

Titan AI™

Competency Mastery Report

Adaptive conversational assessment · Individual

How this person worked through real job situations in their own words:
what each competency shows, where the evidence is still thin, and what to
check in the interview.

Candidate C-2291 (pseudonymous) · Customer Success Manager

9 of 14 bank scenarios answered · 84 pages · Sample v3.0

Free-text + follow-up probes

Descriptive score per competency

No ranking · no composite

Thin evidence stays visible

- | | |
|--|--|
| 01 Executive summary | 11 Follow-up probe analysis |
| 02 30-second summary | 12 Adaptive path |
| 03 Competency mastery profile | 13 Precision in depth |
| 04 Competency deep dive 1 | 14 Strengths · development |
| 05 Competency deep dive 2 | 15 Position in the cohort |
| 06 Conversation evidence | 16 Interview · onboarding guide |
| 07 Quote evidence board 1 | 17 Interview kit |
| 08 Quote evidence board 2 | 18 Development plan 30/60/90 |
| 09 Answer quality · four traits | 19 Campaign comparison · note |
| 10 Depth · consistency | |



AI Psychometrics Inc.

Titan AI™ · Individual report

Adaptive scenarios · synthetic 40-person, 14-scenario bank

Synthetic individual source

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Synthetic scenario coverage

Completed scenarios

9 of 14 bank scenarios

First read: a contrast to test

In this generated sample, Stakeholder communication is 68.63/100 versus 42.69/100 for Prioritizing under pressure. The four eligible lenses are separate descriptive views; Data-informed judgment has only one observation, so do not rank it. Use the pressure scenarios as an interview question, not a selection verdict.

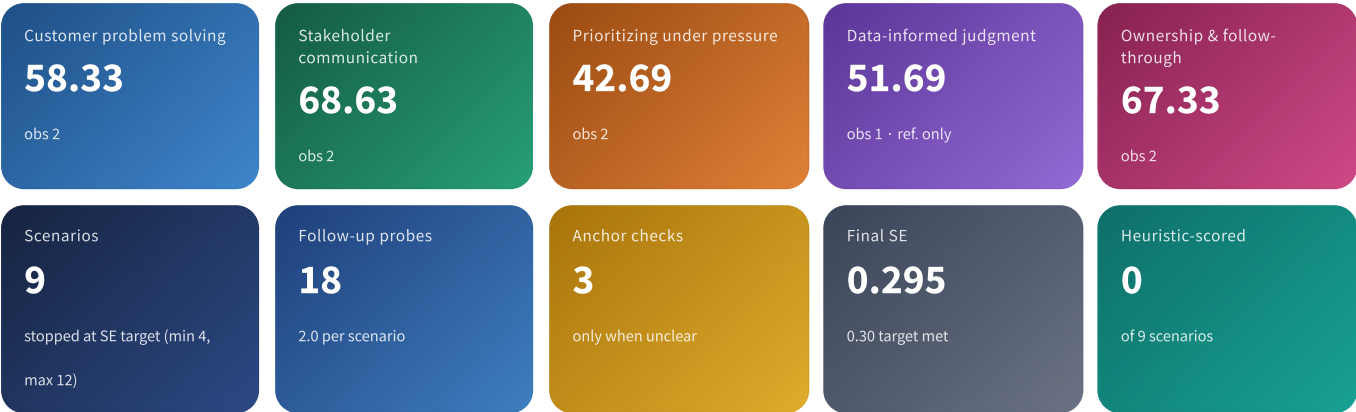
Within this synthetic cohort of 40 completers, Stakeholder communication is higher than 37 of the 40 cohort scores, while Prioritizing under pressure is lower than 26 of the 40. Neither position is a norm, a validated hiring rank, or evidence of workplace performance.

Source: synthetic adaptive 40-person sample, 14-scenario bank

Reading path

01	Summary and interview focus	p. 4
02	Candidate comparison evidence	p. 17
03	Five descriptive competency rows	p. 27
04	Masked answer evidence	p. 39
05	Interview questions and record	p. 73
06	Campaign comparison and measurement	p. 82

If you read one page before the interview, read this one.
Stakeholder communication is strong; prioritizing under pressure needs checking.



Four things to know

- Stakeholder communication and ownership & follow-through are the strongest evidence (68.63 and 67.33), the two highest of the five competencies.
- Communication is the highest trait (mean 3.56/5) but dipped to 2 in scenario 8.
- Prioritizing under pressure is lowest at 42.69; both of its 2 scenarios needed an anchor check, and risk awareness averages 1.50 there.
- Data-informed judgment has one observation, so 51.69 is reference only. Don't read it as a score; check it in the interview.

30-minute interview plan

- 0-5' Rapport · motivation
 - 5-15' Prioritizing under pressure · choosing what to defer
 - 15-23' Data-informed judgment · data vs field
 - 23-28' Stakeholder communication · confirm strength
 - 28-30' Candidate questions
- Time goes first where evidence is thinnest.

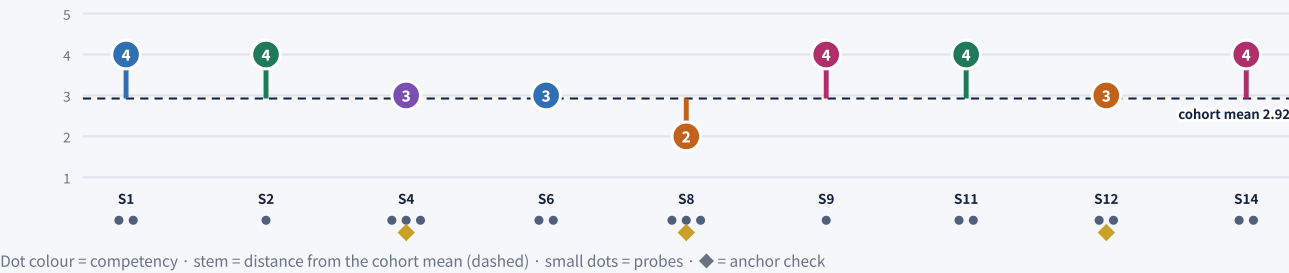
What this report says

How the candidate worked through job situations in their own words, how much evidence each competency has, and what to check in the interview.

Evidence ledger

What	Evidence
Competency scores	scored answers from 9 of the 14 bank scenarios · 1-2 observations per competency
Quotes	verbatim excerpts (synthetic in this sample)
Comparisons	40 completers in the same campaign · not a rank

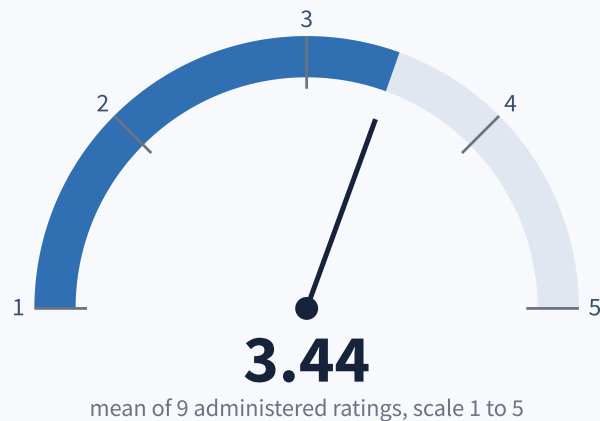
9-scenario lollipop · effectiveness vs cohort mean



Overall effectiveness at a glance

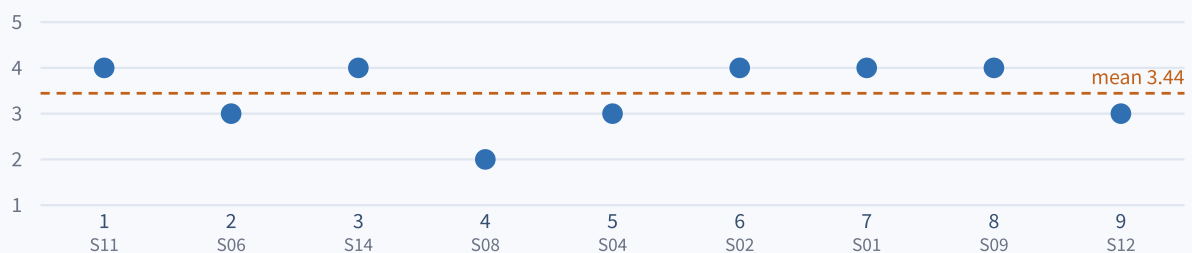
This gauge shows the mean of the 9 generated-answer effectiveness ratings (1 to 5) that the engine administered. It is a descriptive summary of this synthetic run, not a competency score, a calibrated measure or a hiring verdict.

Mean effectiveness rating



How to read. The needle marks the arithmetic mean (3.44) of the 9 administered ratings on the 1 to 5 input scale. The scale has no pass mark; ratings from different scenarios are different prompts.

Rating per administered scenario

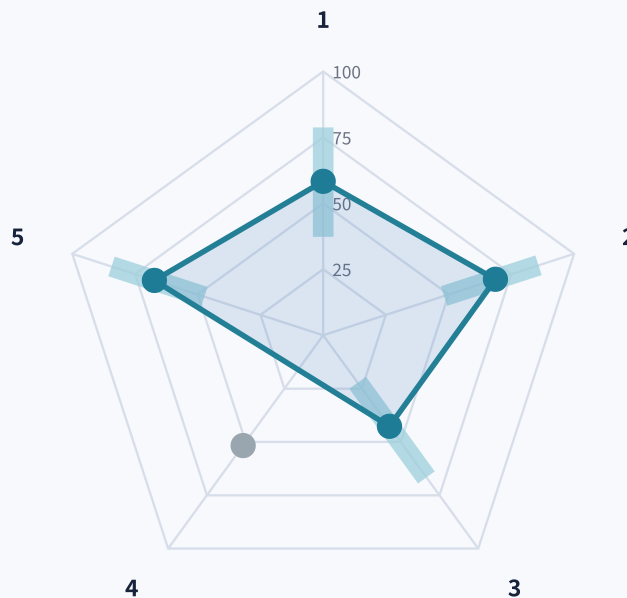


How to read. Dots follow the administered order (1 to 9) with scenario IDs below. The dashed line is the mean. Scenarios not administered have no rating.

Competency profile radar

Descriptive generated-answer scores on a 0 to 100 scale. The pale bars are 90% bands from each competency's own posterior, not calibrated intervals. Competencies with one observation are reference only and stay out of the polygon.

Competency profile



How to read. Axes follow the numbered list below, rings at 25, 50, 75 and 100. The filled polygon joins the 4 competencies with two or more observations; the grey dot is reference only. This is a profile shape, not a ranking.

Axis	Competency	Score	90% band	Obs.
1	Customer problem solving	58.33	37.3–78.8	2
2	Stakeholder communication	68.63	48.1–85.8	2
3	Prioritizing under pressure	42.69	22.2–66.6	2
4	Data-informed judgment*	51.69	none (reference only)	1
5	Ownership & follow-through	67.33	47.4–84.4	2

* One observation: reference only, no band, no comparison.

Executive summary / 1

Read this before the interview

Stakeholder communication is the highest adequately observed descriptive competency (68.63/100, two observations); prioritizing under pressure is lowest (42.69/100, two observations). These are generated-answer descriptions, not demonstrated workplace strengths or deficits.

Ask the candidate for real examples of priority trade-offs before interpreting the pattern.

Answer inputs	5/9 generated scenario answers rated 4 or 5; 3 anchor checks
Thin lens	Data-informed judgment: one observation; show as reference only

30-minute interview plan

The timing is a suggested interviewer workflow, not a validated prediction or automated hiring recommendation. Record independent evidence separately.

0–5 min	Context and motivation
5–15 min	Priority trade-off under pressure
15–23 min	Data versus field evidence
23–28 min	Stakeholder communication example
28–30 min	Candidate questions

What the evidence includes

The descriptive competency numbers come from generated answers across 9 of the 14 bank scenarios, with one to two observations per competency. Comparison plots use the same cohort of 40 completers but do not establish a hiring rank. Raw dialogue is masked in this employer-facing proof.

