrast, postsmoothing cannot be separated from equating because postsmoothing is performed on equated equivalents. Therefore, choosing a "proper" S based on the optimum C_Y and C_X does not guarantee that the resulting equating relationship represents the observed (unsmoothed) equated equivalents well.

For postsmoothing, currently a main tool that is widely used to choose a "proper" S is visual inspection of smoothed functions relative to unsmoothed equivalents and a standard error band. Future research should consider developing a statistical criterion to choose an appropriate S for direct use in cubic-spline postsmoothing.

Table 1

Generating Distributions for b for Each Form Difference in Difficulty

Form Difference	New Form	Old Form
0.05	$N(0.05, 1)$	$N(0, 1)$
0.10	$N(0.10, 1)$	$N(0, 1)$

Note: b is a difficulty parameter in item response theory.

Table 2
Overall Weighted Absolute Differences (WAD)

						S				
C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
	2	0.117	0.108	0.099	0.083	0.071	0.064	0.060	0.054	0.051
	3	0.248	0.243	0.239	0.235	0.233	0.230	0.229	0.227	0.227
	4	0.383	0.379	0.377	0.376	0.376	0.375	0.373	0.372	0.372
2	5	0.386	0.383	0.380	0.380	0.380	0.379	0.377	0.376	0.376
	6	0.386	0.382	0.381	0.380	0.380	0.379	0.377	0.376	0.376
	7	0.387	0.383	0.382	0.381	0.381	0.380	0.378	0.377	0.377
	8	0.387	0.384	0.383	0.382	0.382	0.381	0.379	0.378	0.378
	2	0.252	0.250	0.248	0.244	0.244	0.245	0.246	0.247	0.248
	3	0.094	0.083	0.074	0.059	0.052	0.050	0.049	0.051	0.054
	4	0.248	0.244	0.241	0.239	0.240	0.239	0.239	0.240	0.242
3	5	0.250	0.246	0.243	0.242	0.243	0.243	0.242	0.243	0.245
	6	0.251	0.247	0.245	0.243	0.244	0.244	0.243	0.244	0.246
	7	0.251	0.248	0.245	0.244	0.245	0.245	0.245	0.246	0.247
	8	0.251	0.247	0.245	0.244	0.245	0.245	0.244	0.245	0.247
	2	0.330	0.331	0.331	0.332	0.332	0.334	0.336	0.338	0.340
	3	0.207	0.206	0.207	0.207	0.209	0.211	0.214	0.218	0.222
	4	0.055	0.046	**0.039**	0.043	0.053	0.062	0.070	0.083	0.092
4	5	0.052	0.043	0.037	0.042	0.054	0.063	0.070	0.084	0.093
	6	0.050	0.041	0.037	0.043	0.054	0.063	0.071	0.084	0.093
	7	0.050	0.042	0.038	0.045	0.056	0.065	0.072	0.086	0.095
	8	0.049	0.042	0.040	0.046	0.057	0.066	0.073	0.087	0.095
	2	0.331	0.332	0.333	0.334	0.334	0.336	0.339	0.341	0.342
	3	0.209	0.209	0.209	0.210	0.211	0.214	0.217	0.221	0.224
	4	0.052	0.044	0.040	0.045	0.056	0.065	0.072	0.085	0.095
5	5	0.046	0.037	0.033	0.041	0.053	0.064	0.072	0.085	0.094
	6	0.044	0.035	0.033	0.042	0.055	0.065	0.073	0.086	0.095
	7	0.043	0.035	0.033	0.043	0.056	0.066	0.074	0.087	0.096
	8	0.043	0.035	0.034	0.044	0.057	0.067	0.075	0.088	0.097
	2	0.331	0.331	0.332	0.333	0.334	0.336	0.338	0.341	0.342
	3	0.208	0.208	0.209	0.210	0.211	0.214	0.217	0.221	0.225
	4	0.051	0.043	0.040	0.046	0.057	0.066	0.073	0.087	0.096
6	5	0.045	0.036	0.033	0.042	0.055	0.066	0.074	0.087	0.096
	6	0.042	0.033	**0.032**	0.042	0.055	0.066	0.074	0.088	0.096
	7	0.040	0.032	0.031	0.043	0.056	0.067	0.075	0.088	0.097
	8	0.040	0.033	0.033	0.045	0.058	0.069	0.077	0.089	0.098
	2	0.330	0.331	0.332	0.333	0.333	0.336	0.338	0.340	0.342
	3	0.209	0.209	0.209	0.210	0.212	0.214	0.217	0.221	0.225
	4	0.050	0.043	0.040	0.047	0.059	0.068	0.076	0.089	0.098
7	5	0.044	0.035	0.033	0.043	0.057	0.068	0.075	0.089	0.097
	6	0.040	0.032	0.032	0.044	0.058	0.068	0.076	0.089	0.098
	7	0.038	0.031	0.032	0.045	0.058	0.069	0.077	0.090	0.099
	8	0.038	0.032	0.033	0.047	0.060	0.070	0.078	0.091	0.099
	2	0.332	0.333	0.334	0.335	0.336	0.338	0.340	0.342	0.344
	3	0.209	0.209	0.210	0.212	0.213	0.216	0.219	0.222	0.226
	4	0.050	0.043	0.043	0.050	0.062	0.071	0.078	0.090	0.099
8	5	0.043	0.035	0.036	0.046	0.059	0.069	0.077	0.090	0.099
	6	0.040	0.033	0.035	0.046	0.059	0.069	0.077	0.089	0.098
	7	0.038	0.032	0.035	0.047	0.060	0.070	0.078	0.091	0.100
	8	0.034	0.028	0.032	0.046	0.060	0.070	0.078	0.091	0.099

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD within a block for the same C_Y. Each dark-gray cell refers to the smallest WAD across all combinations of C_Y and C_X for each postsmoothing degree S. For $C_Y = C_X = 4$ and $C_Y = C_X = 6$, boldface numbers refer to postsmoothing degrees with the smallest WADs.

