

RATIO · WHITE PAPER · SEPTEMBER 2026

The RATE Framework

One scoring rubric for every interview question

RATE is a four-part rubric for scoring an interview answer: Relevance, Approach, Trade-offs, Evidence. This paper explains why it exists, what each part measures, the research it draws on, and how one rubric holds up across every question type and every level of a role.

Who it is for

HR and talent leaders choosing a scoring method, and hiring managers who want to know what they are being asked to rate.

What's inside

Eight parts, one scored answer, a framework comparison table, five questions and answers, and references.

Written by

Michael Day, Co-Founder and CEO. 17 years in agency recruiting at Robert Half, Adecco, and Frank Recruitment Group.

PART 1

The scoring problem

Most interviews are unstructured conversations. The interviewer asks whatever comes to mind, judges the answers on feel, and the hiring decision rests on impressions nobody wrote down. The research is clear about what that costs. The largest meta-analysis of selection methods to date puts the validity of an unstructured interview at .19 and a structured one at .42, and ranks the structured interview first among every method it examined.¹

Structure has two halves: the same questions for every candidate, and the same scoring of every answer. Most teams that call their interviews structured have the first half. Few have the second, and the second is where most of the validity comes from. The reason is practical. Structured interviewing, as usually taught, asks the interviewer to hold a different scoring method for each kind of question.

Behavioral questions call for STAR or SOAR (Situation, Task or Obstacle, Action, Result). Hard-skill questions call for Webb's Depth of Knowledge, which tracks cognitive demand from recall up to strategic thinking.² Situational questions call for Bransford and Stein's IDEAL problem-solving model: Identify, Define, Explore, Act, Look.³ Each is well established. Each was designed for a different purpose. Using all three in one interview loop means an interviewer has to remember which applies to which question, hold three sets of criteria in mind while listening, and somehow rate consistently across all of it.

Most teams give up. They fall back to conversation, or apply structure so unevenly that the benefit disappears. RATE exists to make the scoring half practical. It starts from the observation that the three frameworks are measuring the same underlying things with different vocabulary, and it names those things once. Interviewers learn one rubric. The rigor of the research stays intact.

PART 2

What RATE stands for

Each letter names one thing a complete answer contains. The four definitions are fixed. They apply to every question type, every role and every proficiency level. What changes is the question, never the lenses.

R • Relevance

Specific, grounded detail that anchors the answer to the question. Did the candidate cite a real example, name concrete tools or decisions, or correctly diagnose the problem in front of them? An answer without Relevance is a generality: "I always communicate proactively."

A • Approach

The right method, with personal ownership. Did the candidate describe how they would actually do the work, in the first person, about decisions they personally made? "I decided" and "I built," not "the team handled it."

T • Trade-offs

What made it hard. Did the candidate name what they gave up, what constrained them, or which alternative they rejected? This is the lens that separates a senior answer from a competent one, and Part 6 is about why.

E • Evidence

Outcome or learning shown. Did the candidate explain what actually happened, how they know it worked, or what they would do differently next time? A specific, partly negative outcome usually reads as more true than a clean success.

PART 3

The research RATE draws on

RATE does not invent the underlying ideas. It maps onto three established frameworks that arrived independently at similar structures, and it takes its level calibration from a fourth.

WHAT IS MEASURED	WEBB'S DOK (HARD SKILLS)	SOAR (BEHAVIORAL)	IDEAL (SITUATIONAL)	RATE
Recall and application	Levels 1-2	Situation	Identify	Relevance
Procedural skill	Level 2	Action	Act	Approach
Strategic thinking	Level 3	Obstacle	Explore	Trade-offs
Outcomes and learning	Levels 3-4	Result	Look	Evidence

The convergence is not a coincidence. The frameworks were developed across decades for different settings, but they are all trying to describe cognitive and behavioral competence, and competence has a shape: you know the specifics, you can execute the method, you can handle what makes it hard, and you can show it worked. RATE makes that shape explicit and usable in a single interview.