Executive summary / 2

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Scores, observations and 90% bands

Customer problem solving descriptive score	58.33
Customer problem solving observations	2
Customer problem solving 90% band lower	37.33
Customer problem solving 90% band upper	78.82
Stakeholder communication descriptive score	68.63
Stakeholder communication observations	2
Stakeholder communication 90% band lower	48.14
Stakeholder communication 90% band upper	85.85
Prioritizing under pressure descriptive score	42.69
Prioritizing under pressure observations	2
Prioritizing under pressure 90% band lower	22.22
Prioritizing under pressure 90% band upper	66.59

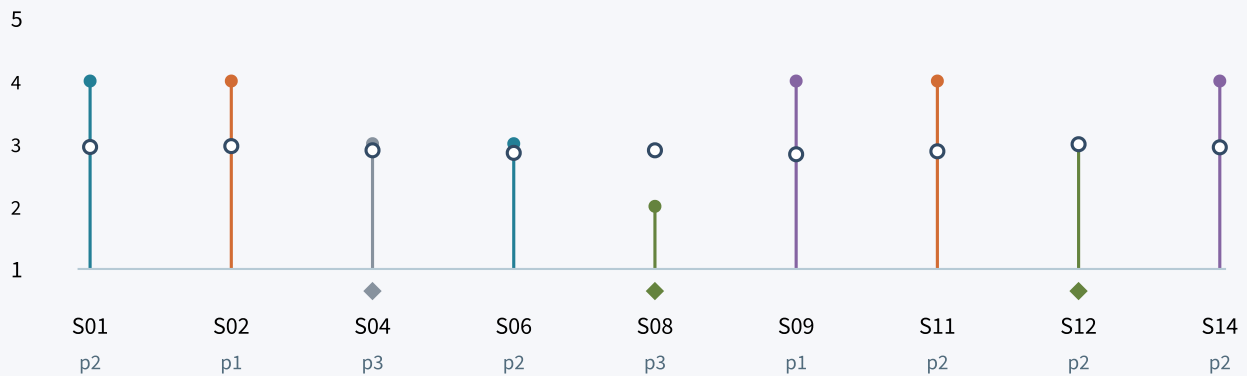
Scores, observations and 90% bands · continued

Data-informed judgment descriptive score	51.69
Data-informed judgment observations	1
Data-informed judgment 90% band lower	no band (reference only)
Data-informed judgment 90% band upper	no band (reference only)
Ownership & follow-through descriptive score	67.33
Ownership & follow-through observations	2
Ownership & follow-through 90% band lower	47.36
Ownership & follow-through 90% band upper	84.36
Follow-up probes	18
Anchor checks	3
Overall theta SE	0.295

Executive summary / 3

Nine administered scenario ratings with source keys

Customer problem solving Stakeholder communication Prioritizing under pressure Data-informed judgment
Ownership & follow-through



Colored lollipops: this generated answer (1–5). Hollow navy: cohort mean for the same scenario, rounded to two decimals; the table shows how many of the 40 completers answered it. ♦ anchor check; p = follow-up count. Colors identify competencies, not ability. Synthetic adaptive records, not a rank or hiring verdict.

Scenario trace: rating, probes and anchor

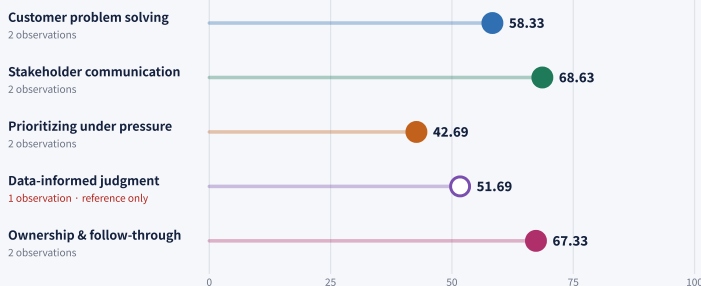
The 1-5 effectiveness rating describes a generated answer. The cohort column is the mean of the completers who answered that same scenario (n shown); it is not a competency score. Not administered in this session: S03, S05, S07, S10, S13. ♦ marks an anchor check.

ID	Competency	1–5	Rating	Cohort mean (n)	Probes	Anchor
S01	Customer problem solving		4	2.95 (n=38)	2	
S02	Stakeholder communication		4	2.96 (n=38)	1	
S04	Data-informed judgment		3	2.89 (n=40)	3	♦
S06	Customer problem solving		3	2.85 (n=39)	2	
S08	Prioritizing under pressure		2	2.89 (n=40)	3	♦
S09	Ownership & follow-through		4	2.83 (n=40)	1	
S11	Stakeholder communication		4	2.88 (n=37)	2	
S12	Prioritizing under pressure		3	2.99 (n=29)	2	♦
S14	Ownership & follow-through		4	2.94 (n=39)	2	

FIRST SCREEN · 30-SECOND SUMMARY

Communication with customers and internal teams is the strongest evidence. Check prioritizing under pressure in the interview.

Mastery by competency (0–100, descriptive)



Dot = candidate score · hollow dot = fewer than 2 observations (reference only)

Strongest evidence

Stakeholder communication, 68.63. Delivering bad news to an executive sponsor, the candidate laid out facts, impact and next steps in order and held up under follow-up probes.

Evidence: scenarios ②⑪

Check in the interview

Prioritizing under pressure, 42.69, the lowest of the five. Twice, when work piled up, the candidate took everything on alone and escalated late.

Evidence: scenarios ⑧⑫ · questions on page 9

Where evidence is still thin

Data-informed judgment was observed in one scenario only, so 51.69 is reference only. Don't read it as a score; check it directly.

Scenarios

9

stopped at SE target

Competencies covered

4/5

1 has a single observation

Final SE

0.295

0.30 target met

Follow-up probes

18

extra questions on answers

Evidence base

9 of 14 scenarios

1–2 observations per competency

Role-fit summary (narrative)

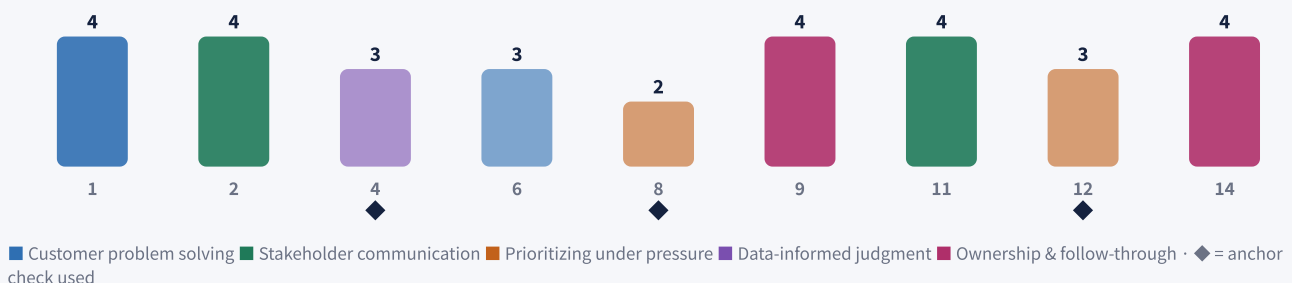
C-2291 is clearest at sorting out a situation between customers and internal teams and putting it into words. When something breaks, the candidate checks the cause first and makes specific promises to the customer about what will happen by when. Under stacked demands, though, the candidate tried to solve it by stretching their own time rather than setting a rule for what to drop. Worth covering in onboarding for the stage when the book of accounts grows.

No composite score

No ranking

People make the decision

All 9 administered conversations at a glance (overall 1–5, colour = competency)



Before the interview

Of the four questions on page 9, start with the prioritizing one.

In the interview

Ask for one more example of data-informed judgment to fill the evidence gap.

If hired

Onboarding idea on page 9: practise deciding what to defer in heavy weeks.

First screen

Synthetic template responses; not real candidate data or calibrated Titan AI™ output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Five generated-answer scores at a glance

Customer problem solving: 58.33

Stakeholder communication: 68.63

Prioritizing under pressure: 42.69

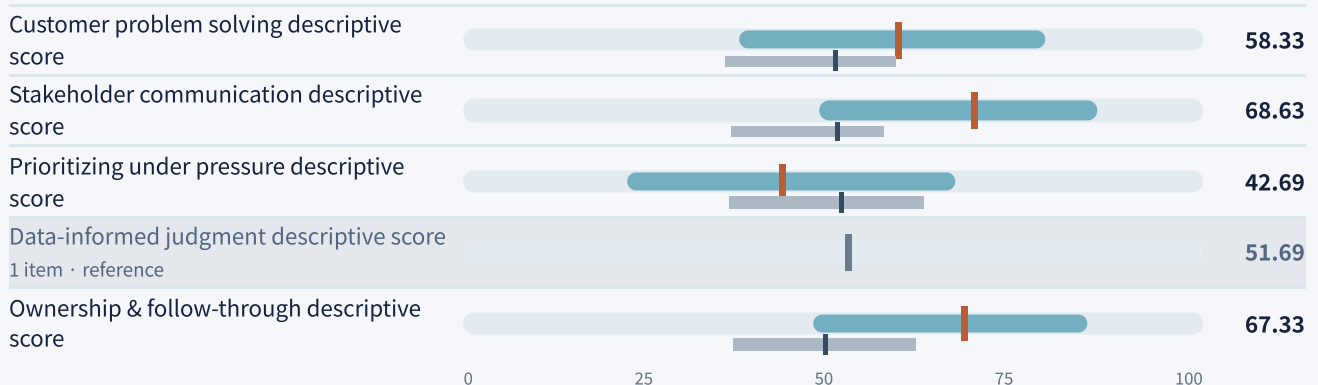
Data-informed judgment: 51.69 (one answer, reference)

Ownership & follow-through: 67.33

Read these patterns, then verify

Among the four comparison-eligible lenses, this generated respondent has the highest observed effectiveness in stakeholder communication (68.63) and the lowest in prioritizing under pressure (42.69). Check the latter with a consistent interview prompt; these answers do not show how the person performs at work.

Case scores, 90% bands and same-sample reference

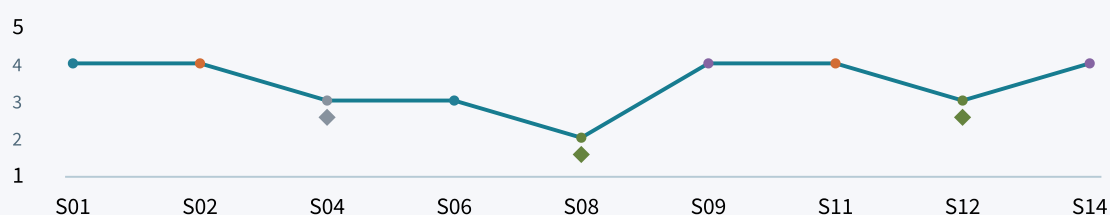


Descriptive generated-answer score. * Data-informed judgment has one item: 51.69/100 is reference-only and has no band; it is not a calibrated interval, eligible relative comparison, or workplace evidence.

Orange: case score. Teal: 90% band from the competency's own posterior, not calibrated. Beneath each eligible rail, gray: middle 50% of 40 person scores (linear p25–p75), navy tick: median. These are descriptive references, not confidence or hiring thresholds.

Nine scenario ratings with competency and anchor keys

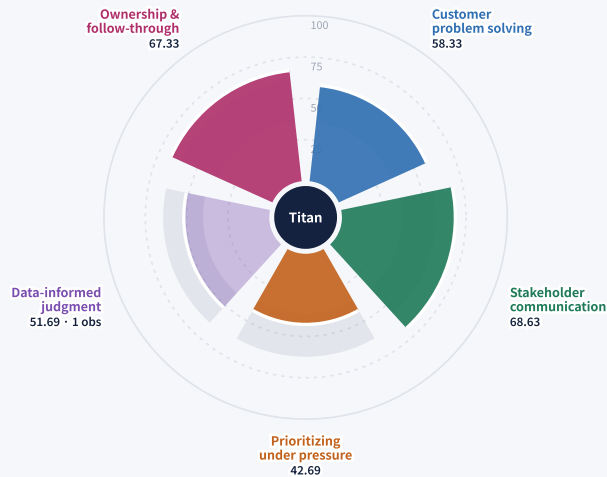
Customer problem solving Stakeholder communication Prioritizing under pressure Data-informed judgment
Ownership & follow-through



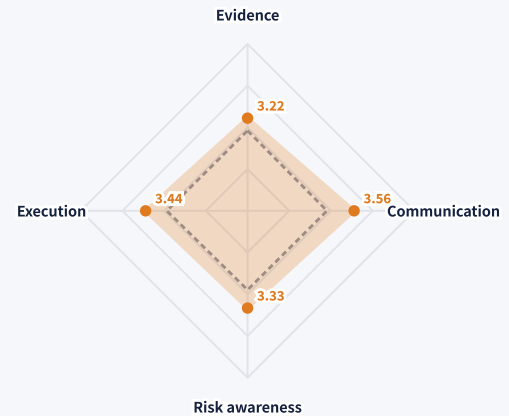
Color labels identify the competency, not ability. Diamond marks the 3 administered anchors. Each dot is one generated answer rating (1–5), ordered S01–S14, administered scenarios only; joined lines are not a calibrated trend. Data-informed judgment has one answer, reference only.

Stakeholder communication is highest (68.63) and prioritizing under pressure lowest (42.69). Data-informed judgment has one observation, so its score is reference only.

