

*Maintained by: HR Director, People & Culture Last updated: September 2025 Classification: Internal Review cadence: Annual (next review due September 2026 — overdue)*

## 1. Purpose

This document defines the bias categories and review criteria used when assessing HR artifacts for biased language, subjective judgements, and exclusionary patterns. It applies to performance reviews, promotion panel documentation, compensation rationale, and policy drafts.

## 2. Bias Categories

### 2.1 Gendered Language

Language that implicitly associates a behaviour, trait, or expectation with a specific gender. Includes gendered pronouns used generically, adjectives disproportionately applied to one gender in workplace research (e.g., "nurturing", "aggressive", "abrasive"), and role assumptions that imply gender norms.

**Flagging threshold:** Any instance of gendered language in an evaluative statement about an employee's capability, behaviour, or potential.

### 2.2 Subjective Qualifier

Evaluative language that expresses a judgement without supporting evidence or measurable criteria. Examples include "not a culture fit", "lacks gravitas", "shows potential", "strong leadership presence". Subjective qualifiers are acceptable only when accompanied by a specific behavioural example or measurable outcome.

**Flagging threshold:** Subjective qualifiers used in any evaluative or decision-influencing passage without an accompanying evidence citation.

### 2.3 Cultural Bias

Language or standards that assume a single cultural norm as the default. Includes assumptions about communication styles (e.g., equating directness with competence, equating reserved communication with disengagement), social norms, and professional conventions that vary across cultures. This category also covers post-acquisition integration language that privileges legacy Northbridge practices without acknowledging acquired-firm norms.

**Flagging threshold:** Any passage where a cultural norm is presented as a universal standard without qualification, or where employees from acquired firms are described using deficit-framing language.

Note: This criterion was expanded in January 2026 to include post-acquisition language following the Meridian integration. The expanded scope has not been tested against existing artifacts.

### 2.4 Exclusionary Term

Terminology that explicitly or implicitly excludes a protected group or creates an unwelcoming environment. Includes ableist language (e.g., "tone-deaf"), age-related assumptions, and terms with documented exclusionary connotations in professional settings.

**Flagging threshold:** Any instance in an HR artifact, regardless of context.

### 2.5 Vague Justification

Decision rationale that lacks specificity, making it impossible to verify whether the decision was applied consistently and equitably. Includes phrases like "not ready", "needs more seasoning", "use their judgement", and circular reasoning where the conclusion restates the premise.

**Flagging threshold:** Any decision rationale in a promotion panel outcome, compensation recommendation, or policy enforcement explanation that does not cite a specific policy clause, metric, or documented behaviour.

Note: There is acknowledged overlap between Vague Justification and Subjective Qualifier. The distinction is that Subjective Qualifier applies to descriptive language about an individual, while Vague Justification applies to decision rationale. In practice, reviewers have reported difficulty distinguishing the two when a subjective description is used as the basis for a decision.

---

### 3. Severity Framework

Severity	Definition
Critical	Language or justification that directly influences an employment decision (hiring, promotion, compensation, termination) and lacks supporting evidence. Requires immediate correction before the artifact can be approved.
Major	Language that introduces bias risk into a governance artifact but does not directly determine an employment decision in its current context. Requires correction before the next review cycle.
Minor	Language that could be improved for inclusivity or clarity but does not materially affect decision quality. Correction recommended but not blocking.
Informational	Observation about language patterns that may warrant attention in future review cycles. No correction required.

Note: The severity framework references the legacy Mercer grade structure (M1–M5) in its original appendix for mapping severity to employee grade level. Since Northbridge transitioned to a Level 1–5 banding structure in Q2 2025, the grade-level severity mapping has not been updated. Reviewers should apply severity classifications based on the definitions above without reference to grade-level modifiers until the appendix is revised.

---

### 4. Review Scope

The following artifact types are in scope for bias review:

1. Performance review narratives — all goal-level narratives and overall summary
2. Promotion panel outcomes — evidence summaries, panel rationale, and next steps
3. Compensation recommendation rationale — individual adjustment rationale and methodology notes
4. Policy draft redlines — revised language and rationale statements

Manager feedback forms and self-assessments are excluded from direct bias review but may be cited as source material when evaluating the traceability of claims in the above artifacts.

---

### 5. Review Process

1. Assemble all in-scope artifacts for the review period
2. Scan each artifact individually for bias category matches
3. Cross-reference findings across artifacts for systemic patterns
4. Classify each finding by severity
5. Draft corrections and submit for governance approval

The review should be completed within 10 working days of receiving all artifacts. If the review cannot be completed within this window, notify the HR Director.

---

AGASI AiOS — Sample material. For illustrative purposes. All data is fictional.