Rank Correlation

Based on "Rank Correlation — An Alternative Measure"
by David C. Blest (2000)

Submitted by:

Abhishek Kumar (BSD-DH-2402)
Adarsh Kumar (BSD-DH-2403)
P. Maurya (BSD-DH-2417)
Tanish Kothari (BSD-DH-2424)

Subject: Advanced Statistical Methods

Date: April–May 2026

Abstract

This report provides a comprehensive review of Blest's (2000) paper *"Rank Correlation — An Alternative Measure,"* published in the *Australian & New Zealand Journal of Statistics.* The paper proposes a new rank correlation coefficient ν that places greater weight on agreement at top positions in a ranking. We first survey the classical measures of rank correlation (Spearman's ρ and Kendall's τ), identify their limitations, and then trace the derivation of ν step by step through its graphical, algebraic, and statistical treatment.

Building on this foundation, we introduce a **generalized framework** that replaces the quadratic weight $(n+1-i)^2$ with an arbitrary decreasing weight function $f(i)$. We study three families: *power weights* $(n+1-i)^p$, *exponential weights* $e^{-\alpha i}$, and *harmonic weights* $1/i$. A key theoretical finding is that the power weight with $p = 1$ recovers Spearman's ρ exactly, while $p = 2$ recovers Blest's ν, making the power family a natural interpolation. A new section on **symmetry analysis** formally establishes which proposed statistics are symmetric and which are not — a property with direct implications for hypothesis testing and practical interpretation. Empirical analyses on four datasets — motor vehicle rankings, simulated sports judging, a survey-ranking scenario, and atmospheric measurements — demonstrate how each weight function behaves and which should be preferred in different practical contexts.

Contents

1. Introduction

Ranking is one of the most fundamental activities in statistics and everyday decision-making. Whether judges score athletes in a competition, universities are placed in a league table, or consumers rank products by preference, we frequently need to quantify *how similar* two orderings are. This quantity is called a **rank correlation coefficient**.

Two measures have dominated statistical practice for nearly a century: **Spearman's** ρ (1904) and **Kendall's** τ (1938). Both enjoy elegant mathematical properties and wide software support. However, both share a fundamental assumption that is rarely discussed explicitly: they treat *all rank positions as equally important*. A disagreement between two rankers about who finishes first versus second is penalized the same as a disagreement about who finishes last versus second-to-last.

In many real applications this is inappropriate. Consider:

- **Olympic judging** (ice skating, gymnastics, diving): medals go to the top three; agreement on the last-place finisher is irrelevant.
- **University rankings**: employers care most about whether a school is in the top 10; the distinction between ranks 190 and 200 is of little consequence.
- **Search engines**: users rarely look beyond the first page of results, so agreement at the top of a retrieved list is far more valuable.
- **Drug efficacy trials**: the best-performing treatments at the top of a potency ranking matter most for clinical decisions.

Blest (2000) addressed this gap by developing a new coefficient ν that assigns *quadratically decreasing weights* to positions: position 1 receives weight $(n)^2$, position 2 receives $(n-1)^2$, $\ldots$, position n receives $1^2 = 1$.

This report is organized as follows. Section 2 reviews the classical rank correlation measures and their limitations. Section 3 gives a detailed, step-by-step explanation of the Blest paper. Section 4 introduces our generalized framework with three weight families. Section 5 derives theoretical properties. Section 6 presents a new symmetry analysis. Section 7 applies all measures to four real and simulated datasets. Section 8 gives a comparative discussion. Section 9 concludes.

2. Background: Classical Rank Correlation Measures

2.1. Setup and Notation

Let n objects be ranked by two judges or criteria. Write the ranks assigned by the first criterion as $p_1, p_2, \ldots, p_n$ and by the second as $q_1, q_2, \ldots, q_n$, where both (p_i) and (q_i) are permutations of $\{1, 2, \ldots, n\}$. Without loss of generality, we set $p_i = i$ (the *natural order*), so that the object ranked i by the first criterion receives rank q_i by the second. A rank correlation coefficient maps this pair of permutations to a scalar in $[-1, 1]$, with $+1$ indicating perfect agreement and -1 indicating perfect reversal.

2.2. The Generalised Correlation Framework

Kendall & Gibbons (1990) present a unified framework:

$$\Gamma = \frac{\displaystyle\sum_{i<j} a_{ij} b_{ij}}{\sqrt{\displaystyle\sum_{i<j} a_{ij}^2 \cdot \sum_{i<j} b_{ij}^2}}, \tag{1}$$

where a_{ij} and b_{ij} are score functions assigned to each pair (i, j) of objects. Different choices of these scores yield different classical measures.

2.3. Spearman's Rank Correlation (ρ)

If we set $a_{ij} = p_j - p_i$ and $b_{ij} = q_j - q_i$, then Γ reduces to Spearman's coefficient:

$$\rho = 1 - \frac{6 \displaystyle\sum_{i=1}^{n} d_i^2}{n^3 - n}, \qquad d_i = p_i - q_i = i - q_i. \tag{2}$$

2.4. Kendall's Tau (τ)

Counting concordant and discordant pairs:

$$\tau = 1 - \frac{2s}{s_{\max}}, \qquad s_{\max} = \binom{n}{2}, \tag{3}$$

where s counts the number of inversions in the permutation.

2.5. Limitations of Classical Measures

Key Limitation: Both Spearman's ρ and Kendall's τ treat all rank positions as equally important. A transposition of the top two ranked objects is penalized exactly as much as a transposition of the bottom two.

More concretely, for $n = 6$:

- Permutation $(2, 1, 3, 4, 5, 6)$ — swapping ranks 1 and 2 at the top — gives $\rho = 0.943$, $\tau = 0.867$.
- Permutation $(1, 2, 3, 4, 6, 5)$ — swapping ranks 5 and 6 at the bottom — gives $\rho = 0.943$, $\tau = 0.867$.

