

# AI Evaluation Bridge – Dependencies and Weighting

## Dependencies

Some evaluation criteria depend on other criteria being correct first. For example, reasoning cannot be evaluated properly if factual extraction is incorrect.

## Weighting

Different evaluation criteria carry different levels of importance.

## Example Weighting

- Factual Accuracy — 40%
- Reasoning Quality — 30%
- Instruction Compliance — 20%
- Formatting — 10%

## Why Weighting Matters

Weighting ensures that the most important aspects of the response influence the final evaluation score.