Table 3

Weighted Absolute Differences (WAD) for Test Length of 15 Items

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
2	2	0.158	0.145	0.134	0.121	0.108	0.102	0.098	0.092	0.089
	3	0.246	0.236	0.228	0.224	0.222	0.218	0.220	0.218	0.219
	4	0.304	0.297	0.295	0.295	0.296	0.295	0.295	0.295	0.295
	5	0.311	0.305	0.303	0.303	0.305	0.304	0.304	0.304	0.304
	6	0.310	0.304	0.302	0.303	0.304	0.303	0.303	0.303	0.304
	7	0.310	0.303	0.302	0.303	0.304	0.303	0.302	0.302	0.303
	8	0.312	0.306	0.305	0.305	0.306	0.305	0.305	0.304	0.305
3	2	0.226	0.222	0.219	0.213	0.214	0.217	0.218	0.217	0.219
	3	0.110	0.093	0.082	0.071	0.069	0.072	0.073	0.078	0.085
	4	0.174	0.167	0.165	0.163	0.166	0.170	0.169	0.177	0.184
	5	0.176	0.170	0.167	0.166	0.169	0.173	0.173	0.180	0.186
	6	0.174	0.167	0.165	0.165	0.168	0.172	0.173	0.179	0.186
	7	0.175	0.168	0.166	0.166	0.170	0.173	0.174	0.181	0.187
	8	0.177	0.171	0.169	0.168	0.172	0.175	0.176	0.182	0.188
4	2	0.243	0.244	0.246	0.247	0.247	0.250	0.252	0.253	0.256
	3	0.134	0.134	0.134	0.135	0.137	0.140	0.145	0.151	0.160
	4	0.060	0.050	**0.047**	0.057	0.070	0.082	0.091	0.107	0.122
	5	0.054	0.044	0.045	0.058	0.072	0.083	0.091	0.107	0.121
	6	0.051	0.043	0.047	0.060	0.073	0.084	0.093	0.108	0.122
	7	0.051	0.045	0.049	0.062	0.075	0.086	0.094	0.109	0.124
	8	0.048	0.045	0.049	0.063	0.075	0.086	0.094	0.109	0.124
5	2	0.251	0.253	0.254	0.256	0.256	0.259	0.261	0.262	0.264
	3	0.137	0.138	0.137	0.138	0.139	0.142	0.148	0.154	0.162
	4	0.056	0.050	0.052	0.062	0.075	0.086	0.093	0.109	0.123
	5	0.050	0.042	0.045	0.058	0.072	0.083	0.091	0.106	0.121
	6	0.047	0.041	0.047	0.061	0.074	0.085	0.095	0.110	0.125
	7	0.045	0.041	0.047	0.062	0.076	0.087	0.096	0.111	0.126
	8	0.045	0.040	0.048	0.062	0.076	0.087	0.096	0.111	0.126
6	2	0.249	0.250	0.252	0.254	0.254	0.257	0.259	0.261	0.263
	3	0.133	0.133	0.133	0.136	0.137	0.141	0.146	0.152	0.161
	4	0.056	0.050	0.052	0.064	0.076	0.087	0.094	0.110	0.125
	5	0.049	0.042	0.046	0.061	0.074	0.086	0.094	0.110	0.125
	6	0.044	**0.039**	0.047	0.063	0.077	0.088	0.098	0.113	0.128
	7	0.042	0.040	0.046	0.064	0.077	0.089	0.099	0.113	0.128
	8	0.042	0.041	0.049	0.065	0.080	0.091	0.101	0.116	0.130
7	2	0.247	0.250	0.251	0.254	0.253	0.256	0.258	0.260	0.263
	3	0.135	0.136	0.136	0.138	0.140	0.143	0.148	0.153	0.161
	4	0.053	0.050	0.052	0.066	0.080	0.090	0.097	0.112	0.127
	5	0.050	0.043	0.047	0.062	0.076	0.087	0.096	0.111	0.125
	6	0.044	0.040	0.048	0.064	0.078	0.089	0.099	0.113	0.128
	7	0.040	0.042	0.049	0.067	0.081	0.091	0.101	0.114	0.129
	8	0.041	0.043	0.051	0.069	0.083	0.093	0.103	0.116	0.131
8	2	0.249	0.252	0.253	0.256	0.255	0.257	0.260	0.261	0.263
	3	0.134	0.136	0.137	0.140	0.142	0.145	0.151	0.154	0.162
	4	0.053	0.048	0.054	0.067	0.080	0.090	0.097	0.112	0.126
	5	0.047	0.041	0.048	0.064	0.078	0.089	0.097	0.112	0.127
	6	0.043	0.041	0.050	0.066	0.078	0.088	0.097	0.111	0.126
	7	0.040	0.043	0.052	0.068	0.081	0.090	0.100	0.113	0.127
	8	0.037	0.040	0.050	0.069	0.082	0.091	0.101	0.115	0.129

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD within a block for the same C_Y. Each dark-gray cell refers to the smallest WAD across all combinations of C_Y and C_X for each postsmoothing degree S. For $C_Y = C_X = 4$ and $C_Y = C_X = 6$, boldface numbers refer to postsmoothing degrees with the smallest WADs.

Table 4

Weighted Absolute Differences (WAD) for Test Length of 40 Items

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
	2	0.103	0.096	0.088	0.071	0.059	0.052	0.047	0.039	0.036
	3	0.253	0.251	0.248	0.245	0.242	0.241	0.238	0.236	0.234
	4	0.401	0.399	0.397	0.397	0.395	0.395	0.393	0.390	0.388
2	5	0.403	0.401	0.399	0.399	0.398	0.397	0.395	0.393	0.391
	6	0.405	0.402	0.401	0.400	0.399	0.399	0.397	0.395	0.393
	7	0.406	0.404	0.402	0.402	0.401	0.400	0.398	0.396	0.394
	8	0.405	0.403	0.401	0.401	0.400	0.399	0.398	0.395	0.394
	2	0.267	0.265	0.263	0.259	0.258	0.256	0.257	0.259	0.260
	3	0.094	0.084	0.075	0.058	0.050	0.044	0.043	0.042	0.043
	4	0.260	0.258	0.255	0.254	0.252	0.249	0.249	0.246	0.246
3	5	0.262	0.260	0.257	0.256	0.255	0.253	0.252	0.249	0.249
	6	0.266	0.263	0.261	0.261	0.259	0.257	0.256	0.253	0.253
	7	0.266	0.264	0.262	0.261	0.260	0.258	0.257	0.254	0.253
	8	0.264	0.263	0.260	0.261	0.259	0.257	0.256	0.254	0.253
	2	0.360	0.361	0.361	0.360	0.361	0.362	0.364	0.367	0.369
	3	0.222	0.222	0.223	0.223	0.225	0.227	0.229	0.232	0.234
	4	0.053	0.043	**0.036**	0.037	0.048	0.056	0.063	0.076	0.082
4	5	0.048	0.038	0.031	0.033	0.046	0.056	0.062	0.076	0.082
	6	0.048	0.037	0.030	0.035	0.048	0.057	0.064	0.078	0.084
	7	0.046	0.037	0.030	0.036	0.049	0.057	0.065	0.079	0.085
	8	0.047	0.038	0.034	0.039	0.051	0.059	0.067	0.080	0.085
	2	0.358	0.359	0.360	0.360	0.361	0.362	0.364	0.367	0.370
	3	0.226	0.226	0.227	0.227	0.228	0.231	0.232	0.236	0.237
	4	0.049	0.040	0.033	0.038	0.049	0.059	0.066	0.080	0.086
5	5	0.044	0.033	0.026	0.035	0.048	0.059	0.067	0.080	0.086
	6	0.042	0.030	0.024	0.035	0.049	0.060	0.068	0.081	0.086
	7	0.041	0.030	0.025	0.037	0.051	0.062	0.070	0.083	0.088
	8	0.042	0.031	0.028	0.038	0.052	0.062	0.069	0.082	0.087
	2	0.359	0.360	0.360	0.360	0.362	0.364	0.366	0.369	0.371
	3	0.227	0.227	0.227	0.228	0.230	0.232	0.234	0.237	0.239
	4	0.049	0.039	0.033	0.039	0.051	0.061	0.068	0.081	0.088
6	5	0.043	0.033	0.027	0.036	0.051	0.062	0.069	0.083	0.088
	6	0.039	0.029	**0.024**	0.035	0.050	0.061	0.069	0.083	0.088
	7	0.038	0.027	0.023	0.036	0.050	0.061	0.069	0.083	0.088
	8	0.038	0.027	0.025	0.038	0.052	0.063	0.070	0.082	0.088
	2	0.357	0.358	0.359	0.359	0.361	0.363	0.365	0.368	0.370
	3	0.226	0.226	0.227	0.227	0.229	0.231	0.233	0.236	0.238
	4	0.047	0.038	0.033	0.040	0.053	0.063	0.071	0.084	0.090
7	5	0.040	0.030	0.027	0.039	0.053	0.064	0.072	0.085	0.091
	6	0.038	0.027	0.024	0.039	0.054	0.064	0.072	0.085	0.091
	7	0.036	0.026	0.024	0.039	0.053	0.065	0.072	0.086	0.091
	8	0.037	0.027	0.025	0.041	0.055	0.065	0.073	0.085	0.090
	2	0.360	0.361	0.362	0.363	0.364	0.366	0.368	0.371	0.374
	3	0.227	0.227	0.228	0.228	0.230	0.233	0.234	0.238	0.239
	4	0.048	0.041	0.038	0.046	0.058	0.067	0.074	0.087	0.093
8	5	0.043	0.035	0.033	0.044	0.057	0.068	0.075	0.088	0.093
	6	0.039	0.029	0.029	0.042	0.056	0.067	0.074	0.087	0.092
	7	0.037	0.028	0.028	0.042	0.056	0.067	0.075	0.088	0.093
	8	0.032	0.023	0.025	0.041	0.056	0.066	0.074	0.087	0.092

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD within a block for the same C_Y. Each dark-gray cell refers to the smallest WAD across all combinations of C_Y and C_X for each postsmoothing degree S. For $C_Y = C_X = 4$ and $C_Y = C_X = 6$, boldface numbers refer to postsmoothing degrees with the smallest WADs.