The logic that adjusts question difficulty by proficiency comes from the Dreyfus model of skill acquisition, which describes how practitioners move from rule-following novices to context-sensitive experts.⁴ Part 5 covers how that shows up in the questions.

PART 4

One rubric, three question types

Structured interviews use three kinds of question. RATE applies to all three through two fixed sets of lead-ins: one for applied, in-the-moment reasoning, and one for real past experience.

Applied Skills & Knowledge (ASK) questions

Test whether the candidate can do the work. Usually present tense, built around a problem to solve. Lead-ins: give specific detail, explain their approach, discuss the trade-off, describe what good looks like.

Behavioral questions

Ask for a real past experience. The most studied question type for predicting future performance.⁵ Lead-ins: provide a real example, explain what they did, discuss the trade-off, share the outcome.

Situational questions

Put the candidate in a realistic scenario and ask how they would handle it. Useful where direct experience cannot be assumed. Because they test applied reasoning, they use the same lead-ins as ASK questions.

The four letters never change. Behavioral questions use the past-experience set; ASK and Situational questions use the applied set. The interviewer listens for one pattern in every question, which is what lets someone who has run three of these interviews run the fourth without retraining.

PART 5

The same lenses at every level

A junior engineer and a senior engineer can both answer a technical question correctly, and the interviewer should expect different things from each. The junior should show they can do the work. The senior should show they can work through complexity, weigh trade-offs and check the result. If the criteria do not scale, the bar is either unfair to juniors or too low for seniors.

RATE scales by changing the question, not the lenses. Ratio sets a proficiency level for each skill from the job description, and the question gets harder with the level.

Beginner. Can perform with guidance. Junior roles, or a skill that is nice to have.

Intermediate. Works independently. The standard expectation for most mid-level requirements.

Advanced. Handles the hard cases and owns the outcome. Senior individual contributors.

In the screening round, the level shapes the register of the question so a generalist recruiter can ask it naturally. In the hiring-manager round, higher levels bring more ambiguity, consequence and judgment into the scenario. The four lenses stay the same, and the interviewer rates against the level the role needs, not against the other candidates.

PART 6

Why Trade-offs does the most work

Junior professionals solve the problem. Senior professionals solve the problem while managing technical debt, team dynamics, timeline pressure or a business constraint they did not choose. That difference is what the T lens is built to bring out, and it is why we chose the word. Trade-offs shows up in every kind of question: technical trade-offs (speed against accuracy, consistency against availability), interpersonal ones

(stakeholder conflict, competing priorities), and cognitive ones (acting on incomplete information). An interview that never asks for it misses the signal that most often separates senior performance from competent performance.

Coached and AI-assisted answers

Trade-offs is also where a rehearsed answer works hardest. AI tools and over-prepared candidates tend to produce polished, balanced answers that keep every option open. The T lens asks for a committed call: pick a side, name what you would give up, say why. Fence-sitting scores low on that line. Lived experience gives a candidate a reason to commit; a prepared answer usually does not.

Three parts of the design carry that work.

The Trade-offs lens itself, under every question in every round, with a prepared follow-up that asks for the specific detail a real experience carries: "What did that cost you?"

A specificity floor on Relevance in the hiring-manager round. The R anchor names what an empty answer sounds like: "Give specific detail, not just 'check the logs.'" A, T and E stay as bare lead-ins. Screening rounds keep all four bare so the round stays light.

An operational constraint in the question, for the most critical skill in the hiring-manager round. Instead of "How do you handle month-end close?", the question is "The warehouse hasn't finished physical counts because of staffing shortages. The board report is due Friday. How do you handle the missing data?" The constraint has to be one the role would genuinely meet, never a gotcha that blocks a legitimate answer. It removes the option of reciting standard procedure.

None of this guarantees that coaching or AI assistance is detected, and Ratio does not claim it does. A capable candidate with a capable tool can produce a passable answer to most questions. What RATE gives the interviewer is a concrete benchmark for telling lived expertise from prepared performance, without turning the interview into a trivia test.