Competency constellation

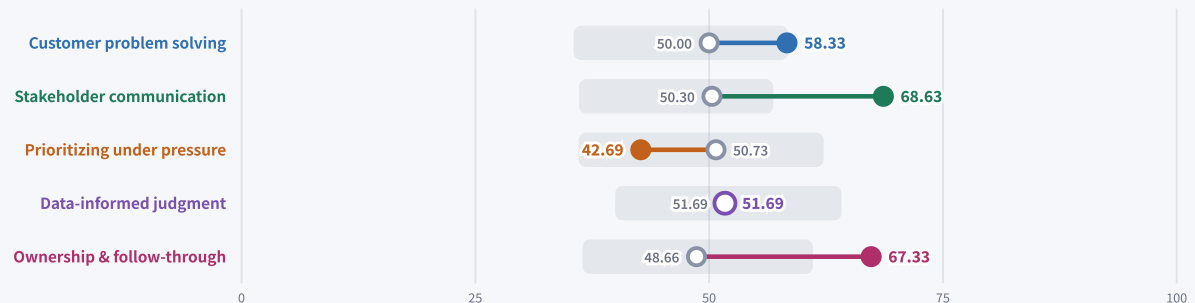


Trait radar · candidate vs cohort



Orange = this candidate's scenario mean · dashed = cohort mean (40 completers).

Dumbbell · candidate score ↔ cohort median



Coloured dot = candidate · white dot = completer median · grey = middle 50% of completers. Descriptive comparison only; a dot inside the grey band is not called a difference.

Competency × trait heatmap

	Evidence	Communication	Risk awareness	Execution
Customer problem solving · 2 obs	3.50	3.50	3.50	3.00
Stakeholder communication · 2 obs	3.50	4.50	4.50	4.50
Prioritizing under pressure · 2 obs	2.50	3.00	1.50	2.50
Data-informed judgment · 1 obs	2.00	3.00	3.00	3.00
Ownership & follow-through · 2 obs	4.00	3.50	4.00	4.00

Mean trait rating (1–5) across that competency's scenarios. With few observations, one or two answers set the whole row.

Generated-answer competency profile / 1

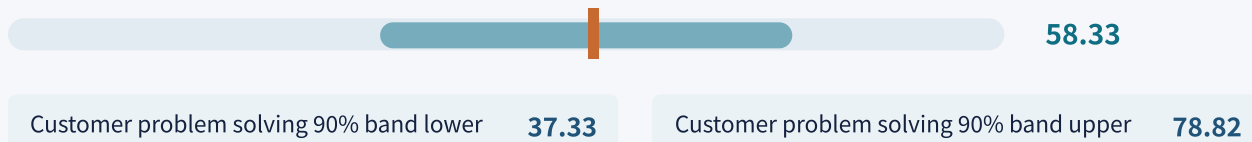
Synthetic template responses; not real candidate data or calibrated Titan AI™ output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Scores and 90% bands by competency

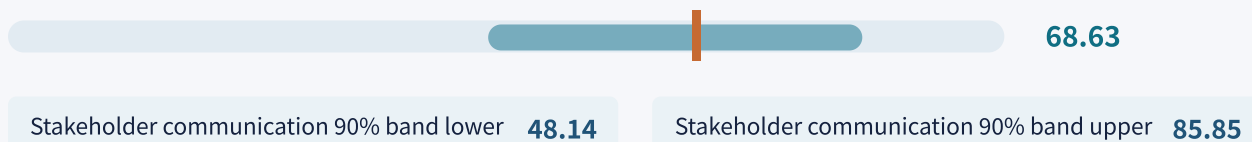
Customer problem solving descriptive score

Customer problem solving observations 2



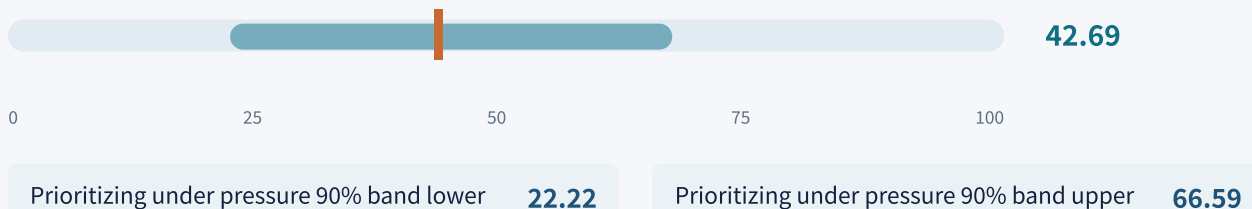
Stakeholder communication descriptive score

Stakeholder communication observations 2



Prioritizing under pressure descriptive score

Prioritizing under pressure observations 2



Rails span a descriptive 0–100 score scale.

Generated-answer competency profile / 2

Computed evidence · continued

Data-informed judgment descriptive score

one observation · reference

Data-informed judgment observations: 1

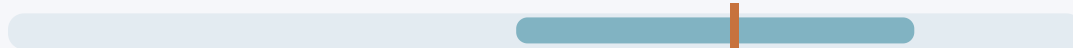
No band: one observation is reference only



51.69

Ownership & follow-through descriptive score

Ownership & follow-through observations: 2



67.33

Ownership & follow-through 90% band lower:

47.36

Ownership & follow-through 90% band upper:

84.36

0

25

50

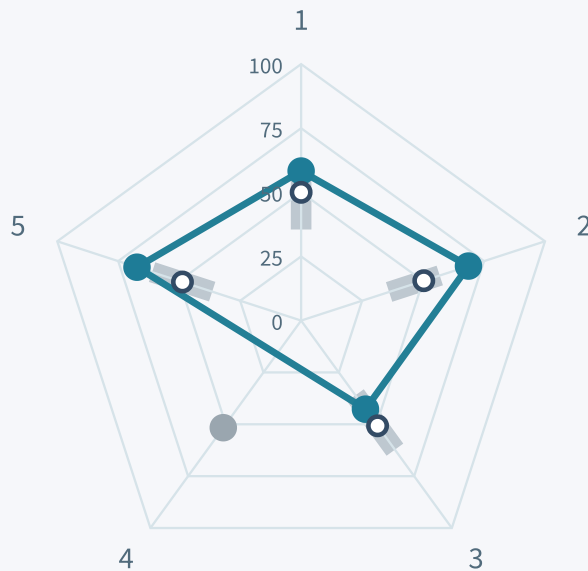
75

100

Rails span a descriptive 0–100 score scale.

Generated-answer competency profile / 3

Case scores with same-sample median and middle 50%



1 Customer problem solving: 58.33

2 Stakeholder communication: 68.63

3 Prioritizing under pressure: 42.69

4 Data-informed judgment: 51.69*

5 Ownership & follow-through: 67.33

Synthetic answer data; no personality or calibrated competency inference.

Generated-answer descriptive scores, not mastery. * Data-informed judgment: one item, 51.69/100 reference-only; a disconnected gray marker separates it from the four other descriptive axes. Do not infer job performance.

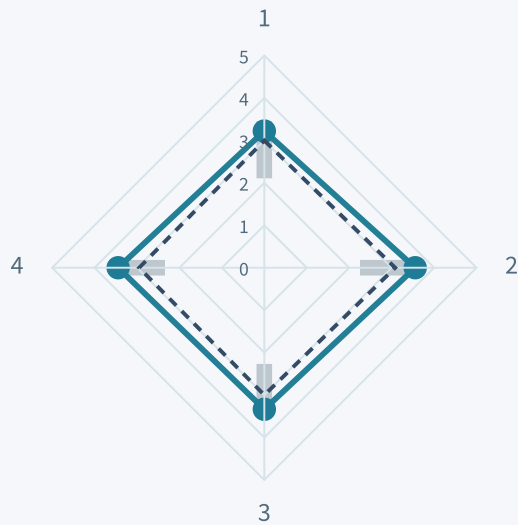
Numbered axes match the list. Radial scale: 0–100 descriptive score; rings at 25, 50, 75 and 100.

Teal: case descriptive score. Hollow navy: median of 40 scores. Gray spoke segments: p25–p75 of those scores (linear interpolation), not confidence intervals. Data-informed judgment is reference-only; no comparison mark. 90% bands appear on the preceding sheet; none is a calibrated native interval.

Competency	Case	Median	Middle 50%
Customer problem solving	58.33	50.00	35.51 to 58.44
Stakeholder communication	68.63	50.30	36.08 to 56.85
Prioritizing under pressure	42.69	50.73	36.03 to 62.23
Ownership & follow-through	67.33	48.66	36.47 to 61.11

Generated-answer competency profile / 4

Trait ratings: synthetic case and same-sample reference



1 evidence: 3.22

2 communication: 3.56

3 risk awareness: 3.33

4 execution: 3.44

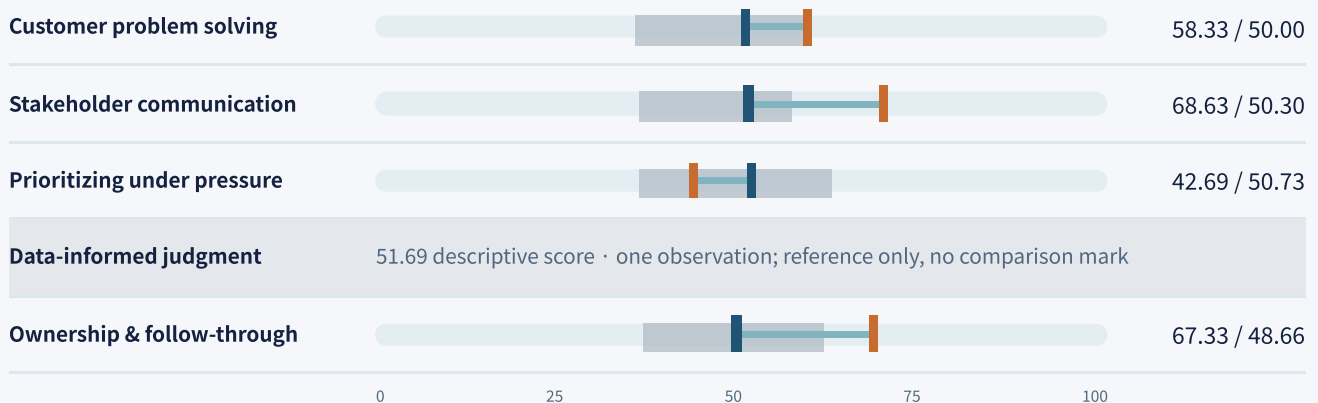
Synthetic answer data; no personality or calibrated competency inference.

Synthetic adaptive source, trait radar. Masked employer view; figures describe generated answers, not job performance.

Teal: case mean. Dashed navy: median of 40 person means. Gray spoke segments: middle 50% (linear-interpolated p25–p75 of 40 person means), not confidence intervals. Numbered axes match the list. Radius 0–5; observed trait ratings 1–5.

Axis	Case mean	40-person median	Middle 50%
1 evidence	3.22	3.00	2.11 to 3.45
2 communication	3.56	3.08	2.25 to 3.41
3 risk awareness	3.33	3.00	2.27 to 3.37
4 execution	3.44	2.95	2.34 to 3.51

Case scores and same-sample reference



Orange: case. Navy: 40-person median. Gray: middle 50% of 40 scores (linear p25–p75), not a confidence interval. Descriptive 0–100.

Synthetic adaptive source. Four eligible comparison marks; Data-informed judgment has one observation and its 51.69 descriptive score is reference-only. Masked employer view; generated answers do not establish job performance.

Generated-answer competency profile / 5

Competency trait heatmap

	Evidence	Communication	Risk awareness	Execution
Customer problem solving	3.50	3.50	3.50	3.00
Stakeholder communication	3.50	4.50	4.50	4.50
Prioritizing under pressure	2.50	3.00	1.50	2.50
Data-informed judgment <i>one observation · reference only</i>	2.00	3.00	3.00	3.00
Ownership & follow-through	4.00	3.50	4.00	4.00

Trait rating (1–5):



Synthetic descriptive chart; verify the underlying data before use.

Darker teal means a higher mean rating; numbers show the exact mean. Color follows the existing linear fill, with the low end clamped at zero. Reference-only cells are not comparison evidence.

Candidate comparison evidence / 1

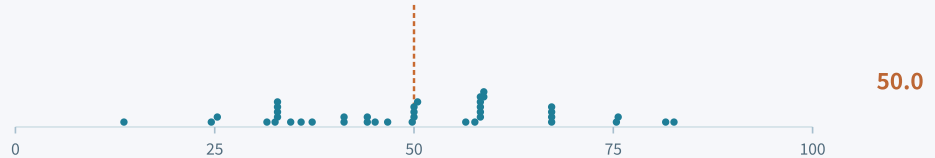
Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Score spread with every candidate point

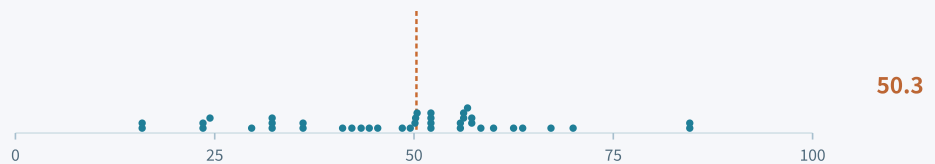
Same-assessment competency distributions (n=40 per eligible lens)

One dot per person, 40 dots per row, placed at the raw score; dots within a few points stack. Orange dashed line = median.

