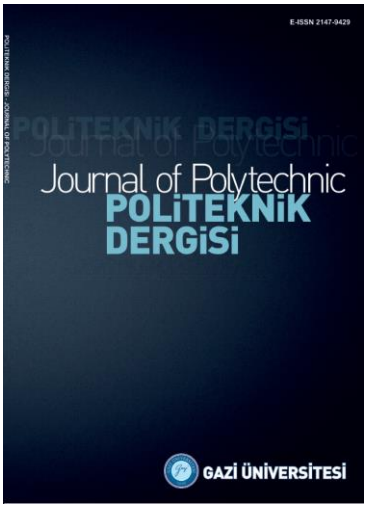


**POLİTEKNİK DERGİSİ**

***JOURNAL of POLYTECHNIC***

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)
URL: http://dergipark.org.tr/politeknik



# Effect of weighting schemes on weighted kappa coefficients in multi-rater agreement studies with ordinal categories

## *Çoklu değerlendiriciye sahip sıralanabilir düzeyli uyum çalışmalarında ağırlıklandırma şemalarının ağırlıklı kappa katsayıları üzerindeki etkisi*

***Yazar(lar) (Author(s)):*** *Ayfer Ezgi YILMAZ[1]*

*ORCID[1]: 0000-0002-6214-8014*



**To cite to this article**: Yilmaz A.E., "Effect of weighting schemes on weighted kappa coefficients in multi-rater agreement studies with ordinal categories", ***Journal of Polytechnic***, *(*): *, (*).

**Bu makaleye şu şekilde atıfta bulunabilirsiniz:** Yilmaz A.E., "Effect of weighting schemes on weighted kappa coefficients in multi-rater agreement studies with ordinal categories", ***Politeknik Dergisi***, *(*): *, (*).

**Erişim linki (To link to this article):** http://dergipark.org.tr/politeknik/archive

**DOI:** 10.2339/politeknik.1568563

# Effect of Weighting Schemes on Weighted Kappa Coefficients in Multi-Rater Agreement Studies with Ordinal Categories

## *Highlights*

- *The extensions of the quadratic, ridit linear, ridit quadratic, exponential linear, and exponential quadratic weighting schemes are proposed.*
- *The accuracy of the multi-rater weighted kappa coefficients and weighting schemes is discussed.*
- *A combination of coefficients and weighting schemes for use with various data compositions is discussed.*

## *Graphical Abstract*

*The results are presented in three parts: 1) Comparison of the weighted kappa coefficients, 2) Comparison of the weighting schemes, and 3) Effect of the number of categories on weighting schemes.*

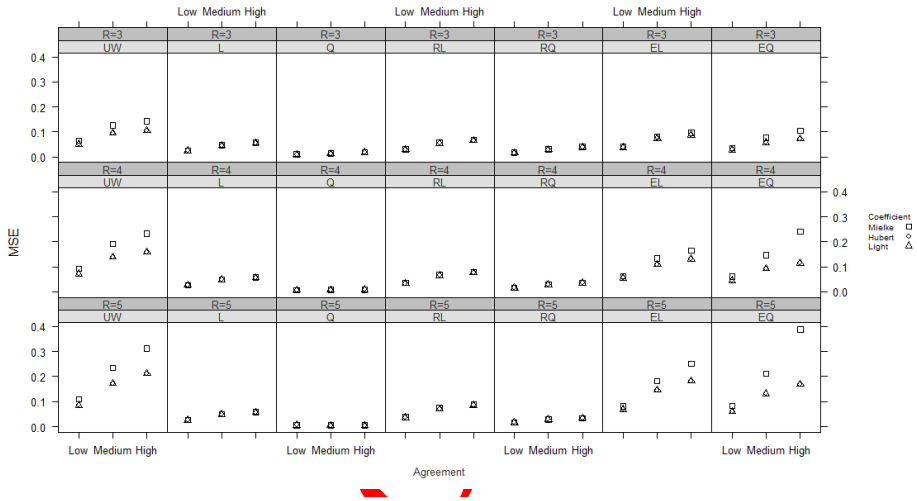


**Figure.** The MSE results of the balanced tables with three raters and $n = 100$ by weighted kappas

## *Aim*

*The aim of this study is to investigate the accuracy of weighted kappa coefficients and the effects of linear, quadratic, ridit type, and exponential type weighting schemes on these coefficients in the multi-rater agreement studies with ordinal categories.*

## *Design & Methodology*

*A Monte Carlo simulation is performed to compare three weighted kappa coefficients and the weighting schemes for multi-raters with ordinal scales. The simulation study uses 216 different combinations.*

## *Originality*

*This study provides an extended comparation of weighted kappa coefficients in terms of weighting scheme, the level of true agreement, and the structure of the table. Also, the extension of the quadratic, ridit linear, ridit quadratic, exponential linear, and exponential quadratic weighting schemes are proposed.*

## *Findings*

*Hubert's and Light's weighted kappas have similar results in most situations. The unweighted kappa coefficients yield poorer results compared to other weighting schemes.*

## *Conclusion*

*The results indicate that the structure of the table, the level of agreement, and the number of categories also have an effect on the coefficients.*

## *Declaration of Ethical Standards*

*The author of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.*

# Effect of Weighting Schemes on Weighted Kappa Coefficients in Multi-Rater Agreement Studies with Ordinal Categories

*Araştırma Makalesi / Research Article*

**Ayfer Ezgi YILMAZ[1]***

[1]Faculty of Science, Department of Statistics, Hacettepe University, Türkiye

(Geliş/Received : 16.10.2024 ; Kabul/Accepted : 20.02.2025 ; Erken Görünüm/Early View : 27.02.2025 )

**ABSTRACT**

Weighted kappa and kappa-like coefficients are used for the calculation of inter-rater agreement in cases where raters classify objects into ordinal categories. Weighted kappa coefficients are extended for use in studies with multiple raters. It is crucial to select appropriate weighting schemes as they can significantly impact the value of the coefficient. In this study, the accuracy of weighted kappa coefficients and the effects of linear, quadratic, ridit type, and exponential type weighting schemes on these coefficients are discussed in the multi-rater agreement studies with ordinal categories. The accuracy of the coefficients is investigated by an illustrative data and a simulation study.

**Keywords: Inter-rater agreement, multi-raters, ordinal scales, weighted kappa, weighting schemes.**

# Çoklu Değerlendiriciye Sahip Sıralanabilir Düzeyli Uyum Çalışmalarında Ağırlıklandırma Şemalarının Ağırlıklı Kappa Katsayıları Üzerindeki Etkisi

**ÖZ**

Ağırlıklı kappa ve kappa benzeri katsayılar, değerlendiricilerin gözlemleri sıralanabilir düzeyler halinde sınıflandırdığı durumlarda, değerlendiriciler arası uyumun hesaplanmasında kullanılır. Ağırlıklı kappa katsayıları, çoklu değerlendiriciye sahip çalışmalarda kullanılmak üzere genişletilmiştir. Katsayının değerini doğrudan etkileyebileceğinden dolayı, uygun ağırlıklandırma şemalarının seçilmesi çok önemlidir. Bu çalışmada, çoklu değerlendiriciye sahip sıralanabilir düzeyli çalışmalarda ağırlıklı kappa katsayılarının doğruluğu ve doğrusal, karesel, ridit tipi ve üstel tipi ağırlıklandırma şemalarının bu katsayılara etkisi tartışılmıştır. Katsayıların doğruluğu, örnek bir veri ve simülasyon çalışması üzerinden araştırılmıştır.

**Anahtar Kelimeler: Değerlendiriciler arası uyum, çoklu değerlendiriciler, sıralanabilir ölçekler, ağırlıklı kappa, ağırlıklandırma şemaları.**



## 1. INTRODUCTION

Inter-rater agreement is essential in the medical, social, and behavioral sciences. For example, when pathologists categorize patients based on cancer stage, the agreement between their categorizations is examined. A high level of agreement is expected to yield more reliable results if there is no discrepancy among the raters. Agreement coefficients are employed to assess the level of agreement. Different coefficients may be used depending on the type of scale and the number of raters. While the kappa coefficient is well-known for inter-rater agreement on nominal scales [1,2], the weighted kappa is often preferred for ordinal scales [3].

Recent studies aim to increase the reliability of results by involving more raters. In diagnostic tests, using more than two doctors is recommended to classify patients accurately according to stage and disease severity. These studies evaluate the agreement level among raters. Despite the various options available for two rater studies, there is a lack of literature examining the agreement of more than two raters who have classified subjects into ordinal categories. Agreement among multiple raters can be defined as 2-agreement or m-agreement [4,5,6,7]. In a 2-agreement or pairwise agreement, the agreement is accepted if the ratings of two of the raters or the agreement of all possible pairs of raters are consistent [6]. A weighted version of the kappa coefficients by Light [8] and Hubert [9] are the weighted kappa coefficients based on 2-agreement. The weighted kappa coefficients by Mielke et al. [10] are proposed for analyzing the agreement between multiple raters' ordered classifications based on 3-agreement. Additionally, several more specific weighted kappas for multiple raters have been proposed by [6,11,12,13].

The weight matrices employed by classical methods are symmetric and are used for continuousordinal scale data. According to Kvalseth [14], cumulative probabilities can be utilized to compute weighted kappa coefficients and provide additional information from ordinal variables.

*Sorumlu Yazar (Corresponding Author)
e-posta : ezgiyilmaz@hacettepe.edu.tr

Yilmaz & Aktas [15] recommends the use of ridit and exponential type weighting schemes for ordinal data, given the prevalent applications of ordinal scales in practice. These weighting schemes are based on the score values and are adapted to the weighted kappa coefficients to conduct weights for ordinal scales. The authors analyze the impact of different weighting schemes through various numerical examples. They conclude that the ridit weighted kappa coefficient is more accurate when the true agreement is slight. The coefficient with quadratic exponential weights is more accurate when the true agreement is moderate, and the coefficient with classical linear and quadratic weights is more accurate when the true agreement is almost perfect. Warrens [16] examines the practical applications of kappa and finds that, in most cases, the level of weighted kappa increases when comparing unweighted and weighted (linear and quadratic) kappa applied to the same data. Specifically, the weighted kappa increases from unweighted to linearly weighted and from linearly weighted to quadratically weighted. Tran et al. [17] provide an extensive comparison of weighted kappa-like coefficients and weighting schemes for the two-rater scenario. They compare six coefficients, including unweighted, linear, quadratic, ordinal, radical, and ratio weighting schemes. Their simulation study demonstrates that the precision of the coefficients is influenced by the weights used when the table is unbalanced and the true agreement is low. When the table is balanced, they perform similarly. Vanbelle et al. [18] provide an information about how to choice inter-rater agreement coefficients for ordinal raters. Demirhan & Yilmaz [19] discuss the effects of gray zones on the ordinal agreement coefficients. Yilmaz & Demirhan [20] discuss the effects of weighted kappa-like coefficients and weighting schemes in machine learning studies. de Raadt et al. [21] provides an extended comparison of the unweighted, linearly weighted, and quadratically weighted kappa coefficients, as well as Pearson, Spearman, and Kendall correlations, and inter-class correlation coefficients. Their findings reveal that the unweighted kappa and linearly weighted kappa coefficients generally yield lower values compared to the other coefficients. Also, the simulation study indicates that when there is moderate agreement, these coefficients tend to have closer values than if the agreement is high. The results support that the quadratically weighted kappa functions as a correlation coefficient. Additionally, the unweighted kappa and linearly weighted kappa generally have lower values than the other coefficients.

Even there are several weighting schemes in the literature for two-raters, there are only a few study for the multi-rater case. Mielke et al. [10,22] and Mielke & Berry [23] discuss the extended linear and quadratic weights for the three raters, and Warrens [6,24] for the multi-rater case.

The previous studies have demonstrated the significance of choosing an appropriate weighting scheme for weighted kappas. The weights have a direct impact on the agreement coefficients. In this study, we focus on the coefficients derived from combinations of 2- and m-agreement and weighting schemes to yield accurate and reliable inferences for multi-rater studies involving ordinal categories. We consider the Light's, Hubert's and Mielke's weighted kappa coefficients, exhibiting linear and quadratic versions of classical, ridit type, and exponential type weighting schemes. We examine a total of 21 combinations and compare them based on the mean square error (MSE) and mean absolute error (MAE), incorporating differing sample sizes, number of categories, number of raters, agreement levels, and table structure (balanced and unbalanced). The study's contributions include: i) the extension of the quadratic, ridit linear, ridit quadratic, exponential linear, and exponential quadratic weighting schemes, ii) an investigation into the accuracy of the multi-rater weighted kappa coefficients, iii) an investigation into the accuracy of the weighting schemes, iv) identification of a combination of coefficients and weighting schemes for use with various data compositions.

The general information on weighted inter-rater agreement coefficients is presented in Section 2. Linear and the proposed weighting schemes are presented in Section 3. The results of the illustrative data and synthetic data are presented in Sections 4 and 5. General conclusions are given in Section 6.

## 2. INTER-RATER AGREEMENT COEFFICIENTS

Cohen [1] unweighted kappa and Cohen [3] weighted kappa coefficients are suggested to calculate the inter-rater agreemen for nominal and ordinal classifications of two rates, respectively. Suppose a square contingency table with $R$ categories in the row and column variables. $n_{ij}$ denote the number of subjects and $\pi_{ij}$ denote the cell probabilities where $i, j = 1, \dots, R$. $\pi_{i.}$ and $\pi_{.j}$ are the row and column totals. Cohen's unweighted kappa coefficient is

$$\kappa = \frac{\sum_{i=1}^{R} \pi_{ii} - \sum_{i=1}^{R} \pi_{i.}\,\pi_{.i}}{1 - \sum_{i=1}^{R} \pi_{i.}\,\pi_{.i}}, \qquad (1)$$

weighted kappa coefficient is

$$\kappa_w = \frac{\sum_{i=1}^{R}\sum_{j=1}^{R} w_{ij}\pi_{ij} - \sum_{i=1}^{R}\sum_{j=1}^{R} w_{ij}\pi_{i.}\,\pi_{.j}}{1 - \sum_{i=1}^{R}\sum_{j=1}^{R} w_{ij}\pi_{i.}\,\pi_{.j}}. \qquad (2)$$

For the classifications of more than two raters, Light [8], Hubert [9], and Mielke et al. [10] weighted kappa coefficients can be refereed. Hubert's $\kappa$ coefficient is reformulated for ordinal tables [25]. Suppose three raters rate $n$ subjects into $R$ categories. $n_{ijk}$ is the number of subjects and $\pi_{ijk}$ is the cell probabilities. The marginal probabilities $p_i$, $q_j$, and $r_k$ are

$$p_i = \sum_{j=1}^{R}\sum_{k=1}^{R} \pi_{ijk}\,, \qquad q_i = \sum_{j=1}^{R}\sum_{k=1}^{R} \pi_{jik}, \qquad (3)$$

$$r_i = \sum_{j=1}^{R}\sum_{k=1}^{R} \pi_{jki}.$$

The sub-tables $A = \{a_{ij}\}$, $B = \{b_{ij}\}$, and $C = \{c_{ij}\}$ are

$$a_{ij} = \sum_{k=1}^{R} \pi_{ijk}, \qquad b_{ij} = \sum_{k=1}^{R} \pi_{ikj}, \qquad c_{ij} = \sum_{k=1}^{R} \pi_{kij}. \tag{4}$$

Hubert's weighted kappa coefficient can be calculated by Eq.5 and $w_{ij}$ is the weight, ranges $0 \le w_{ij} \le 1$.

Hubert's weighted kappa coefficient can be extended for more than three raters. Let $n_{(1i)(2i)\dots(M_i)}$ be the number of subjects at category $i$ where $M$ is the number of raters and $i = 1, \dots, R$. $\pi_{(1i)(2i)\dots(M_i)}$ represents the cell probabilities. $a_{mij}$ represents the $R \times R$ sub-tables and $p_{mi}$ represents the marginal probabilities where $m = 1, \dots, M$ and $i, j = 1, \dots, R$. Then Hubert's generalized weighted kappa coefficient based on 2-agreement is given in Eq. 6.

Mielke et al. [9] suggest a weighted kappa coefficient for three rater studies with ordinal categories. Mielke's weighted kappa coefficient as 3-agreement is given in Eq. 7. Here $w_{ijk}$ is the weight, ranges $0 \le w_{ijk} \le 1$.

Light's [8] kappa coefficient is an arithmetic mean of the kappa coefficients calculated for every possible rater pair. The weighted version of Light's $\kappa$ coefficient is calculated using weighted kappas. Suppose there are $M$ raters and $\kappa_w^{ij}$ is the weighted kappa coefficient among $i$th and $j$th raters where $i < j$. $\kappa_w^{ij}$ can be calculated from Eq. 2 for each sub-table in Eq. 4. Then, Light's weighted kappa coefficient is given in Eq. 8.

In the literature, numerous interpretations of the kappa statistic have been proposed. Landis & Koch [26] propose the following levels of interpretation. Additionally, Altman [27]'s and Fleiss et al. [28]'s interpretation levels are available (see [29], Table 5).

$$H.\kappa_w = \frac{\sum_{i=1}^{R}\sum_{j=1}^{R} w_{ij}(a_{ij} + b_{ij} + c_{ij}) - \sum_{i=1}^{R}\sum_{j=1}^{R} w_{ij}(p_i q_j + p_i r_j + q_i r_j)}{3 - \sum_{i=1}^{R}\sum_{j=1}^{R} w_{ij}(p_i q_j + p_i r_j + q_i r_j)}, \tag{5}$$

$$H.\kappa_w = \frac{\sum_{m=1}^{M}\sum_{i,j=1}^{R} w_{ij} a_{mij} - \sum_{m<m`=1}^{M}\sum_{i,j=1}^{R} w_{ij} p_{mi} p_{m`j}}{\frac{M(M-1)}{2} - \sum_{m<m`=1}^{M}\sum_{i,j=1}^{R} w_{ij} p_{mi} p_{m`j}}. \tag{6}$$

$$M.\kappa_w = \frac{\sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} w_{ijk}\pi_{ijk} - \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} w_{ijk} p_i q_j r_k}{1 - \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} w_{ijk} p_i q_j r_k}. \tag{7}$$

$$L.\kappa_w = \frac{2}{M(M-1)} \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} \kappa_w^{ij}. \tag{8}$$

## 3. WEIGHTING SCHEMES

According to the weighting scheme, each cell is attributed with a defined value of weight. The most commonly used weight schemes are the linear weights [30] in Eq. 10 and the quadratic weights [31] in Eq. 11. The linear weights consist of equally spaced numbers. Furthermore, the quadratic weights are a function that increases quadratically as the number of steps increases, whereas the linear weights increase linearly with the number of steps.

- Unweighted:

$$w_{ij} = \begin{cases} 1, & i = j \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

- Linear weights:

$$w_{ij} = 1 - \frac{|i-j|}{R-1}. \tag{10}$$

- Quadratic weights:

$$w_{ij} = 1 - \frac{(i-j)^2}{(R-1)^2}. \tag{11}$$

Although the literature offers numerous suggestions for weighting schemes in two-rater studies, there is a lack of options for multi-rater studies. The linear weight version for three raters, in its generalized form, is discussed by [23].

$$w_{ijk} = 1 - \frac{|i-j| + |i-k| + |j-k|}{2(R-1)}, \tag{12}$$

where $w_{iii} = 1$ for $i, j, k = 1, \dots, R$.

Mielke's weighted kappa coefficient is calculated using the 3-agreement and requires a three-dimensional matrix. This study expands on the classical quadratic weights and introduces the linear and quadratic versions of the ridit and exponential weights to compute Mielke's weighted kappa coefficient.

The unweighted weighting scheme is the matrix of one. The unweighted scheme is

$$w_{ijk} = \begin{cases} 1, & i = j = k \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

The three-dimensional generalized quadratic weights matrix to calculate Mielke's weighted kappa coefficient is

$$w_{ijk} = 1 - \frac{(i-j)^2 + (i-k)^2 + (j-k)^2}{2(R-1)^2}. \tag{14}$$

The weighting matrices for classical methods are symmetric and applicable to continuous-ordinal scale data. Yilmaz & Aktas [15] suggest using ridit and exponential weighting schemes for ordinal scale data. These schemes consider cumulative probabilities rather than constant values of $i$ and $j$. These schemes consider cumulative probabilities rather than constant values of $i$ and $j$. As a result, they exhibit an asymmetric structure.

In this study, Light's, Hubert's, and Mielke's weighted kappa coefficients are discussed with the weighting schemes for three or more raters. Since Mielke's weighted kappa coefficient is proposed for three raters, the ridit and exponential weights are proposed for the cases of three raters. On the other hand, a two-dimensional weighting matrix is used to calculate Hubert and Light's kappa coefficients. Thus, two-dimensional ridit and exponential weights are also suggested.

Ridit type scores are suggested by Bross [32], adapted for square contingency tables by Iki et al. [33], and used as weighing schemes in Yilmaz & Aktas [15]. The three-dimensional ridit weights matrix is extended to calculate Mielke's weighted kappa coefficient. Suppose there are the classifications of three raters named A, B, C. The $i$th ridit score of rater A ($r_i^A$), the $j$th ridit score of rater B ($r_j^B$), and the $k$th ridit score of rater C ($r_k^C$) are shown in Eq. 15.

$$\begin{aligned} r_i^A &= \frac{F_{i-1}^A - F_i^A}{2}, \quad r_j^B = \frac{F_{j-1}^B - F_j^B}{2}, \\ r_k^C &= \frac{F_{k-1}^C - F_k^C}{2}, \end{aligned} \tag{15}$$

where the cumulative distribution function of A, B, and C are

$$\begin{aligned} F_i^A &= \sum_{m \le i} \pi_{m..}, \quad F_j^B = \sum_{m \le j} \pi_{.m.}, \\ F_k^C &= \sum_{m \le k} \pi_{..m}. \end{aligned} \tag{16}$$

Then, score values are

$$\begin{aligned} u_{ij}^{ABr} &= \frac{r_i^A + r_j^B}{2}, \quad u_{ik}^{ACr} = \frac{r_i^A + r_k^C}{2}, \\ u_{jk}^{BCr} &= \frac{r_j^B + r_k^C}{2}. \end{aligned} \tag{17}$$

The ridit type linear and quadratic weights are calculated from Eqs. 18 and 19, respectively.

- Ridit linear weights:

$$w_{ijk} = 1 - \frac{\frac{|r_i^A - r_j^B|}{u_{ij}^{ABr}} + \frac{|r_i^A - r_k^C|}{u_{ik}^{ACr}} + \frac{|r_j^B - r_k^C|}{u_{jk}^{BCr}}}{3(R-1)}, \tag{18}$$

where $i, j, k = 1, \ldots, R$.

- Ridit quadratic weights:

$$w_{ijk} = 1 - \frac{\frac{(r_i^A - r_j^B)^2}{(u_{ij}^{ABr})^2} + \frac{(r_i^A - r_k^C)^2}{(u_{ik}^{ACr})^2} + \frac{(r_j^B - r_k^C)^2}{(u_{jk}^{BCr})^2}}{3(R-1)^2}, \tag{19}$$

where $i, j, k = 1, \ldots, R$.