16

Table 5
Weighted Absolute Differences (WAD) for Test Length of 60 Items

C_Y	C_X	S .01	.05	.10	.20	.30	.40	.50	.75	1.00
2	2	0.088	0.082	0.074	0.058	0.047	0.039	0.034	0.029	0.028
	3	0.246	0.244	0.241	0.236	0.235	0.232	0.230	0.227	0.228
	4	0.442	0.441	0.438	0.437	0.436	0.434	0.432	0.431	0.431
	5	0.443	0.442	0.439	0.438	0.437	0.435	0.433	0.432	0.432
	6	0.442	0.441	0.439	0.437	0.437	0.435	0.432	0.431	0.431
	7	0.444	0.443	0.441	0.439	0.439	0.437	0.435	0.433	0.434
	8	0.445	0.444	0.441	0.440	0.439	0.438	0.436	0.434	0.435
3	2	0.264	0.263	0.262	0.260	0.259	0.261	0.262	0.265	0.264
	3	0.078	0.072	0.063	0.048	0.039	0.034	0.033	0.032	0.034
	4	0.309	0.308	0.304	0.302	0.301	0.299	0.297	0.296	0.296
	5	0.311	0.309	0.306	0.304	0.304	0.302	0.301	0.300	0.299
	6	0.312	0.310	0.307	0.305	0.304	0.303	0.301	0.300	0.300
	7	0.312	0.311	0.308	0.306	0.306	0.304	0.303	0.302	0.301
	8	0.310	0.309	0.306	0.304	0.304	0.302	0.301	0.300	0.299
4	2	0.387	0.387	0.387	0.387	0.388	0.390	0.393	0.395	0.395
	3	0.263	0.262	0.263	0.264	0.264	0.266	0.268	0.270	0.272
	4	0.053	0.044	0.035	**0.034**	0.042	0.049	0.055	0.066	0.072
	5	0.054	0.045	0.037	0.036	0.043	0.051	0.058	0.068	0.074
	6	0.051	0.042	0.034	0.033	0.041	0.049	0.056	0.067	0.072
	7	0.052	0.043	0.036	0.036	0.043	0.051	0.058	0.069	0.075
	8	0.051	0.042	0.037	0.037	0.044	0.053	0.060	0.070	0.076
5	2	0.385	0.384	0.385	0.385	0.386	0.388	0.391	0.393	0.393
	3	0.264	0.263	0.264	0.265	0.266	0.267	0.270	0.273	0.274
	4	0.051	0.042	0.036	0.035	0.043	0.050	0.057	0.068	0.074
	5	0.045	0.035	0.027	0.030	0.041	0.049	0.057	0.068	0.074
	6	0.043	0.033	0.027	0.029	0.041	0.050	0.057	0.068	0.074
	7	0.044	0.033	0.028	0.031	0.042	0.050	0.057	0.069	0.075
	8	0.043	0.032	0.028	0.032	0.044	0.052	0.059	0.070	0.076
6	2	0.384	0.384	0.384	0.385	0.386	0.387	0.390	0.392	0.392
	3	0.265	0.264	0.266	0.267	0.267	0.269	0.271	0.274	0.275
	4	0.049	0.041	0.034	0.035	0.043	0.050	0.057	0.069	0.075
	5	0.043	0.033	0.026	0.028	0.040	0.050	0.057	0.068	0.075
	6	0.042	0.031	**0.024**	0.027	0.039	0.049	0.056	0.067	0.073
	7	0.040	0.030	0.024	0.028	0.041	0.050	0.057	0.069	0.075
	8	0.040	0.029	0.024	0.030	0.043	0.052	0.059	0.070	0.076
7	2	0.385	0.384	0.385	0.386	0.386	0.388	0.391	0.393	0.393
	3	0.264	0.264	0.265	0.266	0.267	0.268	0.271	0.274	0.275
	4	0.049	0.041	0.035	0.036	0.044	0.052	0.059	0.070	0.077
	5	0.042	0.032	0.026	0.029	0.041	0.051	0.059	0.070	0.076
	6	0.039	0.029	0.025	0.028	0.042	0.051	0.058	0.070	0.076
	7	0.038	0.027	0.023	0.028	0.041	0.051	0.058	0.070	0.076
	8	0.037	0.026	0.022	0.030	0.044	0.052	0.059	0.070	0.076
8	2	0.386	0.386	0.387	0.387	0.388	0.390	0.393	0.395	0.395
	3	0.265	0.264	0.266	0.267	0.268	0.270	0.272	0.275	0.276
	4	0.048	0.039	0.036	0.038	0.047	0.054	0.062	0.072	0.078
	5	0.040	0.030	0.027	0.031	0.043	0.052	0.059	0.069	0.076
	6	0.038	0.028	0.025	0.031	0.043	0.052	0.060	0.070	0.077
	7	0.037	0.026	0.024	0.031	0.044	0.053	0.060	0.071	0.078
	8	0.034	0.022	0.020	0.029	0.043	0.053	0.060	0.071	0.077

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD within a block for the same C_Y. Each dark-gray cell refers to the smallest WAD across all combinations of C_Y and C_X for each postsmoothing degree S. For $C_Y = C_X = 4$ and $C_Y = C_X = 6$, boldface numbers refer to postsmoothing degrees with the smallest WADs.

Table 6

Weighted Absolute Differences (WAD) for Test Length of 15 Items (for $C_Y = C_X$)

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
2	2	0.158	0.145	0.134	0.121	0.108	0.102	0.098	0.092	**0.089**
3	3	0.110	0.093	0.082	0.071	**0.069**	0.072	0.073	0.078	0.085
4	4	0.060	0.050	**0.047**	0.057	0.070	0.082	0.091	0.107	0.122
5	5	0.050	**0.042**	0.045	0.058	0.072	0.083	0.091	0.106	0.121
6	6	0.044	**0.039**	0.047	0.063	0.077	0.088	0.098	0.113	0.128
7	7	**0.040**	0.042	0.049	0.067	0.081	0.091	0.101	0.114	0.129
8	8	**0.037**	0.040	0.050	0.069	0.082	0.091	0.101	0.115	0.129

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD among considered combinations of C_Y and C_X. For each presmoothing degree such that $C_Y = C_X$, a boldface number refers to the smallest WAD among nine postsmoothing degrees.

Table 7

Weighted Absolute Differences (WAD) for Test Length of 40 Items (for $C_Y = C_X$)

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
2	2	0.103	0.096	0.088	0.071	0.059	0.052	0.047	0.039	**0.036**
3	3	0.094	0.084	0.075	0.058	0.050	0.044	0.043	**0.042**	0.043
4	4	0.053	0.043	**0.036**	0.037	0.048	0.056	0.063	0.076	0.082
5	5	0.044	0.033	**0.026**	0.035	0.048	0.059	0.067	0.080	0.086
6	6	0.039	0.029	**0.024**	0.035	0.050	0.061	0.069	0.083	0.088
7	7	0.036	0.026	**0.024**	0.039	0.053	0.065	0.072	0.086	0.091
8	8	0.032	**0.023**	0.025	0.041	0.056	0.066	0.074	0.087	0.092

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD among considered combinations of C_Y and C_X. For each presmoothing degree such that $C_Y = C_X$, a boldface number refers to the smallest WAD among nine postsmoothing degrees.