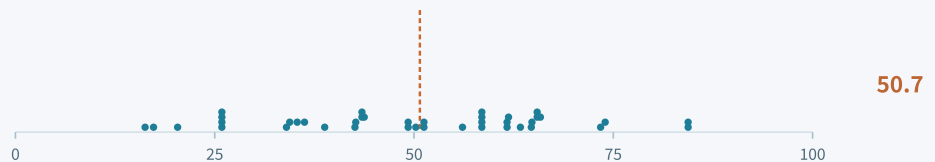
Customer problem solving



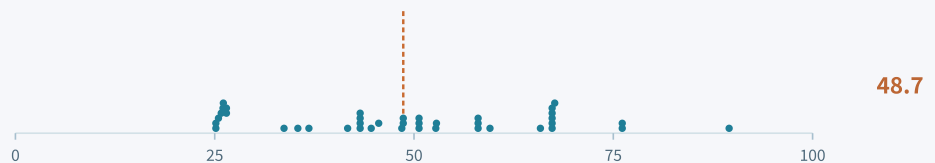
Stakeholder communication



Prioritizing under pressure



Ownership & follow-through

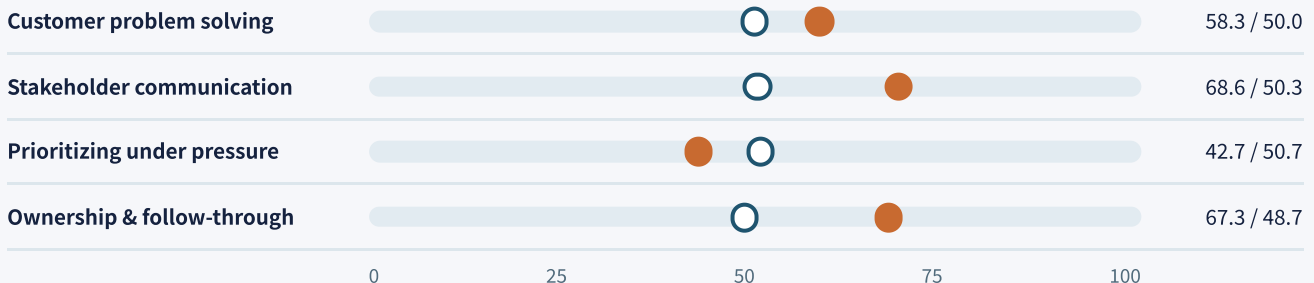


Data-informed judgment: one observation per person; reference only, no comparison.

Points are generated descriptive scores; the median tick uses the same 40 values. No inferred density or external norm.

Source: sorted generated scores and medians, 40 people per competency

Four paired, eligible competency comparisons



Both marks use the same descriptive 0–100 score scale.

Filled = candidate; ring = synthetic cohort median; no overall rank.

Each lens is separate. Tie-aware counts are available in the source, but these are not hiring ranks.

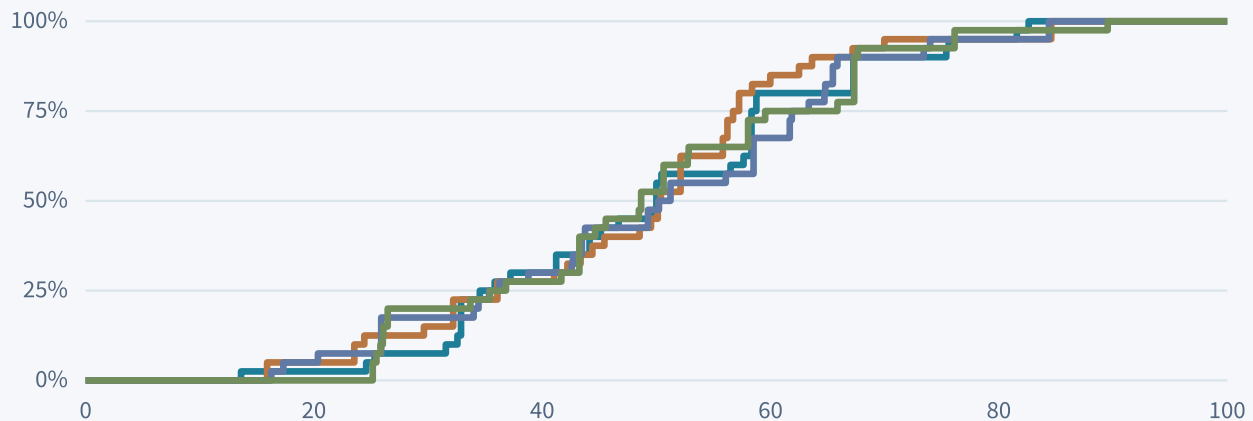
Source: C-2291 compared with the 40-person sample

Candidate comparison evidence / 2

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Where do the same 40 scores accumulate and cluster?

Empirical cumulative score curves



Descriptive score (0–100)

Customer problem solving

Stakeholder communication

Prioritizing under pressure

Ownership & follow-through

Share at or below a score, within the 40 generated people for that competency. Each step is one person. No external percentile.

Source: sorted generated scores, 40 people per competency

Ten-point score frequency grid by competency

	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–99	
Customer problem solving	0	1	2	9	7	13	4	2	2	0	40
Stakeholder communication	0	2	4	5	7	16	4	0	2	0	40
Prioritizing under pressure	0	2	5	5	7	8	9	2	2	0	40
Ownership & follow-through	0	0	8	3	10	9	7	2	1	0	40

People per score band: 0 3 6 10 13 16

Each number is the exact number of people in its ten-point score band. Tile fill intensity rises with count (0–16); each row totals 40. The one-answer Data-informed judgment lens is excluded. These scores are not hiring norms or evidence grades.

Source: generated scores in 10-point bins, 40 people per competency

Each card holds the score, the administered scenarios, traits and one key quote.

Customer problem solving

58.33 · obs 2

within the cohort middle 50%

4 probes

Effectiveness by scenario (1-5)



Four traits · line = cohort



Key quote · ① Escalated renewal at risk 4/5

"I'd book a call within 24 hours and lead with what changed and the fix timeline, not excuses."

Scoring saw: Checks cause → talks to customer → limits promises

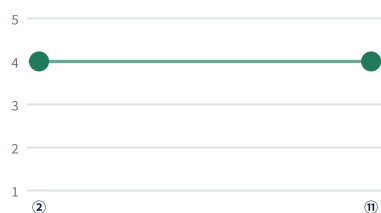
Stakeholder communication

68.63 · obs 2

above the cohort middle 50%

3 probes

Effectiveness by scenario (1-5)



Four traits · line = cohort



Key quote · ② Conflicting asks from sales and product 4/5

"I'd put both asks on one page side by side and ask the sales and product leads for 15 minutes together, framed on customer impact."

Scoring saw: Puts the conflict on one page; uses customer impact as the test

Prioritizing under pressure

42.69 · obs 2

within the cohort middle 50%

5 probes

2 anchor checks

Effectiveness by scenario (1-5)



Four traits · line = cohort



Key quote · ⑧ Quarter-end workload spike 2/5

"They're all important, so I'd put in evenings to hit every one."

Scoring saw: Stretches own time; doesn't choose what to defer

Grey band = cohort middle 50% · grey line = cohort median · gold ring = anchor check used

Competency deep dive 1 / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Competency deep dive · reading route

The full competency-value, observation-count and 90%-band source table remains on page 7. This section separates the roles of the charts and their inputs instead of repeating that table.

Competency values

Pages 21, 23, 24

The dot is a descriptive score and the bar is an 90% band. It is not a calibrated confidence interval or a range of job ability.

Matching scenario answers

Pages 22, 23, 25

These show generated answer ratings for the matching competency. Keep the source input scale separate from other measures.

Answer-quality inputs

Pages 22, 24, 25

These axes are inputs recorded for the answers, not personality traits or validated competency scores.

Interpretation boundary

Do not combine different scales into an overall rank. The 90%-band width is not an evidence grade validated against independent human ratings. Relative comparison follows the separate comparison pages and their same-source tie rules.

Competency deep dive 1 / 2

58.33

DESCRIPTIVE SCORE

37.33–78.82

90%
BAND

2

OBSERVATIONS

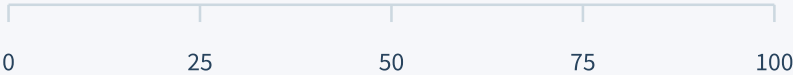
Customer problem solving

Only this chart's competency is shown here, from the page 7 source table. Other competency rows are not repeated.

Customer problem solving descriptive score	58.33
Customer problem solving observations	2
Customer problem solving 90% band lower	37.33
Customer problem solving 90% band upper	78.82

Competency value and 90% band

Customer problem solving



Descriptive scale · 0–100

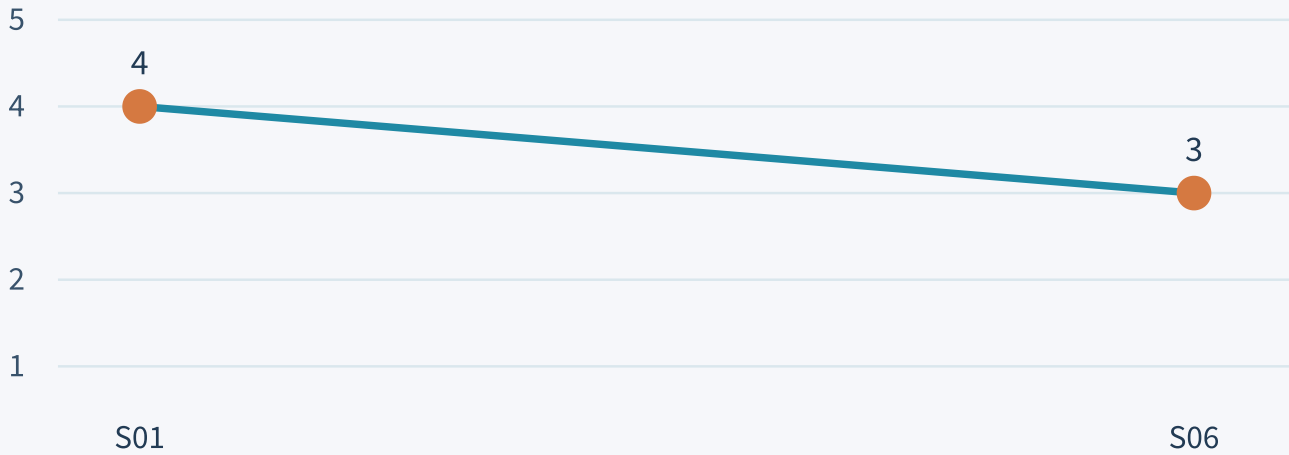
Dot: descriptive value. Bar: separate 90% band, not a calibrated confidence interval. These synthetic marks do not establish applicant or job ability.

Keep the scales separate

The observation count is the number of scenarios linked to this competency. This 0–100 descriptive scale is not the same unit as a scenario's 1–5 answer rating. Do not turn 90%-band width into an evidence grade.

Competency deep dive 1 / 3

Customer problem solving · answer ratings

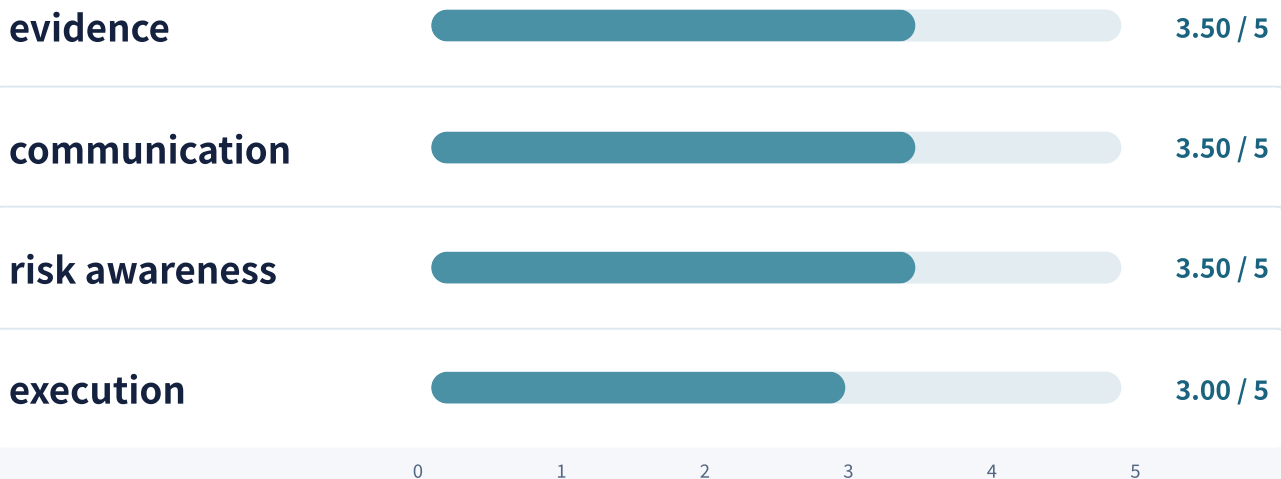


S01: Escalated renewal at risk

S06: Customer demands an unsupported feature

2 administered scenarios for this competency (1 of 3 bank scenarios not asked in this adaptive run); * means an anchor option was checked. Ordinal effectiveness input (1–5), not a calibrated competency estimate. No raw candidate dialogue shown.