The exponential scores of Bagheban & Zayeri [34] are used as exponential weights [15]. The exponential score values of multi-rater tables are

$$\begin{aligned} u_{ij}^{ABe} &= \frac{i^a + j^b}{2}, \quad u_{ik}^{ACe} = \frac{i^a + k^c}{2}, \\ u_{jk}^{BCe} &= \frac{j^b + k^c}{2}. \end{aligned} \tag{20}$$

Here, $a$, $b$, and $c$ are the power parameters ($a, b, c > 0$) which are calculated directly from the cumulative distribution functions. The power parameters are

$$\begin{aligned} a &= \left[\prod_{i=1}^{R-1} \alpha_i\right]^{R-1}, \quad b = \left[\prod_{j=1}^{R-1} \beta_j\right]^{R-1}, \\ c &= \left[\prod_{k=1}^{R-1} \gamma_k\right]^{R-1}, \end{aligned} \tag{21}$$

where $\alpha_i = F_{i+1}^A / F_i^A$, $\beta_j = F_{j+1}^B / F_j^B$, and $\gamma_k = F_{k+1}^C / F_k^C$ where $i, j, k = 1, \ldots, R$. The exponential type linear and quadratic weights can be calculated from Eqs. 22 and 23.

- Exponential linear weights:

$$w_{ijk} = 1 - \frac{\frac{|i^a - j^b|}{u_{ij}^{ABe}} + \frac{|i^a - k^c|}{u_{ik}^{ACe}} + \frac{|j^b - k^c|}{u_{jk}^{BCe}}}{3(R-1)}, \tag{22}$$

where $i, j, k = 1, \ldots, R$.

- Exponential quadratic weights:

$$w_{ijk} = 1 - \frac{\frac{(i^a - j^b)^2}{(u_{ij}^{ABe})^2} + \frac{(i^a - k^c)^2}{(u_{ik}^{ACe})^2} + \frac{(j^b - k^c)^2}{(u_{jk}^{BCe})^2}}{3(R-1)^2}, \tag{23}$$

where $i, j, k = 1, \ldots, R$.

Hubert and Light's weighted kappa coefficients are calculated using two-dimensional weighting matrices based on 2-agreement, as shown in Eqs. 10 and 11. Hubert and Light's weighted kappa coefficients are

calculated using two-dimensional weighting matrices based on 2-agreement, as shown in Eqs. 10 and 11. These matrices utilize the information from the three-way table to compute ridit and exponential weights for the coefficients. Eqs. 24 and 25 present the two-dimensional versions of the ridit weights.

- Ridit linear weights:

$$w_{ij} = 1 - \frac{\frac{|r_i^A - r_j^B|}{u_{ij}^{ABr}} + \frac{|r_i^A - r_j^C|}{u_{ij}^{ACr}} + \frac{|r_i^B - r_j^C|}{u_{ij}^{BCr}}}{3(R-1)}, \quad (24)$$

where $i, j, k = 1, \ldots, R$.

- Ridit quadratic weights:

$$w_{ij} = 1 - \frac{\frac{(r_i^A - r_j^B)^2}{(u_{ij}^{ABr})^2} + \frac{(r_i^A - r_j^C)^2}{(u_{ij}^{ACr})^2} + \frac{(r_i^B - r_j^C)^2}{(u_{ij}^{BCr})^2}}{3(R-1)^2}, \quad (25)$$

where $i, j, k = 1, \ldots, R$.

The two-dimensional versions of the exponential weights are given in Eqs. 26–27.

- Exponential linear weights:

$$w_{ij} = 1 - \frac{\frac{|i^a - j^b|}{u_{ij}^{ABe}} + \frac{|i^a - j^c|}{u_{ij}^{ACe}} + \frac{|i^b - j^c|}{u_{ij}^{BCe}}}{3(R-1)}, \quad (26)$$

where $i, j, k = 1, \ldots, R$.

- Exponential quadratic weights:

$$w_{ij} = 1 - \frac{\frac{(i^a - j^b)^2}{(u_{ij}^{ABe})^2} + \frac{(i^a - j^c)^2}{(u_{ij}^{ACe})^2} + \frac{(i^b - j^c)^2}{(u_{ij}^{BCe})^2}}{3(R-1)^2}, \quad (27)$$

where $i, j, k = 1, \ldots, R$.

For all the two-way weighting schemes, $w_{ii} = 1$ for $i, j = 1, \ldots, R$ where $0 \leq w_{ij} \leq 1$ and for all the three-way weighting schemes, $w_{iii} = 1$ for $i, j, k = 1, \ldots, R$ where $0 \leq w_{ijk} \leq 1$.

## 4. ILLUSTRATIVE EXAMPLE

The illustrative data are taken from [35]. The patients who were either discharged from the hospital or treated and released from the emergency room (ER) for suspected or confirmed ischemic heart disease in the past three months were examined [35]. An interview was conducted with each patient, where they were asked about their perceptions of the hospital, what information they received regarding their heart condition, the treatment options offered, etc. Then, three evaluators (A, B, C) independently reviewed transcribed audiotapes of patients 765 patients and classified them into three categories: (1) Not a factor, (2) Minor factor, (3) Major factor. The data are given in Table 1 and contains $M = 3$ raters, $R = 3$ categories, and $n = 765$ subjects.

**Table 1.** The classifications of three raters

| A | B | C | Observed Frequency |
|---|---|---|---|
| 1 | 1 | 1 | 266 |
| 1 | 1 | 2 | 59 |
| 1 | 1 | 3 | 164 |
| 1 | 2 | 3 | 47 |
| 2 | 1 | 1 | 14 |
| 2 | 2 | 3 | 29 |
| 3 | 3 | 1 | 68 |
| 3 | 3 | 2 | 44 |
| 3 | 3 | 3 | 74 |

The overall agreement coefficients computed by the raters are detailed in Table 2, with values range between 0.259 and 0.418. The agreement level is similar when Mielke's and Hubert's weighted kappas with linear and quadratic weights are used. Light's kappa coefficient indicates greater agreement values than the other coefficients, except for ridit linear weights. The lowest agreement is observed with exponential quadratic weights when using Mielke's weighted kappa. When Hubert's and Lights's kappa coefficients are used, the highest agreement is observed with ridit quadratic weights. When Mielke's kappa coefficient is used, the highest agreement is observed with ridit linear weights.

This illustrative example demonstrates that different combinations of coefficients and weighting schemes give different results. These results strongly indicate the need for a detailed simulation study to discuss the accuracy of the coefficients and also the weighting schemes.

**Table 2.** The calculated overall agreement coefficients among the raters

| | Weighting Schemes | | | | | | |
|---|---|---|---|---|---|---|---|
| **Coefficients** | **UW** | **L** | **Q** | **RL** | **RQ** | **EL** | **EQ** |
| Mielke | 0.279 | 0.320 | 0.337 | 0.394 | 0.342 | 0.313 | 0.259 |
| Hubert | 0.295 | 0.320 | 0.337 | 0.370 | 0.405 | 0.318 | 0.319 |
| Light | 0.318 | 0.353 | 0.377 | 0.389 | 0.418 | 0.356 | 0.366 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

## 5. NUMERICAL EXPERIMENTS WITH SYNTHETIC DATA

In this section, we discuss the numerical studies with synthetic data. We conduct a simulation study to compare the accuracy of weighted kappa coefficients with different weighting schemes introduced in Sections 2 and 3. In this section, simulation design and its results are given.

### 5.1. The Simulation Design

We perform a Monte Carlo simulation to compare three weighted kappa coefficients and the weighting schemes for multi-raters with ordinal scales. The generalized versions of the classical linear and quadratic weights are considered. The simulation study uses 216 different combinations (see Table 3).

**Table 3.** The simulation design and the abridgments

| | **Level** |
|---|---|
| Coefficient | Light (L), Hubert (H), Mielke (M) |
| Weighting scheme | Unweighed (UW), Linear (Li), Quadratic (Q), Ridit Linear (RL), Ridit Quadratic (RQ), Exponential Linear (EL), Exponential Quadratic (EQ) |
| Sample size | 100, 500, 1000, 5000 |
| Number of categories | 3, 4, 5 |
| Number of raters | 3, 4, 5 |
| True level of agreement | Low, Medium, High |

Tran et al. [17] used the correlation coefficient as the true value of the inter-rater agreement. To generate ordinal squared contingency tables, they generated ordinal variables by underlying normally distributed continuous variables [36]. They use quantiles to categorize the generated values. In this study, we used the method presented in Sertdemir et al. [37] to generate the ordinal tables. The intraclass correlation coefficient (ICC) is used as the true value of agreement. Normal and skewed distribution settings are used. The data are generated from the following model equation.

$$Y_{ij} = Z_i + \varepsilon_{ij}, \tag{28}$$

where $i = 1, \dots, n$ and $j = 1, \dots, M$. Here, $Z_i \sim N(0, \sigma_p^2)$ indicates the true value of object $i$ and $\varepsilon_{ij} \sim N(0, \sigma_e^2)$ indicates the stochastic measurement error. The distribution of $Z_i$ and $\varepsilon_{ij}$ are summarized in Table 4 [37]. While the normal distribution produces a balanced table, the skewed distribution produces an unbalanced table. From this point on, the normal distribution will be referred to as the balanced table structure and the skewed distribution will be referred to as the unbalanced table structure.

**Table 4.** The simulation settings by distributions

| | | $\varepsilon_{ij}$ | | |
|---|---|---|---|---|
| **Distribution** | $Z_i$ | **Low** | **Medium** | **High** |
| Normal | $N(0,1)$ | $N(0,1.5)$ | $N(0,0.6667)$ | $N(0,0.25)$ |
| Skewed | $\chi_3^2$ | $N(0,9)$ | $N(0,4)$ | $N(0,1.5)$ |

The true agreement level is considered as 0.40, 0.60, and 0.80 which refer to low, medium, and high agreement among the raters. The variance of stochastic measurement error term ($\sigma_e^2$) for each agreement level is calculated by $ICC = \sigma_p^2/(\sigma_p^2 + \sigma_e^2)$ where $\sigma_p^2 = 1$. Then, $\sigma_e^2$ is calculated as 1.5, 0.6667, and 0.25, respectively. The continuous $Y_{ij}$ is transformed to $X_{ij}$ by using Eq. 29 which is an ordinal scale. Here, RND is the nearest integer [36]

$$X_{ij} = RND\left(\frac{R + 0.5\,exp(Y_{ij}) + 0.5}{1 + exp\,(Y_{ij})}\right). \tag{29}$$

The accuracies of the agreement coefficient-weighting scheme combinations are assessed using MAE and MSE as

$$MAE = \frac{1}{r}\sum_{i=1}^{r}|\kappa - \hat{\kappa}_i| \quad and \quad MSE = \frac{1}{r}\sum_{i=1}^{r}(\kappa - \hat{\kappa}_i)^2\,. \tag{30}$$

Here, $r$ is the number of replications, $\kappa$ is the true agreement, and $\hat{\kappa}_i$ is the inter-rater agreement estimation in the $i$th replication. In the scenarios that involve three raters, three weighted kappas are calculated.

In three-rater scenarios, three of the weighted kappa coefficients are calculated. As Mielke's weighted kappa coefficient is only proposed for cases with three raters, Hubert's and Light's coefficients are calculated for cases with more than three raters.

The simulation program is developed by the author in R version 3.6.1. The results are based on 1000 replications. The lattice package is used to visualize the results [38].

### 5.2. Results

The results are discussed in terms of the weighted kappa coefficients, the weighting scheme, the level of true agreement, and the structure of the table. The results are presented in three parts: 1) Comparison of the weighted kappa coefficients, 2) Comparison of the weighting schemes, and 3) Effect of the number of categories on weighting schemes.

The MSE results are summarized in Tables 5-13 by the number of raters, number of categories, sample size, true level of agreement, coefficient, and weighting scheme. Tables 5-7 show the results for the three raters, Tables 8-10 for the four raters, and Tables 11-13 for the five raters. Not all results are tabulated here due to space limitations. The MAE results can be found in Tables A1-A9 of the Appendix.