Table 8

Weighted Absolute Differences (WAD) for Test Length of 60 Items (for $C_Y = C_X$)

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
2	2	0.088	0.082	0.074	0.058	0.047	0.039	0.034	0.029	**0.028**
3	3	0.078	0.072	0.063	0.048	0.039	0.034	0.033	**0.032**	0.034
4	4	0.053	0.044	0.035	**0.034**	0.042	0.049	0.055	0.066	0.072
5	5	0.045	0.035	**0.027**	0.030	0.041	0.049	0.057	0.068	0.074
6	6	0.042	0.031	**0.024**	0.027	0.039	0.049	0.056	0.067	0.073
7	7	0.038	0.027	**0.023**	0.028	0.041	0.051	0.058	0.070	0.076
8	8	0.034	**0.022**	0.020	0.029	0.043	0.053	0.060	0.071	0.077

Note: For each postsmoothing degree S, each light-gray cell refers to the smallest WAD among considered combinations of C_Y and C_X. For each presmoothing degree such that $C_Y = C_X$, a boldface number refers to the smallest WAD among nine postsmoothing degrees.

Table 9

Decreases in WAD for the Change in Test Length from 15 to 40

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
	2	**0.055**	**0.049**	**0.046**	**0.049**	**0.050**	**0.050**	**0.052**	**0.053**	**0.053**
	3	-0.007	-0.015	-0.020	-0.021	-0.020	-0.023	-0.019	-0.018	-0.015
	4	-0.098	-0.102	-0.102	-0.102	-0.099	-0.100	-0.098	-0.095	-0.093
2	5	-0.093	-0.096	-0.096	-0.096	-0.093	-0.093	-0.092	-0.089	-0.087
	6	-0.094	-0.099	-0.098	-0.098	-0.095	-0.096	-0.094	-0.092	-0.089
	7	-0.096	-0.100	-0.100	-0.099	-0.097	-0.097	-0.096	-0.094	-0.092
	8	-0.094	-0.097	-0.097	-0.096	-0.094	-0.094	-0.093	-0.091	-0.089
	2	-0.040	-0.043	-0.045	-0.045	-0.044	-0.039	-0.039	-0.042	-0.041
	3	**0.016**	**0.009**	**0.007**	**0.013**	**0.019**	**0.028**	**0.031**	**0.036**	**0.042**
	4	-0.085	-0.091	-0.090	-0.090	-0.085	-0.080	-0.079	-0.070	-0.062
3	5	-0.086	-0.090	-0.091	-0.090	-0.086	-0.080	-0.079	-0.070	-0.063
	6	-0.091	-0.097	-0.096	-0.096	-0.091	-0.085	-0.083	-0.074	-0.067
	7	-0.091	-0.095	-0.095	-0.095	-0.090	-0.084	-0.083	-0.073	-0.066
	8	-0.088	-0.092	-0.092	-0.092	-0.087	-0.082	-0.080	-0.071	-0.065
	2	-0.117	-0.117	-0.115	-0.113	-0.114	-0.113	-0.112	-0.114	-0.114
	3	-0.088	-0.089	-0.089	-0.088	-0.088	-0.087	-0.084	-0.081	-0.073
	4	**0.007**	**0.007**	**0.011**	**0.020**	**0.023**	**0.026**	**0.028**	**0.031**	**0.040**
4	5	**0.006**	**0.006**	**0.014**	**0.024**	**0.026**	**0.027**	**0.029**	**0.030**	**0.040**
	6	**0.003**	**0.007**	**0.017**	**0.025**	**0.025**	**0.027**	**0.028**	**0.030**	**0.038**
	7	**0.005**	**0.008**	**0.018**	**0.026**	**0.027**	**0.028**	**0.028**	**0.030**	**0.039**
	8	**0.001**	**0.007**	**0.015**	**0.023**	**0.024**	**0.026**	**0.027**	**0.029**	**0.039**
	2	-0.107	-0.106	-0.105	-0.103	-0.104	-0.103	-0.103	-0.105	-0.106
	3	-0.089	-0.088	-0.089	-0.089	-0.090	-0.089	-0.085	-0.082	-0.075
	4	**0.007**	**0.010**	**0.019**	**0.024**	**0.026**	**0.027**	**0.027**	**0.029**	**0.038**
5	5	**0.006**	**0.009**	**0.019**	**0.024**	**0.024**	**0.024**	**0.024**	**0.026**	**0.035**
	6	**0.005**	**0.010**	**0.023**	**0.026**	**0.025**	**0.025**	**0.027**	**0.029**	**0.039**
	7	**0.004**	**0.011**	**0.022**	**0.025**	**0.025**	**0.025**	**0.027**	**0.028**	**0.038**
	8	**0.003**	**0.009**	**0.020**	**0.024**	**0.024**	**0.024**	**0.026**	**0.029**	**0.039**
	2	-0.110	-0.109	-0.108	-0.106	-0.108	-0.107	-0.107	-0.108	-0.108
	3	-0.094	-0.094	-0.094	-0.092	-0.093	-0.092	-0.087	-0.085	-0.078
	4	**0.007**	**0.011**	**0.018**	**0.025**	**0.025**	**0.026**	**0.026**	**0.028**	**0.037**
6	5	**0.006**	**0.009**	**0.018**	**0.025**	**0.024**	**0.024**	**0.025**	**0.028**	**0.037**
	6	**0.004**	**0.010**	**0.023**	**0.028**	**0.026**	**0.027**	**0.028**	**0.031**	**0.040**
	7	**0.004**	**0.013**	**0.024**	**0.028**	**0.027**	**0.027**	**0.030**	**0.031**	**0.040**
	8	**0.004**	**0.015**	**0.024**	**0.027**	**0.027**	**0.028**	**0.031**	**0.033**	**0.043**
	2	-0.110	-0.108	-0.108	-0.106	-0.108	-0.107	-0.106	-0.108	-0.107
	3	-0.092	-0.091	-0.091	-0.088	-0.089	-0.088	-0.085	-0.083	-0.076
	4	**0.006**	**0.011**	**0.019**	**0.026**	**0.027**	**0.027**	**0.027**	**0.028**	**0.037**
7	5	**0.009**	**0.012**	**0.020**	**0.023**	**0.023**	**0.023**	**0.024**	**0.025**	**0.035**
	6	**0.005**	**0.013**	**0.024**	**0.025**	**0.025**	**0.025**	**0.027**	**0.028**	**0.037**
	7	**0.004**	**0.016**	**0.026**	**0.029**	**0.027**	**0.027**	**0.029**	**0.029**	**0.038**
	8	**0.004**	**0.016**	**0.026**	**0.028**	**0.028**	**0.028**	**0.030**	**0.031**	**0.041**
	2	-0.111	-0.109	-0.109	-0.107	-0.109	-0.109	-0.108	-0.111	-0.110
	3	-0.094	-0.091	-0.091	-0.088	-0.088	-0.088	-0.083	-0.083	-0.077
	4	**0.005**	**0.008**	**0.016**	**0.021**	**0.021**	**0.022**	**0.023**	**0.025**	**0.033**
8	5	**0.004**	**0.007**	**0.015**	**0.020**	**0.021**	**0.021**	**0.022**	**0.024**	**0.033**
	6	**0.004**	**0.011**	**0.021**	**0.023**	**0.022**	**0.021**	**0.023**	**0.024**	**0.034**
	7	**0.003**	**0.015**	**0.023**	**0.026**	**0.024**	**0.023**	**0.025**	**0.026**	**0.034**
	8	**0.006**	**0.017**	**0.026**	**0.028**	**0.026**	**0.025**	**0.028**	**0.028**	**0.037**

Note: Boldface number refers to a decrease in the WAD as a test length increases.

