SPUR Score Verification Protocol for Peer Reviewers

You are verifying an author's self-assessed SPUR (Scientific Paper Uniqueness Ranking) score for a submitted manuscript. Your role is to validate the claimed score, check for gaming, and ensure proper application of the framework's three critical enhancements.

Manuscript Information Needed

1. **Author's claimed SPUR score**: [Insert score]
2. **Author's dimensional breakdown**: [Insert 7 scores]
3. **Research field**: [Insert primary/secondary discipline]
4. **Author's field baseline used**: [Insert baseline mean]
5. **Manuscript title and abstract**: [Paste here]
6. **Author's SPUR justification**: [Paste their explanation]

Verification Steps

STEP 1: Field-Specific Baseline Verification

Check if author used appropriate baseline for their discipline:

Expected baselines (mean SPUR scores for "typical" papers):
- Biomedical Sciences: 45-50/100
- Physical Sciences: 42-48/100
- Social Sciences: 48-52/100
- Mathematics: 40-45/100
- Humanities: 38-45/100

Verification questions:
1. Did the author correctly identify their primary research field?
2. Is the baseline they used within the expected range?
3. For interdisciplinary work, did they justify baseline selection?
4. If they used field normalization (z-score), was the calculation correct?

Red flags:
- Using baseline from wrong discipline to inflate percentile ranking
- Cherry-picking lowest baseline to appear more unique
- Ignoring field context entirely

STEP 2: Conditional Weighting Verification

Check if conditional weighting was properly applied:

Rule: IF Methodological Innovation score ≥ 80/100, THEN:

- Methodological Innovation weight: 20% → 25%
- Replicability & Transparency weight: 10% → 5%

Rationale: Highly innovative methods are inherently harder to replicate, so the framework automatically reduces replicability penalties for breakthrough methodologies.

Verification questions:
1. What is the author's Methodological Innovation score? [Check]
2. If ≥80, did they apply 25% weight to Method Innovation? [Yes/No]
3. If ≥80, did they apply 5% weight to Replicability? [Yes/No]
4. If <80, did they use standard weights (20% and 10%)? [Yes/No]
5. Is the conditional trigger justified by truly innovative methodology?

Red flags:
- Inflating Method Innovation score to 80+ just to trigger weighting shift
- Applying conditional weights when Method Innovation <80
- Not applying conditional weights when Method Innovation legitimately ≥80

STEP 3: Replicability Barrier Assessment

Check if replicability scoring distinguishes transparency from practical barriers:

The framework now recognizes two types of replication challenges:

Type A: Transparency Issues (penalized normally)
- Missing data
- Inadequate method descriptions
- Proprietary tools without alternatives
- Selective reporting
- **These should lower replicability scores**

Type B: Legitimate Barriers (not penalized when Method Innovation ≥80)
- Novel equipment not yet widely available
- Cutting-edge techniques requiring specialized training
- Unique observational opportunities (eclipses, rare events)
- Computational requirements beyond standard resources
- **These trigger conditional weighting protection**

Verification questions:
1. Did the author distinguish between transparency and legitimate barriers?
2. If claiming high Method Innovation, are replication barriers justified?
3. Are there any transparency issues being hidden behind "innovation" claims?
4. Is the replicability score reasonable given the methodology?

Red flags:
- Claiming "innovative so can't replicate" to hide poor documentation

- Conflating legitimate barriers with sloppy methodology
- Over-penalizing genuinely innovative work for inherent barriers

---

STEP 4: Calculate Independent SPUR Score

Using the author's manuscript, score each dimension independently:

1. Methodological Innovation (0-100): [Your score]
- Novel techniques, tools, or approaches
- Advancement beyond current methods
- Transferability to other research

2. Conceptual Originality (0-100): [Your score]
- New theoretical frameworks
- Unexpected connections
- Paradigm-challenging insights

3. Empirical Scope & Scale (0-100): [Your score]
- Data comprehensiveness
- Temporal/spatial coverage
- Population/sample representativeness

4. Societal Impact Potential (0-100): [Your score]
- Policy relevance
- Economic implications
- Social/environmental benefit
- Stakeholder reach