**Table 5.** MSE values computed for low agreement tables with three raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Mielke | 0.0646 | 0.0250 | 0.0098 | 0.0300 | 0.0166 | 0.0411 | 0.0351 | 0.0287 | 0.0274 | 0.0210 | 0.0190 | 0.0182 | 0.0257 | 0.0198 |
| | | Hubert | 0.0509 | 0.0250 | 0.0098 | 0.0290 | 0.0163 | 0.0376 | 0.0282 | 0.0355 | 0.0274 | 0.0210 | 0.0198 | 0.0179 | 0.0249 | 0.0196 |
| | | Light | 0.0509 | 0.0250 | 0.0098 | 0.0290 | 0.0164 | 0.0376 | 0.0282 | 0.0358 | 0.0276 | 0.0213 | 0.0200 | 0.0181 | 0.0251 | 0.0198 |
| | 500 | Mielke | 0.0620 | 0.0224 | 0.0066 | 0.0270 | 0.0131 | 0.0366 | 0.0259 | 0.0242 | 0.0224 | 0.0142 | 0.0137 | 0.0119 | 0.0199 | 0.0133 |
| | | Hubert | 0.0484 | 0.0224 | 0.0066 | 0.0264 | 0.0131 | 0.0350 | 0.0244 | 0.0311 | 0.0224 | 0.0142 | 0.0140 | 0.0118 | 0.0195 | 0.0132 |
| | | Light | 0.0484 | 0.0224 | 0.0066 | 0.0264 | 0.0131 | 0.0350 | 0.0244 | 0.0311 | 0.0224 | 0.0142 | 0.0140 | 0.0119 | 0.0196 | 0.0133 |
| | 1000 | Mielke | 0.0618 | 0.0222 | 0.0063 | 0.0266 | 0.0128 | 0.0358 | 0.0247 | 0.0236 | 0.0217 | 0.0133 | 0.0130 | 0.0111 | 0.0191 | 0.0124 |
| | | Hubert | 0.0482 | 0.0222 | 0.0063 | 0.0262 | 0.0128 | 0.0348 | 0.0240 | 0.0305 | 0.0217 | 0.0133 | 0.0132 | 0.0110 | 0.0188 | 0.0124 |
| | | Light | 0.0482 | 0.0222 | 0.0063 | 0.0262 | 0.0128 | 0.0348 | 0.0240 | 0.0305 | 0.0217 | 0.0133 | 0.0132 | 0.0111 | 0.0188 | 0.0124 |
| | 5000 | Mielke | 0.0616 | 0.0220 | 0.0059 | 0.0262 | 0.0125 | 0.0351 | 0.0238 | 0.0233 | 0.0212 | 0.0126 | 0.0127 | 0.0106 | 0.0185 | 0.0118 |
| | | Hubert | 0.0481 | 0.0220 | 0.0059 | 0.0260 | 0.0125 | 0.0346 | 0.0237 | 0.0301 | 0.0212 | 0.0126 | 0.0127 | 0.0106 | 0.0184 | 0.0118 |
| | | Light | 0.0481 | 0.0220 | 0.0059 | 0.0260 | 0.0125 | 0.0346 | 0.0237 | 0.0301 | 0.0212 | 0.0126 | 0.0127 | 0.0106 | 0.0184 | 0.0118 |
| 4 | 100 | Mielke | 0.0923 | 0.0263 | 0.0068 | 0.0351 | 0.0161 | 0.0626 | 0.0611 | 0.0226 | 0.0176 | 0.0100 | 0.0085 | 0.0074 | 0.0154 | 0.0088 |
| | | Hubert | 0.0713 | 0.0263 | 0.0068 | 0.0336 | 0.0158 | 0.0542 | 0.0447 | 0.0317 | 0.0176 | 0.0100 | 0.0089 | 0.0072 | 0.0147 | 0.0087 |
| | | Light | 0.0712 | 0.0263 | 0.0069 | 0.0336 | 0.0159 | 0.0542 | 0.0447 | 0.0319 | 0.0178 | 0.0101 | 0.0090 | 0.0073 | 0.0148 | 0.0088 |
| | 500 | Mielke | 0.0906 | 0.0238 | 0.0034 | 0.0321 | 0.0123 | 0.0555 | 0.0440 | 0.0197 | 0.0142 | 0.0049 | 0.0046 | 0.0027 | 0.0114 | 0.0042 |
| | | Hubert | 0.0694 | 0.0238 | 0.0034 | 0.0311 | 0.0122 | 0.0515 | 0.0397 | 0.0288 | 0.0142 | 0.0049 | 0.0048 | 0.0027 | 0.0111 | 0.0042 |
| | | Light | 0.0694 | 0.0238 | 0.0034 | 0.0312 | 0.0122 | 0.0515 | 0.0397 | 0.0288 | 0.0143 | 0.0049 | 0.0048 | 0.0027 | 0.0111 | 0.0042 |
| | 1000 | Mielke | 0.0906 | 0.0236 | 0.0030 | 0.0317 | 0.0119 | 0.0543 | 0.0413 | 0.0192 | 0.0137 | 0.0042 | 0.0040 | 0.0020 | 0.0107 | 0.0035 |
| | | Hubert | 0.0693 | 0.0236 | 0.0030 | 0.0310 | 0.0119 | 0.0516 | 0.0392 | 0.0282 | 0.0137 | 0.0042 | 0.0041 | 0.0020 | 0.0105 | 0.0035 |
| | | Light | 0.0693 | 0.0236 | 0.0030 | 0.0310 | 0.0119 | 0.0516 | 0.0392 | 0.0282 | 0.0137 | 0.0042 | 0.0041 | 0.0020 | 0.0105 | 0.0035 |
| | 5000 | Mielke | 0.0904 | 0.0234 | 0.0027 | 0.0312 | 0.0117 | 0.0528 | 0.0393 | 0.0190 | 0.0134 | 0.0037 | 0.0037 | 0.0016 | 0.0102 | 0.0031 |
| | | Hubert | 0.0690 | 0.0234 | 0.0027 | 0.0308 | 0.0117 | 0.0516 | 0.0388 | 0.0279 | 0.0134 | 0.0037 | 0.0038 | 0.0015 | 0.0101 | 0.0031 |
| | | Light | 0.0690 | 0.0234 | 0.0027 | 0.0308 | 0.0117 | 0.0516 | 0.0388 | 0.0279 | 0.0134 | 0.0037 | 0.0038 | 0.0015 | 0.0101 | 0.0031 |
| 5 | 100 | Mielke | 0.1097 | 0.0265 | 0.0053 | 0.0383 | 0.0161 | 0.0812 | 0.0837 | 0.0205 | 0.0121 | 0.0057 | 0.0048 | 0.0045 | 0.0103 | 0.0047 |
| | | Hubert | 0.0848 | 0.0265 | 0.0053 | 0.0364 | 0.0160 | 0.0687 | 0.0604 | 0.0304 | 0.0121 | 0.0057 | 0.0047 | 0.0049 | 0.0097 | 0.0046 |
| | | Light | 0.0848 | 0.0265 | 0.0053 | 0.0364 | 0.0160 | 0.0687 | 0.0604 | 0.0305 | 0.0122 | 0.0057 | 0.0048 | 0.0049 | 0.0097 | 0.0046 |
| | 500 | Mielke | 0.1090 | 0.0245 | 0.0022 | 0.0357 | 0.0126 | 0.0721 | 0.0631 | 0.0182 | 0.0103 | 0.0019 | 0.0016 | 0.0010 | 0.0076 | 0.0015 |
| | | Hubert | 0.0840 | 0.0245 | 0.0022 | 0.0346 | 0.0126 | 0.0650 | 0.0544 | 0.0281 | 0.0103 | 0.0019 | 0.0016 | 0.0011 | 0.0073 | 0.0015 |
| | | Light | 0.0840 | 0.0245 | 0.0022 | 0.0346 | 0.0126 | 0.0650 | 0.0545 | 0.0281 | 0.0103 | 0.0019 | 0.0016 | 0.0011 | 0.0073 | 0.0015 |
| | 1000 | Mielke | 0.1089 | 0.0243 | 0.0019 | 0.0353 | 0.0123 | 0.0701 | 0.0584 | 0.0179 | 0.0098 | 0.0014 | 0.0011 | 0.0005 | 0.0071 | 0.0010 |
| | | Hubert | 0.0839 | 0.0243 | 0.0019 | 0.0345 | 0.0123 | 0.0652 | 0.0539 | 0.0276 | 0.0098 | 0.0014 | 0.0011 | 0.0005 | 0.0068 | 0.0010 |
| | | Light | 0.0839 | 0.0243 | 0.0019 | 0.0345 | 0.0123 | 0.0652 | 0.0539 | 0.0276 | 0.0098 | 0.0014 | 0.0011 | 0.0005 | 0.0069 | 0.0010 |
| | 5000 | Mielke | 0.1088 | 0.0242 | 0.0016 | 0.0349 | 0.0121 | 0.0682 | 0.0547 | 0.0178 | 0.0095 | 0.0009 | 0.0008 | 0.0001 | 0.0067 | 0.0007 |
| | | Hubert | 0.0839 | 0.0242 | 0.0016 | 0.0345 | 0.0121 | 0.0659 | 0.0536 | 0.0275 | 0.0095 | 0.0009 | 0.0008 | 0.0001 | 0.0066 | 0.0007 |
| | | Light | 0.0839 | 0.0242 | 0.0016 | 0.0345 | 0.0121 | 0.0659 | 0.0536 | 0.0275 | 0.0095 | 0.0009 | 0.0008 | 0.0001 | 0.0066 | 0.0007 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 6.** MSE values computed for medium agreement tables with three raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Mielke | 0.1264 | 0.0462 | 0.0136 | 0.0566 | 0.0300 | 0.0818 | 0.0762 | 0.0714 | 0.0680 | 0.0552 | 0.0530 | 0.0520 | 0.0651 | 0.0541 |
| | | Hubert | 0.0952 | 0.0462 | 0.0136 | 0.0548 | 0.0293 | 0.0737 | 0.0572 | 0.0792 | 0.0680 | 0.0552 | 0.0545 | 0.0511 | 0.0636 | 0.0537 |
| | | Light | 0.0952 | 0.0463 | 0.0136 | 0.0548 | 0.0294 | 0.0737 | 0.0572 | 0.0796 | 0.0684 | 0.0555 | 0.0548 | 0.0515 | 0.0639 | 0.0540 |
| | 500 | Mielke | 0.1234 | 0.0435 | 0.0108 | 0.0532 | 0.0262 | 0.0744 | 0.0563 | 0.0666 | 0.0641 | 0.0515 | 0.0507 | 0.0487 | 0.0603 | 0.0503 |
| | | Hubert | 0.0923 | 0.0435 | 0.0108 | 0.0521 | 0.0261 | 0.0708 | 0.0525 | 0.0733 | 0.0641 | 0.0515 | 0.0513 | 0.0484 | 0.0596 | 0.0501 |
| | | Light | 0.0923 | 0.0435 | 0.0108 | 0.0521 | 0.0261 | 0.0708 | 0.0525 | 0.0734 | 0.0642 | 0.0517 | 0.0514 | 0.0486 | 0.0597 | 0.0503 |
| | 1000 | Mielke | 0.1231 | 0.0432 | 0.0104 | 0.0526 | 0.0258 | 0.0728 | 0.0537 | 0.0661 | 0.0636 | 0.0509 | 0.0503 | 0.0479 | 0.0595 | 0.0495 |
| | | Hubert | 0.0920 | 0.0432 | 0.0104 | 0.0518 | 0.0257 | 0.0704 | 0.0519 | 0.0728 | 0.0636 | 0.0509 | 0.0506 | 0.0478 | 0.0590 | 0.0494 |
| | | Light | 0.0920 | 0.0432 | 0.0104 | 0.0518 | 0.0257 | 0.0704 | 0.0519 | 0.0728 | 0.0637 | 0.0510 | 0.0507 | 0.0479 | 0.0591 | 0.0495 |
| | 5000 | Mielke | 0.1227 | 0.0429 | 0.0101 | 0.0519 | 0.0254 | 0.0713 | 0.0518 | 0.0656 | 0.0632 | 0.0504 | 0.0500 | 0.0474 | 0.0587 | 0.0489 |
| | | Hubert | 0.0917 | 0.0429 | 0.0101 | 0.0515 | 0.0254 | 0.0702 | 0.0515 | 0.0722 | 0.0632 | 0.0504 | 0.0501 | 0.0473 | 0.0585 | 0.0489 |
| | | Light | 0.0917 | 0.0429 | 0.0101 | 0.0515 | 0.0254 | 0.0702 | 0.0515 | 0.0722 | 0.0632 | 0.0504 | 0.0501 | 0.0473 | 0.0585 | 0.0489 |
| 4 | 100 | Mielke | 0.1910 | 0.0482 | 0.0072 | 0.0673 | 0.0283 | 0.1329 | 0.1463 | 0.0590 | 0.0423 | 0.0207 | 0.0210 | 0.0164 | 0.0374 | 0.0200 |
| | | Hubert | 0.1394 | 0.0482 | 0.0072 | 0.0641 | 0.0280 | 0.1096 | 0.0935 | 0.0696 | 0.0423 | 0.0207 | 0.0218 | 0.0159 | 0.0358 | 0.0196 |
| | | Light | 0.1394 | 0.0482 | 0.0072 | 0.0642 | 0.0280 | 0.1096 | 0.0935 | 0.0698 | 0.0425 | 0.0209 | 0.0219 | 0.0161 | 0.0360 | 0.0198 |
| | 500 | Mielke | 0.1892 | 0.0458 | 0.0046 | 0.0637 | 0.0244 | 0.1192 | 0.1025 | 0.0563 | 0.0412 | 0.0198 | 0.0211 | 0.0155 | 0.0354 | 0.0188 |
| | | Hubert | 0.1377 | 0.0458 | 0.0046 | 0.0619 | 0.0242 | 0.1077 | 0.0873 | 0.0652 | 0.0412 | 0.0198 | 0.0214 | 0.0154 | 0.0347 | 0.0187 |
| | | Light | 0.1377 | 0.0459 | 0.0046 | 0.0619 | 0.0242 | 0.1077 | 0.0873 | 0.0652 | 0.0412 | 0.0199 | 0.0214 | 0.0154 | 0.0347 | 0.0188 |
| | 1000 | Mielke | 0.1891 | 0.0457 | 0.0043 | 0.0632 | 0.0241 | 0.1165 | 0.0945 | 0.0554 | 0.0405 | 0.0194 | 0.0205 | 0.0149 | 0.0344 | 0.0182 |
| | | Hubert | 0.1376 | 0.0457 | 0.0043 | 0.0618 | 0.0240 | 0.1085 | 0.0871 | 0.0641 | 0.0405 | 0.0194 | 0.0207 | 0.0148 | 0.0339 | 0.0181 |
| | | Light | 0.1376 | 0.0457 | 0.0043 | 0.0618 | 0.0240 | 0.1086 | 0.0871 | 0.0641 | 0.0406 | 0.0194 | 0.0207 | 0.0148 | 0.0340 | 0.0181 |
| | 5000 | Mielke | 0.1888 | 0.0454 | 0.0041 | 0.0622 | 0.0237 | 0.1127 | 0.0886 | 0.0551 | 0.0401 | 0.0188 | 0.0202 | 0.0143 | 0.0337 | 0.0176 |
| | | Hubert | 0.1372 | 0.0454 | 0.0041 | 0.0616 | 0.0237 | 0.1091 | 0.0870 | 0.0635 | 0.0401 | 0.0188 | 0.0202 | 0.0143 | 0.0334 | 0.0176 |
| | | Light | 0.1372 | 0.0454 | 0.0041 | 0.0616 | 0.0237 | 0.1091 | 0.0870 | 0.0635 | 0.0401 | 0.0188 | 0.0203 | 0.0143 | 0.0334 | 0.0176 |
| 5 | 100 | Mielke | 0.2340 | 0.0493 | 0.0051 | 0.0748 | 0.0289 | 0.1815 | 0.2113 | 0.0592 | 0.0312 | 0.0098 | 0.0099 | 0.0058 | 0.0268 | 0.0096 |
| | | Hubert | 0.1730 | 0.0493 | 0.0051 | 0.0709 | 0.0287 | 0.1456 | 0.1321 | 0.0689 | 0.0312 | 0.0098 | 0.0103 | 0.0057 | 0.0251 | 0.0093 |
| | | Light | 0.1730 | 0.0493 | 0.0052 | 0.0710 | 0.0288 | 0.1456 | 0.1321 | 0.0690 | 0.0313 | 0.0098 | 0.0104 | 0.0057 | 0.0252 | 0.0093 |
| | 500 | Mielke | 0.2330 | 0.0469 | 0.0026 | 0.0710 | 0.0247 | 0.1612 | 0.1554 | 0.0559 | 0.0309 | 0.0084 | 0.0096 | 0.0041 | 0.0249 | 0.0078 |
| | | Hubert | 0.1715 | 0.0469 | 0.0026 | 0.0687 | 0.0246 | 0.1388 | 0.1210 | 0.0651 | 0.0309 | 0.0084 | 0.0097 | 0.0040 | 0.0243 | 0.0078 |
| | | Light | 0.1715 | 0.0469 | 0.0026 | 0.0687 | 0.0247 | 0.1388 | 0.1210 | 0.0651 | 0.0309 | 0.0085 | 0.0097 | 0.0040 | 0.0243 | 0.0078 |
| | 1000 | Mielke | 0.2326 | 0.0467 | 0.0023 | 0.0703 | 0.0244 | 0.1564 | 0.1404 | 0.0553 | 0.0305 | 0.0081 | 0.0094 | 0.0037 | 0.0244 | 0.0075 |
| | | Hubert | 0.1712 | 0.0467 | 0.0023 | 0.0686 | 0.0244 | 0.1400 | 0.1208 | 0.0643 | 0.0305 | 0.0081 | 0.0095 | 0.0037 | 0.0239 | 0.0074 |
| | | Light | 0.1712 | 0.0467 | 0.0023 | 0.0686 | 0.0244 | 0.1400 | 0.1208 | 0.0643 | 0.0305 | 0.0081 | 0.0095 | 0.0037 | 0.0239 | 0.0074 |
| | 5000 | Mielke | 0.2327 | 0.0465 | 0.0021 | 0.0693 | 0.0240 | 0.1510 | 0.1270 | 0.0552 | 0.0302 | 0.0076 | 0.0092 | 0.0034 | 0.0239 | 0.0071 |
| | | Hubert | 0.1711 | 0.0465 | 0.0021 | 0.0685 | 0.0240 | 0.1431 | 0.1222 | 0.0641 | 0.0302 | 0.0076 | 0.0092 | 0.0034 | 0.0236 | 0.0071 |
| | | Light | 0.1711 | 0.0465 | 0.0021 | 0.0685 | 0.0240 | 0.1431 | 0.1222 | 0.0641 | 0.0302 | 0.0076 | 0.0092 | 0.0034 | 0.0236 | 0.0071 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 7.** MSE values computed for high agreement tables with three raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Mielke | 0.1423 | 0.0559 | 0.0177 | 0.0672 | 0.0396 | 0.0959 | 0.1051 | 0.1008 | 0.0918 | 0.0739 | 0.0737 | 0.0725 | 0.0883 | 0.0742 |
| | | Hubert | 0.1050 | 0.0559 | 0.0177 | 0.0656 | 0.0386 | 0.0867 | 0.0716 | 0.1047 | 0.0918 | 0.0739 | 0.0753 | 0.0713 | 0.0864 | 0.0736 |
| | | Light | 0.1050 | 0.0560 | 0.0178 | 0.0656 | 0.0386 | 0.0867 | 0.0716 | 0.1053 | 0.0924 | 0.0744 | 0.0760 | 0.0719 | 0.0870 | 0.0741 |
| | 500 | Mielke | 0.1381 | 0.0531 | 0.0157 | 0.0637 | 0.0359 | 0.0875 | 0.0736 | 0.0955 | 0.0927 | 0.0851 | 0.0847 | 0.0841 | 0.0909 | 0.0848 |
| | | Hubert | 0.1013 | 0.0531 | 0.0157 | 0.0627 | 0.0357 | 0.0834 | 0.0673 | 0.0970 | 0.0927 | 0.0851 | 0.0853 | 0.0838 | 0.0901 | 0.0846 |
| | | Light | 0.1013 | 0.0531 | 0.0157 | 0.0627 | 0.0357 | 0.0834 | 0.0673 | 0.0972 | 0.0928 | 0.0852 | 0.0855 | 0.0840 | 0.0902 | 0.0848 |
| | 1000 | Mielke | 0.1370 | 0.0524 | 0.0153 | 0.0627 | 0.0352 | 0.0852 | 0.0691 | 0.0934 | 0.0914 | 0.0854 | 0.0851 | 0.0844 | 0.0898 | 0.0850 |
| | | Hubert | 0.1003 | 0.0524 | 0.0153 | 0.0620 | 0.0351 | 0.0825 | 0.0663 | 0.0947 | 0.0914 | 0.0854 | 0.0854 | 0.0843 | 0.0892 | 0.0849 |
| | | Light | 0.1003 | 0.0524 | 0.0153 | 0.0620 | 0.0351 | 0.0825 | 0.0663 | 0.0948 | 0.0915 | 0.0855 | 0.0855 | 0.0843 | 0.0893 | 0.0850 |
| | 5000 | Mielke | 0.1368 | 0.0522 | 0.0151 | 0.0622 | 0.0350 | 0.0836 | 0.0667 | 0.0922 | 0.0904 | 0.0850 | 0.0848 | 0.0839 | 0.0886 | 0.0844 |
| | | Hubert | 0.1001 | 0.0522 | 0.0151 | 0.0619 | 0.0350 | 0.0824 | 0.0661 | 0.0934 | 0.0904 | 0.0850 | 0.0849 | 0.0839 | 0.0884 | 0.0844 |
| | | Light | 0.1001 | 0.0522 | 0.0151 | 0.0619 | 0.0350 | 0.0824 | 0.0661 | 0.0934 | 0.0905 | 0.0850 | 0.0849 | 0.0839 | 0.0884 | 0.0844 |
| 4 | 100 | Mielke | 0.2320 | 0.0570 | 0.0079 | 0.0794 | 0.0353 | 0.1643 | 0.2398 | 0.0855 | 0.0576 | 0.0266 | 0.0292 | 0.0225 | 0.0499 | 0.0265 |
| | | Hubert | 0.1595 | 0.0570 | 0.0079 | 0.0761 | 0.0347 | 0.1306 | 0.1147 | 0.0883 | 0.0576 | 0.0266 | 0.0296 | 0.0219 | 0.0480 | 0.0260 |
| | | Light | 0.1594 | 0.0570 | 0.0079 | 0.0761 | 0.0347 | 0.1306 | 0.1147 | 0.0884 | 0.0577 | 0.0268 | 0.0297 | 0.0220 | 0.0482 | 0.0262 |
| | 500 | Mielke | 0.2306 | 0.0549 | 0.0064 | 0.0763 | 0.0320 | 0.1482 | 0.1470 | 0.0787 | 0.0565 | 0.0310 | 0.0323 | 0.0258 | 0.0488 | 0.0295 |
| | | Hubert | 0.1576 | 0.0549 | 0.0064 | 0.0743 | 0.0319 | 0.1309 | 0.1116 | 0.0793 | 0.0565 | 0.0310 | 0.0325 | 0.0257 | 0.0480 | 0.0294 |
| | | Light | 0.1576 | 0.0549 | 0.0064 | 0.0743 | 0.0319 | 0.1309 | 0.1116 | 0.0794 | 0.0566 | 0.0311 | 0.0326 | 0.0257 | 0.0480 | 0.0294 |
| | 1000 | Mielke | 0.2313 | 0.0550 | 0.0063 | 0.0759 | 0.0318 | 0.1450 | 0.1304 | 0.0774 | 0.0559 | 0.0312 | 0.0332 | 0.0267 | 0.0484 | 0.0300 |
| | | Hubert | 0.1581 | 0.0550 | 0.0063 | 0.0745 | 0.0317 | 0.1325 | 0.1125 | 0.0777 | 0.0559 | 0.0312 | 0.0333 | 0.0266 | 0.0479 | 0.0299 |
| | | Light | 0.1581 | 0.0550 | 0.0063 | 0.0745 | 0.0318 | 0.1325 | 0.1125 | 0.0777 | 0.0559 | 0.0313 | 0.0333 | 0.0266 | 0.0479 | 0.0300 |
| | 5000 | Mielke | 0.2309 | 0.0547 | 0.0061 | 0.0749 | 0.0315 | 0.1393 | 0.1169 | 0.0768 | 0.0553 | 0.0310 | 0.0335 | 0.0270 | 0.0479 | 0.0300 |
| | | Hubert | 0.1578 | 0.0547 | 0.0061 | 0.0742 | 0.0315 | 0.1335 | 0.1131 | 0.0768 | 0.0553 | 0.0310 | 0.0336 | 0.0270 | 0.0477 | 0.0300 |
| | | Light | 0.1578 | 0.0547 | 0.0061 | 0.0742 | 0.0315 | 0.1335 | 0.1131 | 0.0768 | 0.0553 | 0.0310 | 0.0336 | 0.0270 | 0.0477 | 0.0300 |
| 5 | 100 | Mielke | 0.3118 | 0.0580 | 0.0045 | 0.0894 | 0.0340 | 0.2503 | 0.3883 | 0.0536 | 0.0234 | 0.0056 | 0.0059 | 0.0012 | 0.0525 | 0.0525 |
| | | Hubert | 0.2123 | 0.0580 | 0.0045 | 0.0849 | 0.0337 | 0.1832 | 0.1693 | 0.0525 | 0.0234 | 0.0056 | 0.0060 | 0.0012 | 0.0525 | 0.0525 |
| | | Light | 0.2123 | 0.0580 | 0.0046 | 0.0849 | 0.0338 | 0.1832 | 0.1692 | 0.0524 | 0.0232 | 0.0054 | 0.0059 | 0.0012 | 0.0524 | 0.0524 |
| | 500 | Mielke | 0.3097 | 0.0558 | 0.0032 | 0.0852 | 0.0302 | 0.2206 | 0.2643 | 0.1289 | 0.0903 | 0.0636 | 0.0675 | 0.0593 | 0.0958 | 0.0754 |
| | | Hubert | 0.2098 | 0.0558 | 0.0032 | 0.0825 | 0.0301 | 0.1760 | 0.1578 | 0.1261 | 0.0903 | 0.0636 | 0.0674 | 0.0592 | 0.0950 | 0.0753 |
| | | Light | 0.2098 | 0.0558 | 0.0032 | 0.0826 | 0.0301 | 0.1760 | 0.1578 | 0.1262 | 0.0903 | 0.0636 | 0.0674 | 0.0592 | 0.0950 | 0.0753 |
| | 1000 | Mielke | 0.3091 | 0.0555 | 0.0031 | 0.0842 | 0.0297 | 0.2119 | 0.2214 | 0.0873 | 0.0440 | 0.0139 | 0.0188 | 0.0098 | 0.0369 | 0.0141 |
| | | Hubert | 0.2094 | 0.0555 | 0.0031 | 0.0822 | 0.0297 | 0.1793 | 0.1603 | 0.0842 | 0.0440 | 0.0139 | 0.0188 | 0.0097 | 0.0362 | 0.0140 |
| | | Light | 0.2094 | 0.0555 | 0.0031 | 0.0822 | 0.0297 | 0.1793 | 0.1603 | 0.0842 | 0.0440 | 0.0139 | 0.0188 | 0.0098 | 0.0362 | 0.0140 |
| | 5000 | Mielke | 0.3093 | 0.0554 | 0.0030 | 0.0832 | 0.0295 | 0.2019 | 0.1824 | 0.0863 | 0.0433 | 0.0135 | 0.0187 | 0.0097 | 0.0360 | 0.0138 |
| | | Hubert | 0.2095 | 0.0554 | 0.0030 | 0.0822 | 0.0295 | 0.1854 | 0.1660 | 0.0831 | 0.0433 | 0.0135 | 0.0187 | 0.0097 | 0.0357 | 0.0137 |
| | | Light | 0.2095 | 0.0554 | 0.0030 | 0.0822 | 0.0295 | 0.1854 | 0.1660 | 0.0831 | 0.0433 | 0.0135 | 0.0187 | 0.0097 | 0.0357 | 0.0137 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 8.** MSE values computed for low agreement tables with four raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.0500 | 0.0244 | 0.0090 | 0.0281 | 0.0154 | 0.0367 | 0.0271 | 0.0334 | 0.0247 | 0.0174 | 0.0166 | 0.0145 | 0.0220 | 0.0162 |
| | | Light | 0.0500 | 0.0245 | 0.0090 | 0.0282 | 0.0154 | 0.0367 | 0.0272 | 0.0337 | 0.0251 | 0.0177 | 0.0169 | 0.0148 | 0.0223 | 0.0165 |
| | 500 | Hubert | 0.0482 | 0.0223 | 0.0064 | 0.0262 | 0.0128 | 0.0348 | 0.0241 | 0.0306 | 0.0218 | 0.0134 | 0.0134 | 0.0113 | 0.0190 | 0.0125 |
| | | Light | 0.0482 | 0.0223 | 0.0064 | 0.0262 | 0.0129 | 0.0348 | 0.0241 | 0.0307 | 0.0219 | 0.0135 | 0.0135 | 0.0113 | 0.0190 | 0.0126 |
| | 1000 | Hubert | 0.0482 | 0.0222 | 0.0062 | 0.0261 | 0.0127 | 0.0347 | 0.0239 | 0.0302 | 0.0214 | 0.0129 | 0.0129 | 0.0108 | 0.0185 | 0.0120 |
| | | Light | 0.0482 | 0.0222 | 0.0062 | 0.0261 | 0.0127 | 0.0347 | 0.0239 | 0.0302 | 0.0214 | 0.0129 | 0.0130 | 0.0108 | 0.0186 | 0.0121 |
| | 5000 | Hubert | 0.0480 | 0.0219 | 0.0059 | 0.0258 | 0.0123 | 0.0345 | 0.0235 | 0.0302 | 0.0213 | 0.0127 | 0.0128 | 0.0106 | 0.0184 | 0.0118 |
| | | Light | 0.0480 | 0.0219 | 0.0059 | 0.0258 | 0.0123 | 0.0345 | 0.0235 | 0.0302 | 0.0213 | 0.0127 | 0.0128 | 0.0106 | 0.0184 | 0.0119 |
| 4 | 100 | Hubert | 0.0705 | 0.0258 | 0.0059 | 0.0328 | 0.0149 | 0.0531 | 0.0435 | 0.0301 | 0.0156 | 0.0072 | 0.0067 | 0.0049 | 0.0126 | 0.0063 |
| | | Light | 0.0705 | 0.0259 | 0.0059 | 0.0328 | 0.0149 | 0.0531 | 0.0435 | 0.0303 | 0.0158 | 0.0074 | 0.0068 | 0.0050 | 0.0127 | 0.0064 |
| | 500 | Hubert | 0.0690 | 0.0237 | 0.0033 | 0.0310 | 0.0121 | 0.0511 | 0.0394 | 0.0283 | 0.0139 | 0.0044 | 0.0044 | 0.0023 | 0.0107 | 0.0038 |
| | | Light | 0.0690 | 0.0237 | 0.0033 | 0.0310 | 0.0121 | 0.0511 | 0.0394 | 0.0283 | 0.0139 | 0.0045 | 0.0044 | 0.0023 | 0.0107 | 0.0038 |
| | 1000 | Hubert | 0.0689 | 0.0235 | 0.0030 | 0.0309 | 0.0118 | 0.0513 | 0.0390 | 0.0280 | 0.0135 | 0.0040 | 0.0040 | 0.0018 | 0.0104 | 0.0034 |
| | | Light | 0.0689 | 0.0235 | 0.0030 | 0.0309 | 0.0119 | 0.0513 | 0.0390 | 0.0280 | 0.0136 | 0.0040 | 0.0040 | 0.0019 | 0.0104 | 0.0034 |
| | 5000 | Hubert | 0.0689 | 0.0233 | 0.0027 | 0.0307 | 0.0116 | 0.0515 | 0.0387 | 0.0278 | 0.0134 | 0.0037 | 0.0037 | 0.0015 | 0.0101 | 0.0031 |
| | | Light | 0.0689 | 0.0233 | 0.0027 | 0.0307 | 0.0116 | 0.0515 | 0.0387 | 0.0278 | 0.0134 | 0.0037 | 0.0037 | 0.0015 | 0.0101 | 0.0031 |
| 5 | 100 | Hubert | 0.0852 | 0.0266 | 0.0047 | 0.0364 | 0.0155 | 0.0690 | 0.0607 | 0.0291 | 0.0110 | 0.0040 | 0.0036 | 0.0038 | 0.0084 | 0.0034 |
| | | Light | 0.0852 | 0.0266 | 0.0047 | 0.0365 | 0.0155 | 0.0690 | 0.0607 | 0.0292 | 0.0111 | 0.0041 | 0.0036 | 0.0038 | 0.0085 | 0.0035 |
| | 500 | Hubert | 0.0839 | 0.0246 | 0.0022 | 0.0348 | 0.0126 | 0.0651 | 0.0546 | 0.0279 | 0.0100 | 0.0015 | 0.0014 | 0.0008 | 0.0071 | 0.0012 |