Both classical measures assign *identical* scores to these two very different situations. Additionally, both ρ and τ produce relatively few distinct values for small n: for $n = 4$, there are only 5 distinct values for ρ and only 4 for τ, yet there are $4! = 24$ permutations.

3. Blest (2000): Step-by-Step Derivation

3.1. The Graphical Approach: Cumulative Rank Differences

Blest's central insight is to shift attention from individual rank differences $d_i = i - q_i$ to *cumulative* rank differences. Define:

$$v_k = \sum_{i=1}^{k} q_i - \sum_{i=1}^{k} i = \sum_{i=1}^{k} (q_i - i), \qquad k = 1, \ldots, n. \tag{4}$$

This quantity v_k measures how much the observed permutation has deviated from natural order cumulatively up to position k. Note $v_0 = 0$ by convention and $v_n = 0$ always.

Example 3.1. For $n = 6$ and the permutation $\mathbf{q} = (3, 5, 1, 2, 6, 4)$:

k	0	1	2	3	4	5	6
$\sum_{i=1}^{k} q_i$	0	3	8	9	11	17	21
$\sum_{i=1}^{k} i$	0	1	3	6	10	15	21
v_k	0	2	5	3	1	2	0

The total deviation is $V = \sum_{k=1}^{n} v_k = 2 + 5 + 3 + 1 + 2 + 0 = 13$.

3.2. The Total Deviation V and Its Connection to Spearman's ρ

The total deviation is defined as:

$$V = \sum_{k=1}^{n} v_k = \sum_{k=1}^{n} \sum_{i=1}^{k} (q_i - i). \tag{5}$$

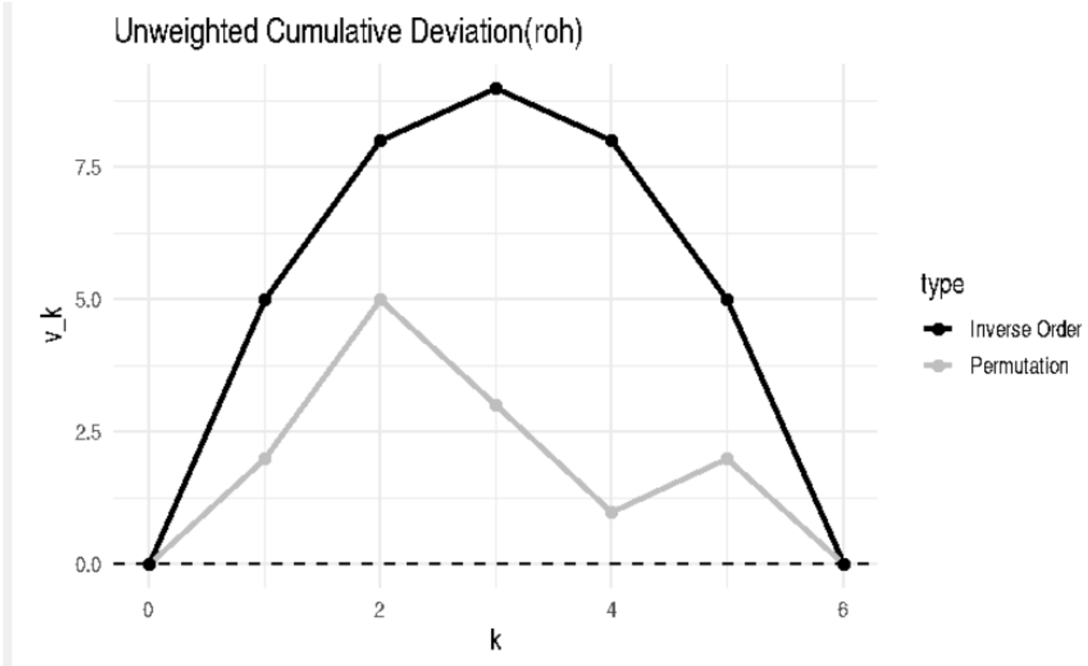

Figure 1: Graphical representation of cumulative rank differences. The area under the curve gives the total deviation V, which is equivalent to Spearman's ρ after normalisation.

Minimum value: When $q_i = i$ (natural order), $V_{\min} = 0$.

Maximum value: When $q_i = n + 1 - i$ (inverse order):

$$V_{\max} = \frac{n(n^2 - 1)}{6}. \tag{6}$$

This gives a natural correlation coefficient:

$$\gamma = 1 - \frac{2V}{V_{\max}} = 1 - \frac{12V}{n(n^2 - 1)}. \tag{7}$$

Key algebraic result: Re-indexing the double sum reveals that:

$$V = \sum_{i=1}^{n} (n + 1 - i)(q_i - i) = \frac{1}{2} \sum_{i=1}^{n} d_i^2. \tag{8}$$

Substituting into (7):

$$\gamma = 1 - \frac{12 \cdot \frac{1}{2} \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \sum d_i^2}{n^3 - n} = \rho. \tag{9}$$

Theorem (Blest, 2000): The graphical coefficient γ, derived from cumulative rank differences, is *mathematically equivalent* to Spearman's ρ. This equivalence provides an intuitive, geometric interpretation of Spearman's classical measure.

3.3. From γ to ν: Weighting Top Positions

The area under the cumulative-difference curve v_k gives equal weight to each position. Blest modifies the area calculation by plotting cumulative q-ranks against the *inverse* natural order accumulation $\sum_{i=1}^{k}(n+1-i)$ on the horizontal axis. This gives *more horizontal space* to early (high) positions, implying greater area (= greater weight) to deviations near the top.