Table 10

Decreases in WAD for the Change in Test Length from 40 to 60

C_Y	C_X	.01	.05	.10	.20	.30	.40	.50	.75	1.00
						S				
2	2	**0.015**	**0.014**	**0.014**	**0.013**	**0.011**	**0.013**	**0.013**	**0.010**	**0.008**
	3	**0.007**	**0.007**	**0.007**	**0.009**	**0.007**	**0.009**	**0.009**	**0.008**	**0.006**
	4	-0.041	-0.042	-0.041	-0.040	-0.041	-0.039	-0.039	-0.041	-0.043
	5	-0.040	-0.040	-0.040	-0.039	-0.040	-0.038	-0.038	-0.039	-0.041
	6	-0.037	-0.038	-0.038	-0.037	-0.037	-0.036	-0.035	-0.036	-0.038
	7	-0.038	-0.039	-0.039	-0.038	-0.038	-0.037	-0.036	-0.037	-0.039
	8	-0.039	-0.040	-0.040	-0.039	-0.039	-0.038	-0.038	-0.039	-0.041
3	2	**0.003**	**0.002**	**0.001**	-0.002	-0.002	-0.005	-0.005	-0.006	-0.004
	3	**0.016**	**0.013**	**0.013**	**0.010**	**0.011**	**0.010**	**0.010**	**0.010**	**0.009**
	4	-0.050	-0.050	-0.049	-0.048	-0.049	-0.050	-0.049	-0.049	-0.050
	5	-0.049	-0.050	-0.049	-0.048	-0.049	-0.050	-0.049	-0.051	-0.050
	6	-0.046	-0.047	-0.046	-0.044	-0.045	-0.046	-0.045	-0.047	-0.047
	7	-0.046	-0.047	-0.046	-0.045	-0.046	-0.046	-0.046	-0.048	-0.048
	8	-0.046	-0.046	-0.046	-0.043	-0.044	-0.045	-0.044	-0.046	-0.046
4	2	-0.027	-0.026	-0.026	-0.027	-0.028	-0.028	-0.029	-0.028	-0.026
	3	-0.041	-0.040	-0.040	-0.041	-0.039	-0.038	-0.039	-0.039	-0.038
	4	-0.001	-0.001	**0.001**	**0.003**	**0.005**	**0.007**	**0.008**	**0.010**	**0.011**
	5	-0.006	-0.007	-0.006	-0.002	**0.003**	**0.005**	**0.005**	**0.008**	**0.008**
	6	-0.004	-0.005	-0.004	**0.002**	**0.007**	**0.008**	**0.009**	**0.012**	**0.012**
	7	-0.006	-0.006	-0.006	**0.000**	**0.005**	**0.006**	**0.008**	**0.010**	**0.010**
	8	-0.003	-0.004	-0.003	**0.002**	**0.007**	**0.006**	**0.007**	**0.010**	**0.009**
5	2	-0.026	-0.025	-0.025	-0.026	-0.025	-0.026	-0.026	-0.026	-0.023
	3	-0.038	-0.037	-0.038	-0.039	-0.037	-0.036	-0.037	-0.037	-0.036
	4	-0.002	-0.002	-0.003	**0.003**	**0.006**	**0.009**	**0.009**	**0.012**	**0.011**
	5	-0.002	-0.001	-0.002	**0.005**	**0.008**	**0.010**	**0.010**	**0.012**	**0.012**
	6	-0.001	-0.003	-0.002	**0.005**	**0.009**	**0.011**	**0.011**	**0.013**	**0.012**
	7	-0.003	-0.003	-0.003	**0.006**	**0.009**	**0.012**	**0.012**	**0.014**	**0.013**
	8	-0.001	-0.001	**0.000**	**0.007**	**0.008**	**0.010**	**0.010**	**0.012**	**0.011**
6	2	-0.025	-0.024	-0.024	-0.024	-0.024	-0.024	-0.024	-0.024	-0.021
	3	-0.038	-0.038	-0.038	-0.039	-0.037	-0.036	-0.038	-0.037	-0.036
	4	0.000	-0.001	0.000	**0.005**	**0.008**	**0.011**	**0.011**	**0.013**	**0.013**
	5	0.000	0.000	**0.001**	**0.008**	**0.010**	**0.012**	**0.012**	**0.014**	**0.014**
	6	-0.002	-0.002	**0.000**	**0.008**	**0.011**	**0.013**	**0.013**	**0.015**	**0.015**
	7	-0.002	-0.003	-0.001	**0.008**	**0.009**	**0.011**	**0.012**	**0.014**	**0.014**
	8	-0.002	-0.003	**0.001**	**0.008**	**0.009**	**0.011**	**0.011**	**0.012**	**0.012**
7	2	-0.028	-0.026	-0.026	-0.026	-0.026	-0.026	-0.026	-0.026	-0.023
	3	-0.038	-0.037	-0.038	-0.039	-0.038	-0.037	-0.038	-0.038	-0.037
	4	-0.002	-0.002	-0.003	**0.004**	**0.008**	**0.011**	**0.012**	**0.014**	**0.014**
	5	-0.002	-0.001	**0.001**	**0.010**	**0.011**	**0.013**	**0.013**	**0.016**	**0.015**
	6	-0.001	-0.002	-0.001	**0.011**	**0.012**	**0.013**	**0.013**	**0.016**	**0.015**
	7	-0.002	-0.001	**0.000**	**0.010**	**0.012**	**0.014**	**0.014**	**0.016**	**0.016**
	8	-0.001	**0.001**	**0.003**	**0.010**	**0.011**	**0.013**	**0.013**	**0.015**	**0.014**
8	2	-0.026	-0.024	-0.025	-0.024	-0.024	-0.024	-0.025	-0.024	-0.022
	3	-0.038	-0.037	-0.038	-0.039	-0.038	-0.037	-0.038	-0.037	-0.037
	4	**0.000**	**0.002**	**0.002**	**0.008**	**0.011**	**0.013**	**0.012**	**0.015**	**0.015**
	5	**0.003**	**0.004**	**0.006**	**0.013**	**0.015**	**0.016**	**0.016**	**0.019**	**0.018**
	6	**0.001**	**0.002**	**0.004**	**0.012**	**0.013**	**0.014**	**0.014**	**0.016**	**0.016**
	7	0.000	**0.002**	**0.004**	**0.011**	**0.012**	**0.014**	**0.014**	**0.016**	**0.016**
	8	-0.002	**0.000**	**0.005**	**0.012**	**0.012**	**0.014**	**0.014**	**0.016**	**0.015**

Note: Boldface number refers to a decrease in the WAD as a test length increases.