| | | Light | 0.0839 | 0.0246 | 0.0022 | 0.0348 | 0.0126 | 0.0651 | 0.0546 | 0.0279 | 0.0100 | 0.0015 | 0.0014 | 0.0008 | 0.0071 | 0.0012 |
| | 1000 | Hubert | 0.0840 | 0.0244 | 0.0019 | 0.0346 | 0.0123 | 0.0653 | 0.0540 | 0.0277 | 0.0097 | 0.0011 | 0.0011 | 0.0004 | 0.0068 | 0.0009 |
| | | Light | 0.0840 | 0.0245 | 0.0019 | 0.0346 | 0.0123 | 0.0653 | 0.0540 | 0.0277 | 0.0097 | 0.0011 | 0.0011 | 0.0004 | 0.0068 | 0.0009 |
| | 5000 | Hubert | 0.0839 | 0.0241 | 0.0016 | 0.0344 | 0.0120 | 0.0658 | 0.0536 | 0.0275 | 0.0096 | 0.0009 | 0.0008 | 0.0001 | 0.0066 | 0.0007 |
| | | Light | 0.0839 | 0.0242 | 0.0016 | 0.0344 | 0.0120 | 0.0658 | 0.0536 | 0.0275 | 0.0096 | 0.0009 | 0.0008 | 0.0001 | 0.0066 | 0.0007 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 9.** MSE values computed for medium agreement tables with four raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.0747 | 0.0628 | 0.0484 | 0.0483 | 0.0448 | 0.0579 | 0.0471 | 0.0962 | 0.0472 | 0.0140 | 0.0556 | 0.0298 | 0.0745 | 0.0578 |
| | | Light | 0.0755 | 0.0635 | 0.0492 | 0.0489 | 0.0455 | 0.0586 | 0.0478 | 0.0962 | 0.0473 | 0.0140 | 0.0557 | 0.0299 | 0.0745 | 0.0578 |
| | 500 | Hubert | 0.0731 | 0.0639 | 0.0512 | 0.0508 | 0.0480 | 0.0593 | 0.0497 | 0.0923 | 0.0436 | 0.0108 | 0.0521 | 0.0261 | 0.0708 | 0.0526 |
| | | Light | 0.0732 | 0.0640 | 0.0513 | 0.0510 | 0.0482 | 0.0594 | 0.0498 | 0.0923 | 0.0436 | 0.0108 | 0.0522 | 0.0262 | 0.0708 | 0.0526 |
| | 1000 | Hubert | 0.0725 | 0.0633 | 0.0504 | 0.0503 | 0.0474 | 0.0586 | 0.0490 | 0.0919 | 0.0432 | 0.0104 | 0.0517 | 0.0257 | 0.0704 | 0.0519 |
| | | Light | 0.0725 | 0.0633 | 0.0505 | 0.0503 | 0.0475 | 0.0587 | 0.0490 | 0.0919 | 0.0432 | 0.0105 | 0.0517 | 0.0257 | 0.0704 | 0.0519 |
| | 5000 | Hubert | 0.0723 | 0.0632 | 0.0504 | 0.0502 | 0.0474 | 0.0585 | 0.0489 | 0.0916 | 0.0428 | 0.0101 | 0.0514 | 0.0253 | 0.0701 | 0.0514 |
| | | Light | 0.0723 | 0.0632 | 0.0504 | 0.0502 | 0.0474 | 0.0586 | 0.0489 | 0.0916 | 0.0428 | 0.0101 | 0.0514 | 0.0253 | 0.0701 | 0.0514 |
| 4 | 100 | Hubert | 0.0656 | 0.0397 | 0.0184 | 0.0187 | 0.0130 | 0.0328 | 0.0169 | 0.1403 | 0.0486 | 0.0069 | 0.0643 | 0.0276 | 0.1100 | 0.0934 |
| | | Light | 0.0659 | 0.0400 | 0.0186 | 0.0189 | 0.0131 | 0.0330 | 0.0172 | 0.1403 | 0.0486 | 0.0069 | 0.0644 | 0.0277 | 0.1100 | 0.0934 |
| | 500 | Hubert | 0.0638 | 0.0402 | 0.0192 | 0.0205 | 0.0146 | 0.0336 | 0.0179 | 0.1376 | 0.0459 | 0.0046 | 0.0620 | 0.0244 | 0.1077 | 0.0875 |
| | | Light | 0.0639 | 0.0403 | 0.0192 | 0.0205 | 0.0147 | 0.0337 | 0.0180 | 0.1376 | 0.0459 | 0.0046 | 0.0621 | 0.0244 | 0.1078 | 0.0875 |
| | 1000 | Hubert | 0.0638 | 0.0402 | 0.0190 | 0.0205 | 0.0146 | 0.0337 | 0.0178 | 0.1374 | 0.0458 | 0.0044 | 0.0618 | 0.0240 | 0.1084 | 0.0872 |
| | | Light | 0.0638 | 0.0403 | 0.0190 | 0.0205 | 0.0146 | 0.0337 | 0.0179 | 0.1374 | 0.0458 | 0.0044 | 0.0618 | 0.0240 | 0.1084 | 0.0872 |
| | 5000 | Hubert | 0.0636 | 0.0401 | 0.0189 | 0.0203 | 0.0143 | 0.0335 | 0.0176 | 0.1369 | 0.0453 | 0.0040 | 0.0614 | 0.0236 | 0.1088 | 0.0868 |
| | | Light | 0.0636 | 0.0401 | 0.0189 | 0.0203 | 0.0144 | 0.0335 | 0.0176 | 0.1369 | 0.0453 | 0.0041 | 0.0614 | 0.0236 | 0.1088 | 0.0868 |
| 5 | 100 | Hubert | 0.0669 | 0.0300 | 0.0082 | 0.0090 | 0.0042 | 0.0237 | 0.0078 | 0.1736 | 0.0498 | 0.0048 | 0.0715 | 0.0287 | 0.1459 | 0.1325 |
| | | Light | 0.0670 | 0.0302 | 0.0083 | 0.0090 | 0.0042 | 0.0238 | 0.0079 | 0.1736 | 0.0499 | 0.0048 | 0.0715 | 0.0288 | 0.1459 | 0.1325 |
| | 500 | Hubert | 0.0646 | 0.0304 | 0.0080 | 0.0094 | 0.0037 | 0.0239 | 0.0074 | 0.1712 | 0.0470 | 0.0026 | 0.0689 | 0.0248 | 0.1388 | 0.1211 |
| | | Light | 0.0646 | 0.0305 | 0.0081 | 0.0094 | 0.0037 | 0.0239 | 0.0074 | 0.1712 | 0.0470 | 0.0026 | 0.0689 | 0.0249 | 0.1388 | 0.1211 |
| | 1000 | Hubert | 0.0644 | 0.0305 | 0.0080 | 0.0095 | 0.0037 | 0.0239 | 0.0074 | 0.1714 | 0.0468 | 0.0023 | 0.0687 | 0.0243 | 0.1400 | 0.1210 |
| | | Light | 0.0645 | 0.0305 | 0.0080 | 0.0095 | 0.0037 | 0.0239 | 0.0074 | 0.1714 | 0.0468 | 0.0023 | 0.0687 | 0.0243 | 0.1400 | 0.1210 |
| | 5000 | Hubert | 0.0641 | 0.0302 | 0.0077 | 0.0093 | 0.0034 | 0.0237 | 0.0071 | 0.1712 | 0.0465 | 0.0021 | 0.0684 | 0.0239 | 0.1430 | 0.1221 |
| | | Light | 0.0641 | 0.0302 | 0.0077 | 0.0093 | 0.0034 | 0.0237 | 0.0071 | 0.1712 | 0.0465 | 0.0021 | 0.0684 | 0.0239 | 0.1430 | 0.1221 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 10.** MSE values computed for high agreement tables with four raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.1040 | 0.0558 | 0.0178 | 0.0653 | 0.0386 | 0.0861 | 0.0713 | 0.1009 | 0.0881 | 0.0699 | 0.0719 | 0.0679 | 0.0826 | 0.0697 |
| | | Light | 0.1040 | 0.0558 | 0.0178 | 0.0653 | 0.0386 | 0.0861 | 0.0713 | 0.1019 | 0.0889 | 0.0706 | 0.0727 | 0.0688 | 0.0835 | 0.0704 |
| | 500 | Hubert | 0.1004 | 0.0526 | 0.0155 | 0.0622 | 0.0355 | 0.0827 | 0.0668 | 0.0939 | 0.0899 | 0.0826 | 0.0826 | 0.0812 | 0.0872 | 0.0819 |
| | | Light | 0.1004 | 0.0526 | 0.0155 | 0.0622 | 0.0355 | 0.0827 | 0.0668 | 0.0942 | 0.0901 | 0.0828 | 0.0828 | 0.0814 | 0.0874 | 0.0822 |
| | 1000 | Hubert | 0.1002 | 0.0524 | 0.0153 | 0.0619 | 0.0351 | 0.0825 | 0.0663 | 0.0942 | 0.0910 | 0.0850 | 0.0850 | 0.0839 | 0.0888 | 0.0844 |
| | | Light | 0.1002 | 0.0524 | 0.0153 | 0.0619 | 0.0351 | 0.0825 | 0.0663 | 0.0943 | 0.0911 | 0.0851 | 0.0851 | 0.0840 | 0.0889 | 0.0845 |
| | 5000 | Hubert | 0.1002 | 0.0522 | 0.0152 | 0.0619 | 0.0350 | 0.0825 | 0.0662 | 0.0937 | 0.0907 | 0.0851 | 0.0851 | 0.0841 | 0.0886 | 0.0845 |
| | | Light | 0.1002 | 0.0522 | 0.0152 | 0.0619 | 0.0350 | 0.0825 | 0.0662 | 0.0937 | 0.0907 | 0.0852 | 0.0851 | 0.0841 | 0.0886 | 0.0845 |
| 4 | 100 | Hubert | 0.1628 | 0.0586 | 0.0082 | 0.0786 | 0.0363 | 0.1330 | 0.1169 | 0.0738 | 0.0495 | 0.0258 | 0.0218 | 0.0107 | 0.0552 | 0.0459 |
| | | Light | 0.1628 | 0.0586 | 0.0082 | 0.0786 | 0.0363 | 0.1329 | 0.1168 | 0.0738 | 0.0495 | 0.0260 | 0.0218 | 0.0107 | 0.0552 | 0.0460 |
| | 500 | Hubert | 0.1584 | 0.0554 | 0.0065 | 0.0751 | 0.0324 | 0.1316 | 0.1125 | 0.0783 | 0.0547 | 0.0287 | 0.0307 | 0.0236 | 0.0467 | 0.0272 |
| | | Light | 0.1584 | 0.0554 | 0.0065 | 0.0751 | 0.0324 | 0.1316 | 0.1125 | 0.0784 | 0.0547 | 0.0287 | 0.0307 | 0.0237 | 0.0467 | 0.0272 |
| | 1000 | Hubert | 0.1578 | 0.0550 | 0.0063 | 0.0745 | 0.0318 | 0.1324 | 0.1126 | 0.0790 | 0.0567 | 0.0315 | 0.0344 | 0.0275 | 0.0489 | 0.0306 |
| | | Light | 0.1578 | 0.0550 | 0.0063 | 0.0745 | 0.0318 | 0.1324 | 0.1126 | 0.0790 | 0.0568 | 0.0315 | 0.0344 | 0.0276 | 0.0489 | 0.0307 |
| | 5000 | Hubert | 0.1574 | 0.0546 | 0.0061 | 0.0740 | 0.0314 | 0.1332 | 0.1129 | 0.0764 | 0.0551 | 0.0309 | 0.0333 | 0.0268 | 0.0474 | 0.0298 |
| | | Light | 0.1574 | 0.0546 | 0.0061 | 0.0740 | 0.0314 | 0.1332 | 0.1129 | 0.0764 | 0.0551 | 0.0309 | 0.0333 | 0.0268 | 0.0474 | 0.0298 |
| 5 | 100 | Hubert | 0.2131 | 0.0586 | 0.0045 | 0.0858 | 0.0342 | 0.1835 | 0.1694 | 0.1579 | 0.0916 | 0.0364 | 0.1034 | 0.1130 | 0.0341 | 0.0126 |
| | | Light | 0.2131 | 0.0586 | 0.0045 | 0.0858 | 0.0342 | 0.1835 | 0.1694 | 0.1360 | 0.0768 | 0.0286 | 0.0587 | 0.0477 | 0.0341 | 0.0127 |
| | 500 | Hubert | 0.2098 | 0.0559 | 0.0032 | 0.0828 | 0.0303 | 0.1763 | 0.1582 | 0.1214 | 0.0677 | 0.0251 | 0.0590 | 0.0583 | 0.0316 | 0.0110 |
| | | Light | 0.2097 | 0.0559 | 0.0032 | 0.0828 | 0.0303 | 0.1763 | 0.1582 | 0.1109 | 0.0609 | 0.0216 | 0.0374 | 0.0266 | 0.0316 | 0.0110 |
| | 1000 | Hubert | 0.2095 | 0.0557 | 0.0031 | 0.0824 | 0.0298 | 0.1792 | 0.1604 | 0.1579 | 0.0916 | 0.0364 | 0.1034 | 0.1130 | 0.0341 | 0.0126 |
| | | Light | 0.2095 | 0.0557 | 0.0031 | 0.0825 | 0.0298 | 0.1792 | 0.1604 | 0.1360 | 0.0768 | 0.0286 | 0.0587 | 0.0477 | 0.0341 | 0.0127 |
| | 5000 | Hubert | 0.2093 | 0.0554 | 0.0029 | 0.0820 | 0.0293 | 0.1853 | 0.1659 | 0.0834 | 0.0427 | 0.0129 | 0.0185 | 0.0095 | 0.0353 | 0.0133 |
| | | Light | 0.2093 | 0.0554 | 0.0030 | 0.0820 | 0.0293 | 0.1853 | 0.1659 | 0.0834 | 0.0427 | 0.0129 | 0.0185 | 0.0095 | 0.0353 | 0.0133 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 11.** MSE values computed for low agreement tables with five raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.0494 | 0.0238 | 0.0084 | 0.0277 | 0.0149 | 0.0360 | 0.0265 | 0.0334 | 0.0247 | 0.0174 | 0.0166 | 0.0145 | 0.0220 | 0.0162 |
| | | Light | 0.0494 | 0.0238 | 0.0084 | 0.0278 | 0.0150 | 0.0360 | 0.0265 | 0.0337 | 0.0251 | 0.0177 | 0.0169 | 0.0148 | 0.0223 | 0.0165 |
| | 500 | Hubert | 0.0484 | 0.0223 | 0.0063 | 0.0262 | 0.0128 | 0.0347 | 0.0240 | 0.0306 | 0.0218 | 0.0134 | 0.0134 | 0.0113 | 0.0190 | 0.0125 |
| | | Light | 0.0484 | 0.0223 | 0.0064 | 0.0262 | 0.0128 | 0.0348 | 0.0241 | 0.0307 | 0.0219 | 0.0135 | 0.0135 | 0.0113 | 0.0190 | 0.0126 |
| | 1000 | Hubert | 0.0482 | 0.0221 | 0.0061 | 0.0260 | 0.0126 | 0.0346 | 0.0238 | 0.0302 | 0.0214 | 0.0129 | 0.0129 | 0.0108 | 0.0185 | 0.0120 |
| | | Light | 0.0482 | 0.0221 | 0.0061 | 0.0260 | 0.0126 | 0.0346 | 0.0238 | 0.0302 | 0.0214 | 0.0129 | 0.0130 | 0.0108 | 0.0186 | 0.0121 |
| | 5000 | Hubert | 0.0480 | 0.0219 | 0.0059 | 0.0259 | 0.0123 | 0.0345 | 0.0235 | 0.0302 | 0.0213 | 0.0127 | 0.0128 | 0.0106 | 0.0184 | 0.0118 |
| | | Light | 0.0480 | 0.0219 | 0.0059 | 0.0259 | 0.0123 | 0.0345 | 0.0235 | 0.0302 | 0.0213 | 0.0127 | 0.0128 | 0.0106 | 0.0184 | 0.0119 |
| 4 | 100 | Hubert | 0.0702 | 0.0250 | 0.0051 | 0.0324 | 0.0142 | 0.0526 | 0.0428 | 0.0301 | 0.0156 | 0.0072 | 0.0067 | 0.0049 | 0.0126 | 0.0063 |
| | | Light | 0.0702 | 0.0251 | 0.0051 | 0.0324 | 0.0142 | 0.0526 | 0.0430 | 0.0303 | 0.0158 | 0.0074 | 0.0068 | 0.0050 | 0.0127 | 0.0064 |
| | 500 | Hubert | 0.0691 | 0.0236 | 0.0032 | 0.0310 | 0.0121 | 0.0511 | 0.0393 | 0.0283 | 0.0139 | 0.0044 | 0.0044 | 0.0023 | 0.0107 | 0.0038 |
| | | Light | 0.0691 | 0.0237 | 0.0032 | 0.0310 | 0.0121 | 0.0511 | 0.0394 | 0.0283 | 0.0139 | 0.0045 | 0.0044 | 0.0023 | 0.0107 | 0.0038 |
| | 1000 | Hubert | 0.0691 | 0.0235 | 0.0029 | 0.0309 | 0.0118 | 0.0513 | 0.0390 | 0.0280 | 0.0135 | 0.0040 | 0.0040 | 0.0018 | 0.0104 | 0.0034 |
| | | Light | 0.0691 | 0.0235 | 0.0029 | 0.0309 | 0.0118 | 0.0513 | 0.0391 | 0.0280 | 0.0136 | 0.0040 | 0.0040 | 0.0019 | 0.0104 | 0.0034 |
| | 5000 | Hubert | 0.0690 | 0.0234 | 0.0027 | 0.0307 | 0.0116 | 0.0515 | 0.0387 | 0.0278 | 0.0134 | 0.0037 | 0.0037 | 0.0015 | 0.0101 | 0.0031 |
| | | Light | 0.0690 | 0.0234 | 0.0027 | 0.0307 | 0.0116 | 0.0515 | 0.0388 | 0.0278 | 0.0134 | 0.0037 | 0.0037 | 0.0015 | 0.0101 | 0.0031 |
| 5 | 100 | Hubert | 0.0846 | 0.0257 | 0.0040 | 0.0359 | 0.0148 | 0.0680 | 0.0595 | 0.0291 | 0.0110 | 0.0040 | 0.0036 | 0.0038 | 0.0084 | 0.0024 |
| | | Light | 0.0846 | 0.0257 | 0.0040 | 0.0360 | 0.0148 | 0.0680 | 0.0596 | 0.0292 | 0.0111 | 0.0041 | 0.0036 | 0.0038 | 0.0085 | 0.0025 |
| | 500 | Hubert | 0.0840 | 0.0244 | 0.0020 | 0.0347 | 0.0125 | 0.0650 | 0.0543 | 0.0279 | 0.0100 | 0.0015 | 0.0014 | 0.0008 | 0.0071 | 0.0012 |
| | | Light | 0.0840 | 0.0244 | 0.0020 | 0.0347 | 0.0125 | 0.0650 | 0.0544 | 0.0279 | 0.0100 | 0.0015 | 0.0014 | 0.0008 | 0.0071 | 0.0012 |
| | 1000 | Hubert | 0.0840 | 0.0243 | 0.0018 | 0.0346 | 0.0122 | 0.0652 | 0.0539 | 0.0277 | 0.0097 | 0.0011 | 0.0011 | 0.0004 | 0.0068 | 0.0009 |
| | | Light | 0.0840 | 0.0243 | 0.0018 | 0.0346 | 0.0123 | 0.0652 | 0.0539 | 0.0277 | 0.0097 | 0.0011 | 0.0011 | 0.0004 | 0.0068 | 0.0009 |
| | 5000 | Hubert | 0.0839 | 0.0241 | 0.0016 | 0.0344 | 0.0120 | 0.0658 | 0.0535 | 0.0275 | 0.0096 | 0.0009 | 0.0008 | 0.0001 | 0.0066 | 0.0007 |
| | | Light | 0.0839 | 0.0241 | 0.0016 | 0.0344 | 0.0120 | 0.0658 | 0.0535 | 0.0275 | 0.0096 | 0.0009 | 0.0008 | 0.0001 | 0.0066 | 0.0007 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 12.** MSE values computed for medium agreement tables with five raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.0938 | 0.0452 | 0.0126 | 0.0539 | 0.0285 | 0.0720 | 0.0553 | 0.0755 | 0.0643 | 0.0504 | 0.0498 | 0.0464 | 0.0593 | 0.0488 |
| | | Light | 0.0938 | 0.0453 | 0.0126 | 0.0540 | 0.0286 | 0.0720 | 0.0553 | 0.0764 | 0.0651 | 0.0512 | 0.0507 | 0.0473 | 0.0602 | 0.0496 |
| | 500 | Hubert | 0.0919 | 0.0432 | 0.0105 | 0.0517 | 0.0258 | 0.0702 | 0.0518 | 0.0732 | 0.0641 | 0.0515 | 0.0511 | 0.0483 | 0.0595 | 0.0499 |
| | | Light | 0.0919 | 0.0432 | 0.0106 | 0.0517 | 0.0258 | 0.0702 | 0.0520 | 0.0734 | 0.0643 | 0.0517 | 0.0513 | 0.0485 | 0.0597 | 0.0502 |
| | 1000 | Hubert | 0.0918 | 0.0430 | 0.0103 | 0.0516 | 0.0255 | 0.0701 | 0.0516 | 0.0732 | 0.0641 | 0.0513 | 0.0510 | 0.0482 | 0.0594 | 0.0498 |
| | | Light | 0.0918 | 0.0430 | 0.0103 | 0.0516 | 0.0256 | 0.0701 | 0.0516 | 0.0733 | 0.0642 | 0.0514 | 0.0511 | 0.0483 | 0.0595 | 0.0499 |
| | 5000 | Hubert | 0.0916 | 0.0428 | 0.0101 | 0.0514 | 0.0253 | 0.0700 | 0.0513 | 0.0723 | 0.0633 | 0.0505 | 0.0502 | 0.0474 | 0.0586 | 0.0489 |
| | | Light | 0.0916 | 0.0428 | 0.0101 | 0.0514 | 0.0253 | 0.0700 | 0.0513 | 0.0724 | 0.0633 | 0.0505 | 0.0502 | 0.0474 | 0.0586 | 0.0489 |
| 4 | 100 | Hubert | 0.1391 | 0.0479 | 0.0065 | 0.0640 | 0.0273 | 0.1087 | 0.0921 | 0.0653 | 0.0389 | 0.0173 | 0.0182 | 0.0124 | 0.0321 | 0.0162 |
| | | Light | 0.1391 | 0.0480 | 0.0065 | 0.0641 | 0.0274 | 0.1087 | 0.0923 | 0.0657 | 0.0392 | 0.0176 | 0.0184 | 0.0126 | 0.0324 | 0.0166 |
| | 500 | Hubert | 0.1374 | 0.0458 | 0.0045 | 0.0619 | 0.0243 | 0.1072 | 0.0868 | 0.0639 | 0.0404 | 0.0193 | 0.0207 | 0.0148 | 0.0338 | 0.0181 |
| | | Light | 0.1374 | 0.0458 | 0.0045 | 0.0619 | 0.0243 | 0.1073 | 0.0869 | 0.0640 | 0.0405 | 0.0194 | 0.0207 | 0.0148 | 0.0339 | 0.0182 |
| | 1000 | Hubert | 0.1372 | 0.0456 | 0.0043 | 0.0616 | 0.0239 | 0.1081 | 0.0867 | 0.0641 | 0.0406 | 0.0194 | 0.0208 | 0.0149 | 0.0340 | 0.0182 |
| | | Light | 0.1372 | 0.0456 | 0.0043 | 0.0616 | 0.0239 | 0.1081 | 0.0868 | 0.0641 | 0.0407 | 0.0194 | 0.0209 | 0.0149 | 0.0341 | 0.0182 |
| | 5000 | Hubert | 0.1369 | 0.0453 | 0.0040 | 0.0613 | 0.0235 | 0.1086 | 0.0865 | 0.0635 | 0.0401 | 0.0189 | 0.0203 | 0.0144 | 0.0335 | 0.0176 |
| | | Light | 0.1369 | 0.0453 | 0.0040 | 0.0613 | 0.0235 | 0.1086 | 0.0865 | 0.0635 | 0.0401 | 0.0189 | 0.0203 | 0.0144 | 0.0335 | 0.0176 |
| 5 | 100 | Hubert | 0.1731 | 0.0488 | 0.0042 | 0.0712 | 0.0282 | 0.1453 | 0.1318 | 0.0656 | 0.0282 | 0.0066 | 0.0077 | 0.0032 | 0.0220 | 0.0064 |
| | | Light | 0.1731 | 0.0489 | 0.0042 | 0.0713 | 0.0283 | 0.1454 | 0.1318 | 0.0658 | 0.0284 | 0.0067 | 0.0078 | 0.0032 | 0.0222 | 0.0066 |
| | 500 | Hubert | 0.1715 | 0.0469 | 0.0025 | 0.0688 | 0.0247 | 0.1388 | 0.1207 | 0.0646 | 0.0303 | 0.0079 | 0.0093 | 0.0035 | 0.0238 | 0.0073 |
| | | Light | 0.1715 | 0.0469 | 0.0025 | 0.0689 | 0.0247 | 0.1388 | 0.1208 | 0.0646 | 0.0304 | 0.0079 | 0.0093 | 0.0036 | 0.0238 | 0.0073 |
| | 1000 | Hubert | 0.1714 | 0.0467 | 0.0023 | 0.0687 | 0.0244 | 0.1399 | 0.1208 | 0.0643 | 0.0303 | 0.0078 | 0.0094 | 0.0035 | 0.0238 | 0.0073 |
| | | Light | 0.1714 | 0.0467 | 0.0023 | 0.0687 | 0.0244 | 0.1399 | 0.1208 | 0.0643 | 0.0303 | 0.0078 | 0.0094 | 0.0036 | 0.0238 | 0.0073 |
| | 5000 | Hubert | 0.1712 | 0.0464 | 0.0021 | 0.0683 | 0.0239 | 0.1430 | 0.1220 | 0.0641 | 0.0302 | 0.0077 | 0.0092 | 0.0034 | 0.0237 | 0.0071 |
| | | Light | 0.1712 | 0.0464 | 0.0021 | 0.0683 | 0.0239 | 0.1430 | 0.1220 | 0.0641 | 0.0302 | 0.0077 | 0.0092 | 0.0034 | 0.0237 | 0.0071 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table 13.** MSE values computed for high agreement tables with five raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.1030 | 0.0548 | 0.0171 | 0.0646 | 0.0381 | 0.0846 | 0.0697 | 0.0944 | 0.0824 | 0.0650 | 0.0659 | 0.0620 | 0.0767 | 0.0642 |
| | | Light | 0.1030 | 0.0548 | 0.0172 | 0.0647 | 0.0382 | 0.0846 | 0.0698 | 0.0956 | 0.0834 | 0.0657 | 0.0669 | 0.0630 | 0.0777 | 0.0659 |
| | 500 | Hubert | 0.1002 | 0.0524 | 0.0155 | 0.0620 | 0.0353 | 0.0823 | 0.0662 | 0.0945 | 0.0904 | 0.0830 | 0.0831 | 0.0816 | 0.0877 | 0.0824 |
| | | Light | 0.1002 | 0.0524 | 0.0155 | 0.0620 | 0.0353 | 0.0823 | 0.0664 | 0.0948 | 0.0906 | 0.0833 | 0.0833 | 0.0819 | 0.0880 | 0.0827 |
| | 1000 | Hubert | 0.1002 | 0.0524 | 0.0153 | 0.0620 | 0.0353 | 0.0825 | 0.0662 | 0.0946 | 0.0913 | 0.0854 | 0.0854 | 0.0843 | 0.0892 | 0.0848 |
| | | Light | 0.1002 | 0.0524 | 0.0154 | 0.0620 | 0.0353 | 0.0825 | 0.0663 | 0.0947 | 0.0914 | 0.0855 | 0.0855 | 0.0844 | 0.0893 | 0.0850 |
| | 5000 | Hubert | 0.1000 | 0.0521 | 0.0151 | 0.0618 | 0.0349 | 0.0823 | 0.0659 | 0.0938 | 0.0907 | 0.0852 | 0.0852 | 0.0842 | 0.0887 | 0.0846 |
| | | Light | 0.1000 | 0.0521 | 0.0151 | 0.0618 | 0.0349 | 0.0823 | 0.0660 | 0.0938 | 0.0908 | 0.0852 | 0.0852 | 0.0842 | 0.0887 | 0.0846 |
| 4 | 100 | Hubert | 0.1620 | 0.0580 | 0.0078 | 0.0783 | 0.0360 | 0.1323 | 0.1161 | 0.0753 | 0.0533 | 0.0281 | 0.0322 | 0.0257 | 0.0460 | 0.0283 |
| | | Light | 0.1620 | 0.0580 | 0.0079 | 0.0784 | 0.0361 | 0.1323 | 0.1159 | 0.0753 | 0.0533 | 0.0282 | 0.0322 | 0.0257 | 0.0460 | 0.0280 |
| | 500 | Hubert | 0.1581 | 0.0552 | 0.0064 | 0.0746 | 0.0321 | 0.1309 | 0.1114 | 0.0753 | 0.0533 | 0.0281 | 0.0322 | 0.0257 | 0.0460 | 0.0283 |
| | | Light | 0.1581 | 0.0552 | 0.0064 | 0.0746 | 0.0321 | 0.1309 | 0.1115 | 0.0753 | 0.0533 | 0.0282 | 0.0322 | 0.0257 | 0.0460 | 0.0280 |
| | 1000 | Hubert | 0.1579 | 0.0549 | 0.0063 | 0.0744 | 0.0318 | 0.1322 | 0.1122 | 0.0816 | 0.0584 | 0.0323 | 0.0369 | 0.0311 | 0.0478 | 0.0296 |
| | | Light | 0.1579 | 0.0549 | 0.0063 | 0.0744 | 0.0318 | 0.1322 | 0.1122 | 0.0802 | 0.0572 | 0.0315 | 0.0347 | 0.0278 | 0.0479 | 0.0297 |
| | 5000 | Hubert | 0.1573 | 0.0545 | 0.0061 | 0.0739 | 0.0313 | 0.1330 | 0.1126 | 0.0763 | 0.0551 | 0.0310 | 0.0336 | 0.0271 | 0.0475 | 0.0301 |
| | | Light | 0.1573 | 0.0545 | 0.0061 | 0.0739 | 0.0313 | 0.1330 | 0.1126 | 0.0763 | 0.0551 | 0.0310 | 0.0336 | 0.0271 | 0.0475 | 0.0301 |
| 5 | 100 | Hubert | 0.2126 | 0.0579 | 0.0042 | 0.0855 | 0.0339 | 0.1826 | 0.1683 | 0.0809 | 0.0405 | 0.0118 | 0.0091 | 0.0016 | 0.0809 | 0.0809 |
| | | Light | 0.2126 | 0.0580 | 0.0042 | 0.0855 | 0.0339 | 0.1826 | 0.1688 | 0.0810 | 0.0406 | 0.0120 | 0.0092 | 0.0016 | 0.0810 | 0.0818 |
| | 500 | Hubert | 0.2097 | 0.0558 | 0.0032 | 0.0825 | 0.0301 | 0.1759 | 0.1574 | 0.0836 | 0.0428 | 0.0129 | 0.0192 | 0.0085 | 0.0825 | 0.0814 |
| | | Light | 0.2097 | 0.0558 | 0.0032 | 0.0826 | 0.0301 | 0.1759 | 0.1576 | 0.0837 | 0.0429 | 0.0130 | 0.0192 | 0.0085 | 0.0823 | 0.0810 |
| | 1000 | Hubert | 0.2094 | 0.0555 | 0.0030 | 0.0823 | 0.0298 | 0.1789 | 0.1599 | 0.1128 | 0.0621 | 0.0226 | 0.0489 | 0.0515 | 0.0325 | 0.0111 |
| | | Light | 0.2094 | 0.0555 | 0.0030 | 0.0823 | 0.0298 | 0.1789 | 0.1600 | 0.1034 | 0.0554 | 0.0187 | 0.0322 | 0.0222 | 0.0325 | 0.0108 |
| | 5000 | Hubert | 0.2094 | 0.0554 | 0.0029 | 0.0820 | 0.0293 | 0.1852 | 0.1656 | 0.1120 | 0.0625 | 0.0234 | 0.0498 | 0.0536 | 0.0350 | 0.0134 |
| | | Light | 0.2094 | 0.0554 | 0.0029 | 0.0820 | 0.0293 | 0.1852 | 0.1656 | 0.1015 | 0.0551 | 0.0191 | 0.0319 | 0.0230 | 0.0350 | 0.0134 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