5. Cross-Disciplinary Integration (0-100): [Your score]
- Methods from multiple fields
- Synthetic theoretical frameworks
- Bridges between disciplines

6. Replicability & Transparency (0-100): [Your score]
- **Assess transparency issues** (data, code, methods)
- **Note legitimate barriers separately**
- Consider whether conditional weighting applies

7. Theoretical Advancement (0-100): [Your score]
- New predictions or explanations
- Generalizable principles
- Theoretical coherence

STEP 5: Apply SPUR Formula with Conditional Logic

Calculate base score with proper weights:

```
IF Methodological_Innovation ≥ 80:
    Method_Weight = 0.25
    Replicability_Weight = 0.05
ELSE:
    Method_Weight = 0.20
    Replicability_Weight = 0.10

Base_SPUR = (Methodological × Method_Weight) +
        (Conceptual × 0.18) +
        (Empirical × 0.15) +
        (Societal × 0.15) +
        (CrossDisciplinary × 0.12) +
        (Replicability × Replicability_Weight) +
        (Theoretical × 0.10)
```

Apply impact multiplier:

```
Multiplier = 1 + (0.3 × Societal_Impact / 100)
Final_SPUR = Base_SPUR × Multiplier
```

Your calculated SPUR score: [Insert]

STEP 6: Gaming Detection Checks

Examine the manuscript and author's justification for these gaming patterns:

1. Vocabulary Manipulation Gaming
- Using jargon complexity to appear novel
- **Check**: Is complexity justified by actual innovation?

2. Method Combination Gaming
- Combining standard methods and claiming innovation
- **Check**: Is synthesis genuinely novel or just procedural?

3. False Interdisciplinary Claims
- Superficial citations across fields
- **Check**: Are methods truly integrated or just referenced?

4. Impact Exaggeration

- Overstating policy/societal relevance
- **Check**: Is impact plausible and well-justified?

5. Complexity Obfuscation
- Obscuring simple methods to appear sophisticated
- **Check**: Could methods be described more simply?

6. Baseline Shopping
- Selecting field baseline to maximize percentile
- **Check**: Is discipline classification accurate?

7. Conditional Weighting Gaming
- Inflating Method Innovation to ≥80 just for weight shift
- **Check**: Is ≥80 score genuinely warranted?

STEP 7: Comparison and Recommendation

Score Comparison:
- Author's SPUR: [X]
- Your verified SPUR: [Y]
- Difference: [Y - X]

Acceptable differences: ±5 points (subjective interpretation range)

Recommendation:

ACCEPT if:
- Difference ≤5 points
- No gaming indicators detected
- Field baseline appropriate
- Conditional weighting correctly applied
- Replicability barriers properly assessed

ADJUST if:
- Difference 6-10 points
- Minor gaming indicators (vocabulary, impact slight exaggeration)
- Baseline defensible but suboptimal
- Conditional logic applied but marginal (Method Innovation 78-82 range)
- Suggest revised score with justification

REJECT if:
- Difference >10 points
- Clear gaming detected (false interdisciplinary, baseline shopping)
- Wrong baseline intentionally used
- Conditional weighting misapplied
- Transparency issues hidden behind innovation claims

- Require author to reassess with your feedback

STEP 8: Provide Feedback to Author

Your verification report should include:

1. **Verified SPUR score** (accept author's, or provide adjusted score)
2. **Dimensional score adjustments** (which dimensions changed and why)
3. **Field baseline confirmation** (correct or needs revision)
4. **Conditional weighting check** (properly applied or not)
5. **Replicability assessment** (transparency vs. barriers clarification)
6. **Gaming concerns** (if any detected)
7. **Justification** (2-3 sentences explaining verification outcome)

Example feedback format:

> "The author's claimed SPUR score of 87.3 has been verified with minor adjustment to 84.6. The Methodological Innovation score of 82 appropriately triggers conditional weighting (25%/5%), and the replication barriers are legitimate given the novel neuroimaging technique. However, the Societal Impact score was reduced from 88 to 78 as the policy applications, while promising, are not yet empirically demonstrated. The field baseline for Biomedical Sciences (47.5) is appropriate. No gaming indicators detected.
> **Recommendation: ACCEPT with adjusted score of 84.6/100 (93rd percentile).**"