PART 7

RATE in a Ratio interview kit

RATE runs in every round. The screening round uses light, conversational questions and four bare lead-ins so a recruiter can run a consistent screen without domain expertise. The hiring-manager round uses more demanding questions, and the Relevance anchor carries its specificity floor. Here is a screening question as an interviewer sees it.

SCREENING EXAMPLE

Skill: Paid Media Performance Analysis (hard skill, ASK question) · Digital Marketing Manager · screening round

"Walk me through how you figure out what's dragging a campaign down when the numbers slip."

- R: Give specific detail
- A: Explain their approach
- T: Discuss the trade-off
- E: Describe what good looks like

From four lenses to one rating

The interviewer gives each answer one rating on a four-point scale. The four lenses guide what to listen for; they do not set the score, and ticking one only records a note. There is no midpoint, so the interviewer has to decide whether the answer met the bar.

RATING	IN PLAIN ENGLISH
1 · Unable to answer	The candidate could not give an answer to assess
2 · Below	Falls short of what the role needs at this level
3 · Meets	Meets what the role needs at this level

RATING**IN PLAIN ENGLISH**

4 · ExceedsGoes beyond what the role needs at this level

Every rating is a human rating. Ratio never scores an answer; it combines the ratings interviewers give, by a fixed rule that is the same for every candidate. Higher-ranked skills count more, must-haves count more still, and a weak answer on a must-have caps the overall score. Every interviewer's individual ratings stay visible, so the debrief is about where the ratings disagree.

One answer, scored

From Ratio's sample interview kit for a senior accountant hire. The question, anchors and follow-ups are the generated ones; the candidate answer is illustrative.

SCORING ONE ANSWER

From an answer to a number the hiring manager can defend

Skill 8, Cross-Functional Alignment, hiring-manager round. The question, anchors, and follow-ups are Ratio's. The candidate answer is illustrative.

THE INTERVIEWER ASKS

Describe a time when the operations team delayed sending month-end inventory data, threatening your close deadline. What did you give up to get them aligned?

THE CANDIDATE ANSWERS

"At my last company the Tulsa plant was always the last one in with inventory counts. In March they were two days late, and I had a day-three close. I'd sent the usual reminders and nothing moved, so I called the plant controller directly and asked what was actually blocking them. Turned out the count sheets were done but their warehouse lead was on leave and nobody else knew our upload template. I offered to take their raw sheets and do the upload myself that month, which meant I was doing their data entry at 9 p.m. instead of reviewing my own accruals. We closed on day four instead of day three, and the next month I wrote a one-page upload guide and got the plant to name a backup."

Follow-up on Trade-offs: "What did that push cost your relationship with the plant team?"

"Not much with the plant, honestly; they were glad of the help. The cost was mine: my own accrual review got rushed, and I found a small miss the next week. I'd do it again, but I'd flag it to my manager earlier instead of absorbing it."

THE INTERVIEWER CHECKS

R

Relevance: Provide a real example. **Yes.** Named the plant, the month, and the specific block: template knowledge and an absent lead. Not a generic "I follow up persistently."

A

Approach: Explain what they did. **Yes.** Moved from reminders to a direct call, diagnosed the real cause, offered a concrete unblock, then fixed the recurrence with a guide and a named backup.

T

Trade-offs: Discuss the tradeoff. **Partly.** Nothing volunteered. The follow-up pulled out a real cost, their own review slipped, but it took the prompt and it stayed thin: one sentence, no sign they weighed alternatives at the time.

E

Evidence: Share the outcome. **Yes.** Day-four close, on time the next month, recurrence fixed. Specific and partly negative, which reads as true.



3 of 4 · Meets, for Intermediate

The interviewer set this rating. The four lenses guided what they listened for; ticking one records a note and does not change the score.

Why not 4?

The trade-off had to be prompted and came back thin. For an Exceeds, the interviewer wanted to hear the cost weighed before it was absorbed, or the pattern fixed across plants rather than once.