### 5.2.1 Comparison of the weighted kappa coefficients

The coefficients are compared across weighting schemes, levels of true agreement and table structure. The results in Tables 5-12 and Tables A1-A8 show that:

- In the balanced tables with three raters, the unweighted versions of three kappa coefficients are compared, Hubert's and Light's coefficients performing better than Mielke's. For example, for the tables with three categories where $n = 100$ and the agreement level is low, the MSE values of the unweighted version of Hubert's and Light's coefficients were around 0.050, whereas Mielke's was 0.064. A similar outcome is observed when exponential linear and exponential quadratic weights are employed (Table 5).
- In the balanced tables with three raters where the agreement level is low or medium, all the MAE/MSE values of three coefficients are similar when linear and quadratic, ridit linear and ridit quadratic weights are used. For instance, when linear weights were employed for tables with three categories where $n = 100$, the MSE values of the three coefficients were found around 0.025 for low agreement and 0.046 for medium agreement (Table 5 and Table 6). When ridit quadratic weights were employed for the tables comprising four categories, where $n = 1000$, the MSE values of the three coefficients were around 0.012 for low agreement and 0.024 for medium agreement (Table 5 and Table 6).
- In the balanced tables with three raters where there is high agreement, all the MAE/MSE values of the three coefficients are similar when linear and quadratic weights are used. For example, when linear and quadratic weights were used for the tables with three categories where $n = 100$, the MSE values of all three coefficients were around 0.060 and 0.017, respectively (Table 7).
- In the balanced tables with three raters, all the MAE/MSE values of Hubert'a and Light's coefficients are similar.
- In the balanced tables with three raters, when using the unweighted version of kappas, exponential linear and exponential quadratic weights, the difference between the coefficients becomes more noticeable. When using the unweighted version of kappas, exponential linear or exponential quadratic weights, Hubert's and Light's weighted kappa coefficients perform similarly and better than Mielke's. For example, when exponential quadratic weights were used for the tables with five categories where $n = 5000$ and the agreement level is medium, the MSE values of Hubert's and Light's coefficients were around 0.209 where Mielke's was 0.309 (Table 6).
- In the unbalanced tables with three raters, three coefficients perform similarly. The difference occurs when the unweighted version of the kappa coefficients is used. The unweighted version of Mielke's kappa coefficient performs better than the other coefficients. For example, for the tables with three categories where $n = 100$ and the agreement level is low, the MSE values of unweighted version of Hubert's and Light's coefficients were around 0.035 where Mielke's was 0.028 (Table 5).
- In the balanced tables with four or five raters, the MAE/MSE values of Hubert's and Light's weighted kappa coefficients are similar.
- In the unbalanced tables with four or five raters, the MAE/MSE values of these two coefficients are also mostly similar. The performance of Hubert's and Light's weighted kappa coefficients is similar when there is low or medium agreement. When there is high agreement among four raters' classifications, Light's kappa, calculated using linear, quadratic, ridit linear and ridit quadratic weights, shows slightly better performance than Hubert's kappa when there are five categories and $n = 100, 500, 1000$. For instance, in the unbalanced, high agreement tables with four raters where $R = 5$ and $n = 100$, the MSE value of Light's ridit linear weighted coefficient is around 0.058, whereas Hubert's is 0.103 (Table 10).
- When there is high agreement among five raters' classifications, Light's kappa, calculated using linear, quadratic, ridit linear and ridit quadratic weights, shows slightly better performance than Hubert's kappa when there are five categories and $n = 1000, 5000$. For instance, in the unbalanced, high agreement tables with five raters where $R = 5$ and $n = 1000$, the MSE value of Light's ridit quadratic weighted coefficient is around 0.022, whereas Hubert's is 0.051 (Table 13).

The coefficients in relation to their best performance when there are three raters are summarized in Table 14.

### 5.2.2. Comparison of the weighting schemes

The comparative analysis of weighting schemes is performed based on the weighted kappa coefficients, level of true agreement, and table structure (Tables 5-12 and Tables A1-A8).

The accuracy of the coefficients depends on the weights used, the number of categories, and the structure of the table. In the balanced tables, the kappa coefficients with quadratic weights, followed by the ridit quadratic weights, perform better than the other weights in every scenario. Unweighted kappa coefficients generate poorer results than the other weighting schemes. For instance, in the balanced, low agreement tables with three raters where $R = 3$ and $n = 100$, the MSE values of Hubert's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficients were found to be 0.050, 0.025, 0.009, 0.029, 0.016, 0.037, and 0.028, respectively (Table 5). This result also indicates that the ordinal structure of the categories should not be ignored and that the coefficients for ordinal tables should be used in applications.

**Table 14.** The summary of coefficients related to MSE in three rater studies

| Structure | $n$ | UW | L | Q | RL | RQ | EL | EQ |
|---|---|---|---|---|---|---|---|---|
| Balanced | 100, 500, 1000 | Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Hubert, Light | Hubert, Light |
| | 5000 | Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light |
| Unbalanced | All | Mielke | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light | Mielke, Hubert, Light |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

In the case of unbalanced tables, the results vary depending on the kappa coefficients, the number of categories, and the degree of true agreement:

- When three raters are involved, the coefficients employing ridit quadratic weights perform better than other weighting schemes. The performance of exponential quadratic weights follows that of ridit quadratic weights in most scenarios. However, it fails to estimate true agreement only in scenarios where $R = 5$ and $n = 100$, which result in very sparse tables. In contrast, the ridit quadratic and quadratic weights perform better than the others even in such sparse tables. For instance, in the unbalanced, medium agreement tables with three raters where $R = 5$ and $n = 100$, the MSE values of Mielke's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficients were found to be 0.0592, 0.0312, 0.0098, 0.0099, 0.0058, 0.0268, and 0.0096, respectively (Table 6).
- In scenarios where four raters categorise subjects into three or four categories and the agreement levels are low or high, the Hubert's and Light's kappa coefficients with ridit quadratic weights demonstrate better performance in comparison to alternative weighting schemes. For instance, in the unbalanced, low agreement tables with four raters where $R = 4$ and $n = 500$, the MSE values of Hubert's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficients were found to be 0.028, 0.013, 0.004, 0.004, 0.023, 0.010, and 0.003, respectively (Table 8). In the same conditions but the agreement level is medium, the Hubert's and Light's kappa coefficients with quadratic, followed by ridit quadratic weights demonstrate better performance. For instance, in the unbalanced, medium agreement tables with four raters where $R = 4$ and $n = 500$, the MSE values of Hubert's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficients were found to be 0.137, 0.045, 0.004, 0.062, 0.024, 0.107, and 0.087, respectively (Table 9).
- In scenarios where four raters categorise subjects into five categories, Hubert's and Light's kappa coefficients with exponential quadratic weights perform slightly better than the other weighting schemes in the low agreement tables where $n = 100$ and in the high agreement tables where $n = 100, 500, 1000$. For instance, in the unbalanced, high agreement tables with four raters where $R = 5$ and $n = 100$, the MSE values of Hubert's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficiens twere found to be 0.157, 0.091, 0.036, 0.103, 0.113, 0.034, and 0.012, respectively (Table 10). However, if $R = 5$ and $n = 5000$, the MSE values of Hubert's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficient were found to be 0.083, 0.042, 0.012, 0.018, 0.009, 0.035, and 0.013, respectively (Table 10). In this case, ridit quadratic weights result in lower MSE values than other weighting schemes.
- When five raters are present, Hubert's and Light's kappa coefficients using ridit quadratic weights, followed by exponential quadratic weights, tend to outperform other weight schemes. In cases of low agreement scenarios with $n = 100$ and $R = 5$, Hubert's and Light's kappa coefficients with exponential quadratic weights are performed slightly better than the other weights. The MSE values of Hubert's unweighted, linearly, quadratically, ridit linearly, ridit quadraticaly, exponential linearly, and exponential quadraticaly weighted coefficients were found to be 0.0291, 0.0110, 0.0040, 0.0036, 0.0038, 0.0084, and 0.0024, respectively (Table 11). Besides, exponential quadratic weights are performed better than the other weights in high agreement scenarios with $n = 1000, 5000$.

### 5.2.3. Effect of the number of categories on weighting schemes

In this section, we aim to examine the impact of the number of categories on weighting schemes (Tables 5-12 and Tables A1-A8).

Depending on the weighting scheme and distribution, the effect of number of categories on the MAE/MSE values differ:

- In the balanced tables, it is observed that the MSE values of the unweighted kappas and the weighted kappas with linear, ridit linear, and exponential linear weights increase as the number of categories increases. Conversely, the MSE values of the kappas with quadratic, ridit quadratic, and exponential quadratic weights decrease as the number of categories increases. For instance, in the balanced, low agreement tables with three raters where $n = 100$, the MSE values of unweighted Mielke's coefficient were found as 0.064, 0.092, and 0.109 where $R = 3$, $R = 4$, and $R = 5$, respectively (Table 5). The MSE values of quadraticaly weighted Mielke's coefficient were found as 0.009, 0.006, and 0.005 where $R = 3$, $R = 4$, and $R = 5$, respectively (Table 5).
- In the unbalanced tables, it is observed that the MSE values of the Mielke's unweighted kappa and the weighted kappa with linear, quadratic, ridit linear, and ridit quadratic weights decrease as the number of categories increases. For instance, in the unbalanced, medium agreement tables with three raters where $n = 500$, the MSE values of Mielke's coefficient with ridit linear weights were found as 0.050, 0.021, and 0.009 where $R = 3$, $R = 4$, and $R = 5$, respectively (Table 6).
- In the unbalanced tables with low or medium agreement, the MSE values of the Mielke's kappa with exponential linear and exponential quadratic weights also decrease as the number of categories increases. For instance, in the unbalanced, medium agreement tables with three raters where $n = 1000$, the MSE values of Mielke's coefficient with exponential quadratic weights were found as 0.049, 0.018, and 0.007 where $R = 3$, $R = 4$, and $R = 5$, respectively (Table 6).
- When there is low agreement in the unbalanced tables, the increase in the number of categories has a minimal effect on MSE. However, as the level of agreement rises, so does the impact of MSE. For instance, in the unbalanced, low agreement tables with three raters where $n = 100$, the MSE values of Light's coefficient with quadratic weights were found as 0.021, 0.010, and 0.005 where $R = 3$, $R = 4$, and $R = 5$, respectively (Table 5). However, in high agreement tables, these values were found as 0.074, 0.026, 0.005, respectively (Table 7).

## 6. CONCLUSIONS

This study can guide researchers in choosing a weighting scheme and the weighted kappa coefficient when more than two raters rate the ordinal categories. The best combination of weights and weighted kappa coefficients is summarized in Table 15.

**Table 15.** The summary of **weights** and *coefficients*

| | | R = 3 | | | R = 4 | | | R = 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **M** | **Structure** | **Low** | **Medium** | **High** | **Low** | **Medium** | **High** | **Low** | **Medium** | **High** |
| 3 | Balanced | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* | **Q** *Mielke, Hubert, Light* |
| | Unbalanced | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke, Hubert, Light* | **RQ** *Mielke* |
| 4 | Balanced | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Mielke, Hubert, Light* |
| | Unbalanced | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ,EQ*** *Hubert, Light* | **RQ** *Hubert, Light L* | **RQ,EQ⁺** *Hubert, Light* |
| 5 | Balanced | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Hubert, Light* | **Q** *Mielke, Hubert, Light* |
| | Unbalanced | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ,EQ*** *Hubert, Light* | **RQ** *Hubert, Light* | **RQ,EQ#** *Mielke* |

Abbr.: Quadratic (Q), Ridit (R), Exponential (E). *: EQ can be used when $n = 100$ while RQ is appropriate when $n = 500, 1000, 5000$. +: EQ can be used when $n = 100, 500, 1000$ while RQ is appropriate when $n = 5000$. #: RQ can be used when $n = 100, 500$ while RQ is appropriate when $n = 1000, 5000$.

The main results of the study are summarized as follows:

- The simulation study results show that the coefficients can show different behavior when different weights are used. It can be said that the different weights used have more influence on the coefficients. In terms of estimating the true agreement, the study has shown that all three coefficients give consistent results when the correct weight is used.
- The results indicate that the structure of the table, the level of agreement, and the number of categories also have an effect on the coefficients. The MSE of low agreement is found to be smaller than the MSE of medium agreement and the MSE of medium agreement is found to be smaller than the MSE of high agreement, except for the situations with Mielke's weighted kappa with linear, quadratic, ridit linear, and ridit quadratic weights in the unbalanced tables when there are five categories. The impact of the level of agreement on MSE and MAE values is less than the other weighing schemes for quadratic and ridit quadratic weights.
- Hubert's and Light's weighted kappas have similar results in most situations.
- The unweighted kappa coefficients yield poorer results compared to other weighting schemes. This result shows the importance of using weighted kappa coefficients when dealing with ordinal categories in a table.
- In the balanced tables, any of these three coefficients can be used with the quadratic weighting scheme.
- In the case of tables with three or four categories, the ridit quadratic weighting scheme can be implemented, using any of these three coefficients. For tables with three raters who assess subjects in five categories, Mielke's kappa coefficients can be used with the ridit quadratic weighting scheme in the unbalanced and high agreement tables. In the low or high agreement tables with four or five raters rating subjects in five categories, ridit quadratic or exponential quadratic weights can be utilized depending on the sample size.

The illustrative example in Section 4 is classified by three raters into three ordered categories when $n = 765$, with an imbalanced sample size design. Based on the results from the simulation study, the inter-rater agreement is summarized by the ridit quadratic weighting scheme. The agreement among the raters ranges from 0.342 to 0.418, depending on the coefficient used.

## DECLARATION OF ETHICAL STANDARDS

The author of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

## CONFLICT OF INTEREST

There is no conflict of interest in this study.

## REFERENCES


[1] Cohen J., "A coefficient of agreement for nominal scales", ***Educational and Psychological Measurement***, 20(1): 37–46, (1960).

[2] Aydin E.A., "EEG sinyalleri kullanılarak zihinsel iş yükü seviyelerinin sınıflandırılması", ***Politeknik Dergisi***, 24(2): 681–689, (2021).

[3] Cohen J., "Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit", ***Psychological Bulletin***, 70(4): 213–220, (1968).

[4] Conger A.J., "Integration and generalization of kappas for multiple raters", ***Psychological Bulletin***, 88(2): 322–328, (1960).

[5] Warrens M.J., "A family of multi-rater kappas that can always be increased and decreased by combining categories", ***Statistical Methodology***, 9(3): 330–340, (2012).

[6] Warrens M.J., "Equivalences of weighted kappas for multiple raters", ***Statistical Methodology***, 9(3): 407–422, (2012).

[7] Moss J., "Measures of agreement with multiple raters: Fréchet variances and inference", ***Psychometrika***, 89(2): 517–541, (2024).

[8] Light R.J., "Measures of response agreement for qualitative data: some generalizations and Alternatives", ***Psychological Bulletin***, 76(5): 365–377, (1971).

[9] Hubert L., "Kappa revisited", ***Psychological Bulletin***, 84(2): 289–297, (1977).

[10] Mielke P.W., Berry K.J. and Johnston J.E., "The exact variance of weighted kappa with multiple raters", ***Psychological Reports***, 101(2): 655–660, (2007).

[11] Abraira V. and de Vargas A.P., "Generalization of the kappa coeficient for ordinal categorical data, multiple observers and incomplete designs", ***Questiio***, 23(3): 561–571, (1999).

[12] Schuster C. and Smith D.A., "Dispersion-weighted kappa: An integrative framework for metric and nominal scale agreement coefficients", ***Psychometrika***, 70(1): 135–146, (2005).

[13] Vanbelle S. and Albert A., "Agreement between an isolated rater and a group of raters", ***Statistica Neerlandica***, 63(1): 82–100, (2009).

[14] Kvalseth T.O., "An alternative interpretation of the linearly weighted Kappa coefficients for ordinal data", ***Psychometrika***, 83(3): 618–627, (2018).

[15] Yilmaz A.E. and Aktas S., "Ridit and exponential type scores for estimating the kappa statistic", ***Kuwait Journal of Science***, 45 (1): 89-99, (2018).

[16] Warrens M.J., "Conditional inequalities between Cohen's kappa and weighted kappas", ***Statistical Methodology***, 10(1): 14–22, (2013).

[17] Tran D., Dolgun A. and Demirhan H., "Weighted inter-rater agreement measures for ordinal outcomes", ***Communications in Statistics-Simulation and Computation***, 49(4): 989–1003, (2020).