Customer problem solving · answer-quality means



Answer-quality means: ratings 1–5, drawn from a zero baseline. Bar lengths use the raw means; labels show 2 decimals.

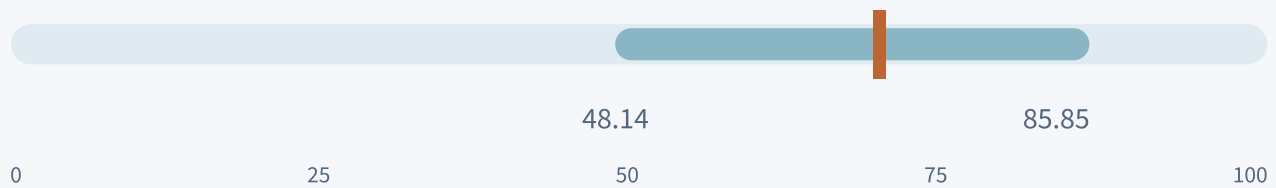
Synthetic descriptive chart; verify the underlying data before use.

Competency deep dive 1 / 4

Stakeholder communication · descriptive score

Stakeholder communication

68.63 / 100



Descriptive score scale: 0–100. Numbers above this axis label the 90% band endpoints.

Synthetic descriptive chart; verify the underlying data before use. The light band is a 90% interval from the competency posterior, not a calibrated confidence interval.

Stakeholder communication · answer ratings



S02: Conflicting asks from sales and product

S11: Team disagreement on an account plan

2 administered scenarios for this competency (1 of 3 bank scenarios not asked in this adaptive run); * means an anchor option was checked. Ordinal effectiveness input (1–5), not a calibrated competency estimate. No raw candidate dialogue shown.

Competency deep dive 1 / 5

Stakeholder communication · answer-quality means

evidence



3.50 / 5

communication



4.50 / 5

risk awareness



4.50 / 5

execution



4.50 / 5

0 1 2 3 4 5

Stakeholder communication trait means: ratings 1–5, drawn from a zero baseline. Bar lengths use the raw means; labels show 2 decimals.

Synthetic descriptive chart; verify the underlying data before use.

Prioritizing under pressure · descriptive score

Prioritizing under pressure

42.69 / 100



22.22

66.59

0

25

50

75

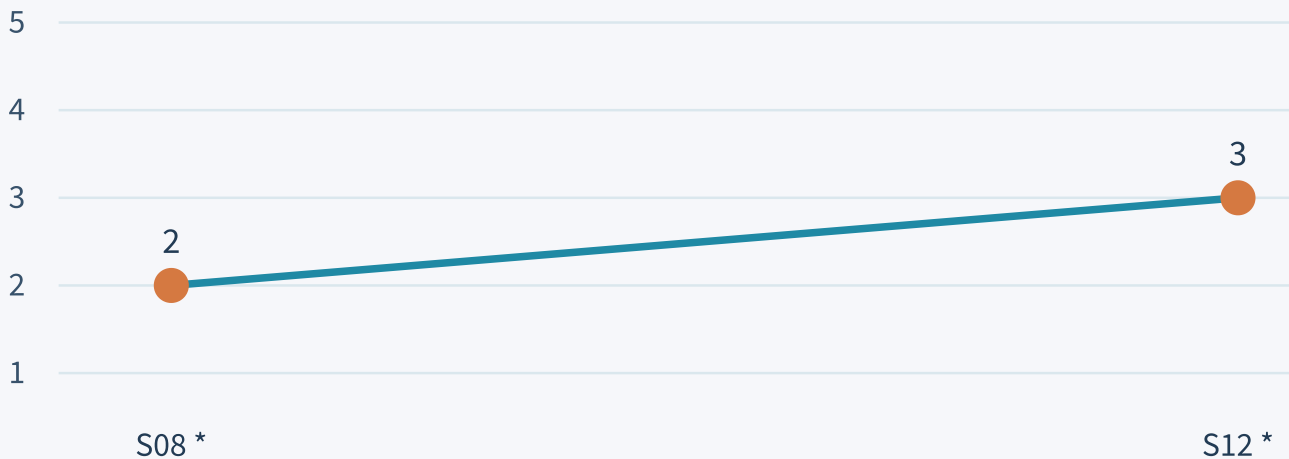
100

Descriptive score scale: 0–100. Numbers above this axis label the 90% band endpoints.

Synthetic descriptive chart; verify the underlying data before use. Light band = 90% interval from the competency posterior, not a calibrated confidence interval.

Competency deep dive 1 / 6

Prioritizing under pressure · answer ratings

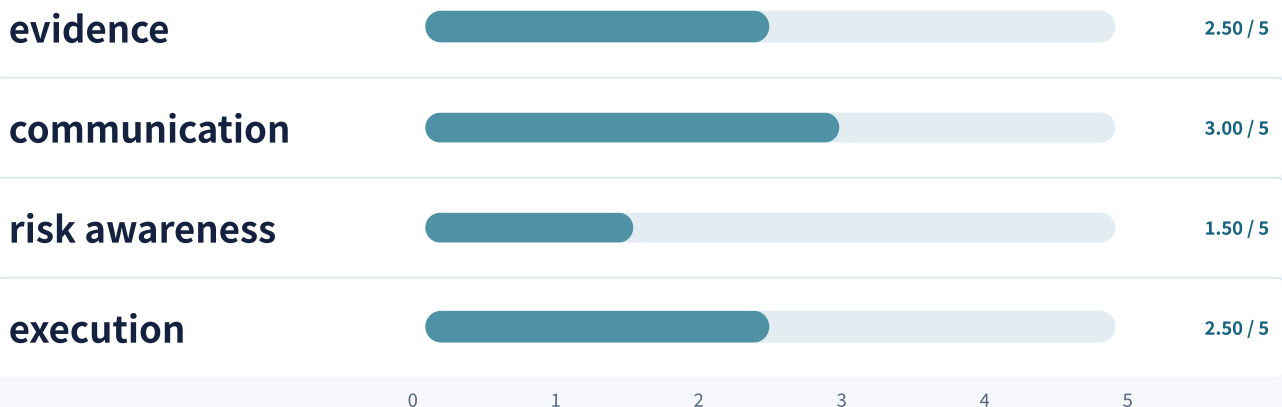


S08 *: Quarter-end workload spike

S12 *: Deadline vs quality trade-off

2 administered scenarios for this competency (2 of 4 bank scenarios not asked in this adaptive run); * means an anchor option was checked. Ordinal effectiveness input (1–5), not a calibrated competency estimate. No raw candidate dialogue shown.

Prioritizing under pressure · answer-quality means



Answer-quality means: ratings 1–5, drawn from a zero baseline. Bar lengths use the raw means; labels show 2 decimals.

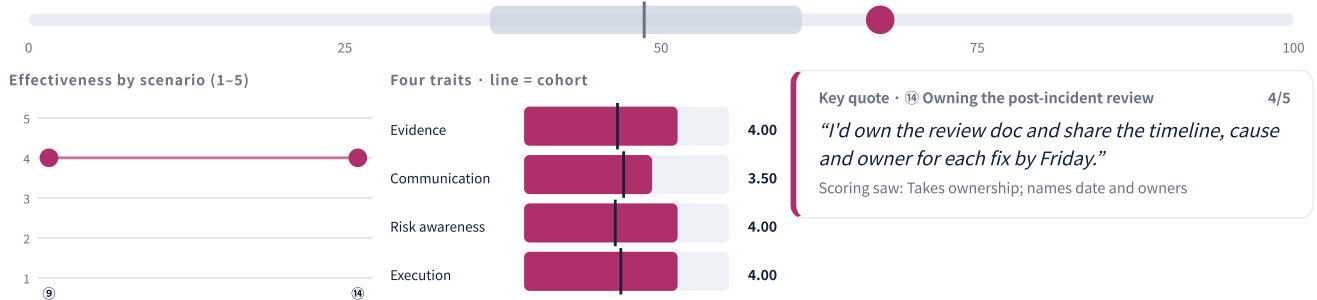
Synthetic descriptive chart; verify the underlying data before use.

Ownership & follow-through scored 4 both times. Data-informed judgment has a single observation, so 51.69 is reference only.

Ownership & follow-through

67.33 · obs 2

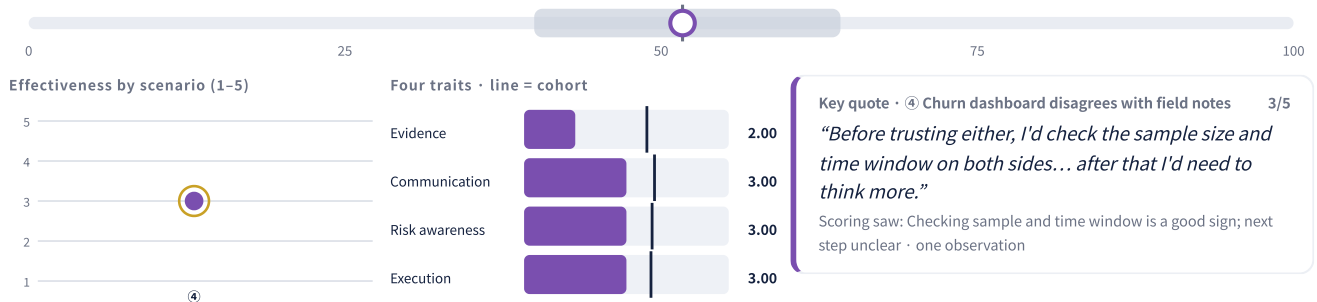
above the cohort middle 50% 3 probes



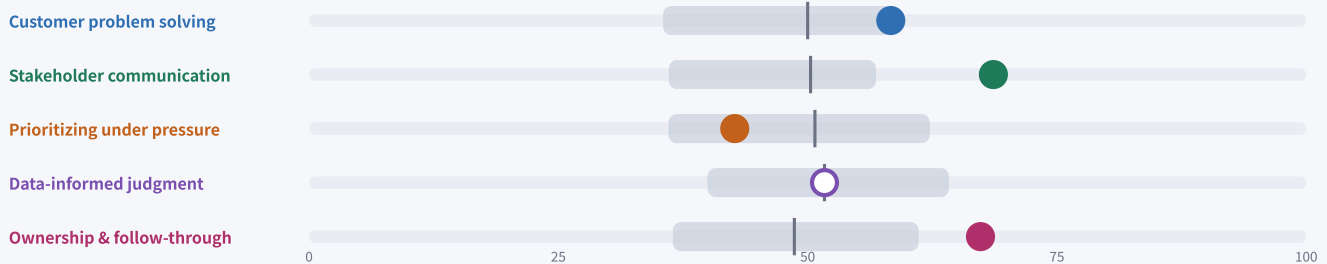
Data-informed judgment

51.69 · obs 1

Reference only · 1 observation 3 probes 1 anchor checks



All five on one scale



All five at a glance

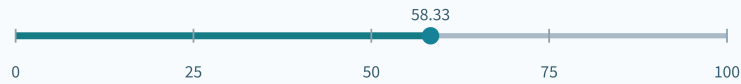
Competency	Score	Obs	Probes	Anchors	Mean eff.	Cohort median [middle 50%]	Position vs cohort
Customer problem solving	58.33	2	4	0	3.50	50.00 [35.51–58.44]	within the cohort middle 50%
Stakeholder communication	68.63	2	3	0	4.00	50.30 [36.08–56.85]	above the cohort middle 50%
Prioritizing under pressure	42.69	2	5	2	2.50	50.73 [36.03–62.23]	within the cohort middle 50%
Data-informed judgment	51.69	1	3	1	3.00	51.69 [39.96–64.17]	reference only
Ownership & follow-through	67.33	2	3	0	4.00	48.66 [36.47–61.11]	above the cohort middle 50%

Fewer than 2 observations is reference only, not a score.

Five descriptive competency rows

Each row uses the same descriptive 0–100 scale derived from matching 1–5 answer-effectiveness inputs. Bands appear on the profile pages. No evidence grade. The one-observation data lens is reference only.

