Shipley Webinar
Top Weaknesses That Drive Lower Proposal Ratings and Mitigation Measures

Presenter: Nia Mahallati
Senior Vice President, Shipley Associates

Agenda

- Government Evaluation and Proposal Ratings (2024–2025 Trends)
- Understanding How Evaluation Criteria Drive Proposal Scoring Outcomes
- FAR Requirements Driving Evaluation Scores
- Where Low Evaluation Score Risks Actually Begin (BD → Capture → Pricing → Proposal)
- Top Drivers of Low Evaluation Scores & Why They Occur
- Protests and Case Law Review: Kauffman, Criterion, WBD
- What Can Help Prevent Low Evaluation Scores (Corrective Actions Missed)
- Data, Patterns, and Trends from Shipley's Internal Analytics (2023–2025)
- Mitigation Measures Across BD, Capture, Pricing & Proposal Development to Increase Evaluation Scores
- Tools & Techniques to Strengthen Evaluation Scores

A Little About Shipley

A Global Training (40 offices worldwide), Consulting, and Professional Services firm established in 1972

- Business development, sales, capture management, proposal management, proposal writing, and price to win public and corporate training
- Full life-cycle BD, capture, sales, and proposal management and writing consulting services
- Business winning process improvement
- Training (averaging 15 courses per month), certification, and career development
  - Free webinars, industry insights, blogs, business winning tips, guidebooks, playbooks

To learn more visit us at _www.shipleywins.com_

Shipley's data is drawn from thousands of evaluator comments, score sheets, and debrief findings across federal agencies.

- Thousands of proposals supported annually across DoD, FedCiv, IC, State, DHS, HHS

- Direct insight into evaluator comments, strengths, weaknesses, BOE mismatches

- Internal analytics (2023–2025) reveal quantification's impact:

  — Quantified claims → +61% green or blue ratings

  — Narrative generalities → 72.3% of all documented weaknesses

  — Missing traceability (RFP Requirements → Your Approach → Features/Methods → Evidence/Proof →)= 3.4× lower scores and protest vulnerability

  — Subfactor scores below 70% reliably predict GAO later confirming weaknesses in those same areas

Strengthening Capture, Pricing, and Proposal Development to Improve Evaluation Scores

Strengthening Capture, Pricing, and Proposal Development to Improve Evaluation Scores

- Upstream BD, Capture, Pricing, and Proposal decisions shape evaluation scores
- Quantification, traceability, and BOE alignment amount to higher evaluation scores

Why This Matters

- Many proposals **score lower due to weak linkage between claims and evaluation criteria**
- High evaluation scores reflect early positioning, focused BD, intentional strategy, and strong capture aligned with pricing and proposal execution
- Low scores commonly result from inadequate substantiation of strengths and discriminators
- Contractors directly influence evaluation outcomes through how clearly proposals present evidence

FAR Evaluation Standards and Protest Sustainments

- FAR 15.305(a)(2): Agencies must evaluate solely on stated factors/**subfactors**

- FAR 15.308: SSA must document rationale for tradeoffs and decisions

- FAR 15.404-1(d): Cost realism requires alignment of proposed cost with technical approach

- Most GAO protest sustainments cite failure to document evaluation logic at **subfactor level (prioritize subfactors & reqs. through requirements prioritization and/or evaluation factor analysis)**

Government Adjectival/Color Bands (Illustrative 2025 Format)

Adjectival	Color	Score Range (Example)	Summary Meaning
Unsatisfactory	Red	0–39 / 100	Fails to meet requirements; high performance risk
Marginal	Yellow	40–69 / 100	Partially meets requirements; significant performance risk
Acceptable	Green	70–89 / 100	Meets requirements; low to moderate risk
Outstanding (Far Exceeds)	Blue	90–100 / 100	Exceeds requirements; very low performance risk

Illustrative Government Evaluation Scorecard (Roll-Up View)

Factor	Weight	Adjectival	Color	Numeric Score
Technical	40%	Acceptable	Green	82 / 100
Past Performance	25%	Substantial Confidence	Blue	90 / 100
Cost Realism	20%	Reasonable	Green	78 / 100
Management	15%	Marginal	Yellow	62 / 100

Performance Confidence Ratings (Gov. also uses Relevancy Ratings)

Satisfactory Confidence	Based on the offeror's recent/relevant performance record, the Government has a reasonable expectation that the offeror will successfully perform the required effort.
Neutral Confidence	No recent/relevant performance record is available or the offeror's performance record is so sparse that no meaningful confidence assessment rating can be reasonably assigned. The offeror may not be evaluated favorably or unfavorably on the factor of past performance
Limited Confidence	Based on the offeror's recent/relevant performance record, the Government has a low expectation that the offeror will successfully perform the required effort.
No Confidence	Based on the offeror's recent/relevant performance record, the Government has no expectation that the offeror will be able to successfully perform the required effort.

Why does the Government Place so Much Weight on Scoring? Evaluation Weaknesses Drive Sustained Protests (2023–2025)

- Evaluation Weaknesses Driving Sustained Protests (2023–2025)
- Unreasonable Technical Evaluation (unsupported or inconsistent evaluation scores)
- Flawed or undocumented Best-Value Tradeoff
- Unreasonable Price Realism Evaluation
- These evaluation failures account for the majority of sustained protests over the past three years

Where Protest Risk Originates

- BD: Misinterpreting customer mission drivers and evaluation priorities
- Capture: Failure to develop measurable discriminators or quantified benefits
- Pricing: BOE disconnected from technical tasks and staffing justification
- Proposal Management: Missing L/M traceability → evaluator confusion
- Proposal Writing: Qualitative claims (subjective/adjectives) instead of quantified, verifiable evidence

Example: Kauffman (HHS- SB Training Contract- 2024): What Happened

Citation: B-421917.2, .3 (2024).