[18] Vanbelle S., Engelhart C.H. and Blix E., "A comprehensive guide to study the agreement and

reliability of multi-observer ordinal data", ***BMC Medical Research Methodology***, 24(1): 1–14, (2024).

[19] Demirhan H. and Yilmaz A. E., "Detection of grey zones in inter-rater agreement studies", ***BMC Medical Research Methodology***, 23(1): 1–15, (2023).

[20] Yilmaz A.E. and Demirhan H., "Weighted kappa measures for ordinal multi-class classification performance", ***Applied Soft Computing***, 134(110020): 1–16, (2023).

[21] de Raadt A., Warrens M.J., Bosker R.J. and Kiers H.A., "A comparison of reliability coefficients for ordinal rating scales", ***Journal of Classification***, 38(3): 519–543, (2021).

[22] Mielke P.W., Berry K.J. and Johnston J.E., "Resampling probability values for weighted kappa with multiple raters", ***Psychological Reports***, 102(2): 606–613, (2008).

[23] [23] Mielke P.W. and Berry, K.J., "A note on Cohen's weighted kappa coefficient of agreement with linear weights", ***Statistical Methodology***, 6(5): 439–446, (2009).

[24] Warrens M.J., "Corrected Zegers-ten Berge coefficients are special cases of Cohen's weighted kappa", ***Journal of Classification***, 31(2): 179–193, (2014).

[25] Warrens M.J., "Cohen's linearly weighted kappa is a weighted average of 2x2 kappas", ***Psychometrika***, 76(3): 471–486, (2011).

[26] Landis J.R. and Koch G.G., "The measurement of observer agreement for categorical data.", ***Biometrics***, 33(1): 159–174, (1977).

[27] Altman D.G., "***Practical Statistics for Medical Research***", Chapman & Hall, London, (1991).

[28] Fleiss J.L., Levin B. and Paik M.C., "***Statistical Methods for Rates & Proportions***", Wiley & Sons, New York, (2003).

[29] Yilmaz A.E. and Saracbasi T., "Assessing agreement between raters from the point of coefficients and log-linear models", ***Journal of Data Science***, 15(1), 1–24, (2017).

[30] Cicchetti D.V. and Allison T., "A new procedure for assessing reliability of scoring EEG sleep recordings", ***American Journal of EEG Technology***, 11(3): 101–110, (1971).

[31] Fleiss J.L. and Cohen J., "The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability", ***Educational and Psychological Measurement***, 33(3): 613–619, (1973).

[32] Bross I.D.J., "How to use ridit analysis", ***Journal of Applied Statistics***, 14 (1): 18-38, (1958).

[33] Iki K., Tahata K. and Tomizawa S., "Ridit score type quasi-symmetry and decomposition of symmetry for square contingency tables with ordered categories", ***Austrian Journal of Statistics***, 38(3): 183–192, (2009).

[34] Bagheban A.A. and Zayeri F., "A generalization of the uniform association model for assessing rater agreement in ordinal scales", ***Journal of Applied Statistics***, 37(8): 1265–1273, (2010).

[35] Weinberger M., Ferguson J.A., Westmoreland G., Mamlin L.A., Segar D.S., Eckert G.J., Greene J.Y., Martin D.K. and Tierney W.M., "Can raters consistently evaluate the content of focus groups?", ***Social Science & Medicine***, 46(7): 929–933, (1998).

[36] Muthen B., "A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators", ***Psychometrika***, 49(1): 115–132, (1984).

[37] Sertdemir Y., Burgut H.R., Alparslan Z.N., Unal I. and Gunasti S., "Comparing the methods of measuring multi-rater agreement on an ordinal rating scale: a simulation study with an application to real data", ***Journal of Applied Statistics***, 40(7): 1506–1519, (2013).

[38] Sarkar D., "***Lattice: Multivariate Data Visualization with R***", (2008).

## APPENDIX

The MAE results are summarized in Tables A1-A9 by the number of raters, number of categories, sample size, true level of agreement, coefficient, and weighting scheme. Tables A1-A3 show the results for the three raters, Tables A4-A6 for the four raters, and Tables A7-A9 for the five raters.