Customer problem solving



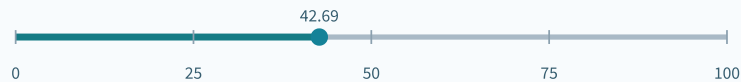
2 observations (S01, S06). Descriptive, not a hiring threshold.

Stakeholder communication



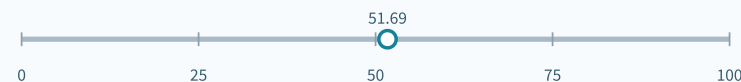
2 observations (S02, S11). Descriptive, not a hiring threshold.

Prioritizing under pressure



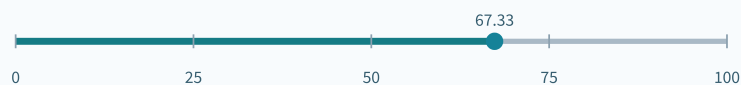
2 observations (S08, S12). Descriptive, not a hiring threshold.

Data-informed judgment



1 observation (S04) · reference only; no relative rank.

Ownership & follow-through



2 observations (S09, S14). Descriptive, not a hiring threshold.

Competency deep dive 2 / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Competency deep dive · reading route

The full competency-value, observation-count and 90%-band source table remains on page 7. This section separates the roles of the charts and their inputs instead of repeating that table.

Competency values

Pages 29, 31, 34

Pages 29 and 34 show descriptive scores and separate 90% bands. Page 31 shows the one-answer Data-informed judgment score for reference only, without a comparison band. None is a calibrated confidence interval or a range of job ability.

Matching scenario answers

Pages 30, 31

These show generated answer ratings for the matching competency. Keep the source input scale separate from other measures.

Answer-quality inputs

Pages 30, 32

These axes are inputs recorded for the answers, not personality traits or validated competency scores.

Masked records

Pages 32–34

These are restricted displays of generated answer evidence. They are not raw dialogue or a grant of access.

Interpretation boundary

Do not combine different scales into an overall rank. The 90%-band width is not an evidence grade validated against independent human ratings. Relative comparison follows the separate comparison pages and their same-source tie rules.

Competency deep dive 2 / 2

67.33

DESCRIPTIVE SCORE

47.36–84.36

90% BAND

2

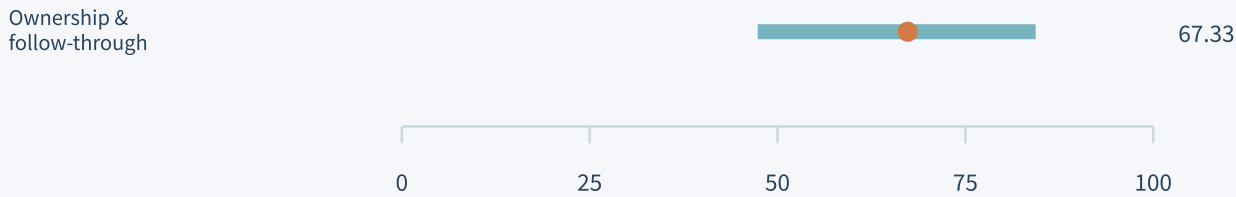
OBSERVATIONS

Ownership & follow-through

Only this chart's competency is shown here, from the page 7 source table. Other competency rows are not repeated.

Ownership & follow-through descriptive score	67.33
Ownership & follow-through observations	2
Ownership & follow-through 90% band lower	47.36
Ownership & follow-through 90% band upper	84.36

Competency value and 90% band



Descriptive scale · 0–100

Dot: descriptive value. Bar: separate 90% band, not a calibrated confidence interval. These synthetic marks do not establish applicant or job ability.

Keep the scales separate

The observation count is the number of scenarios linked to this competency. This 0–100 descriptive scale is not the same unit as a scenario's 1–5 answer rating. Do not turn 90%-band width into an evidence grade.

Competency deep dive 2 / 3

Ownership & follow-through · answer ratings



2 administered scenarios for this competency (1 of 3 bank scenarios not asked in this adaptive run); * means anchor check. Answer-effectiveness input 1–5, not a calibrated competency estimate. No raw dialogue shown.

Ownership & follow-through · answer-quality means



Synthetic descriptive chart; verify the underlying data before use.

Trait means: ratings 1-5, drawn from a zero baseline.

Reading guide: the top chart plots the 1-5 answer rating for each administered scenario, and the lower chart shows the mean of each answer-quality trait over those scenarios. Different scales are not combined.

Competency deep dive 2 / 4

Data-informed judgment · reference score

Data-informed judgment* 51.69*

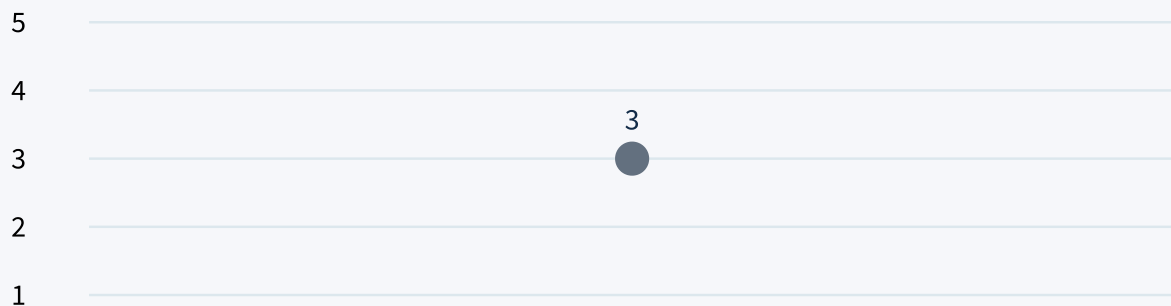
51.69 / 100

One generated answer, not a calibrated estimate



Descriptive generated-answer score. * Data-informed judgment has one observation: 51.69/100 is reference only and has no band; it is not a calibrated estimate, eligible relative comparison, or workplace evidence.

Data-informed judgment · one answer



S04 *

1 administered scenario (S04); * means anchor check. Answer-effectiveness input 1–5, not a calibrated competency estimate. One observation: reference only; no person-to-person rank.

Competency deep dive 2 / 5

Data-informed judgment · reference inputs

evidence2.00 / 5

communication3.00 / 5

risk awareness3.00 / 5

execution3.00 / 5

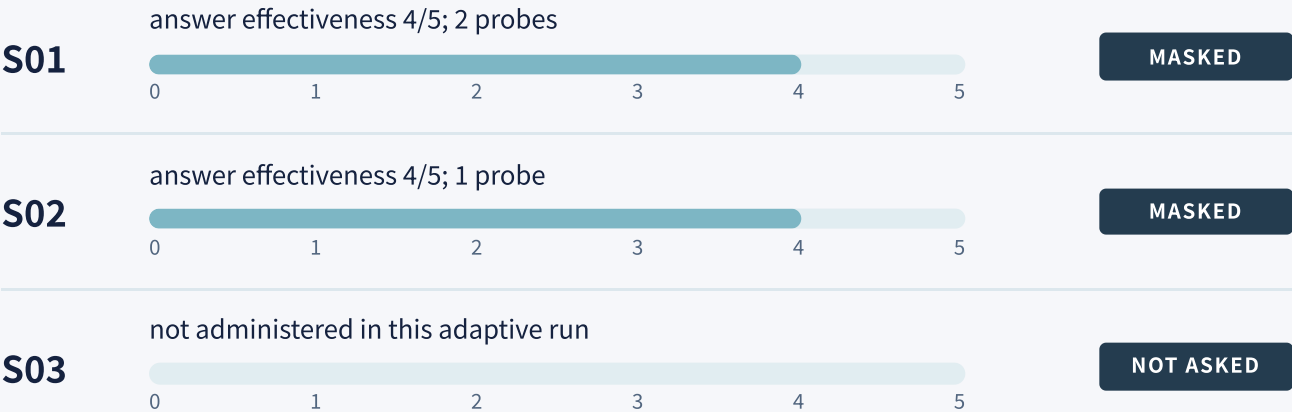
012345

One observation only. Reference input ratings 1–5, drawn from a zero baseline; no relative comparison.

Synthetic descriptive chart; verify the underlying data before use.

S01–S03 · masked answer records

Restricted: masked response evidence, not verbatim quotes



Answer-effectiveness ratings 1–5; rail baseline 0.

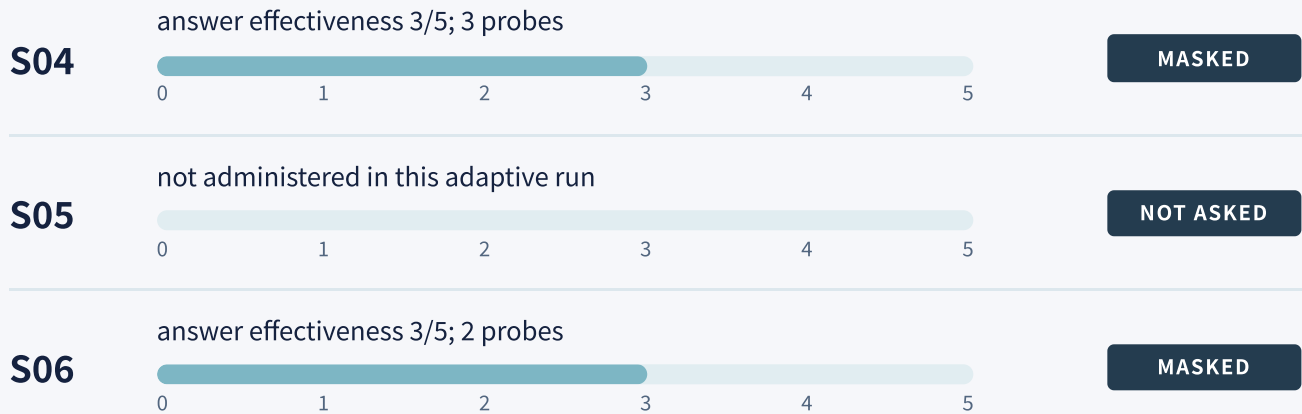
Full dialogue requires separate tenant-bound restricted access.

Synthetic adaptive source, scenario quote matrix. Masked employer view; figures describe generated answers, not job performance.

Competency deep dive 2 / 6

S04–S06 · masked answer records

Restricted: masked response evidence, not verbatim quotes



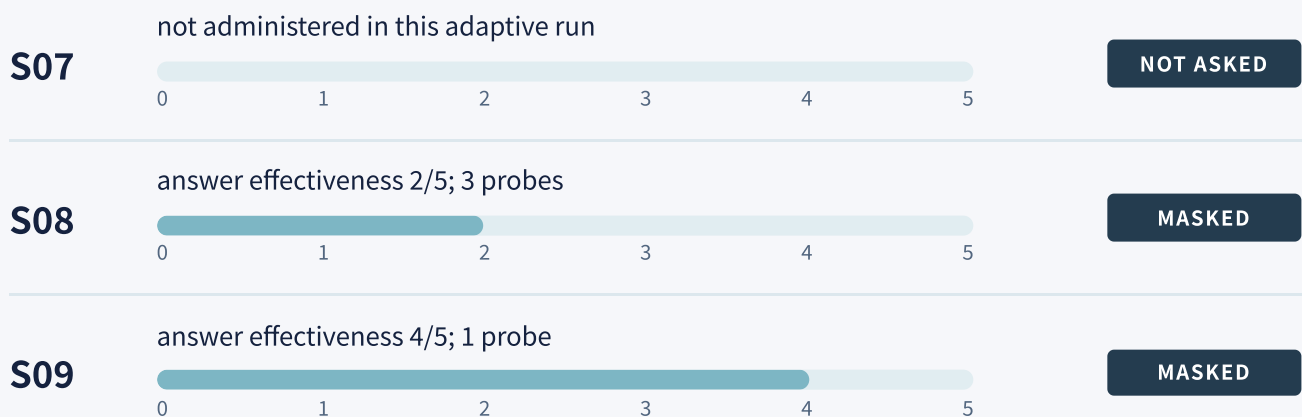
Answer-effectiveness ratings 1-5; rail baseline 0.

Full dialogue requires separate tenant-bound restricted access.

Synthetic adaptive source, scenario quote matrix. Masked employer view; figures describe generated answers, not job performance.

S07–S09 · masked answer records

Restricted: masked response evidence, not verbatim quotes



Answer-effectiveness ratings 1-5; rail baseline 0.