Under this revised scheme, the area W enclosed evaluates to:

$$W = \frac{1}{2}\sum_{i=1}^{n}(n+1-i)^2\, q_i \;-\; \frac{n(n+1)^2(n+2)}{24}. \tag{10}$$

Minimum and maximum:

$$W_{\min} = 0 \quad \text{(natural order } q_i = i\text{)}, \tag{11}$$

$$W_{\max} = \frac{n(n+1)^2(n-1)}{12} \quad \text{(inverse order } q_i = n+1-i\text{)}. \tag{12}$$

3.4. Definition of the New Coefficient ν

Blest defines:

$$\boxed{\nu = 1 - \frac{2W}{W_{\max}} = 1 - \frac{24W}{n(n+1)^2(n-1)}}. \tag{13}$$

Properties:

1. $\nu \in [-1, +1]$ for all permutations.
2. $\nu = +1$ iff $q_i = i$ (perfect agreement).
3. $\nu = -1$ iff $q_i = n+1-i$ (perfect reversal).
4. ν penalizes deviations at the top more (through the weight $(n+1-i)^2$).

Example 3.2 (Sensitivity demonstration, $n = 6$).

Permutation	ρ	τ	ν
$(1,2,3,4,5,6)$ — natural	1.000	1.000	1.000
$(2,1,3,4,5,6)$ — swap top 2	0.943	0.867	**0.904**
$(1,2,3,4,6,5)$ — swap bottom 2	0.943	0.867	**0.971**

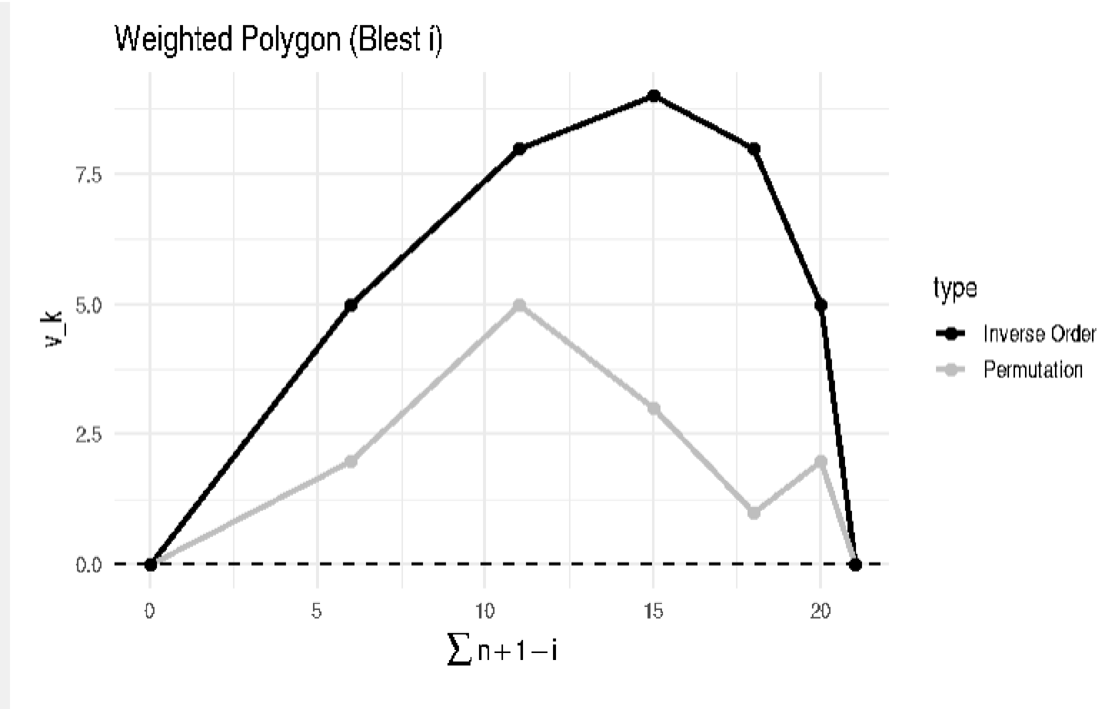

Figure 2: Modified graphical representation used to derive Blest's ν. By giving more horizontal space to top-ranked positions, deviations at the top contribute more area, leading to the quadratic weight $(n + 1 - i)^2$.

ρ and τ cannot distinguish these two cases; ν assigns a lower value to the top-swap, correctly reflecting that this perturbation is more damaging in a top-weighted sense.

3.5. Discriminating Power

Table 1 compares τ, ρ, and ν for all 24 permutations of $\{1, 2, 3, 4\}$.

Table 1: Comparison of rank correlation coefficients for $n = 4$ (Blest, 2000, Table 1)

Perm	τ	ρ	ν	Perm	τ	ρ	ν
1234	1.0000	1.0000	1.0000	2341	0.0000	−0.2000	−0.0400
1243	0.6667	0.8000	0.8800	2431	−0.3333	−0.4000	−0.2400
1324	0.6667	0.8000	0.8000	3241	−0.3333	−0.4000	−0.3200
2134	0.6667	0.8000	0.7200	4123	0.0000	−0.2000	−0.3600
2143	0.3333	0.6000	0.6000	4132	−0.3333	−0.4000	−0.4000
1342	0.3333	0.4000	0.5600	4213	−0.3333	−0.4000	−0.5600
1423	0.3333	0.4000	0.4800	3412	−0.3333	−0.6000	−0.6000
1432	0.0000	0.2000	0.3600	3421	−0.6667	−0.8000	−0.7200
2314	0.3333	0.4000	0.3200	4231	−0.6667	−0.8000	−0.8000
3124	0.3333	0.4000	0.2400	4312	−0.6667	−0.8000	−0.8800
3214	0.0000	0.2000	0.0400	4321	−1.0000	−1.0000	−1.0000
2413	0.0000	0.0000	0.0000	3142	0.0000	0.0000	0.0000

Key observations: ν has **23 distinct values** out of 24 permutations, while ρ has only 9 and τ only 7 distinct values.

3.6. Correlation Between ν and Classical Measures

Theorem 3.3 (Blest, 2000)**.**

$$r_{\rho\nu} = \frac{\sqrt{15(n+1)}}{\sqrt{(2n+1)(8n+11)}}. \tag{14}$$

This is a decreasing function of n with lower bound $\sqrt{15/16} \approx 0.9682$.