**Table A1.** MAE values computed for low agreement tables with three raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Mielke | 0.2495 | 0.1491 | 0.0838 | 0.1643 | 0.1149 | 0.1960 | 0.1780 | 0.1561 | 0.1501 | 0.1246 | 0.1188 | 0.1152 | 0.1443 | 0.1214 |
| | | Hubert | 0.2202 | 0.1491 | 0.0838 | 0.1612 | 0.1134 | 0.1868 | 0.1582 | 0.1761 | 0.1501 | 0.1246 | 0.1218 | 0.1138 | 0.1413 | 0.1206 |
| | | Light | 0.2202 | 0.1492 | 0.0840 | 0.1613 | 0.1135 | 0.1869 | 0.1583 | 0.1768 | 0.1510 | 0.1258 | 0.1225 | 0.1146 | 0.1421 | 0.1215 |
| | 500 | Mielke | 0.2481 | 0.1479 | 0.0763 | 0.1625 | 0.1113 | 0.1898 | 0.1589 | 0.1526 | 0.1460 | 0.1127 | 0.1116 | 0.1029 | 0.1373 | 0.1093 |
| | | Hubert | 0.2189 | 0.1479 | 0.0763 | 0.1606 | 0.1110 | 0.1856 | 0.1542 | 0.1737 | 0.1460 | 0.1127 | 0.1128 | 0.1025 | 0.1359 | 0.1091 |
| | | Light | 0.2189 | 0.1479 | 0.0763 | 0.1606 | 0.1110 | 0.1856 | 0.1542 | 0.1739 | 0.1462 | 0.1130 | 0.1130 | 0.1027 | 0.1361 | 0.1093 |
| | 1000 | Mielke | 0.2482 | 0.1481 | 0.0766 | 0.1623 | 0.1114 | 0.1885 | 0.1561 | 0.1523 | 0.1455 | 0.1122 | 0.1115 | 0.1021 | 0.1363 | 0.1084 |
| | | Hubert | 0.2191 | 0.1481 | 0.0766 | 0.1609 | 0.1113 | 0.1858 | 0.1538 | 0.1733 | 0.1455 | 0.1122 | 0.1122 | 0.1019 | 0.1353 | 0.1083 |
| | | Light | 0.2191 | 0.1481 | 0.0766 | 0.1609 | 0.1113 | 0.1858 | 0.1539 | 0.1734 | 0.1456 | 0.1123 | 0.1123 | 0.1020 | 0.1353 | 0.1084 |
| | 5000 | Mielke | 0.2481 | 0.1482 | 0.0766 | 0.1616 | 0.1113 | 0.1871 | 0.1540 | 0.1522 | 0.1454 | 0.1118 | 0.1121 | 0.1022 | 0.1356 | 0.1081 |
| | | Hubert | 0.2192 | 0.1482 | 0.0766 | 0.1610 | 0.1113 | 0.1858 | 0.1536 | 0.1733 | 0.1454 | 0.1118 | 0.1124 | 0.1021 | 0.1352 | 0.1081 |
| | | Light | 0.2192 | 0.1482 | 0.0766 | 0.1610 | 0.1113 | 0.1858 | 0.1536 | 0.1733 | 0.1454 | 0.1119 | 0.1124 | 0.1022 | 0.1352 | 0.1081 |
| 4 | 100 | Mielke | 0.3013 | 0.1541 | 0.0675 | 0.1803 | 0.1118 | 0.2459 | 0.2403 | 0.1407 | 0.1190 | 0.0821 | 0.0749 | 0.0686 | 0.1096 | 0.0765 |
| | | Hubert | 0.2636 | 0.1541 | 0.0675 | 0.1758 | 0.1108 | 0.2280 | 0.2049 | 0.1694 | 0.1190 | 0.0821 | 0.0771 | 0.0680 | 0.1062 | 0.0759 |
| | | Light | 0.2636 | 0.1543 | 0.0676 | 0.1759 | 0.1110 | 0.2280 | 0.2050 | 0.1697 | 0.1195 | 0.0828 | 0.0776 | 0.0683 | 0.1067 | 0.0764 |
| | 500 | Mielke | 0.3004 | 0.1527 | 0.0518 | 0.1777 | 0.1074 | 0.2347 | 0.2083 | 0.1383 | 0.1160 | 0.0614 | 0.0606 | 0.0432 | 0.1029 | 0.0570 |
| | | Hubert | 0.2628 | 0.1527 | 0.0518 | 0.1750 | 0.1072 | 0.2260 | 0.1979 | 0.1678 | 0.1160 | 0.0614 | 0.0617 | 0.0429 | 0.1013 | 0.0567 |
| | | Light | 0.2628 | 0.1527 | 0.0518 | 0.1751 | 0.1072 | 0.2260 | 0.1979 | 0.1678 | 0.1161 | 0.0616 | 0.0618 | 0.0430 | 0.1014 | 0.0569 |
| | 1000 | Mielke | 0.3008 | 0.1527 | 0.0513 | 0.1773 | 0.1075 | 0.2326 | 0.2024 | 0.1375 | 0.1154 | 0.0597 | 0.0597 | 0.0390 | 0.1016 | 0.0551 |
| | | Hubert | 0.2629 | 0.1527 | 0.0513 | 0.1753 | 0.1074 | 0.2267 | 0.1973 | 0.1670 | 0.1154 | 0.0597 | 0.0605 | 0.0388 | 0.1005 | 0.0550 |
| | | Light | 0.2629 | 0.1527 | 0.0513 | 0.1753 | 0.1074 | 0.2267 | 0.1973 | 0.1670 | 0.1154 | 0.0599 | 0.0605 | 0.0388 | 0.1005 | 0.0551 |
| | 5000 | Mielke | 0.3006 | 0.1527 | 0.0515 | 0.1763 | 0.1076 | 0.2298 | 0.1981 | 0.1377 | 0.1153 | 0.0599 | 0.0603 | 0.0381 | 0.1009 | 0.0549 |
| | | Hubert | 0.2627 | 0.1527 | 0.0515 | 0.1754 | 0.1076 | 0.2270 | 0.1969 | 0.1667 | 0.1153 | 0.0599 | 0.0605 | 0.0380 | 0.1004 | 0.0549 |
| | | Light | 0.2627 | 0.1527 | 0.0515 | 0.1754 | 0.1076 | 0.2270 | 0.1969 | 0.1667 | 0.1153 | 0.0599 | 0.0605 | 0.0380 | 0.1004 | 0.0549 |
| 5 | 100 | Mielke | 0.3297 | 0.1556 | 0.0594 | 0.1896 | 0.1125 | 0.2817 | 0.2844 | 0.1354 | 0.0976 | 0.0596 | 0.0560 | 0.0538 | 0.0883 | 0.0545 |
| | | Hubert | 0.2889 | 0.1556 | 0.0594 | 0.1844 | 0.1120 | 0.2586 | 0.2412 | 0.1678 | 0.0976 | 0.0596 | 0.0549 | 0.0563 | 0.0847 | 0.0541 |
| | | Light | 0.2889 | 0.1558 | 0.0596 | 0.1845 | 0.1121 | 0.2587 | 0.2413 | 0.1681 | 0.0979 | 0.0598 | 0.0550 | 0.0564 | 0.0850 | 0.0543 |
| | 500 | Mielke | 0.3299 | 0.1550 | 0.0406 | 0.1878 | 0.1091 | 0.2678 | 0.2498 | 0.1333 | 0.0981 | 0.0360 | 0.0321 | 0.0262 | 0.0835 | 0.0317 |
| | | Hubert | 0.2894 | 0.1550 | 0.0406 | 0.1847 | 0.1089 | 0.2543 | 0.2324 | 0.1663 | 0.0981 | 0.0360 | 0.0327 | 0.0263 | 0.0816 | 0.0316 |
| | | Light | 0.2894 | 0.1551 | 0.0406 | 0.1848 | 0.1089 | 0.2543 | 0.2324 | 0.1663 | 0.0982 | 0.0360 | 0.0328 | 0.0263 | 0.0817 | 0.0316 |
| | 1000 | Mielke | 0.3298 | 0.1552 | 0.0397 | 0.1874 | 0.1093 | 0.2645 | 0.2411 | 0.1329 | 0.0974 | 0.0309 | 0.0281 | 0.0180 | 0.0821 | 0.0268 |
| | | Hubert | 0.2894 | 0.1552 | 0.0397 | 0.1852 | 0.1092 | 0.2550 | 0.2317 | 0.1655 | 0.0974 | 0.0309 | 0.0286 | 0.0181 | 0.0808 | 0.0266 |
| | | Light | 0.2894 | 0.1552 | 0.0397 | 0.1852 | 0.1092 | 0.2550 | 0.2317 | 0.1656 | 0.0975 | 0.0310 | 0.0286 | 0.0180 | 0.0809 | 0.0267 |
| | 5000 | Mielke | 0.3299 | 0.1554 | 0.0394 | 0.1866 | 0.1095 | 0.2611 | 0.2337 | 0.1333 | 0.0973 | 0.0280 | 0.0266 | 0.0096 | 0.0813 | 0.0240 |
| | | Hubert | 0.2896 | 0.1554 | 0.0394 | 0.1855 | 0.1095 | 0.2566 | 0.2315 | 0.1656 | 0.0973 | 0.0280 | 0.0268 | 0.0097 | 0.0808 | 0.0240 |
| | | Light | 0.2896 | 0.1554 | 0.0394 | 0.1855 | 0.1095 | 0.2566 | 0.2315 | 0.1656 | 0.0973 | 0.0280 | 0.0268 | 0.0097 | 0.0808 | 0.0240 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A2.** MAE values computed for medium agreement tables with three raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Mielke | 0.3511 | 0.2083 | 0.1032 | 0.2311 | 0.1620 | 0.2803 | 0.2668 | 0.2565 | 0.2492 | 0.2184 | 0.2149 | 0.2117 | 0.2426 | 0.2172 |
| | | Hubert | 0.3039 | 0.2083 | 0.1032 | 0.2271 | 0.1598 | 0.2658 | 0.2317 | 0.2708 | 0.2492 | 0.2184 | 0.2181 | 0.2097 | 0.2395 | 0.2161 |
| | | Light | 0.3039 | 0.2084 | 0.1034 | 0.2272 | 0.1600 | 0.2658 | 0.2318 | 0.2714 | 0.2498 | 0.2191 | 0.2188 | 0.2105 | 0.2401 | 0.2168 |
| | 500 | Mielke | 0.3505 | 0.2073 | 0.1007 | 0.2293 | 0.1597 | 0.2717 | 0.2357 | 0.2557 | 0.2506 | 0.2235 | 0.2219 | 0.2171 | 0.2429 | 0.2209 |
| | | Hubert | 0.3030 | 0.2073 | 0.1007 | 0.2269 | 0.1593 | 0.2650 | 0.2277 | 0.2683 | 0.2506 | 0.2235 | 0.2232 | 0.2166 | 0.2414 | 0.2206 |
| | | Light | 0.3030 | 0.2073 | 0.1007 | 0.2269 | 0.1593 | 0.2650 | 0.2278 | 0.2686 | 0.2509 | 0.2238 | 0.2234 | 0.2169 | 0.2416 | 0.2209 |
| | 1000 | Mielke | 0.3504 | 0.2070 | 0.1005 | 0.2286 | 0.1594 | 0.2693 | 0.2308 | 0.2560 | 0.2510 | 0.2239 | 0.2226 | 0.2172 | 0.2426 | 0.2208 |
| | | Hubert | 0.3028 | 0.2070 | 0.1005 | 0.2268 | 0.1592 | 0.2648 | 0.2270 | 0.2686 | 0.2510 | 0.2239 | 0.2234 | 0.2169 | 0.2415 | 0.2206 |
| | | Light | 0.3028 | 0.2071 | 0.1005 | 0.2268 | 0.1592 | 0.2648 | 0.2270 | 0.2687 | 0.2511 | 0.2241 | 0.2235 | 0.2170 | 0.2416 | 0.2208 |
| | 5000 | Mielke | 0.3502 | 0.2070 | 0.1004 | 0.2276 | 0.1591 | 0.2668 | 0.2275 | 0.2558 | 0.2511 | 0.2241 | 0.2233 | 0.2173 | 0.2420 | 0.2207 |
| | | Hubert | 0.3027 | 0.2070 | 0.1004 | 0.2268 | 0.1591 | 0.2648 | 0.2267 | 0.2685 | 0.2511 | 0.2241 | 0.2236 | 0.2172 | 0.2416 | 0.2207 |
| | | Light | 0.3027 | 0.2070 | 0.1004 | 0.2268 | 0.1591 | 0.2648 | 0.2267 | 0.2685 | 0.2511 | 0.2242 | 0.2236 | 0.2172 | 0.2416 | 0.2207 |
| 4 | 100 | Mielke | 0.4344 | 0.2138 | 0.0710 | 0.2538 | 0.1566 | 0.3601 | 0.3736 | 0.2361 | 0.1961 | 0.1250 | 0.1287 | 0.1102 | 0.1825 | 0.1247 |
| | | Hubert | 0.3703 | 0.2138 | 0.0710 | 0.2474 | 0.1552 | 0.3269 | 0.3002 | 0.2568 | 0.1961 | 0.1250 | 0.1313 | 0.1081 | 0.1780 | 0.1234 |
| | | Light | 0.3703 | 0.2139 | 0.0711 | 0.2475 | 0.1554 | 0.3269 | 0.3002 | 0.2571 | 0.1966 | 0.1257 | 0.1317 | 0.1085 | 0.1785 | 0.1239 |
| | 500 | Mielke | 0.4344 | 0.2130 | 0.0633 | 0.2514 | 0.1537 | 0.3445 | 0.3183 | 0.2357 | 0.2009 | 0.1367 | 0.1414 | 0.1196 | 0.1857 | 0.1332 |
| | | Hubert | 0.3705 | 0.2130 | 0.0633 | 0.2476 | 0.1533 | 0.3274 | 0.2945 | 0.2537 | 0.2009 | 0.1367 | 0.1422 | 0.1191 | 0.1837 | 0.1328 |
| | | Light | 0.3705 | 0.2130 | 0.0634 | 0.2477 | 0.1534 | 0.3274 | 0.2945 | 0.2538 | 0.2010 | 0.1369 | 0.1423 | 0.1192 | 0.1838 | 0.1330 |
| | 1000 | Mielke | 0.4346 | 0.2132 | 0.0634 | 0.2508 | 0.1539 | 0.3409 | 0.3067 | 0.2347 | 0.2004 | 0.1372 | 0.1415 | 0.1195 | 0.1844 | 0.1328 |
| | | Hubert | 0.3707 | 0.2132 | 0.0634 | 0.2481 | 0.1537 | 0.3291 | 0.2947 | 0.2524 | 0.2004 | 0.1372 | 0.1421 | 0.1193 | 0.1831 | 0.1327 |
| | | Light | 0.3707 | 0.2132 | 0.0634 | 0.2481 | 0.1538 | 0.3291 | 0.2947 | 0.2525 | 0.2004 | 0.1373 | 0.1421 | 0.1193 | 0.1831 | 0.1327 |
| | 5000 | Mielke | 0.4345 | 0.2130 | 0.0634 | 0.2493 | 0.1537 | 0.3357 | 0.2975 | 0.2347 | 0.1999 | 0.1368 | 0.1418 | 0.1193 | 0.1832 | 0.1323 |
| | | Hubert | 0.3704 | 0.2130 | 0.0634 | 0.2480 | 0.1537 | 0.3302 | 0.2948 | 0.2519 | 0.1999 | 0.1368 | 0.1419 | 0.1192 | 0.1826 | 0.1322 |
| | | Light | 0.3704 | 0.2130 | 0.0634 | 0.2480 | 0.1537 | 0.3302 | 0.2948 | 0.2519 | 0.1999 | 0.1369 | 0.1419 | 0.1193 | 0.1827 | 0.1322 |
| 5 | 100 | Mielke | 0.4822 | 0.2169 | 0.0584 | 0.2687 | 0.1582 | 0.4228 | 0.4534 | 0.2380 | 0.1681 | 0.0811 | 0.0828 | 0.0603 | 0.1535 | 0.0803 |
| | | Hubert | 0.4138 | 0.2169 | 0.0584 | 0.2614 | 0.1576 | 0.3786 | 0.3595 | 0.2574 | 0.1681 | 0.0811 | 0.0848 | 0.0595 | 0.1481 | 0.0789 |
| | | Light | 0.4138 | 0.2170 | 0.0586 | 0.2615 | 0.1578 | 0.3786 | 0.3596 | 0.2576 | 0.1684 | 0.0815 | 0.0851 | 0.0597 | 0.1483 | 0.0792 |
| | 500 | Mielke | 0.4824 | 0.2156 | 0.0456 | 0.2656 | 0.1549 | 0.4009 | 0.3923 | 0.2354 | 0.1739 | 0.0866 | 0.0935 | 0.0562 | 0.1558 | 0.0836 |
| | | Hubert | 0.4137 | 0.2156 | 0.0456 | 0.2612 | 0.1547 | 0.3720 | 0.3471 | 0.2540 | 0.1739 | 0.0866 | 0.0940 | 0.0558 | 0.1535 | 0.0833 |
| | | Light | 0.4137 | 0.2157 | 0.0457 | 0.2612 | 0.1547 | 0.3721 | 0.3472 | 0.2540 | 0.1740 | 0.0867 | 0.0940 | 0.0559 | 0.1536 | 0.0834 |
| | 1000 | Mielke | 0.4822 | 0.2156 | 0.0454 | 0.2647 | 0.1551 | 0.3952 | 0.3738 | 0.2346 | 0.1737 | 0.0872 | 0.0948 | 0.0570 | 0.1551 | 0.0840 |
| | | Hubert | 0.4136 | 0.2156 | 0.0454 | 0.2615 | 0.1549 | 0.3739 | 0.3473 | 0.2531 | 0.1737 | 0.0872 | 0.0951 | 0.0568 | 0.1535 | 0.0838 |
| | | Light | 0.4136 | 0.2156 | 0.0455 | 0.2615 | 0.1550 | 0.3739 | 0.3473 | 0.2531 | 0.1737 | 0.0873 | 0.0951 | 0.0568 | 0.1535 | 0.0839 |
| | 5000 | Mielke | 0.4823 | 0.2155 | 0.0452 | 0.2631 | 0.1548 | 0.3885 | 0.3562 | 0.2349 | 0.1735 | 0.0869 | 0.0956 | 0.0575 | 0.1542 | 0.0839 |
| | | Hubert | 0.4136 | 0.2155 | 0.0452 | 0.2616 | 0.1548 | 0.3782 | 0.3495 | 0.2530 | 0.1735 | 0.0869 | 0.0957 | 0.0574 | 0.1536 | 0.0839 |
| | | Light | 0.4136 | 0.2155 | 0.0452 | 0.2616 | 0.1548 | 0.3782 | 0.3495 | 0.2530 | 0.1735 | 0.0870 | 0.0957 | 0.0574 | 0.1536 | 0.0839 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A3.** MAE values computed for high agreement tables with three raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Mielke | 0.3724 | 0.2314 | 0.1262 | 0.2533 | 0.1906 | 0.3041 | 0.3125 | 0.3055 | 0.2913 | 0.2586 | 0.2567 | 0.2540 | 0.2844 | 0.2586 |
| | | Hubert | 0.3193 | 0.2314 | 0.1262 | 0.2501 | 0.1880 | 0.2892 | 0.2613 | 0.3116 | 0.2913 | 0.2586 | 0.2597 | 0.2517 | 0.2812 | 0.2573 |
| | | Light | 0.3193 | 0.2314 | 0.1263 | 0.2502 | 0.1881 | 0.2892 | 0.2613 | 0.3127 | 0.2922 | 0.2594 | 0.2608 | 0.2528 | 0.2822 | 0.2582 |
| | 500 | Mielke | 0.3708 | 0.2294 | 0.1240 | 0.2512 | 0.1879 | 0.2948 | 0.2698 | 0.3062 | 0.3017 | 0.2889 | 0.2881 | 0.2870 | 0.2986 | 0.2883 |
| | | Hubert | 0.3173 | 0.2294 | 0.1240 | 0.2492 | 0.1874 | 0.2878 | 0.2583 | 0.3088 | 0.3017 | 0.2889 | 0.2892 | 0.2865 | 0.2974 | 0.2880 |
| | | Light | 0.3173 | 0.2294 | 0.1240 | 0.2492 | 0.1874 | 0.2878 | 0.2583 | 0.3090 | 0.3020 | 0.2891 | 0.2894 | 0.2868 | 0.2976 | 0.2882 |
| | 1000 | Mielke | 0.3697 | 0.2283 | 0.1230 | 0.2498 | 0.1868 | 0.2914 | 0.2622 | 0.3042 | 0.3009 | 0.2908 | 0.2901 | 0.2890 | 0.2981 | 0.2900 |
| | | Hubert | 0.3163 | 0.2283 | 0.1230 | 0.2484 | 0.1866 | 0.2868 | 0.2569 | 0.3063 | 0.3009 | 0.2908 | 0.2907 | 0.2887 | 0.2973 | 0.2898 |
| | | Light | 0.3163 | 0.2283 | 0.1230 | 0.2484 | 0.1866 | 0.2868 | 0.2569 | 0.3064 | 0.3010 | 0.2909 | 0.2909 | 0.2889 | 0.2974 | 0.2899 |
| | 5000 | Mielke | 0.3697 | 0.2283 | 0.1229 | 0.2493 | 0.1869 | 0.2891 | 0.2581 | 0.3034 | 0.3005 | 0.2912 | 0.2909 | 0.2894 | 0.2974 | 0.2902 |
| | | Hubert | 0.3163 | 0.2283 | 0.1229 | 0.2486 | 0.1869 | 0.2870 | 0.2570 | 0.3054 | 0.3005 | 0.2912 | 0.2911 | 0.2893 | 0.2971 | 0.2902 |
| | | Light | 0.3163 | 0.2283 | 0.1229 | 0.2486 | 0.1869 | 0.2870 | 0.2570 | 0.3054 | 0.3005 | 0.2912 | 0.2912 | 0.2894 | 0.2971 | 0.2902 |
| 4 | 100 | Mielke | 0.4785 | 0.2346 | 0.0804 | 0.2771 | 0.1795 | 0.4005 | 0.4740 | 0.2861 | 0.2333 | 0.1529 | 0.1579 | 0.1343 | 0.2151 | 0.1518 |
| | | Hubert | 0.3962 | 0.2346 | 0.0804 | 0.2712 | 0.1779 | 0.3575 | 0.3338 | 0.2908 | 0.2333 | 0.1529 | 0.1594 | 0.1323 | 0.2109 | 0.1503 |
| | | Light | 0.3962 | 0.2346 | 0.0805 | 0.2712 | 0.1779 | 0.3575 | 0.3338 | 0.2909 | 0.2335 | 0.1534 | 0.1598 | 0.1327 | 0.2112 | 0.1507 |
| | 500 | Mielke | 0.4796 | 0.2336 | 0.0783 | 0.2753 | 0.1774 | 0.3841 | 0.3801 | 0.2792 | 0.2364 | 0.1741 | 0.1775 | 0.1580 | 0.2193 | 0.1696 |
| | | Hubert | 0.3964 | 0.2336 | 0.0783 | 0.2717 | 0.1769 | 0.3610 | 0.3332 | 0.2802 | 0.2364 | 0.1741 | 0.1780 | 0.1574 | 0.2174 | 0.1693 |
| | | Light | 0.3964 | 0.2336 | 0.0783 | 0.2718 | 0.1769 | 0.3610 | 0.3332 | 0.2804 | 0.2365 | 0.1743 | 0.1781 | 0.1575 | 0.2175 | 0.1694 |
| | 1000 | Mielke | 0.4806 | 0.2341 | 0.0785 | 0.2751 | 0.1775 | 0.3804 | 0.3599 | 0.2775 | 0.2357 | 0.1756 | 0.1809 | 0.1619 | 0.2191 | 0.1720 |
| | | Hubert | 0.3973 | 0.2341 | 0.0785 | 0.2724 | 0.1773 | 0.3636 | 0.3349 | 0.2780 | 0.2357 | 0.1756 | 0.1812 | 0.1617 | 0.2179 | 0.1718 |
| | | Light | 0.3973 | 0.2341 | 0.0785 | 0.2724 | 0.1773 | 0.3636 | 0.3349 | 0.2780 | 0.2357 | 0.1757 | 0.1812 | 0.1617 | 0.2179 | 0.1719 |
| | 5000 | Mielke | 0.4804 | 0.2339 | 0.0782 | 0.2736 | 0.1773 | 0.3731 | 0.3417 | 0.2769 | 0.2350 | 0.1759 | 0.1829 | 0.1640 | 0.2187 | 0.1731 |
| | | Hubert | 0.3971 | 0.2339 | 0.0782 | 0.2724 | 0.1773 | 0.3653 | 0.3362 | 0.2769 | 0.2350 | 0.1759 | 0.1830 | 0.1639 | 0.2181 | 0.1730 |
| | | Light | 0.3971 | 0.2339 | 0.0782 | 0.2724 | 0.1773 | 0.3653 | 0.3362 | 0.2769 | 0.2350 | 0.1759 | 0.1830 | 0.1640 | 0.2181 | 0.1730 |
| 5 | 100 | Mielke | 0.5562 | 0.2370 | 0.0576 | 0.2947 | 0.1757 | 0.4962 | 0.6123 | 0.2275 | 0.1450 | 0.0577 | 0.0662 | 0.0267 | 0.2236 | 0.2236 |
| | | Hubert | 0.4584 | 0.2370 | 0.0576 | 0.2871 | 0.1747 | 0.4247 | 0.4071 | 0.2236 | 0.1450 | 0.0577 | 0.0662 | 0.0272 | 0.2236 | 0.2236 |
| | | Light | 0.4584 | 0.2371 | 0.0578 | 0.2872 | 0.1747 | 0.4247 | 0.4071 | 0.2233 | 0.1449 | 0.0575 | 0.0662 | 0.0270 | 0.2233 | 0.2233 |
| | 500 | Mielke | 0.5561 | 0.2355 | 0.0544 | 0.2911 | 0.1721 | 0.4688 | 0.5091 | 0.3304 | 0.2525 | 0.1688 | 0.1845 | 0.1474 | 0.2494 | 0.1821 |
| | | Hubert | 0.4575 | 0.2355 | 0.0544 | 0.2865 | 0.1717 | 0.4189 | 0.3965 | 0.3255 | 0.2525 | 0.1688 | 0.1843 | 0.1470 | 0.2473 | 0.1818 |
| | | Light | 0.4575 | 0.2355 | 0.0544 | 0.2865 | 0.1717 | 0.4189 | 0.3965 | 0.3255 | 0.2525 | 0.1689 | 0.1844 | 0.1470 | 0.2473 | 0.1818 |
| | 1000 | Mielke | 0.5557 | 0.2353 | 0.0542 | 0.2897 | 0.1716 | 0.4599 | 0.4678 | 0.2950 | 0.2092 | 0.1168 | 0.1361 | 0.0972 | 0.1913 | 0.1174 |
| | | Hubert | 0.4574 | 0.2353 | 0.0542 | 0.2863 | 0.1714 | 0.4231 | 0.4000 | 0.2897 | 0.2092 | 0.1168 | 0.1360 | 0.0970 | 0.1897 | 0.1172 |
| | | Light | 0.4574 | 0.2353 | 0.0542 | 0.2863 | 0.1714 | 0.4231 | 0.4000 | 0.2897 | 0.2092 | 0.1168 | 0.1360 | 0.0970 | 0.1897 | 0.1173 |
| | 5000 | Mielke | 0.5561 | 0.2354 | 0.0542 | 0.2883 | 0.1716 | 0.4492 | 0.4268 | 0.2937 | 0.2080 | 0.1161 | 0.1366 | 0.0983 | 0.1895 | 0.1171 |
| | | Hubert | 0.4576 | 0.2354 | 0.0542 | 0.2867 | 0.1716 | 0.4305 | 0.4074 | 0.2882 | 0.2080 | 0.1161 | 0.1366 | 0.0982 | 0.1888 | 0.1170 |
| | | Light | 0.4576 | 0.2354 | 0.0542 | 0.2867 | 0.1716 | 0.4305 | 0.4074 | 0.2882 | 0.2080 | 0.1161 | 0.1366 | 0.0982 | 0.1888 | 0.1170 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A4.** MAE values computed for low agreement tables with four raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.2202 | 0.1497 | 0.0818 | 0.1615 | 0.1129 | 0.1867 | 0.1579 | 0.1754 | 0.1474 | 0.1164 | 0.1142 | 0.1047 | 0.1372 | 0.1122 |
| | | Light | 0.2202 | 0.1498 | 0.0819 | 0.1616 | 0.1131 | 0.1867 | 0.1580 | 0.1765 | 0.1485 | 0.1177 | 0.1155 | 0.1059 | 0.1385 | 0.1135 |
| | 500 | Hubert | 0.2189 | 0.1479 | 0.0763 | 0.1604 | 0.1107 | 0.1854 | 0.1538 | 0.1734 | 0.1454 | 0.1117 | 0.1124 | 0.1020 | 0.1353 | 0.1082 |
| | | Light | 0.2189 | 0.1479 | 0.0764 | 0.1604 | 0.1108 | 0.1854 | 0.1538 | 0.1736 | 0.1456 | 0.1121 | 0.1126 | 0.1023 | 0.1355 | 0.1085 |
| | 1000 | Hubert | 0.2192 | 0.1482 | 0.0767 | 0.1609 | 0.1112 | 0.1858 | 0.1538 | 0.1730 | 0.1452 | 0.1117 | 0.1120 | 0.1018 | 0.1349 | 0.1079 |
| | | Light | 0.2192 | 0.1483 | 0.0767 | 0.1609 | 0.1112 | 0.1858 | 0.1539 | 0.1731 | 0.1453 | 0.1119 | 0.1122 | 0.1019 | 0.1351 | 0.1081 |
| | 5000 | Hubert | 0.2191 | 0.1480 | 0.0762 | 0.1606 | 0.1108 | 0.1856 | 0.1533 | 0.1735 | 0.1458 | 0.1124 | 0.1127 | 0.1025 | 0.1355 | 0.1085 |
| | | Light | 0.2191 | 0.1480 | 0.0762 | 0.1606 | 0.1108 | 0.1856 | 0.1533 | 0.1735 | 0.1458 | 0.1124 | 0.1127 | 0.1025 | 0.1355 | 0.1085 |
| 4 | 100 | Hubert | 0.2634 | 0.1551 | 0.0636 | 0.1759 | 0.1104 | 0.2274 | 0.2041 | 0.1680 | 0.1149 | 0.0693 | 0.0674 | 0.0563 | 0.1008 | 0.0649 |
| | | Light | 0.2634 | 0.1552 | 0.0637 | 0.1760 | 0.1106 | 0.2274 | 0.2041 | 0.1686 | 0.1158 | 0.0703 | 0.0680 | 0.0567 | 0.1015 | 0.0657 |
| | 500 | Hubert | 0.2622 | 0.1526 | 0.0519 | 0.1748 | 0.1073 | 0.2255 | 0.1974 | 0.1670 | 0.1155 | 0.0604 | 0.0607 | 0.0405 | 0.1006 | 0.0557 |
| | | Light | 0.2622 | 0.1526 | 0.0519 | 0.1749 | 0.1073 | 0.2255 | 0.1975 | 0.1671 | 0.1157 | 0.0607 | 0.0609 | 0.0406 | 0.1008 | 0.0559 |
| | 1000 | Hubert | 0.2624 | 0.1528 | 0.0518 | 0.1751 | 0.1075 | 0.2262 | 0.1971 | 0.1668 | 0.1153 | 0.0598 | 0.0605 | 0.0383 | 0.1004 | 0.0549 |
| | | Light | 0.2624 | 0.1528 | 0.0518 | 0.1751 | 0.1075 | 0.2262 | 0.1971 | 0.1669 | 0.1154 | 0.0599 | 0.0606 | 0.0384 | 0.1005 | 0.0550 |
| | 5000 | Hubert | 0.2624 | 0.1527 | 0.0516 | 0.1751 | 0.1072 | 0.2268 | 0.1967 | 0.1667 | 0.1154 | 0.0601 | 0.0606 | 0.0381 | 0.1004 | 0.0550 |
| | | Light | 0.2624 | 0.1527 | 0.0516 | 0.1751 | 0.1072 | 0.2268 | 0.1967 | 0.1667 | 0.1154 | 0.0601 | 0.0606 | 0.0381 | 0.1004 | 0.0550 |
| 5 | 100 | Hubert | 0.2904 | 0.1579 | 0.0556 | 0.1865 | 0.1130 | 0.2605 | 0.2435 | 0.1661 | 0.0946 | 0.0503 | 0.0480 | 0.0494 | 0.0803 | 0.0469 |
| | | Light | 0.2904 | 0.1581 | 0.0558 | 0.1867 | 0.1133 | 0.2606 | 0.2435 | 0.1665 | 0.0952 | 0.0508 | 0.0482 | 0.0494 | 0.0807 | 0.0472 |
| | 500 | Hubert | 0.2894 | 0.1557 | 0.0409 | 0.1855 | 0.1097 | 0.2547 | 0.2331 | 0.1661 | 0.0975 | 0.0323 | 0.0306 | 0.0221 | 0.0812 | 0.0290 |
| | | Light | 0.2894 | 0.1557 | 0.0410 | 0.1855 | 0.1098 | 0.2547 | 0.2331 | 0.1661 | 0.0977 | 0.0324 | 0.0307 | 0.0221 | 0.0813 | 0.0291 |
| | 1000 | Hubert | 0.2897 | 0.1558 | 0.0401 | 0.1856 | 0.1097 | 0.2553 | 0.2321 | 0.1660 | 0.0975 | 0.0288 | 0.0282 | 0.0162 | 0.0811 | 0.0256 |
| | | Light | 0.2897 | 0.1558 | 0.0401 | 0.1856 | 0.1097 | 0.2553 | 0.2321 | 0.1660 | 0.0975 | 0.0289 | 0.0282 | 0.0161 | 0.0812 | 0.0257 |
| | 5000 | Hubert | 0.2896 | 0.1553 | 0.0392 | 0.1853 | 0.1091 | 0.2566 | 0.2314 | 0.1658 | 0.0976 | 0.0283 | 0.0272 | 0.0085 | 0.0811 | 0.0243 |
| | | Light | 0.2896 | 0.1553 | 0.0393 | 0.1853 | 0.1091 | 0.2566 | 0.2314 | 0.1658 | 0.0977 | 0.0284 | 0.0272 | 0.0085 | 0.0811 | 0.0243 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A5.** MAE values computed for medium agreement tables with four raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.3067 | 0.2119 | 0.1064 | 0.2304 | 0.1634 | 0.2685 | 0.2344 | 0.2666 | 0.2432 | 0.2095 | 0.2095 | 0.2005 | 0.2324 | 0.2068 |
| | | Light | 0.3067 | 0.2120 | 0.1066 | 0.2305 | 0.1636 | 0.2686 | 0.2345 | 0.2681 | 0.2445 | 0.2112 | 0.2110 | 0.2021 | 0.2338 | 0.2083 |
| | 500 | Hubert | 0.3031 | 0.2076 | 0.1013 | 0.2272 | 0.1597 | 0.2652 | 0.2280 | 0.2687 | 0.2509 | 0.2237 | 0.2231 | 0.2166 | 0.2415 | 0.2205 |
| | | Light | 0.3031 | 0.2076 | 0.1013 | 0.2272 | 0.1597 | 0.2652 | 0.2280 | 0.2690 | 0.2512 | 0.2241 | 0.2235 | 0.2169 | 0.2418 | 0.2208 |
| | 1000 | Hubert | 0.3028 | 0.2073 | 0.1009 | 0.2269 | 0.1593 | 0.2649 | 0.2273 | 0.2684 | 0.2507 | 0.2233 | 0.2230 | 0.2165 | 0.2412 | 0.2201 |
| | | Light | 0.3028 | 0.2073 | 0.1009 | 0.2269 | 0.1594 | 0.2649 | 0.2273 | 0.2686 | 0.2508 | 0.2235 | 0.2232 | 0.2167 | 0.2413 | 0.2203 |
| | 5000 | Hubert | 0.3026 | 0.2069 | 0.1003 | 0.2266 | 0.1589 | 0.2647 | 0.2266 | 0.2688 | 0.2513 | 0.2243 | 0.2238 | 0.2174 | 0.2418 | 0.2208 |
| | | Light | 0.3026 | 0.2069 | 0.1003 | 0.2266 | 0.1589 | 0.2647 | 0.2266 | 0.2688 | 0.2513 | 0.2243 | 0.2238 | 0.2175 | 0.2418 | 0.2209 |
| 4 | 100 | Hubert | 0.3726 | 0.2160 | 0.0703 | 0.2494 | 0.1566 | 0.3288 | 0.3018 | 0.2520 | 0.1934 | 0.1231 | 0.1254 | 0.0997 | 0.1739 | 0.1187 |
| | | Light | 0.3726 | 0.2162 | 0.0705 | 0.2495 | 0.1568 | 0.3289 | 0.3018 | 0.2525 | 0.1941 | 0.1243 | 0.1261 | 0.1005 | 0.1746 | 0.1196 |
| | 500 | Hubert | 0.3705 | 0.2134 | 0.0640 | 0.2482 | 0.1542 | 0.3277 | 0.2949 | 0.2516 | 0.1991 | 0.1354 | 0.1404 | 0.1173 | 0.1817 | 0.1310 |
| | | Light | 0.3705 | 0.2134 | 0.0641 | 0.2482 | 0.1542 | 0.3277 | 0.2949 | 0.2518 | 0.1993 | 0.1356 | 0.1405 | 0.1175 | 0.1819 | 0.1312 |
| | 1000 | Hubert | 0.3704 | 0.2134 | 0.0640 | 0.2481 | 0.1539 | 0.3290 | 0.2949 | 0.2520 | 0.1998 | 0.1364 | 0.1418 | 0.1190 | 0.1826 | 0.1321 |
| | | Light | 0.3704 | 0.2134 | 0.0641 | 0.2481 | 0.1539 | 0.3290 | 0.2949 | 0.2521 | 0.1999 | 0.1365 | 0.1419 | 0.1191 | 0.1827 | 0.1322 |
| | 5000 | Hubert | 0.3700 | 0.2127 | 0.0632 | 0.2476 | 0.1533 | 0.3298 | 0.2945 | 0.2520 | 0.2002 | 0.1371 | 0.1421 | 0.1194 | 0.1828 | 0.1324 |
| | | Light | 0.3700 | 0.2127 | 0.0633 | 0.2476 | 0.1533 | 0.3298 | 0.2945 | 0.2520 | 0.2002 | 0.1372 | 0.1421 | 0.1194 | 0.1828 | 0.1325 |
| 5 | 100 | Hubert | 0.4153 | 0.2193 | 0.0569 | 0.2635 | 0.1597 | 0.3799 | 0.3611 | 0.2558 | 0.1680 | 0.0778 | 0.0821 | 0.0517 | 0.1474 | 0.0758 |
| | | Light | 0.4153 | 0.2194 | 0.0572 | 0.2636 | 0.1599 | 0.3799 | 0.3611 | 0.2561 | 0.1684 | 0.0784 | 0.0825 | 0.0519 | 0.1478 | 0.0763 |
| | 500 | Hubert | 0.4135 | 0.2159 | 0.0461 | 0.2617 | 0.1555 | 0.3721 | 0.3474 | 0.2535 | 0.1733 | 0.0862 | 0.0937 | 0.0549 | 0.1530 | 0.0829 |
| | | Light | 0.4135 | 0.2159 | 0.0462 | 0.2617 | 0.1555 | 0.3721 | 0.3474 | 0.2536 | 0.1734 | 0.0865 | 0.0938 | 0.0550 | 0.1531 | 0.0831 |
| | 1000 | Hubert | 0.4139 | 0.2158 | 0.0455 | 0.2617 | 0.1550 | 0.3739 | 0.3476 | 0.2535 | 0.1739 | 0.0874 | 0.0958 | 0.0575 | 0.1539 | 0.0843 |
| | | Light | 0.4139 | 0.2158 | 0.0455 | 0.2617 | 0.1550 | 0.3739 | 0.3476 | 0.2535 | 0.1740 | 0.0875 | 0.0958 | 0.0575 | 0.1540 | 0.0844 |
| | 5000 | Hubert | 0.4137 | 0.2155 | 0.0451 | 0.2614 | 0.1544 | 0.3781 | 0.3494 | 0.2532 | 0.1738 | 0.0874 | 0.0959 | 0.0577 | 0.1538 | 0.0842 |
| | | Light | 0.4137 | 0.2155 | 0.0451 | 0.2614 | 0.1544 | 0.3781 | 0.3494 | 0.2532 | 0.1738 | 0.0874 | 0.0959 | 0.0577 | 0.1538 | 0.0842 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A6.** MAE values computed for high agreement tables with four raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.3189 | 0.2318 | 0.1272 | 0.2507 | 0.1894 | 0.2893 | 0.2619 | 0.3097 | 0.2889 | 0.2554 | 0.2579 | 0.2500 | 0.2788 | 0.2543 |
| | | Light | 0.3188 | 0.2319 | 0.1274 | 0.2508 | 0.1895 | 0.2892 | 0.2619 | 0.3113 | 0.2903 | 0.2566 | 0.2595 | 0.2516 | 0.2803 | 0.2557 |
| | 500 | Hubert | 0.3161 | 0.2284 | 0.1233 | 0.2484 | 0.1869 | 0.2867 | 0.2573 | 0.3047 | 0.2979 | 0.2855 | 0.2854 | 0.2829 | 0.2935 | 0.2843 |
| | | Light | 0.3161 | 0.2284 | 0.1233 | 0.2484 | 0.1869 | 0.2867 | 0.2573 | 0.3050 | 0.2983 | 0.2859 | 0.2857 | 0.2832 | 0.2938 | 0.2846 |
| | 1000 | Hubert | 0.3162 | 0.2284 | 0.1232 | 0.2483 | 0.1865 | 0.2868 | 0.2570 | 0.3061 | 0.3007 | 0.2906 | 0.2906 | 0.2886 | 0.2970 | 0.2896 |
| | | Light | 0.3162 | 0.2284 | 0.1232 | 0.2483 | 0.1866 | 0.2868 | 0.2570 | 0.3062 | 0.3009 | 0.2908 | 0.2908 | 0.2888 | 0.2972 | 0.2898 |
| | 5000 | Hubert | 0.3164 | 0.2284 | 0.1231 | 0.2487 | 0.1869 | 0.2871 | 0.2572 | 0.3059 | 0.3009 | 0.2916 | 0.2915 | 0.2897 | 0.2975 | 0.2905 |
| | | Light | 0.3164 | 0.2284 | 0.1231 | 0.2487 | 0.1869 | 0.2871 | 0.2572 | 0.3059 | 0.3010 | 0.2916 | 0.2916 | 0.2898 | 0.2975 | 0.2905 |
| 4 | 100 | Hubert | 0.4013 | 0.2387 | 0.0831 | 0.2763 | 0.1828 | 0.3617 | 0.3380 | 0.2647 | 0.2149 | 0.1490 | 0.1368 | 0.0865 | 0.2249 | 0.1959 |
| | | Light | 0.4012 | 0.2388 | 0.0833 | 0.2763 | 0.1829 | 0.3616 | 0.3380 | 0.2648 | 0.2150 | 0.1496 | 0.1370 | 0.0866 | 0.2251 | 0.1962 |
| | 500 | Hubert | 0.3976 | 0.2347 | 0.0791 | 0.2732 | 0.1785 | 0.3622 | 0.3345 | 0.2787 | 0.2326 | 0.1676 | 0.1732 | 0.1515 | 0.2145 | 0.1631 |
| | | Light | 0.3976 | 0.2347 | 0.0792 | 0.2732 | 0.1786 | 0.3622 | 0.3345 | 0.2788 | 0.2327 | 0.1678 | 0.1733 | 0.1516 | 0.2146 | 0.1632 |
| | 1000 | Hubert | 0.3971 | 0.2342 | 0.0787 | 0.2725 | 0.1776 | 0.3636 | 0.3351 | 0.2804 | 0.2376 | 0.1766 | 0.1843 | 0.1646 | 0.2204 | 0.1741 |
| | | Light | 0.3971 | 0.2342 | 0.0787 | 0.2725 | 0.1776 | 0.3636 | 0.3351 | 0.2805 | 0.2377 | 0.1768 | 0.1844 | 0.1647 | 0.2205 | 0.1742 |
| | 5000 | Hubert | 0.3967 | 0.2336 | 0.0780 | 0.2720 | 0.1769 | 0.3649 | 0.3360 | 0.2763 | 0.2345 | 0.1756 | 0.1824 | 0.1634 | 0.2175 | 0.1725 |
| | | Light | 0.3967 | 0.2336 | 0.0780 | 0.2720 | 0.1769 | 0.3649 | 0.3360 | 0.2763 | 0.2345 | 0.1756 | 0.1824 | 0.1634 | 0.2176 | 0.1726 |
| 5 | 100 | Hubert | 0.4601 | 0.2391 | 0.0589 | 0.2895 | 0.1772 | 0.4260 | 0.4084 | 0.3691 | 0.2763 | 0.1668 | 0.2479 | 0.2281 | 0.1846 | 0.1125 |
| | | Light | 0.4600 | 0.2392 | 0.0590 | 0.2896 | 0.1773 | 0.4260 | 0.4084 | 0.3514 | 0.2610 | 0.1547 | 0.2078 | 0.1720 | 0.1847 | 0.1126 |
| | 500 | Hubert | 0.4577 | 0.2358 | 0.0548 | 0.2871 | 0.1726 | 0.4195 | 0.3971 | 0.3291 | 0.2432 | 0.1439 | 0.1843 | 0.1484 | 0.1774 | 0.1044 |
| | | Light | 0.4577 | 0.2358 | 0.0548 | 0.2871 | 0.1726 | 0.4195 | 0.3971 | 0.3207 | 0.2361 | 0.1384 | 0.1650 | 0.1212 | 0.1776 | 0.1045 |
| | 1000 | Hubert | 0.4576 | 0.2357 | 0.0547 | 0.2868 | 0.1720 | 0.4231 | 0.4001 | 0.3691 | 0.2763 | 0.1668 | 0.2479 | 0.2281 | 0.1846 | 0.1125 |
| | | Light | 0.4576 | 0.2358 | 0.0547 | 0.2868 | 0.1720 | 0.4231 | 0.4001 | 0.3514 | 0.2610 | 0.1547 | 0.2078 | 0.1720 | 0.1847 | 0.1126 |
| | 5000 | Hubert | 0.4575 | 0.2353 | 0.0541 | 0.2863 | 0.1711 | 0.4304 | 0.4072 | 0.2886 | 0.2066 | 0.1134 | 0.1359 | 0.0972 | 0.1879 | 0.1154 |
| | | Light | 0.4575 | 0.2353 | 0.0541 | 0.2863 | 0.1711 | 0.4304 | 0.4072 | 0.2886 | 0.2066 | 0.1134 | 0.1359 | 0.0972 | 0.1879 | 0.1154 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A7.** MAE values computed for low agreement tables with five raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.2195 | 0.1488 | 0.0796 | 0.1612 | 0.1127 | 0.1857 | 0.1568 | 0.1754 | 0.1474 | 0.1164 | 0.1142 | 0.1047 | 0.1372 | 0.1122 |
| | | Light | 0.2195 | 0.1489 | 0.0797 | 0.1613 | 0.1129 | 0.1858 | 0.1568 | 0.1765 | 0.1485 | 0.1177 | 0.1155 | 0.1059 | 0.1385 | 0.1135 |
| | 500 | Hubert | 0.2193 | 0.1483 | 0.0767 | 0.1608 | 0.1112 | 0.1856 | 0.1538 | 0.1734 | 0.1454 | 0.1117 | 0.1124 | 0.1020 | 0.1353 | 0.1082 |
| | | Light | 0.2193 | 0.1483 | 0.0767 | 0.1608 | 0.1112 | 0.1856 | 0.1540 | 0.1736 | 0.1456 | 0.1121 | 0.1126 | 0.1023 | 0.1355 | 0.1085 |
| | 1000 | Hubert | 0.2193 | 0.1482 | 0.0765 | 0.1609 | 0.1112 | 0.1857 | 0.1536 | 0.1730 | 0.1452 | 0.1117 | 0.1120 | 0.1018 | 0.1349 | 0.1079 |
| | | Light | 0.2193 | 0.1483 | 0.0766 | 0.1609 | 0.1112 | 0.1857 | 0.1537 | 0.1731 | 0.1453 | 0.1119 | 0.1122 | 0.1019 | 0.1351 | 0.1081 |
| | 5000 | Hubert | 0.2191 | 0.1480 | 0.0763 | 0.1607 | 0.1109 | 0.1856 | 0.1532 | 0.1735 | 0.1458 | 0.1124 | 0.1127 | 0.1025 | 0.1355 | 0.1085 |
| | | Light | 0.2191 | 0.1480 | 0.0763 | 0.1607 | 0.1109 | 0.1856 | 0.1532 | 0.1735 | 0.1458 | 0.1124 | 0.1127 | 0.1025 | 0.1355 | 0.1085 |
| 4 | 100 | Hubert | 0.2635 | 0.1537 | 0.0593 | 0.1761 | 0.1100 | 0.2271 | 0.2035 | 0.1680 | 0.1149 | 0.0693 | 0.0674 | 0.0563 | 0.1008 | 0.0649 |
| | | Light | 0.2635 | 0.1539 | 0.0595 | 0.1763 | 0.1103 | 0.2271 | 0.2039 | 0.1686 | 0.1158 | 0.0703 | 0.0680 | 0.0567 | 0.1015 | 0.0657 |
| | 500 | Hubert | 0.2626 | 0.1528 | 0.0520 | 0.1752 | 0.1077 | 0.2256 | 0.1975 | 0.1670 | 0.1155 | 0.0604 | 0.0607 | 0.0405 | 0.1006 | 0.0557 |
| | | Light | 0.2626 | 0.1528 | 0.0520 | 0.1753 | 0.1078 | 0.2256 | 0.1977 | 0.1671 | 0.1157 | 0.0607 | 0.0609 | 0.0406 | 0.1008 | 0.0559 |
| | 1000 | Hubert | 0.2627 | 0.1528 | 0.0517 | 0.1754 | 0.1077 | 0.2263 | 0.1971 | 0.1668 | 0.1153 | 0.0598 | 0.0605 | 0.0383 | 0.1004 | 0.0549 |
| | | Light | 0.2627 | 0.1528 | 0.0518 | 0.1754 | 0.1077 | 0.2263 | 0.1973 | 0.1669 | 0.1154 | 0.0599 | 0.0606 | 0.0384 | 0.1005 | 0.0550 |
| | 5000 | Hubert | 0.2626 | 0.1528 | 0.0516 | 0.1753 | 0.1074 | 0.2269 | 0.1968 | 0.1667 | 0.1154 | 0.0601 | 0.0606 | 0.0381 | 0.1004 | 0.0550 |
| | | Light | 0.2626 | 0.1528 | 0.0516 | 0.1753 | 0.1074 | 0.2269 | 0.1968 | 0.1667 | 0.1154 | 0.0601 | 0.0606 | 0.0381 | 0.1004 | 0.0550 |
| 5 | 100 | Hubert | 0.2899 | 0.1561 | 0.0512 | 0.1860 | 0.1126 | 0.2591 | 0.2417 | 0.1661 | 0.0946 | 0.0503 | 0.0480 | 0.0494 | 0.0803 | 0.0469 |
| | | Light | 0.2898 | 0.1562 | 0.0514 | 0.1862 | 0.1129 | 0.2592 | 0.2418 | 0.1665 | 0.0952 | 0.0508 | 0.0482 | 0.0494 | 0.0807 | 0.0472 |
| | 500 | Hubert | 0.2896 | 0.1554 | 0.0404 | 0.1855 | 0.1098 | 0.2545 | 0.2326 | 0.1661 | 0.0975 | 0.0323 | 0.0306 | 0.0221 | 0.0812 | 0.0290 |
| | | Light | 0.2896 | 0.1555 | 0.0404 | 0.1855 | 0.1099 | 0.2545 | 0.2327 | 0.1661 | 0.0977 | 0.0324 | 0.0307 | 0.0221 | 0.0813 | 0.0291 |
| | 1000 | Hubert | 0.2898 | 0.1555 | 0.0395 | 0.1856 | 0.1097 | 0.2552 | 0.2318 | 0.1660 | 0.0975 | 0.0288 | 0.0282 | 0.0162 | 0.0811 | 0.0256 |
| | | Light | 0.2898 | 0.1555 | 0.0396 | 0.1856 | 0.1097 | 0.2552 | 0.2318 | 0.1660 | 0.0975 | 0.0289 | 0.0282 | 0.0161 | 0.0812 | 0.0257 |
| | 5000 | Hubert | 0.2896 | 0.1553 | 0.0392 | 0.1853 | 0.1091 | 0.2565 | 0.2313 | 0.1658 | 0.0976 | 0.0283 | 0.0272 | 0.0085 | 0.0811 | 0.0243 |
| | | Light | 0.2896 | 0.1553 | 0.0392 | 0.1853 | 0.1091 | 0.2565 | 0.2313 | 0.1658 | 0.0977 | 0.0284 | 0.0272 | 0.0085 | 0.0811 | 0.0243 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A8.** MAE values computed for medium agreement tables with five raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.3038 | 0.2086 | 0.1029 | 0.2279 | 0.1613 | 0.2649 | 0.2304 | 0.2695 | 0.2477 | 0.2160 | 0.2151 | 0.2067 | 0.2370 | 0.2126 |
| | | Light | 0.3038 | 0.2087 | 0.1031 | 0.2280 | 0.1615 | 0.2650 | 0.2302 | 0.2712 | 0.2492 | 0.2178 | 0.2171 | 0.2087 | 0.2387 | 0.2141 |
| | 500 | Hubert | 0.3026 | 0.2071 | 0.1007 | 0.2265 | 0.1591 | 0.2642 | 0.2267 | 0.2694 | 0.2520 | 0.2251 | 0.2244 | 0.2180 | 0.2425 | 0.2217 |
| | | Light | 0.3026 | 0.2071 | 0.1007 | 0.2265 | 0.1591 | 0.2643 | 0.2270 | 0.2698 | 0.2523 | 0.2256 | 0.2248 | 0.2185 | 0.2429 | 0.2221 |
| | 1000 | Hubert | 0.3027 | 0.2070 | 0.1004 | 0.2266 | 0.1591 | 0.2645 | 0.2266 | 0.2700 | 0.2525 | 0.2255 | 0.2251 | 0.2187 | 0.2431 | 0.2222 |
| | | Light | 0.3027 | 0.2070 | 0.1005 | 0.2267 | 0.1591 | 0.2645 | 0.2267 | 0.2702 | 0.2527 | 0.2258 | 0.2253 | 0.2189 | 0.2433 | 0.2224 |
| | 5000 | Hubert | 0.3026 | 0.2068 | 0.1002 | 0.2266 | 0.1588 | 0.2646 | 0.2264 | 0.2688 | 0.2514 | 0.2244 | 0.2238 | 0.2175 | 0.2418 | 0.2209 |
| | | Light | 0.3026 | 0.2068 | 0.1002 | 0.2266 | 0.1588 | 0.2646 | 0.2264 | 0.2689 | 0.2514 | 0.2245 | 0.2239 | 0.2175 | 0.2419 | 0.2210 |
| 4 | 100 | Hubert | 0.3714 | 0.2150 | 0.0691 | 0.2494 | 0.1573 | 0.3272 | 0.3001 | 0.2522 | 0.1918 | 0.1198 | 0.1243 | 0.0979 | 0.1727 | 0.1162 |
| | | Light | 0.3714 | 0.2152 | 0.0694 | 0.2496 | 0.1576 | 0.3273 | 0.3004 | 0.2528 | 0.1926 | 0.1211 | 0.1251 | 0.0987 | 0.1735 | 0.1170 |
| | 500 | Hubert | 0.3703 | 0.2133 | 0.0640 | 0.2480 | 0.1542 | 0.3270 | 0.2940 | 0.2521 | 0.1999 | 0.1366 | 0.1417 | 0.1188 | 0.1827 | 0.1322 |
| | | Light | 0.3703 | 0.2134 | 0.0641 | 0.2480 | 0.1542 | 0.3270 | 0.2942 | 0.2523 | 0.2002 | 0.1370 | 0.1419 | 0.1191 | 0.1829 | 0.1324 |
| | 1000 | Hubert | 0.3702 | 0.2131 | 0.0637 | 0.2479 | 0.1538 | 0.3285 | 0.2941 | 0.2528 | 0.2011 | 0.1382 | 0.1434 | 0.1208 | 0.1839 | 0.1338 |
| | | Light | 0.3702 | 0.2131 | 0.0637 | 0.2479 | 0.1539 | 0.3285 | 0.2943 | 0.2529 | 0.2012 | 0.1384 | 0.1435 | 0.1209 | 0.1840 | 0.1339 |
| | 5000 | Hubert | 0.3699 | 0.2127 | 0.0632 | 0.2474 | 0.1531 | 0.3296 | 0.2941 | 0.2520 | 0.2002 | 0.1372 | 0.1422 | 0.1196 | 0.1829 | 0.1325 |
| | | Light | 0.3699 | 0.2127 | 0.0632 | 0.2474 | 0.1531 | 0.3296 | 0.2941 | 0.2520 | 0.2002 | 0.1372 | 0.1423 | 0.1196 | 0.1829 | 0.1326 |
| 5 | 100 | Hubert | 0.4150 | 0.2176 | 0.0534 | 0.2638 | 0.1603 | 0.3796 | 0.3607 | 0.2540 | 0.1638 | 0.0703 | 0.0774 | 0.0453 | 0.1432 | 0.0700 |
| | | Light | 0.4150 | 0.2178 | 0.0537 | 0.2640 | 0.1605 | 0.3796 | 0.3606 | 0.2542 | 0.1643 | 0.0711 | 0.0777 | 0.0456 | 0.1436 | 0.0710 |
| | 500 | Hubert | 0.4139 | 0.2157 | 0.0456 | 0.2617 | 0.1554 | 0.3722 | 0.3470 | 0.2536 | 0.1732 | 0.0860 | 0.0940 | 0.0550 | 0.1531 | 0.0829 |
| | | Light | 0.4139 | 0.2158 | 0.0457 | 0.2618 | 0.1555 | 0.3722 | 0.3470 | 0.2537 | 0.1734 | 0.0862 | 0.0941 | 0.0552 | 0.1532 | 0.0830 |
| | 1000 | Hubert | 0.4139 | 0.2158 | 0.0455 | 0.2618 | 0.1553 | 0.3739 | 0.3473 | 0.2533 | 0.1736 | 0.0870 | 0.0956 | 0.0572 | 0.1537 | 0.0840 |
| | | Light | 0.4139 | 0.2158 | 0.0455 | 0.2619 | 0.1553 | 0.3739 | 0.3473 | 0.2533 | 0.1737 | 0.0871 | 0.0957 | 0.0573 | 0.1538 | 0.0840 |
| | 5000 | Hubert | 0.4137 | 0.2154 | 0.0450 | 0.2614 | 0.1543 | 0.3781 | 0.3493 | 0.2531 | 0.1736 | 0.0872 | 0.0959 | 0.0577 | 0.1537 | 0.0841 |
| | | Light | 0.4137 | 0.2155 | 0.0450 | 0.2614 | 0.1543 | 0.3781 | 0.3493 | 0.2531 | 0.1736 | 0.0872 | 0.0959 | 0.0577 | 0.1537 | 0.0840 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).