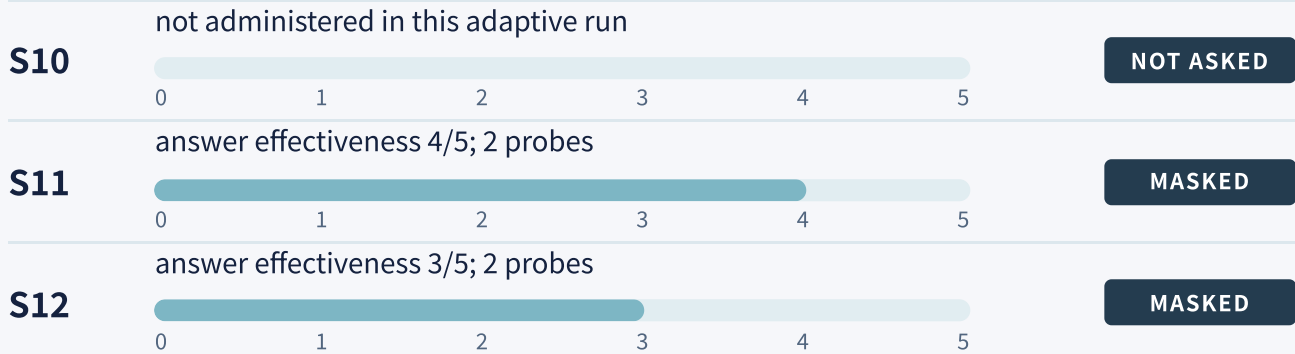
Full dialogue requires separate tenant-bound restricted access.

Synthetic adaptive source, scenario quote matrix. Masked employer view; figures describe generated answers, not job performance.

Competency deep dive 2 / 7

S10–S12 · masked answer records

Restricted: masked response evidence, not verbatim quotes

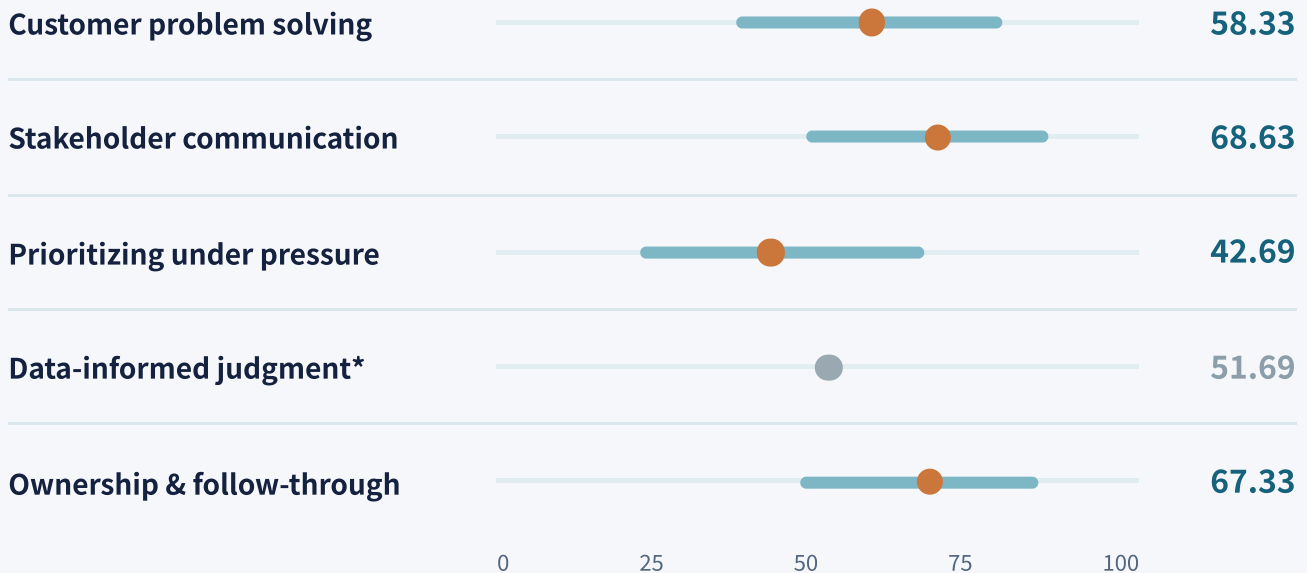


Answer-effectiveness ratings 1-5; rail baseline 0.

Full dialogue requires separate tenant-bound restricted access.

Synthetic adaptive source, scenario quote matrix. Masked employer view; figures describe generated answers, not job performance.

Five descriptive scores and 90% bands



Descriptive score scale 0-100. Dot = score; shaded segment = 90% band; none for a single observation.

Descriptive generated-answer score. * Data-informed judgment has one observation: 51.69/100 is reference only and has no band; it is not a calibrated estimate, eligible relative comparison, or workplace evidence.

CONVERSATION EVIDENCE · JUDGMENT IN THEIR OWN WORDS

A strong and the weakest administered conversation. The difference shows in the follow-up.

① Escalated renewal at risk

Customer problem solving

Scenario

Three weeks before renewal, a key customer says they're evaluating a competitor after two outages. What do you do?

Candidate · first answer (excerpt)

First I'd confirm with engineering what caused both outages and where things stand, then book a call with the customer lead within 24 hours. On the call I'd show what changed and the prevention timeline rather than excuses...

Follow-up probe (generated by the engine)

What if the customer says the decision is basically made?

Candidate · follow-up answer

Then I'd ask what their decision criteria are, pick one we can prove within 30 days and offer a pilot on it. I wouldn't promise what we can't deliver.

What the scoring saw Checks cause before acting, asks back for decision criteria, refuses promises that can't be kept.

Evidence Execution Risk awareness

Effectiveness

4/5

⑧ Quarter-end workload spike

Prioritizing under pressure

Scenario

Quarter end: three renewals, two onboardings and an executive report land in the same week. How do you handle it?

Candidate · first answer (excerpt)

They're all important, so I'd put in evenings to hit every one. Renewals by contract size, and ask onboarding customers to bear with the schedule...

Follow-up probe (generated by the engine)

If one thing still has to slip, who would you talk to and about what?

Candidate · follow-up answer

Hmm... I'd ask my manager whether the report could slip a week. I should have asked that from the start, really.

What the scoring saw Separates urgency but stretches own time instead of deciding what to defer; escalates only after the probe.

Execution Risk awareness

Effectiveness

2/5

How the conversation works

1 · Free text

The candidate answers in their own words, with no options to pick from.

2 · Follow-up probes

Probes the weakest point in the answer. This candidate got 18.

3 · Anchor check

Only when an answer is unclear: pick the best and worst of four actions. Used 3 times.

Conversations are excerpts. The full transcript is available in the admin view.

All 14 bank scenarios · 9 administered

① Escalated renewal at risk

4/5 · probes 2

② Conflicting asks from sales and product

4/5 · probes 1

③ Three urgent tickets, one hour

— NOT ASKED

④ Churn dashboard disagrees with field notes

3/5 · probes 3 · ◆

⑤ A follow-up you promised slipped

— NOT ASKED

⑥ Customer demands an unsupported feature

3/5 · probes 2

⑦ Bad news for an executive sponsor

— NOT ASKED

⑧ Quarter-end workload spike

2/5 · probes 3 · ◆

⑨ Handover from a departing colleague

4/5 · probes 1

⑩ Onboarding stalled at a key account

— NOT ASKED

⑪ Team disagreement on an account plan

4/5 · probes 2

⑫ Deadline vs quality trade-off

3/5 · probes 2 · ◆

⑬ Pricing exception request

— NOT ASKED

⑭ Owning the post-incident review

4/5 · probes 2

Colour = competency · ◆ = anchor check · grey = not administered in this adaptive run. Each conversation's quote is on the quote evidence board.

Trait ratings, strongest vs weakest conversation

Evidence



4 vs 2

Communication



3 vs 2

Risk awareness



4 vs 1

Execution



4 vs 2

Blue = ① Escalated renewal at risk · orange = ⑧ Quarter-end workload spike · scale 1 to 5 · descriptive input ratings, not competency scores.

Conversation evidence

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Two generated ratings to check in interview

S11 Team disagreement on an account plan

Highest generated answer rating among the 9 administered scenarios: 4/5 (S01, S02, S09, S11 and S14 tie). Example S11: two follow-up probes, no anchor check. Input ratings: evidence 3, communication 4, risk awareness 5, execution 5. Dialogue withheld from this employer proof; the numbers do not show why the answer earned its rating.

S08 Quarter-end workload spike

Lowest generated answer rating: 2/5 (S08); three follow-up probes, anchor check administered. Input ratings: evidence 2, communication 2, risk awareness 1, execution 2. Dialogue withheld; do not infer a workplace weakness from this synthetic response.

Conversation path and coverage

Nine of the 14 bank scenarios were asked in this adaptive run (a dash means not asked); the engine stopped when the 0.30 SE target was met. A follow-up count describes the protocol, not the person.

Coverage 9 of 14 bank scenarios

Follow-ups 18

Anchor checks 3

ID	Scenario	Rating / 5	Probes	Anchor
S01	Escalated renewal at risk	4/5	2	
S02	Conflicting asks from sales and product	4/5	1	
S03	Three urgent tickets, one hour	–	–	
S04	Churn dashboard disagrees with field notes	3/5	3	◆
S05	A follow-up you promised slipped	–	–	
S06	Customer demands an unsupported feature	3/5	2	
S07	Bad news for an executive sponsor	–	–	
S08	Quarter-end workload spike	2/5	3	◆
S09	Handover from a departing colleague	4/5	1	
S10	Onboarding stalled at a key account	–	–	
S11	Team disagreement on an account plan	4/5	2	
S12	Deadline vs quality trade-off	3/5	2	◆
S13	Pricing exception request	–	–	
S14	Owning the post-incident review	4/5	2	

QUOTE EVIDENCE BOARD · 1/2 OBSERVED

Every administered score has the candidate's own words behind it. 9 of 14 bank scenarios were asked; the rest are marked.

① Escalated renewal at risk · **Customer problem solving**

Eff. 4/5 · probes 2

““I'd book a call within 24 hours and lead with what changed and the fix timeline, not excuses.””

Scoring saw: Checks cause → talks to customer → limits promises

Eviden 4 Commun 3 Risk a 4 Execut 4

② Conflicting asks from sales and product · **Stakeholder communication**

Eff. 4/5 · probes 1

““I'd put both asks on one page side by side and ask the sales and product leads for 15 minutes together, framed on customer impact.””

Scoring saw: Puts the conflict on one page; uses customer impact as the test

Eviden 4 Commun 5 Risk a 4 Execut 4

③ Three urgent tickets, one hour · **Prioritizing under pressure**

not administered in this adaptive run

“No excerpt: this bank scenario was not asked in this run.”

④ Churn dashboard disagrees with field notes · **Data-informed judgment**

Eff. 3/5 · probes 3 · anchor

““Before trusting either, I'd check the sample size and time window on both sides... after that I'd need to think more.””

Scoring saw: Checking sample and time window is a good sign; next step unclear · one observation

Eviden 2 Commun 3 Risk a 3 Execut 3

⑤ A follow-up you promised slipped · **Ownership & follow-through**

not administered in this adaptive run

“No excerpt: this bank scenario was not asked in this run.”

⑥ Customer demands an unsupported feature · **Customer problem solving**

Eff. 3/5 · probes 2

““I'd be clear it's not supported, then ask what problem they want it to solve and find a workaround that exists today.””

Scoring saw: Says no, then asks what the real problem is

Eviden 3 Commun 4 Risk a 3 Execut 2

⑦ Bad news for an executive sponsor · **Stakeholder communication**

not administered in this adaptive run

“No excerpt: this bank scenario was not asked in this run.”

The four traits the scoring reads

Evidence

mean 3.22

Is there a fact or check behind the claim?

▲ ① 4/5 · Scoring saw: Checks cause → talks to customer → limits promises

▼ ④ 2/5 · Scoring saw: Checking sample and time window is a good sign; next step unclear · one observation

Communication

mean 3.56

Who is told what, in what order?

▲ ② 5/5 · Scoring saw: Puts the conflict on one page; uses customer impact as the test

▼ ⑤ 2/5 · Scoring saw: Stretches own time; doesn't choose what to defer

Risk awareness

mean 3.33

Does the answer see what could go wrong?

▲ ⑩ 5/5 · Scoring saw: Listens, then narrows the disagreement

▼ ⑧ 1/5 · Scoring saw: Stretches own time; doesn't choose what to defer

Execution

mean 3.44

Are dates, owners and next steps concrete?

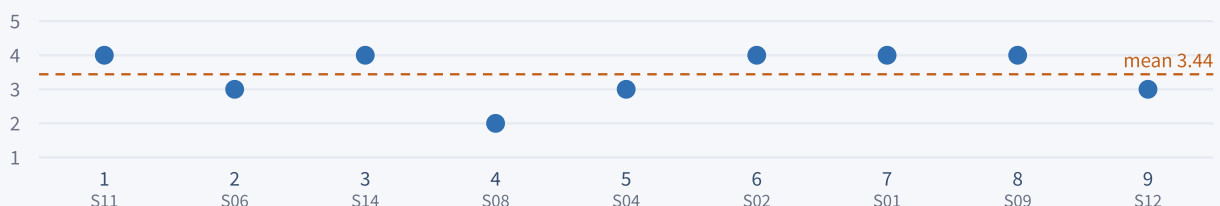
▲ ⑩ 5/5 · Scoring saw: Listens, then narrows the disagreement

▼ ⑥ 2/5 · Scoring saw: Says no, then asks what the real problem is

▲ = scenario where this trait was highest · ▼ = where it was lowest.