Table 11
Overall Results for Summary Variables

C_Y	C_X	Score-Gap			RS-to-lprss			RS-to-hprss			Many-to-One		
		Min	Mode	Max	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max
	2	0	1	4	1	1	1	0	1	2	0	0	8
	3	0	1	3	0	1	1	0	1	3	0	0	8
	4	0	1	3	1	1	1	0	1	4	0	0	7
2	5	0	1	3	1	1	1	0	1	4	0	0	7
	6	0	1	3	1	1	1	0	1	4	0	0	7
	7	0	1	3	1	1	1	0	1	4	0	0	7
	8	0	1	3	1	1	1	0	1	4	0	0	7
	2	0	1	3	1	1	3	0	1	2	0	0	8
	3	0	1	4	1	1	1	0	1	2	0	0	9
	4	0	0	2	1	1	2	0	1	2	0	0	8
3	5	0	0	3	1	1	2	0	1	2	0	0	7
	6	0	0	3	1	1	2	0	1	2	0	0	7
	7	0	0	3	1	1	2	0	1	2	0	0	7
	8	0	0	3	1	1	2	0	1	2	0	0	7
	2	0	1	4	1	1	1	0	1	1	0	0	8
	3	0	2	4	0	1	1	0	1	1	0	0	9
	4	0	1	3	1	1	1	0	1	2	0	0	7
4	5	0	1	3	1	1	1	0	1	1	0	0	7
	6	0	1	3	1	1	1	0	1	1	0	0	7
	7	0	1	3	1	1	1	0	1	1	0	0	7
	8	0	1	3	1	1	1	0	1	1	0	0	7
	2	0	1	4	1	1	1	0	1	1	0	0	8
	3	0	2	4	0	1	1	0	1	1	0	0	9
	4	0	1	3	1	1	1	0	1	1	0	0	7
5	5	0	1	3	1	1	1	0	1	1	0	0	8
	6	0	1	3	1	1	1	0	1	1	0	0	8
	7	0	1	3	1	1	1	0	1	1	0	0	8
	8	0	1	3	1	1	1	0	1	1	0	0	8
	2	0	1	4	1	1	1	0	1	1	0	0	8
	3	0	2	4	0	1	1	0	1	1	0	0	9
	4	0	1	3	1	1	1	0	1	1	0	0	7
6	5	0	1	3	1	1	1	0	1	1	0	0	7
	6	0	1	3	1	1	1	0	1	1	0	0	7
	7	0	1	3	1	1	1	0	1	1	0	0	8
	8	0	1	3	1	1	1	0	1	1	0	0	8
	2	0	1	4	1	1	1	0	1	1	0	0	8
	3	0	2	4	0	1	1	0	1	1	0	0	9
	4	0	1	3	1	1	1	0	1	1	0	0	7
7	5	0	1	3	1	1	1	0	1	1	0	0	7
	6	0	1	3	1	1	1	0	1	1	0	0	7
	7	0	1	3	1	1	1	0	1	1	0	0	8
	8	0	1	3	1	1	1	0	1	1	0	0	8
	2	0	1	4	1	1	1	0	1	1	0	0	7
	3	0	2	4	0	1	1	0	1	1	0	0	9
	4	0	1	3	1	1	1	0	1	1	0	0	7
8	5	0	1	3	1	1	1	0	1	1	0	0	7
	6	0	1	3	1	1	1	0	1	1	0	0	7
	7	0	1	3	1	1	1	0	1	1	0	0	7
	8	0	1	3	1	1	1	0	1	1	0	0	8
	.01	0	1	3	1	1	1	0	1	1	0	0	8
	.05	0	1	3	1	1	1	0	1	1	0	0	8
	.10	0	1	3	1	1	1	0	1	1	0	0	8
	.20	0	1	3	1	1	1	0	1	1	0	0	8
S	.30	0	1	3	1	1	1	0	1	1	0	0	8
	.40	0	1	3	1	1	1	0	1	1	0	0	8
	.50	0	1	3	1	1	1	0	1	1	0	0	8
	.75	0	1	3	1	1	1	0	1	2	0	0	8
	1.00	0	1	4	1	1	1	0	1	2	0	0	8

Table 12
Summary Variables for Test Length of 40 Items

C_Y	C_X	Score-Gap			RS-to-lprss			RS-to-hprss			Many-to-One		
		Min	Mode	Max	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max
	2	0	1	4	1	1	1	0	1	2	0	0	1
	3	0	1	3	0	1	1	0	1	3	0	0	4
	4	0	1	3	1	1	1	0	1	4	0	0	3
2	5	0	1	3	1	1	1	0	1	4	0	0	4
	6	0	1	3	1	1	1	0	1	4	0	0	4
	7	0	1	3	1	1	1	0	1	4	0	0	4
	8	0	1	3	1	1	1	0	1	4	0	0	4
	2	0	1	3	1	1	3	0	1	2	0	0	2
	3	0	1	4	1	1	1	0	1	2	0	0	3
	4	0	0	2	1	1	2	0	1	2	0	0	2
3	5	0	0	3	1	1	2	0	1	2	0	0	2
	6	0	0	3	1	1	2	0	1	2	0	0	2
	7	0	0	3	1	1	2	0	1	2	0	0	2
	8	0	0	3	1	1	2	0	1	2	0	0	2
	2	0	1	4	1	1	1	0	1	1	0	0	2
	3	0	2	4	0	1	1	0	1	1	0	0	3
	4	0	1	3	1	1	1	0	1	2	0	0	2
4	5	0	1	3	1	1	1	0	1	1	0	0	2
	6	0	1	3	1	1	1	0	1	1	0	0	2
	7	0	1	3	1	1	1	0	1	1	0	0	2
	8	0	1	3	1	1	1	0	1	1	0	0	2
	2	0	1	4	1	1	1	0	1	1	0	0	2
	3	0	2	4	0	1	1	0	1	1	0	0	3
	4	0	1	3	1	1	1	0	1	1	0	0	2
5	5	0	1	3	1	1	1	0	1	1	0	0	2
	6	0	1	3	1	1	1	0	1	1	0	0	2
	7	0	1	3	1	1	1	0	1	1	0	0	2
	8	0	1	3	1	1	1	0	1	1	0	0	2
	2	0	1	4	1	1	1	0	1	1	0	0	2
	3	0	2	4	0	1	1	0	1	1	0	0	3
	4	0	1	3	1	1	1	0	1	1	0	0	2
6	5	0	1	3	1	1	1	0	1	1	0	0	2
	6	0	1	3	1	1	1	0	1	1	0	0	2
	7	0	1	3	1	1	1	0	1	1	0	0	2
	8	0	1	3	1	1	1	0	1	1	0	0	2
	2	0	1	4	1	1	1	0	1	1	0	0	2
	3	0	2	4	0	1	1	0	1	1	0	0	3
	4	0	1	3	1	1	1	0	1	1	0	0	2
7	5	0	1	3	1	1	1	0	1	1	0	0	2
	6	0	1	3	1	1	1	0	1	1	0	0	2
	7	0	1	3	1	1	1	0	1	1	0	0	2
	8	0	1	3	1	1	1	0	1	1	0	0	2
	2	0	1	4	1	1	1	0	1	1	0	0	2
	3	0	2	4	0	1	1	0	1	1	0	0	3
	4	0	1	3	1	1	1	0	1	1	0	0	2
8	5	0	1	3	1	1	1	0	1	1	0	0	2
	6	0	1	3	1	1	1	0	1	1	0	0	2
	7	0	1	3	1	1	1	0	1	1	0	0	2
	8	0	1	3	1	1	1	0	1	1	0	0	2
	.01	0	1	3	1	1	1	0	1	1	0	0	2
	.05	0	1	3	1	1	1	0	1	1	0	0	2
	.10	0	1	3	1	1	1	0	1	1	0	0	2
	.20	0	1	3	1	1	1	0	1	1	0	0	2
S	.30	0	1	3	1	1	1	0	1	1	0	0	2
	.40	0	1	3	1	1	1	0	1	1	0	0	2
	.50	0	1	3	1	1	1	0	1	1	0	0	2
	.75	0	1	3	1	1	1	0	1	2	0	0	2
	1.00	0	1	4	1	1	1	0	1	2	0	0	2