**Table A9.** MAE values computed for high agreement tables with five raters

| R | n | Coef. | Balanced | | | | | | | Unbalanced | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | UW | L | Q | RL | RQ | EL | EQ | UW | L | Q | RL | RQ | EL | EQ |
| 3 | 100 | Hubert | 0.3181 | 0.2307 | 0.1259 | 0.2504 | 0.1895 | 0.2877 | 0.2598 | 0.3008 | 0.2808 | 0.2482 | 0.2486 | 0.2406 | 0.2700 | 0.2457 |
| | | Light | 0.3181 | 0.2308 | 0.1261 | 0.2505 | 0.1897 | 0.2876 | 0.2599 | 0.3028 | 0.2824 | 0.2495 | 0.2505 | 0.2425 | 0.2717 | 0.2489 |
| | 500 | Hubert | 0.3159 | 0.2283 | 0.1233 | 0.2481 | 0.1867 | 0.2862 | 0.2564 | 0.3060 | 0.2992 | 0.2866 | 0.2866 | 0.2841 | 0.2947 | 0.2854 |
| | | Light | 0.3159 | 0.2283 | 0.1234 | 0.2482 | 0.1867 | 0.2862 | 0.2566 | 0.3064 | 0.2995 | 0.2870 | 0.2870 | 0.2845 | 0.2951 | 0.2858 |
| | 1000 | Hubert | 0.3163 | 0.2285 | 0.1233 | 0.2487 | 0.1871 | 0.2868 | 0.2569 | 0.3068 | 0.3014 | 0.2913 | 0.2914 | 0.2894 | 0.2978 | 0.2903 |
| | | Light | 0.3163 | 0.2285 | 0.1234 | 0.2487 | 0.1872 | 0.2868 | 0.2570 | 0.3070 | 0.3016 | 0.2916 | 0.2916 | 0.2897 | 0.2980 | 0.2906 |
| | 5000 | Hubert | 0.3161 | 0.2282 | 0.1228 | 0.2485 | 0.1868 | 0.2868 | 0.2567 | 0.3060 | 0.3011 | 0.2917 | 0.2917 | 0.2899 | 0.2976 | 0.2907 |
| | | Light | 0.3161 | 0.2282 | 0.1228 | 0.2485 | 0.1868 | 0.2868 | 0.2567 | 0.3061 | 0.3011 | 0.2918 | 0.2917 | 0.2900 | 0.2977 | 0.2907 |
| 4 | 100 | Hubert | 0.4006 | 0.2378 | 0.0822 | 0.2764 | 0.1832 | 0.3611 | 0.3372 | 0.2737 | 0.2300 | 0.1664 | 0.1773 | 0.1576 | 0.2133 | 0.1663 |
| | | Light | 0.4006 | 0.2379 | 0.0824 | 0.2765 | 0.1835 | 0.3611 | 0.3370 | 0.2737 | 0.2300 | 0.1667 | 0.1774 | 0.1577 | 0.2133 | 0.1657 |
| | 500 | Hubert | 0.3972 | 0.2343 | 0.0788 | 0.2725 | 0.1778 | 0.3613 | 0.3331 | 0.2737 | 0.2300 | 0.1664 | 0.1773 | 0.1576 | 0.2133 | 0.1663 |
| | | Light | 0.3972 | 0.2343 | 0.0788 | 0.2725 | 0.1778 | 0.3613 | 0.3332 | 0.2737 | 0.2300 | 0.1667 | 0.1774 | 0.1577 | 0.2133 | 0.1657 |
| | 1000 | Hubert | 0.3971 | 0.2341 | 0.0785 | 0.2725 | 0.1777 | 0.3633 | 0.3346 | 0.2829 | 0.2391 | 0.1770 | 0.1859 | 0.1668 | 0.2182 | 0.1712 |
| | | Light | 0.3971 | 0.2341 | 0.0785 | 0.2725 | 0.1777 | 0.3633 | 0.3347 | 0.2817 | 0.2379 | 0.1760 | 0.1839 | 0.1638 | 0.2182 | 0.1715 |
| | 5000 | Hubert | 0.3966 | 0.2334 | 0.0779 | 0.2718 | 0.1767 | 0.3646 | 0.3355 | 0.2762 | 0.2346 | 0.1759 | 0.1833 | 0.1644 | 0.2179 | 0.1732 |
| | | Light | 0.3966 | 0.2335 | 0.0779 | 0.2718 | 0.1767 | 0.3646 | 0.3355 | 0.2762 | 0.2346 | 0.1759 | 0.1833 | 0.1644 | 0.2179 | 0.1733 |
| 5 | 100 | Hubert | 0.4599 | 0.2384 | 0.0578 | 0.2898 | 0.1781 | 0.4255 | 0.4078 | 0.2830 | 0.1989 | 0.1039 | 0.0871 | 0.0343 | 0.2830 | 0.2830 |
| | | Light | 0.4599 | 0.2385 | 0.0580 | 0.2899 | 0.1782 | 0.4255 | 0.4083 | 0.2832 | 0.1992 | 0.1044 | 0.0875 | 0.0342 | 0.2832 | 0.2845 |
| | 500 | Hubert | 0.4577 | 0.2357 | 0.0546 | 0.2867 | 0.1722 | 0.4190 | 0.3963 | 0.2890 | 0.2066 | 0.1124 | 0.1375 | 0.0902 | 0.2788 | 0.2702 |
| | | Light | 0.4577 | 0.2357 | 0.0547 | 0.2868 | 0.1722 | 0.4190 | 0.3965 | 0.2891 | 0.2068 | 0.1129 | 0.1376 | 0.0903 | 0.2710 | 0.2698 |
| | 1000 | Hubert | 0.4575 | 0.2354 | 0.0543 | 0.2866 | 0.1719 | 0.4228 | 0.3997 | 0.3218 | 0.2348 | 0.1347 | 0.1804 | 0.1517 | 0.1802 | 0.1054 |
| | | Light | 0.4575 | 0.2354 | 0.0544 | 0.2866 | 0.1719 | 0.4228 | 0.3997 | 0.3135 | 0.2272 | 0.1282 | 0.1628 | 0.1232 | 0.1802 | 0.1039 |
| | 5000 | Hubert | 0.4576 | 0.2352 | 0.0540 | 0.2863 | 0.1710 | 0.4303 | 0.4069 | 0.3195 | 0.2347 | 0.1368 | 0.1796 | 0.1546 | 0.1871 | 0.1155 |
| | | Light | 0.4576 | 0.2352 | 0.0540 | 0.2863 | 0.1710 | 0.4303 | 0.4069 | 0.3104 | 0.2265 | 0.1298 | 0.1611 | 0.1255 | 0.1871 | 0.1156 |

Abbr.: Unweighted (UW), Linear (L), Quadratic (Q), Ridit (R), Exponential (E).