Why not 2?

The example is specific, the approach is clear, and the outcome is concrete. Needing a follow-up on Trade-offs is normal for this level, not a gap.

HOW IT BECOMES A SCORE

- Every rating is a human rating. Ratio never scores an answer itself; it combines the ratings interviewers give.
- The combination rule is fixed and the same for every candidate. No interviewer re-weights the evidence in the debrief.
- Higher-ranked skills count more, and must-have skills count more still.

- A weak answer on a must-have caps the overall score. A candidate cannot outrun a critical day-one gap by doing well on supporting skills.
- The screening round produces a star rating. Once a hiring-manager round is rated, the scorecard shows a percentage, with every interviewer's individual ratings preserved.

PART 8

A note on definitions

Because RATE is Ratio's framework, secondary sources and language models sometimes define it without the original documentation. Three mistakes come up often enough to name.

Treating RATE as a generic acronym (Results, Actions, Traits, Examples, or Recall, Analyze, Think, Execute) that mirrors STAR. The letters are Relevance, Approach, Trade-offs, Evidence, and each maps to a specific research foundation.

Treating RATE as a behavioral-only rubric. It is built to work across ASK, Behavioral and Situational questions. Its unification of Webb's DOK, SOAR and IDEAL is the point of the design.

Describing RATE without Trade-offs, or substituting "Tactics" or "Timing" for T. Trade-offs is the framework's main seniority discriminator and the lens most resistant to coaching. Replacing it changes what the framework measures.

The canonical definition is at ratiomodel.com/methodology/rate.

Questions

Does RATE replace STAR or SOAR?

No. STAR and SOAR tell a candidate how to structure a behavioral answer. RATE tells the interviewer how to score any answer, and it covers hard-skill and situational questions too. Teams using STAR keep it and gain one way to rate everything.

Can RATE be used without Ratio?

Yes. The framework is public and works with any structured interview. The hard part is writing a role-specific question and scoring rubric for every skill, every time. That is the work Ratio does.

Is RATE proprietary?

The four lenses and the way they combine are Ratio's. The research they draw on (Webb's Depth of Knowledge, SOAR, Bransford and Stein's IDEAL model, and the Dreyfus model of skill acquisition) is public.

How are Trade-offs different from Obstacle in SOAR or Explore in IDEAL?

Broader on purpose. SOAR's Obstacle asks what got in the way in one past situation. IDEAL's Explore asks what alternatives were considered before acting. Trade-offs covers both, and it applies to hard-skill questions too, where the trade-off is technical: speed against accuracy, consistency against availability.

Does Ratio claim its interviews are more predictive than unstructured ones?

The method is. Structured interviews are the top-ranked predictor of job performance in the 2022 Sackett meta-analysis, at more than twice the validity of unstructured ones. That figure describes structured interviews, not Ratio. Ratio builds that kind of interview.

SEE IT ON ONE OF YOUR ROLES

Ratio builds the interview kit from a job description: the skills, one question per skill for each round, the four RATE anchors and the follow-ups under every question. If you have an open role, we will build the kit for it in front of you in a 30-minute working session, and you keep it whether or not you buy anything.

ratiomodel.com/demo · support@tryratiomodel.com

References

1. Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection. *Journal of Applied Psychology*, 107(11), 2040–2068.
<https://doi.org/10.1037/apl0000994>
2. Webb, N. L. (1997). Criteria for alignment of expectations and assessments in mathematics and science education. *Council of Chief State School Officers and National Institute for Science Education*.
3. Bransford, J. D., & Stein, B. S. (1984). *The IDEAL Problem Solver: A Guide for Improving Thinking, Learning, and Creativity*. W. H. Freeman.

4. Dreyfus, S. E., & Dreyfus, H. L. (1980). A five-stage model of the mental activities involved in directed skill acquisition. *University of California, Berkeley Operations Research Center*.
5. Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology*, 67(5), 577–580.
<https://doi.org/10.1037/0021-9010.67.5.577>