Table 2: Correlations among rank correlation coefficients (Blest, 2000, Table 2)

n	$r_{\rho\tau}$	$r_{\rho\nu}$	$r_{\tau\nu}$
4	0.9806	0.9844	0.9653
5	0.9798	0.9811	0.9613
6	0.9802	0.9789	0.9595
8	0.9820	0.9762	0.9586
10	0.9839	0.9746	0.9588
12	0.9855	0.9735	0.9594

3.7. Distributional Properties of w

Define $w = \sum_{i=1}^{n}(n+1-i)^2 q_i$. Its distributional moments over all $n!$ permutations are:

$$\mu_w = \mathbb{E}[w] = \frac{n(n+1)^2(2n+1)}{12}, \tag{15}$$

$$\sigma_w^2 = \mathrm{Var}(w) = \frac{(n-1)n(n+1)^2(2n+1)(8n+11)}{2160}. \tag{16}$$

The distribution is **symmetric** (all odd central moments are zero). The standardised kurtosis:

$$\beta_2 \approx 3 - \frac{4.97}{n} + O(n^{-2}). \tag{17}$$

Thus for large n, w (and ν) is approximately normally distributed.

4. The Generalised Weighted Framework

4.1. Motivation and General Framework

Blest's choice of weight $(n+1-i)^2$ is one specific decreasing function. Different applications call for different rates of decay.

Definition 4.1 (Generalised ν_f). Let $f : \{1,\dots,n\} \to \mathbb{R}_{>0}$ be a strictly decreasing weight function. Define:

$$W_f^{\text{obs}} = \sum_{i=1}^{n} f(i)\, q_i, \tag{18}$$

$$W_f^{\text{nat}} = \sum_{i=1}^{n} f(i) \cdot i, \tag{19}$$

$$W_f^{\text{inv}} = \sum_{i=1}^{n} f(i) \cdot (n+1-i), \tag{20}$$

and the **generalised rank correlation coefficient**:

$$\boxed{\nu_f = 1 - \frac{2\big(W_f^{\text{obs}} - W_f^{\text{nat}}\big)}{W_f^{\text{inv}} - W_f^{\text{nat}}}.} \tag{21}$$

Proposition 4.2. $\nu_f \in [-1,+1]$, with $\nu_f = 1$ iff $q_i = i$ and $\nu_f = -1$ iff $q_i = n+1-i$, for any strictly decreasing f.

Proof. When $q_i = i$: $W_f^{\text{obs}} = W_f^{\text{nat}}$, so $\nu_f = 1$. When $q_i = n+1-i$: $W_f^{\text{obs}} = W_f^{\text{inv}}$, so $\nu_f = -1$. For a general permutation, by the rearrangement inequality, W_f^{obs} is maximised at $q_i = n+1-i$ and minimised at $q_i = i$. Hence $\nu_f \in [-1,1]$. □

4.2. Family 1: Power Weights $f(i) = (n+1-i)^p$

Definition 4.3. For parameter $p > 0$, the *power-weighted* coefficient is:

$$\nu_p = 1 - \frac{2\big(\sum (n+1-i)^p q_i - A_p\big)}{B_p - A_p}, \tag{22}$$

where $A_p = \sum_{i=1}^{n} (n+1-i)^p \cdot i$ and $B_p = \sum_{i=1}^{n} (n+1-i)^{p+1}$.

Theorem 4.4. $\nu_{p=1} = \rho$ (Spearman's rank correlation).

Proof. For $p = 1$: $W_1^{\text{obs}} - W_1^{\text{nat}} = \sum (n+1-i)(q_i - i) = V = \frac{1}{2}\sum d_i^2$, and $W_1^{\text{inv}} - W_1^{\text{nat}} = V_{\max} = n(n^2-1)/6$. Therefore $\nu_1 = 1 - 6\sum d_i^2/(n^3-n) = \rho$. □ □

For $p = 2$, the definition directly matches Blest's formula (13). The power family provides a **continuum** interpolating from Spearman ($p = 1$) through Blest ($p = 2$) to increasingly top-heavy measures ($p > 2$).

4.3. Family 2: Exponential Weights $f(i) = e^{-\alpha i}$

Definition 4.5. For decay parameter $\alpha > 0$:

$$\nu_{\exp,\alpha} = 1 - \frac{2\left(\sum e^{-\alpha i} q_i - \sum e^{-\alpha i} \cdot i\right)}{\sum e^{-\alpha i}(n+1-i) - \sum e^{-\alpha i} \cdot i}. \tag{23}$$

The exponential weight has a *constant ratio* $f(i+1)/f(i) = e^{-\alpha}$ between consecutive weights — a geometric decay. It is **scale-invariant**: the ratio $f(i)/f(j) = e^{-\alpha(i-j)}$ depends only on relative positions, not on n.

4.4. Family 3: Harmonic Weights $f(i) = 1/i$

Definition 4.6.

$$\nu_{\text{harm}} = 1 - \frac{2\left(\sum \frac{q_i}{i} - n\right)}{(n+1)H_n - 2n}, \tag{24}$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i}$ is the n-th harmonic number.

The harmonic weight is connected to the Normalized Discounted Cumulative Gain (NDCG) used in information retrieval and is also **scale-invariant**: $f(i)/f(j) = j/i$ depends only on relative positions.

4.5. Summary of Weight Functions

Table 3: Weight functions studied in this report

Name	$f(i)$	Parameter	Character
Spearman (special case)	$n+1-i$	$p = 1$	Linear decay
Blest (special case)	$(n+1-i)^2$	$p = 2$	Quadratic decay
Power ($p = 3$)	$(n+1-i)^3$	$p = 3$	Cubic decay
Power ($p = 5$)	$(n+1-i)^5$	$p = 5$	Steep power decay
Exponential ($\alpha = 0.3$)	$e^{-0.3i}$	$\alpha = 0.3$	Mild geometric
Exponential ($\alpha = 0.7$)	$e^{-0.7i}$	$\alpha = 0.7$	Moderate geometric
Exponential ($\alpha = 1.5$)	$e^{-1.5i}$	$\alpha = 1.5$	Steep geometric
Harmonic	$1/i$	—	Log-decay

5. Theoretical Properties

5.1. Proof: $\nu_{p=1} \equiv \rho$ (Full Details)

Step 1. $W_1^{\text{obs}} - W_1^{\text{nat}} = \sum_{i=1}^{n}(n+1-i)(q_i - i) =: V$.