Table 13
Summary Variables for Test Length of 60 Items

C_Y	C_X	Score-Gap			RS-to-lprss			RS-to-hprss			Many-to-One		
		Min	Mode	Max	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max
	2	0	1	2	1	1	1	0	1	2	0	1	8
	3	0	1	2	0	1	1	0	1	2	1	4	8
	4	0	1	2	1	1	1	1	1	4	1	4	7
2	5	0	1	2	1	1	1	1	1	3	1	4	7
	6	0	1	2	1	1	1	1	1	3	1	4	7
	7	0	1	2	1	1	1	1	1	3	1	4	7
	8	0	1	2	1	1	1	1	1	3	1	4	7
	2	0	1	1	1	1	3	0	1	1	1	3	8
	3	0	1	2	1	1	1	0	1	2	0	1	9
	4	0	0	1	1	1	2	0	1	2	1	4	8
3	5	0	0	1	1	1	2	1	1	2	1	4	7
	6	0	0	1	1	1	2	1	1	2	1	4	7
	7	0	0	1	1	1	2	1	1	2	1	4	7
	8	0	0	2	1	1	2	1	1	2	1	4	7
	2	0	1	2	1	1	1	0	1	1	1	4	8
	3	0	2	2	0	1	1	0	1	1	1	5	9
	4	0	1	2	1	1	1	0	1	1	0	1	7
4	5	0	1	2	1	1	1	0	1	1	0	1	7
	6	0	1	2	1	1	1	0	1	1	0	1	7
	7	0	1	2	1	1	1	0	1	1	0	1	7
	8	0	1	2	1	1	1	0	1	1	0	2	7
	2	0	1	2	1	1	1	0	1	1	1	4	8
	3	0	2	2	0	1	1	0	1	1	1	4	9
	4	0	1	2	1	1	1	0	1	1	0	1	7
5	5	0	1	2	1	1	1	0	1	1	0	1	8
	6	0	1	2	1	1	1	0	1	1	0	1	8
	7	0	1	2	1	1	1	0	1	1	0	1	8
	8	0	1	2	1	1	1	0	1	1	0	1	8
	2	0	1	2	1	1	1	0	1	1	1	4	8
	3	0	2	2	0	1	1	0	1	1	1	4	9
	4	0	1	2	1	1	1	0	1	1	0	1	7
6	5	0	1	2	1	1	1	0	1	1	0	1	7
	6	0	1	2	1	1	1	0	1	1	0	1	7
	7	0	1	2	1	1	1	0	1	1	0	1	8
	8	0	1	2	1	1	1	0	1	1	0	1	8
	2	0	1	2	1	1	1	0	1	1	1	4	8
	3	0	2	2	0	1	1	0	1	1	2	4	9
	4	0	1	2	1	1	1	0	1	1	0	1	7
7	5	0	1	2	1	1	1	0	1	1	0	1	7
	6	0	1	2	1	1	1	0	1	1	0	1	7
	7	0	1	2	1	1	1	0	1	1	0	1	8
	8	0	1	2	1	1	1	0	1	1	0	1	8
	2	0	1	2	1	1	1	0	1	1	1	4	7
	3	0	2	2	0	1	1	0	1	1	2	4	9
	4	0	1	2	1	1	1	0	1	1	0	1	7
8	5	0	1	2	1	1	1	0	1	1	0	1	7
	6	0	1	2	1	1	1	0	1	1	0	1	7
	7	0	1	2	1	1	1	0	1	1	0	1	7
	8	0	1	2	1	1	1	0	1	1	0	1	8
	.01	0	1	2	1	1	1	0	1	1	0	1	8
	.05	0	1	2	1	1	1	0	1	1	0	1	8
	.10	0	1	2	1	1	1	0	1	1	0	1	8
	.20	0	1	2	1	1	1	0	1	1	0	1	8
S	.30	0	1	2	1	1	1	0	1	1	0	1	8
	.40	0	1	2	1	1	1	0	1	1	0	1	8
	.50	0	1	2	1	1	1	0	1	1	0	1	8
	.75	0	1	2	1	1	1	0	1	1	0	1	8
	1.00	0	1	2	1	1	1	0	1	1	0	1	8

Table 14
Postsmoothing Degree with the Most Similar Equating Results given the AIC-selected Presmoothing Degree

Unrounded Scale Scores

$C_Y:C_X$ \ S	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total
2:4	2	2	1							5
2:5				1	1					2
3:3			1							1
3:4	6	4	2						1	13
3:5	3	1								4
3:6	1		1							2
3:7	1		1							2
3:8										
4:2	3	4	1							8
4:3	2	3	1	1						7
4:4	41	112	205	90	19	3	2		1	473
4:5	16	60	84	37	10				1	208
4:6	14	30	42	7	1					94
4:7	5	11	15	2						33
4:8	11	27	28	17	4					87
5:2	1	3	2							6
5:3	2			1						3
5:4	30	55	82	28	4					199
5:5	14	27	44	29	2		1			117
5:6	5	18	11	4						38
5:7	2	12	15	5						34
5:8	5	13	9	6						33
6:2	1	1								2
6:3										
6:4	15	16	29	8	2					70
6:5	5	10	25	6						46
6:6	1	8	6	3						18
6:7	2	5	4	2						13
6:8	6	8	6	2						22
7:3										
7:4	9	12	20	2	1					44
7:5	5	9	12	2						28
7:6	2	5	3	2	1					13
7:7		4	6	1	1					12
7:8	2	3	3							8
8:3										
8:4	22	28	8	9						67
8:5	8	18	18	5	1					50
8:6	4	5	5	1						15
8:7	1	2	1	1						5
8:8	4	8	5	1						18
Total	251	524	697	273	46	3	3		3	1800

Rounded Scale Scores

$C_Y:C_X$ \ S	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total
2:4	2		1				1	1		5
2:5	1						1			2
3:3	1									1
3:4	2	1	6			1		1	2	13
3:5	3								1	4
3:6		2								2
3:7	1		1							2
3:8										
4:2	3							1	4	8
4:3	3		1	1	1				1	7
4:4	134	66	108	62	25	11	7	9	51	473
4:5	57	32	41	28	11	7	5	3	24	208
4:6	32	15	17	9	3	1	3	7	7	94
4:7	13	6	7	3	1	1	1	1	1	33
4:8	23	12	20	6	9	6	1	2	8	87
5:2	2	2		1					1	6
5:3	2			1						3
5:4	55	35	35	23	6	7	7	1	30	199
5:5	32	19	21	20	5	1	2	3	14	117
5:6	13	7	4	10	2		2			38
5:7	9	3	7	7	1	3	1	1	2	34
5:8	8	5	7	3	3		2		5	33
6:2	1	1								2
6:3										
6:4	20	10	16	8	5	2	1		8	70
6:5	14	2	11	7		1	4	2	5	46
6:6	4	3	1	2	2		4		2	18
6:7	3	4	1	4	1					13
6:8	9	4	4	4		1				22
7:3										
7:4	12	7	9	5	2	1	1		7	44
7:5	11	7	7		2		1			28
7:6	3	2	3	1		1	1	1	1	13
7:7	2		4	3		1	1		1	12
7:8	5		2			1				8
8:3										
8:4	32	10	5	4	4	3	4		5	67
8:5	17	9	6	7	2	2	2	1	4	50
8:6	3	4	4	2	1		1			15
8:7	1	1	1		1	1				5
8:8	11			2	2		2		1	18
Total	544	269	350	222	88	56	48	38	185	1800

Note: A shaded cell refers to a postsmoothing degree selected for giving the most similar equating results for a presmoothing degree with $C_{Y_{AIC}}:C_{X_{AIC}}$.