GAO: "The record does not demonstrate the awardee met the requirement."

- Issues:
  - Agency accepted general statements about data protection with no measurable details.
  - Unequal treatment: protester penalized for omissions awardee also had.
  - Missing subfactor-level justification.
  - Evaluation record lacked evidence to support the award decision.

Case: Preventing the Failure

- Provide measurable cybersecurity evidence:
  - AES-256 encryption, MFA, RPO/RTO metrics, SOC-II findings.
- Develop requirement-point justification paragraphs (1 per requirement).
- Use full traceability matrix mapping Section L → Section M → Evidence.
- Ensure strengths are explicitly labeled and evaluator-visible.
- Eliminate narrative generalities before submission.

11

Internal Findings (2023–2025)

Analysis of thousands of proposals for past 2 years (2023-2025):

- +61% higher odds of Green ratings when evidence is quantified
- 72.3% of technical weaknesses stem from narrative generalities
- Traceability gaps → 3.4× increased protest vulnerability
- Subfactor scores under 70% → strongly correlated with GAO sustain findings

Business Development (BD)

- Objective: Identify measurable customer outcomes & evaluation drivers.
- BD Outcome Mapping Chart
- + Pain Point → Required Metric → Evidence
- Example: Backlog → 38% throughput increase needed → Past automation data
- BD identifies measurable outcomes driving all downstream proposal sections

Capture

- Objective: Convert requirements into quantifiable discriminators
- Requirements → Metrics Conversion
- Data Protection: AES-256, RPO 4 hrs, 0 audit findings.
- Capture Risk Register (0–5 index)
- **"Risk" means evaluators view this area as high impact if done poorly, so it requires stronger proof, benchmarks, or evidence to earn strengths and avoid weaknesses.**
- High-risk subfactors must have evidence
- Example: Data Migration risk score high → require error-rate benchmarks

Pricing

- Objective: Align BOE, staffing & technical workflow
- BOE ↔ Technical Alignment Matrix ties hours to workflow steps
- Cost Realism Narrative explains efficiencies with measurable data
- Example: Automation reduces review time 60% → supports lower hours
- Volumes must reconcile exactly across tech, staffing, and BOE

Proposal Management

- Objective: Make evaluation effortless for reviewers
- Full L→M→Page→Evidence traceability matrix required
- Gov-style color scoring (1–10); <7 requires rewrite
- Example: Staffing Plan scoring 6 → add FTE model, ramp timeline, cert matrix

Proposal Writing

- **Objective: Replace adjectives with quantifiable evidence**
- **Metric → Impact → Value strength model**
- Generalities → Metrics conversion table
- Example rewrite: "Automation improves quality" → "22% faster, 31% fewer defects."

Additional Tools

Capture Risk Register

- Risk: Data migration failure → Score 4/5
- Probability × Impact → Risk priority matrix
- Mitigation: Automated validation → reduces error rate 32%
- Impact: Avoids 2-week schedule delay
- Add quantifiable mitigations to reduce evaluation ambiguity

Requirements → Metrics Conversion

- Data protection → AES-256, SC-28 alignment, RPO=4 hrs
- Staffing → Ramp time, FTE levels, certification rates
- Quality → Defect %, audit results, rework reduction metrics
- Cost → Cost avoidance %, variance history, productivity figures

Discriminator Development Worksheet

- Problem → Method → Metric → Impact → Value.
- Example:
  — Backlog accumulation → automation → 38% faster → clears backlog → mission readiness
  — Discriminators must be quantifiable and evaluator-visible

BOE Alignment Matrix

- Workflow Step | Hours | LCAT | Rationale | Technical Link
- Example:
  — Data validation | 420 hrs | Analyst II | Automation reduces manual effort | Workflow Step 3
  — Ensures realism and evaluation defensibility

Traceability Matrix

- Req ID | L Citation | M Factor | Proposal Page | Evidence | Metric
- Evaluators must see exactly where compliance is demonstrated
- Prevents missed requirements and strengthens evaluation record

Color-Team Evaluation Grid

- Gov-style scoring:
- 1–3 Red | 4–6 Yellow | 7–8 Green | 9–10 Blue
- <7 requires mandatory rewrite
- Aligns proposal with government evaluation logic and reduces risk

Generalities → Metrics Conversion

- 'Robust experience' → '97.8% accuracy across 40,000 cases.'
- 'Strong team' → '100% staff certified Level II.'
- 'High quality' → '<1% defect rate across 3-year audit history.'
- **Metrics improve evaluator understanding and defensibility.**

Webinar Summary

- **Quantification strengthens higher evaluation scores**

- **Traceability creates a defensible evaluation record**

- **Alignment across capture, pricing, and proposal phases strengthen proposals**

- **Strong proposals reduce uncertainty/risk → reduce protests → increase awards**

Q&A Session

View Our

Public Training

Calendar

Contact Us

About Anything

Else

Can't scan these? The same links are also available in the chat now.

Email Nia at nmahallati@shipleywins.com or search "*Nia Mahallati*" in LinkedIn to connect