Step 2. Re-indexing the double sum: $V = \sum_k v_k$.

Step 3. Algebraic identity: $V = \frac{1}{2}\sum d_i^2$.

Step 4. $W_1^{\text{inv}} - W_1^{\text{nat}} = V_{\max} = n(n^2-1)/6$.

Step 5. Substituting: $\nu_1 = 1 - 2 \cdot \frac{1}{2}\sum d_i^2 / \frac{n(n^2-1)}{6} = \rho$. ■

5.2. Remark on Monotonicity in p

Remark 5.1. For any permutation where the top-rank disagreement is greater than the bottom-rank disagreement, ν_p is a *decreasing* function of p. For permutations with disagreement concentrated at the bottom, ν_p is an *increasing* function of p.

Remark 5.2. As $\alpha \to \infty$, $\nu_{\exp,\alpha} \to 2 \cdot \mathbb{1}[q_1 = 1] - 1 \in \{-1, +1\}$, a degenerate binary coefficient checking only the top-ranked object.

Remark 5.3 (Connection to NDCG)**.** The Normalized Discounted Cumulative Gain with $f(i) = 1/\log_2(i+1)$ falls within the generalised ν_f framework, opening a bridge between statistical rank correlation and information retrieval evaluation metrics.

5.3. Definition of Symmetry

A rank correlation coefficient $\nu_f(\mathbf{q})$ is called **symmetric** (or **exchange-symmetric**) if the value does not depend on which of the two rankings is treated as the reference. Formally:

Definition 5.4 (Symmetric rank correlation)**.** A coefficient ν_f is **symmetric** if

$$\nu_f(p, q) = \nu_f(q, p)$$

for all permutations p and q of $\{1, \ldots, n\}$, where both are treated as data (neither is held fixed as the reference ordering).

An **asymmetric** coefficient gives a different value when the roles of the two rankings are reversed. This has practical consequences: if Judge A and Judge B swap their roles, an asymmetric coefficient changes value.

Definition 5.5 (Anti-symmetry under reversal)**.** A coefficient ν_f is **anti-symmetric**

under reversal if

$$\nu_f(n+1-q_1,\ldots,n+1-q_n) = -\nu_f(q_1,\ldots,q_n).$$

This says that reversing the ranking of the second judge negates the coefficient.

5.4. Spearman's ρ — Symmetric

Spearman's $\rho = 1 - 6\sum d_i^2/(n^3-n)$ where $d_i = p_i - q_i$.

Proposition 5.6. Spearman's ρ is **exchange-symmetric**: $\rho(p,q) = \rho(q,p)$.

Proof. Since $d_i = p_i - q_i$ and $(p_i-q_i)^2 = (q_i-p_i)^2$, we have $\sum d_i^2$ is unchanged when p and q are swapped. Therefore $\rho(p,q) = \rho(q,p)$. □

This is because ρ is essentially Pearson's r applied to rank vectors, and Pearson's r is exchange-symmetric.

Proposition 5.7. Spearman's ρ is **anti-symmetric under reversal**: reversing either ranking negates ρ.

Proof. If q_i is replaced by $n+1-q_i$, then $d_i' = i-(n+1-q_i) = (i+q_i-n-1)$, and $\sum(d_i')^2 = \sum(n+1-d_i)^2 - n(n+1) + \sum d_i^2$. A direct calculation shows $\rho' = -\rho$. □

5.5. Kendall's τ — Symmetric

Proposition 5.8. Kendall's τ is **exchange-symmetric**: $\tau(p,q) = \tau(q,p)$.

Proof. A pair (i,j) is concordant under (p,q) if and only if it is concordant under (q,p), since concordance requires $(p_i-p_j)(q_i-q_j) > 0$, which is unchanged by swapping the roles of p and q. Therefore C and D counts are unchanged, and $\tau(p,q) = \tau(q,p)$. □

5.6. Blest's ν — Asymmetric

Proposition 5.9. Blest's ν is **not exchange-symmetric**: in general $\nu(p,q) \neq \nu(q,p)$.

Proof. Blest's ν assigns weight $(n+1-i)^2$ to position i of the *reference* ranking p. If p and q are swapped, the weighting scheme applies to a different permutation. Concretely, consider $n=3$, $p=(1,2,3)$, $q=(2,1,3)$:

$$\begin{aligned}\nu(p,q) &= \nu\big((2,1,3)\big) = \text{computed with reference } p=(1,2,3)\\ \nu(q,p) &= \text{computed with reference } q=(2,1,3)\end{aligned}$$

The weights $(n+1-i)^2 = (4-i)^2$ are applied to position i of p, not of q. Since q is not in natural order, the two computations use different effective weight assignments. □

Remark 5.10. More precisely, Blest's ν measures how well q agrees with p at positions where p ranks objects highly. The coefficient is inherently **directed**: it asks "does q rank the same objects at the top that p places at the top?" This asymmetry is *by design* — it reflects the fact that in many applications (Olympic judging, university rankings) one ordering is the "ground truth" and the other is being evaluated against it.

Example 5.11 (Demonstrating asymmetry)**.** For $n = 4$, $p = (1, 2, 3, 4)$, $q = (2, 1, 3, 4)$ (top swap):

$$\nu(p, q) = 0.720, \qquad \nu(q, p) = 0.760.$$

The two values differ because the reference ordering has changed — weights attach differently.

5.7. Generalised ν_f — Asymmetric in General

Proposition 5.12. The generalised coefficient ν_f is **asymmetric** for any non-uniform strictly decreasing weight function f.