Table 15

Postsmoothing Degree with the Most Similar Equating Results given the AIC-selected Presmoothing Degree (Unrounded Scale Scores)

	Test Length 15										Test Length 40										Test Length 60									
	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total
2:4	2	2	1							5																				
2:5		1	1							2																				
3:3			1							1																				
3:4	6	4	2						1	13																				
3:5	3	1								4																				
3:6	1		1							2																				
3:7	1		1							2																				
4:2	3	4	1							8																				
4:3	2	3	1	1						7																				
4:4	39	62	51	9	2					163	2	35	82	49	9				1	178		15	72	32	8	3	1		1	132
4:5	15	33	21	6	1				1	77	1	15	27	20	1					64		12	36	11	8					67
4:6	14	12	12							38		11	19	4						34		7	11	3	1					22
4:7	4	4	1							9		2	6	1						9	1	5	8	1						15
4:8	6	10	3	1	2					22	4	14	12	6	1					37	1	3	13	10	1					28
5:2	1	3	2							6																				
5:3	2			1						3																				
5:4	27	20	8	2						57	3	16	35	10	1					65		19	39	16	3					77
5:5	12	8	12	3	1					36	2	7	12	12						33		12	20	14	1				1	48
5:6	5	6	1							12		9	6	1						16		3	4	3						10
5:7	1		1							2	1	4	6	3						14		8	8	2						18
5:8	3	1								4	2	6	4	2						14		6	5	4						15
6:2	1	1								2																				
6:4	14	8	7							29	1	4	9	3						17		4	13	5	2					24
6:5	5	5	4	1						15			11							11		5	10	5						20
6:6	1	5		1						7		2	3	1						6		1	3	1						5
6:7	2	1								3		1								1		3	4	2						9
6:8	5	1								6		3	3							6	1	4	3	2						10
7:4	7	3	1							11	1	4	12							17	1	5	7	2	1					16
7:5	2	2	1							5	2	3	6	1						12	1	4	5	1						11
7:6	2				1					3												5	3	2						10
7:7												2	2		1					5		2	4	1						7
7:8	2	1								3			2							2	2	1								3
8:4	13	7		1						21	7	12	3	2						24	2	9	5	6						22
8:5	3	6	2		1					12	5	7	4	1						17		5	12	4						21
8:6	3	2								5	1	3	3							7			2	1						3
8:7		1								1	1	1	1	1						4										
8:8	3	1								4	1	4	1	1						7		3	4							7
Total	210	217	136	27	8				2	600	34	165	269	118	13				1	600	7	142	292	128	25	3	2		1	600

Note: A shaded cell refers to a postsmoothing degree selected for giving the most similar equating results for a presmoothing degree with $C_{Y_{AIC}}{:}C_{X_{AIC}}$.

Table 16

Postsmoothing Degree with the Most Similar Equating Results given the AIC-selected Presmoothing Degree (Rounded Scale Scores)

Test Length 15

	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total
2:4	2		1					1	1	5
2:5	1					1				2
3:3	1									1
3:4	2	1	6			1		1	2	13
3:5	3								1	4
3:6		2								2
3:7	1		1							2
4:2	3						1		4	8
4:3	3		1	1	1				1	7
4:4	61	21	23	15	6	2	2	3	30	163
4:5	27	6	12	4	2	3	2	3	18	77
4:6	17	4	3	2	2	1		3	6	38
4:7	7			1					1	9
4:8	6	3	4	1	1	2		1	4	22
5:2	2	2		1					1	6
5:3	2			1						3
5:4	21	6	3	2	3	2	3		17	57
5:5	14	2	1	3	2	1	1	2	10	36
5:6	6	2	2	2						12
5:7				1					1	2
5:8			2						2	4
6:2	1	1								2
6:4	13	3	2	3		2			6	29
6:5	5	1	3				2		4	15
6:6	3		1			2			1	7
6:7	2			1						3
6:8	4	1		1						6
7:4	6	1		1					3	11
7:5	3	1			1					5
7:6	1					1		1		3
7:7										
7:8	1		2							3
8:4	12	1	1	1	1				5	21
8:5	5	1		3			1		2	12
8:6	2	1		1	1					5
8:7		1								1
8:8	3					1				4
Total	240	61	68	42	22	17	12	17	121	600

Test Length 40

	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total
2:4										
2:5										
3:3										
3:4										
3:5										
3:6										
3:7										
4:2										
4:3										
4:4	55	19	37	26	11	7	3	5	15	178
4:5	17	13	15	7	4		3		5	64
4:6	12	6	7	3	1		1	3	1	34
4:7	3	2	3			1				9
4:8	13	4	9	3	4	2			2	37
5:2										
5:3										
5:4	17	12	16	6		3	3		8	65
5:5	9	8	8	4			1		3	33
5:6	6	2	1	4	1			2		16
5:7	3	1	4	2		2	1	1		14
5:8	3	2	3	2	1		1		2	14
6:2										
6:4	4	1	5	2	3				2	17
6:5	4	1	4			2				11
6:6	1	1			1		2		1	6
6:7		1								1
6:8	2	1	1	2						6
7:4	3	3	5	4			1		1	17
7:5	5	4	1			1		1		12
7:6										
7:7	1		1	1		1			1	5
7:8	1					1				2
8:4	11	4	2	2	2	1	2			24
8:5	8	3	1			2		1	2	17
8:6	1	2	2	1			1			7
8:7	1		1		1	1				4
8:8	5					1			1	7
Total	185	89	126	70	29	23	21	14	43	600

Test Length 60

	.01	.05	.10	.20	.30	.40	.50	.75	1.00	Total
2:4										
2:5										
3:3										
3:4										
3:5										
3:6										
3:7										
4:2										
4:3										
4:4	18	26	48	21	8	2	2	1	6	132
4:5	13	13	14	17	5	4			1	67
4:6	3	5	7	4			2	1		22
4:7	3	4	4	2	1				1	15
4:8	4	5	7	2	4	2	1	1	2	28
5:2										
5:3										
5:4	17	17	16	15	3	2	1	1	5	77
5:5	9	9	12	13	3			1	1	48
5:6	1	3	1	4	1					10
5:7	6	2	3	5		1			1	18
5:8	5	3	2	1	2		1		1	15
6:2										
6:4	3	6	9	3	2			1		24
6:5	5	1	4	6			1	2	1	20
6:6			2			2	1			5
6:7	1	3	1	3	1					9
6:8	3	2	3	1		1				10
7:4	3	3	4	1	1	1			3	16
7:5	3	2	6							11
7:6	2	2	3	1			1	1		10
7:7	1		3	2			1			7
7:8	3									3
8:4	9	5	2	1	1	2	2			22
8:5	4	5	5	4	2		1			21
8:6		1	2							3
8:7										
8:8	3			2	2					7
Total	119	119	156	110	37	16	15	7	21	600

Note: A shaded cell refers to a postsmoothing degree selected for giving the most similar equating results for a presmoothing degree with $C_{Y_{AIC}}{:}C_{X_{AIC}}$.

References

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics*, *16*, 3-14.

Antal, J., Kim, Y., & Lee, W. (2011, April). *A comparison of presmoothing and postsmoothing methods for the nonequivalent groups with anchor test (NEAT) design.* A paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans.

Hanson, B. A., Zeng, L., & Colton, D. (1994). *A comparison of presmoothing and postsmoothing methods in equipercentile equating.* (ACT Research Report 94-4). Iowa City, IA: American College Testing.

Kim, H. (2014). *A comparison of smoothing methods for the common item nonequivalent groups design* (Unpublished doctoral dissertation). University of Iowa.

Kolen, M. J. (1984). Effectiveness of analytic smoothing in equipercentile equating. *Journal of Educational Statistics*, *9*, 25-44.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). New York: Springer.

Livingston, S. A., & Feryok, N. J. (1987). *Univariate vs. bivariate smoothing in frequency estimation equating.* (Research Report 87-36). Princeton, NJ: Educational Testing Service.