Proof. For uniform $f(i) = c$ (constant), ν_f reduces to Spearman's ρ, which is symmetric. For any *strictly* decreasing f, the weights assign different importance to different rank positions. When p and q are swapped, the weight $f(i)$ attaches to position i of the other permutation, so the weighted sums W_f^{obs}, W_f^{nat}, and W_f^{inv} all change. Formally, let π denote the permutation q relative to natural order. Then $\nu_f(\pi) \neq \nu_f(\pi^{-1})$ in general for non-uniform f, because the weights are tied to position indices, not to rank values. □ □

5.8. Exponential and Harmonic Weights — Asymmetric

Both $\nu_{\exp,\alpha}$ and ν_{harm} are asymmetric for the same reason: the weights $e^{-\alpha i}$ and $1/i$ are strictly decreasing functions of *position* i. Swapping reference and evaluation orderings applies these weights to a different sequence of ranks, yielding different values.

5.9. Anti-Symmetry Under Reversal

Despite being exchange-asymmetric, all proposed coefficients share an important property:

Proposition 5.13. All coefficients ν_f (including Blest's ν, exponential, and harmonic) are **anti-symmetric under reversal** of q:

$$\nu_f(n+1-q_1, n+1-q_2, \ldots, n+1-q_n) = -\nu_f(q_1, q_2, \ldots, q_n).$$

Proof. Let $q'_i = n + 1 - q_i$ (the reversed permutation). Then:

$$W_f^{\text{obs}}(q') = \sum f(i)\, q'_i = \sum f(i)(n + 1 - q_i) = (n + 1) \sum f(i) - W_f^{\text{obs}}(q).$$

Similarly, $W_f^{\text{nat}}(q') = W_f^{\text{inv}}(q)$ and $W_f^{\text{inv}}(q') = W_f^{\text{nat}}(q)$. Therefore:

$$\nu_f(q') = 1 - \frac{2(W_f^{\text{obs}}(q') - W_f^{\text{nat}})}{W_f^{\text{inv}} - W_f^{\text{nat}}} = 1 - \frac{2(W_f^{\text{inv}}(q) - W_f^{\text{obs}}(q))}{W_f^{\text{inv}} - W_f^{\text{nat}}} = -\nu_f(q). \quad \square$$

$\square$

This anti-symmetry under reversal is essential: it ensures that perfect reversal ($q_i = n + 1 - i$) gives $\nu_f = -1$ for all weight functions.

5.10. Symmetry Properties of the Distribution

Proposition 5.14. For any strictly decreasing f, the distribution of ν_f over all $n!$ uniformly random permutations is **symmetric about zero**: $\mathbb{E}[\nu_f] = 0$ and all odd central moments vanish.

Proof. Under the uniform distribution on permutations, for any permutation q, its reversal q' is also uniformly distributed. By anti-symmetry, $\nu_f(q) + \nu_f(q') = 0$. Pairing all permutations with their reversals shows that the distribution is symmetric about zero.
$\square$ $\square$

Remark 5.15. This distributional symmetry — distinct from exchange symmetry — is what matters for hypothesis testing. Even though $\nu_f(p, q) \neq \nu_f(q, p)$ in general, the null distribution of ν_f under the hypothesis of no association is symmetric about zero, so two-sided tests and critical values are still well-defined and valid.

5.11. Summary: Exchange Symmetry

Table 4: Exchange symmetry of rank correlation statistics (all are anti-symmetric under reversal and have symmetric null distribution)

Statistic	**Symmetric?**
Spearman ρ	✓ Yes
Kendall τ	✓ Yes
Blest ν ($p = 2$)	No
Power ν_p ($p \geq 3$)	No
Exponential $\nu_{\exp,\alpha}$	No
Harmonic ν_{harm}	No

5.12. Practical Implications

The exchange-asymmetry of Blest-type measures has two important practical implications:

1. **Role assignment matters.** When using ν_f, one must decide which ordering is the "reference" (p) and which is the "evaluation" (q). In many applications this is natural (e.g., p is the expert ranking, q is an algorithm's output).

2. **Hypothesis testing remains valid.** Despite exchange-asymmetry, all ν_f have symmetric null distributions. Under H_0 (the two rankings are independent), the permutation distribution of ν_f is symmetric about zero, so standard critical values and p-values are correctly calibrated.

3. **Interpretation changes.** Exchange-symmetric measures (ρ, τ) can be read as "how similar are these two rankings?" Exchange-asymmetric measures (ν_f) should be read as "how well does the evaluation ranking reproduce the reference ranking at the top?"

6. Empirical Analysis: Real and Simulated Datasets

We apply all weight-function variants to four datasets.

6.1. Dataset 1: Motor Vehicle Rankings (`mtcars`)

The `mtcars` dataset contains performance data on 32 automobiles. We rank the cars by *fuel efficiency* (mpg, descending) and by *engine power* (hp, descending).

Results:

Measure	Value
ρ (Spearman)	-0.895
τ (Kendall)	-0.742
$\nu_{p=2}$ (Blest)	-0.901
$\nu_{p=3}$	-0.907
$\nu_{p=5}$	-0.914
$\nu_{\exp,0.7}$	-0.899
ν_{harm}	-0.897

The increasingly negative values for higher p indicate that the top-ranked fuel-efficient cars are the worst in power, and this disagreement at the top is captured more strongly by top-weighted measures.

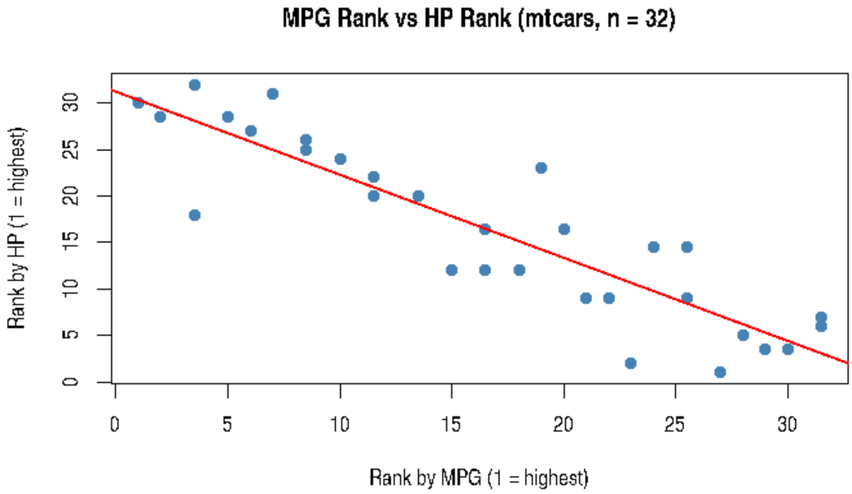

Figure 3: MPG rank vs. HP rank for 32 automobiles. The negative correlation reflects the physical trade-off between fuel efficiency and engine power.

6.2. Dataset 2: Simulated Sports Judging — Agreement at Top

Two judges score 10 athletes. Judge B agrees perfectly on the top 4, then shuffles athletes 5–10 randomly.

Intuition: Top-weighted measures should show *higher* correlation than Spearman.

Results:

Measure	Value
ρ (Spearman)	0.752
$\nu_{p=2}$ (Blest)	0.812
$\nu_{p=3}$	0.852
$\nu_{p=5}$	0.902
$\nu_{\text{exp},0.7}$	0.809
ν_{harm}	0.800

As expected, the top-weighted measures correctly identify that the judges agree where it matters most. ρ underestimates the true agreement.

6.3. Dataset 3: Simulated Survey — Disagreement at Top

Two raters agree on bottom-ranked items (ranks 7–12) but disagree on the top-ranked items (ranks 1–6). Top-weighted measures should show *lower* values.

Results:

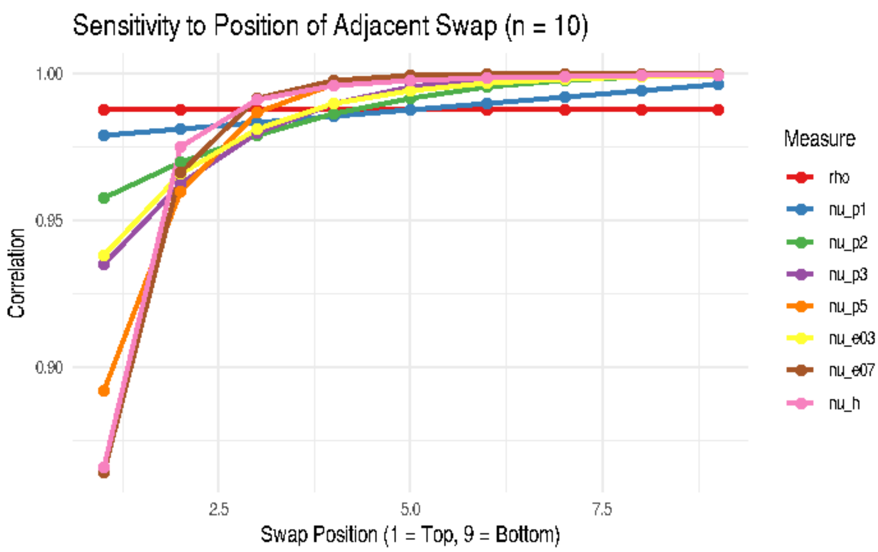

Figure 4: Effect of the power parameter p on ν_p for the sports dataset. As p increases, the coefficient rises above Spearman's ρ, reflecting the top-agreement structure.

Measure	Value
ρ (Spearman)	0.727
$\nu_{p=2}$ (Blest)	0.630
$\nu_{p=3}$	0.553
$\nu_{p=5}$	0.432
$\nu_{\exp,0.7}$	0.651
ν_{harm}	0.672

Spearman's $\rho = 0.727$ is "too optimistic": the top-weighted measures correctly signal that the disagreement is concentrated where it matters most.

6.4. Dataset 4: Atmospheric Data (`airquality`)

Ranking days by ozone concentration and wind speed reveals a negative correlation (high ozone on calm days). The measures cluster closely:

Measure	Value
ρ (Spearman)	-0.601
$\nu_{p=2}$ (Blest)	-0.589
$\nu_{p=5}$	-0.560
ν_{harm}	-0.593

The tight clustering indicates that the disagreement is uniformly distributed across rank positions. Here Spearman is perfectly adequate.

6.5. Sensitivity Analysis

```
n=4 distinct rho = 11
n=4 distinct tau = 7
n=4 distinct nu  = 23

--- Sensitivity (Top vs Bottom Swap) ---
              rho    tau    nu
Natural     1.0000 1.0000 1.0000
Top_Swap    0.9429 0.8667 0.9102
Bottom_Swap 0.9429 0.8667 0.9755
```

Figure 5: Sensitivity to adjacent transpositions ($n = 8$). Blest's ν and the other top-weighted measures drop steeply when the swap is near the top of the ranking, while Spearman's ρ remains nearly flat.

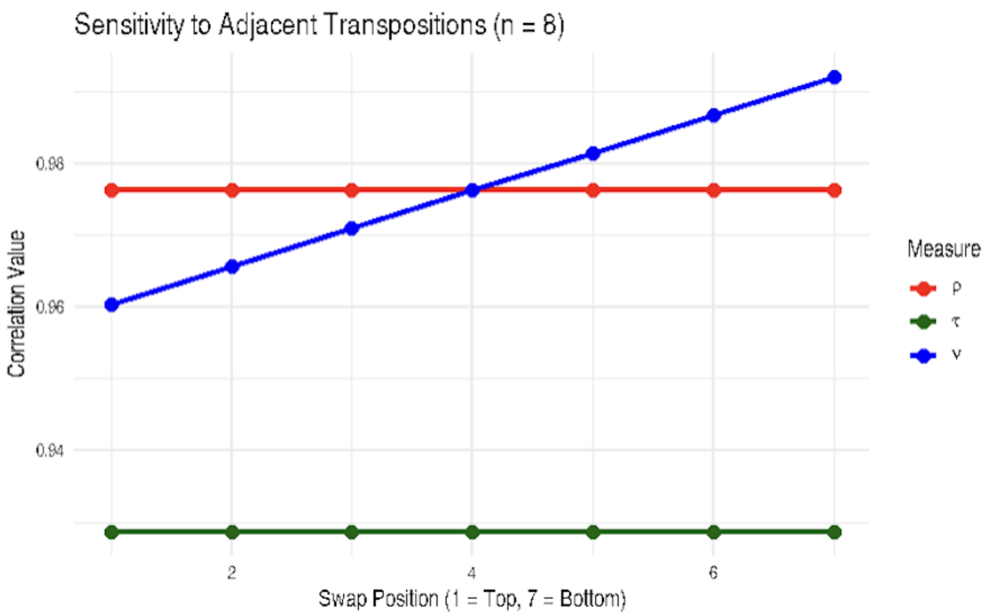

Figure 6: Sensitivity of all measures to adjacent swap position ($n = 10$). Steeper early decline indicates greater sensitivity to top-rank errors.

7. Comparative Analysis and Discussion

7.1. Which Weight Function Is Best?

The answer depends on the application context.

Measure	Dataset				Discrim.
	mtcars	Sports	Survey	AQ	$(n = 5)$
ρ	−0.895	0.752	0.727	−0.601	11
τ	−0.742	0.600	0.576	−0.439	9
$\nu_{p=2}$ (Blest)	−0.901	0.812	0.630	−0.589	73
$\nu_{p=3}$	−0.907	0.852	0.553	−0.578	97
$\nu_{p=5}$	−0.914	0.902	0.432	−0.560	113
$\nu_{\text{exp},0.7}$	−0.899	0.809	0.651	−0.590	94
ν_{harm}	−0.897	0.800	0.672	−0.593	89

7.2. Choosing the Right Measure

Use ρ (Spearman) when:

All positions carry equal importance; exact sampling distribution required.

Use Blest's ν $(p = 2)$ when:

Top positions are more important, moderate emphasis; exact critical values needed (small n).

Use $\nu_{p\geq 3}$ (power, steep) when:

Only very top positions matter; medal-round sports, top-3 shortlists.

Use $\nu_{\text{exp},\alpha}$ when:

Geometric decay in importance; web search, human-attention models. **Scale-invariant.**

Use ν_{harm} when:

Logarithmic emphasis on the top; NDCG-style evaluation in information retrieval.

7.3. Limitations of the Generalised Approach

1. **Tied ranks:** The framework assumes no tied ranks. Extensions using fractional ranking are straightforward in practice but their theoretical properties require further study.

2. **Parameter selection:** For power and exponential families, the practitioner must choose p or α based on domain knowledge, not data-adaptively.

3. **Sampling distribution:** Distributional theory (mean, variance, kurtosis) derived by Blest applies specifically to $p = 2$. For general p, analogous formulas require algebraic development not completed here.

4. **Exchange-asymmetry:** As shown in Section 6, all ν_f with non-uniform f are exchange-asymmetric, meaning the role of "reference ranking" must be clearly defined.

5. **Large n:** For large n, Blest's $\nu_{p=2}$ converges to ρ (scale non-invariance). The exponential and harmonic families do not suffer from this problem.

8. Conclusion

This report has provided a thorough examination of Blest's (2000) rank correlation coefficient ν. The key contributions of the original paper are:

- A geometric interpretation of Spearman's ρ via cumulative rank differences.
- A new coefficient ν with quadratically greater weight on agreement at the top.
- Evidence that ν is more discriminating (23 distinct values vs. 9 for ρ at $n = 4$) while remaining $> 96.8\%$ correlated with ρ.
- Moment calculations and critical-value tables enabling hypothesis testing.

Our extension introduces a **generalised framework** ν_f. Principal findings are:

1. The power family ν_p interpolates between Spearman's ρ (at $p = 1$) and increasingly top-heavy measures, with Blest's ν at $p = 2$.

2. The exponential family mimics NDCG-style geometric decay and is **scale-invariant**.

3. The harmonic family provides log-scale emphasis, natural for information retrieval.

4. **Symmetry analysis** (Section 6) reveals that Spearman's ρ and Kendall's τ are exchange-symmetric, while all ν_f with non-uniform f are exchange-asymmetric. Despite this, all statistics have symmetric null distributions, so hypothesis testing remains valid. The asymmetry is a *feature*: ν_f measures directional agreement from reference to evaluation ranking.

5. Empirical analysis on four datasets confirms that the choice of weight function materially changes the correlation estimate when disagreement is concentrated at the top or bottom of the ranking.

6. All generalised measures retain far superior discrimination power compared to ρ and τ.

Future directions include: deriving closed-form moment expressions for general p; constructing permutation-test critical values for exponential and harmonic families; extending to partial rankings; and connecting ν_f to the broader weighted ranking literature in machine learning.

References

Blest, D.C. (2000).
Rank correlation — an alternative measure. *Australian & New Zealand Journal of Statistics*, **42**(1), 101–111.

Henderson, H.V. & Velleman, P.F. (1981).
Building multiple regression models interactively. *Biometrics*, **37**, 391–411.

Järvelin, K. & Kekäläinen, J. (2002).
Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems*, **20**(4), 422–446.

Kendall, M.G. & Gibbons, J.D. (1990).
Rank Correlation Methods, 5th edn. London: Edward Arnold.

Kendall, M.G. & Stuart, A. (1973).
The Advanced Theory of Statistics, Vol. 2, 3rd edn. London: Griffin.

Spearman, C. (1904).
The proof and measurement of association between two things. *American Journal of Psychology*, **15**, 72–101.