Excerpts in this sample are synthetic. A live report quotes the candidate verbatim; full transcripts are in the admin view.

Effectiveness rating per administered scenario



Dots follow the administered order; the dashed line is the mean of the ratings shown. Not-asked scenarios have no rating. Descriptive, no pass mark.

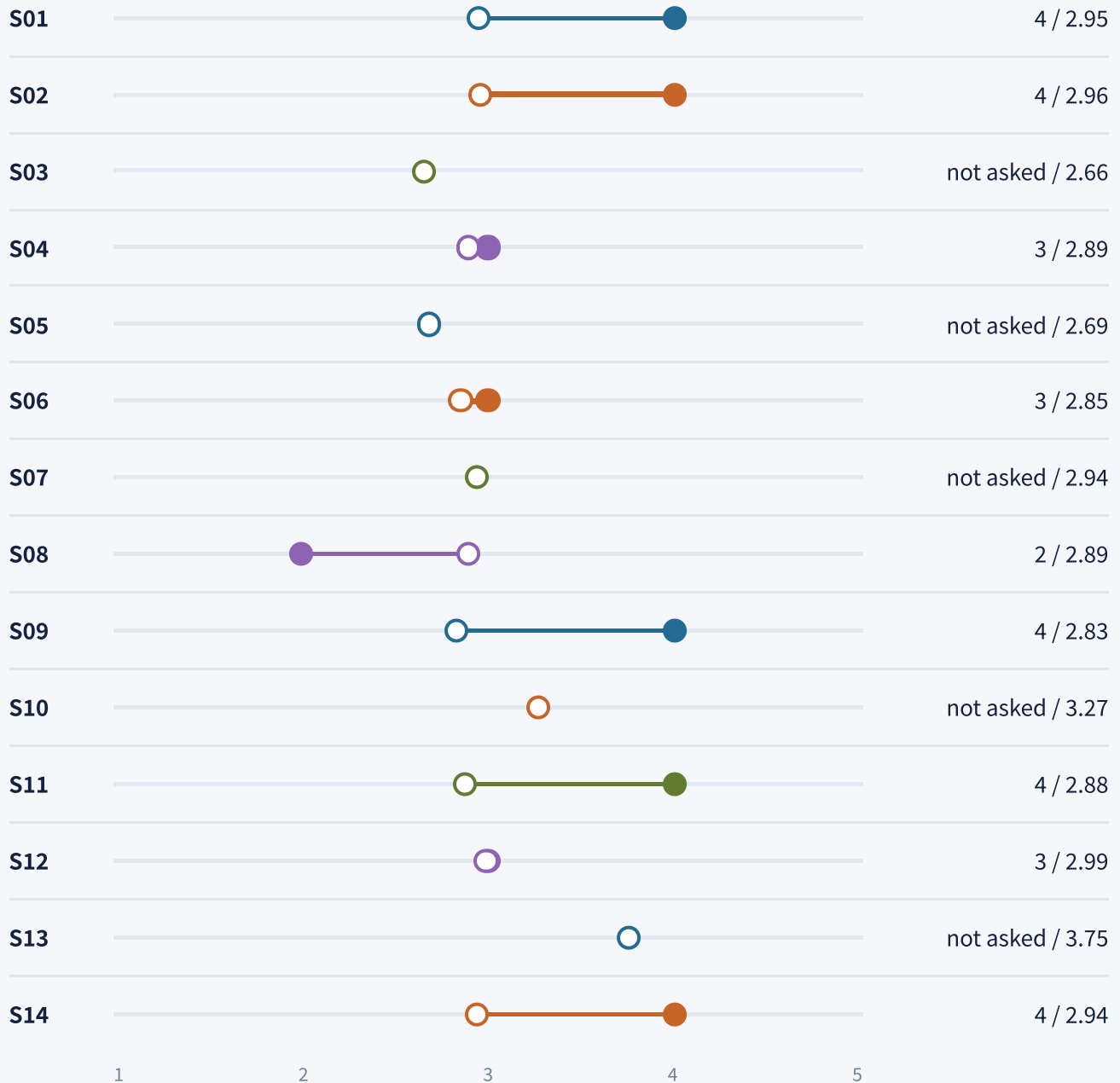
Candidate comparison evidence / 3

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Same-scenario candidate versus cohort mean

● Filled = individual

○ Ring = cohort item mean



Filled is the individual 1–5 rating (administered scenarios only) and the ring is the cohort mean for this same item; n varies by scenario in the adaptive cohort. The difference is not a competency rank.

Source: C-2291 answer ratings and scenario means, administered scenarios

Masked answer evidence board

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

14 bank scenarios · 9 administered

S01

Escalated renewal at risk

Masked answer · generated effectiveness 4/5

S02

Conflicting asks from sales and product

Masked answer · generated effectiveness 4/5

S03

Three urgent tickets, one hour

Not administered in this adaptive run

S04

Churn dashboard disagrees with field notes

Masked answer · generated effectiveness 3/5

S05

A follow-up you promised slipped

Not administered in this adaptive run

S06

Customer demands an unsupported feature

Masked answer · generated effectiveness 3/5

S07

Bad news for an executive sponsor

Not administered in this adaptive run

S08

Quarter-end workload spike

Masked answer · generated effectiveness 2/5

S09

Handover from a departing colleague

Masked answer · generated effectiveness 4/5

S10

Onboarding stalled at a key account

Not administered in this adaptive run

S11

Team disagreement on an account plan

Masked answer · generated effectiveness 4/5

S12

Deadline vs quality trade-off

Masked answer · generated effectiveness 3/5

S13

Pricing exception request

Not administered in this adaptive run

S14

Owning the post-incident review

Masked answer · generated effectiveness 4/5

Masked answer evidence and four input axes

The employer-facing report does not reproduce full candidate dialogue. Restricted access to any raw answer needs separate tenant-bound review and masking. The four numbers below are generated answer-quality inputs, not personality traits.

Four input axes Evidence · communication · risk awareness · execution; 1–5 generated ratings

Four answer-quality inputs

These four independently recorded synthetic 1–5 inputs describe answers, not personality. Highest and lowest examples below are scenario-specific, not predictions about work.

Input	Mean / 5	Selected high-rating scenario	Selected low-rating scenario
Evidence shown	3.22	S01 · 4/5	S04 · 2/5
Communication	3.56	S02 · 5/5	S08 · 2/5
Risk awareness	3.33	S11 · 5/5	S08 · 1/5
Execution detail	3.44	S11 · 5/5	S06 · 2/5

When several scenarios share the same highest or lowest input rating, the table names the first by scenario order. These are selected generated examples, not unique peaks or candidate quotations.

The low-scoring quotes share a pattern: not choosing what to defer, and telling the customer only after the probe.

⑧ Quarter-end workload spike · **Prioritizing under pressure**

Eff. 2/5 · probes 3 · anchor

““They're all important, so I'd put in evenings to hit every one.””

Scoring saw: Stretches own time; doesn't choose what to defer

Eviden 2Commun 2Risk a 1Execut 2

⑨ Handover from a departing colleague · **Ownership & follow-through**

Eff. 4/5 · probes 1

““I'd build a list of open commitments first and send each customer an intro in my name within the first week.””

Scoring saw: Lists commitments; first-week contact

Eviden 4Commun 4Risk a 3Execut 4

⑩ Onboarding stalled at a key account · **Customer problem solving**

not administered in this adaptive run

“No excerpt: this bank scenario was not asked in this run.”

⑪ Team disagreement on an account plan · **Stakeholder communication**

Eff. 4/5 · probes 2

““I'd hear the other view out first, then suggest we test only the part we disagree on with data.””

Scoring saw: Listens, then narrows the disagreement

Eviden 3Commun 4Risk a 5Execut 5

⑫ Deadline vs quality trade-off · **Prioritizing under pressure**

Eff. 3/5 · probes 2 · anchor

““I'd keep the deadline even if quality drops a bit... I suppose I should tell the customer what's being dropped.””

Scoring saw: Deadline first; tells the customer only after the probe

Eviden 3Commun 4Risk a 2Execut 3

⑬ Pricing exception request · **Prioritizing under pressure**

not administered in this adaptive run

“No excerpt: this bank scenario was not asked in this run.”

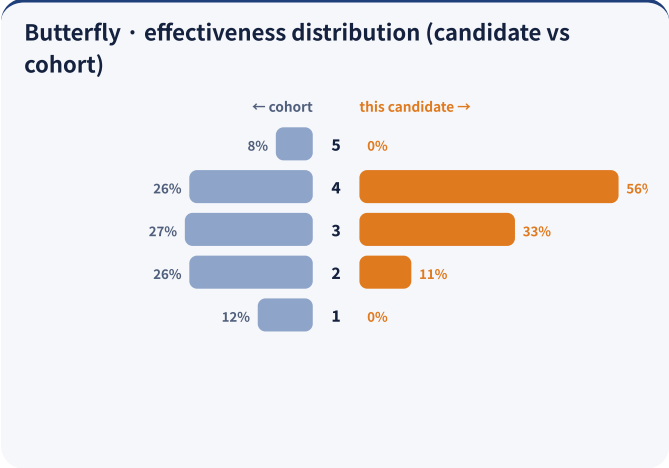
⑭ Owning the post-incident review · **Ownership & follow-through**

Eff. 4/5 · probes 2

““I'd own the review doc and share the timeline, cause and owner for each fix by Friday.””

Scoring saw: Takes ownership; names date and owners

Eviden 4Commun 3Risk a 5Execut 4



What the low scores share

Scenario	Eff.	Probes	Anchor	Lowest trait
④ Churn dashboard disagrees with field notes	3	3	◆	Evidence (2)
⑥ Customer demands an unsupported feature	3	2		Execution (2)
⑧ Quarter-end workload spike	2	3	◆	Risk (1)
⑫ Deadline vs quality trade-off	3	2	◆	Risk (2)

Trait chip colour: red 1 → green 5.

Masked answer records 2 / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Masked scenario ratings · S08–S14

S08 Quarter-end workload spike

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 2/5. Do not use unrelated personal information in a decision.

S09 Handover from a departing colleague

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 4/5. Do not use unrelated personal information in a decision.

S10 Onboarding stalled at a key account

Not administered in this adaptive run; no answer or rating exists for this scenario. Do not use unrelated personal information in a decision.

S11 Team disagreement on an account plan

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 4/5. Do not use unrelated personal information in a decision.

S12 Deadline vs quality trade-off

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 3/5. Do not use unrelated personal information in a decision.

S13 Pricing exception request

Not administered in this adaptive run; no answer or rating exists for this scenario. Do not use unrelated personal information in a decision.

S14 Owning the post-incident review

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 4/5. Do not use unrelated personal information in a decision.

Masked answer records 2 / 2

4 SCENARIO RECORDS

10 FOLLOW-UPS

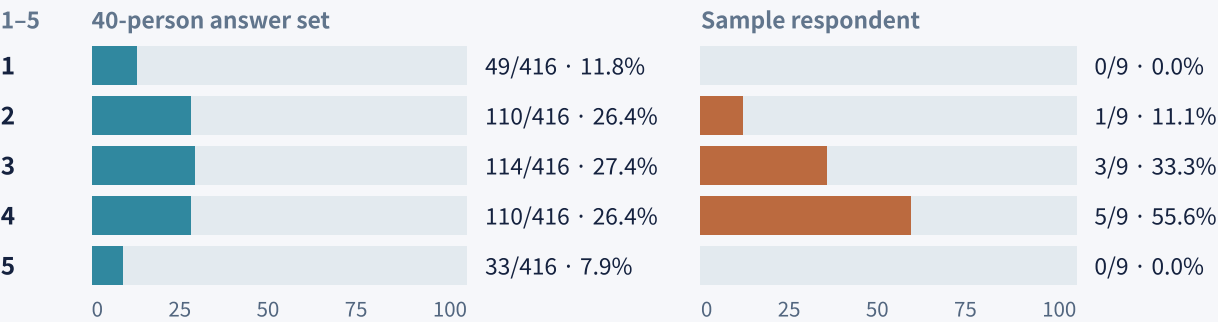
3 of 4 ANCHOR CHECKS YES

Low-scoring scenarios: source cues

Rows describe generated answer ratings only, not a behavioral pattern or cause.

Scenario	Rating	Follow-ups	Anchor	Lowest answer trait
S04	3/5	3	yes	evidence
S06	3/5	2	no	execution
S08	2/5	3	yes	risk awareness
S12	3/5	2	yes	risk awareness

Generated answer ratings · respondent versus 40-person set



Both rail scales: 0-100% of answers within the named set (416 or 9 answers).

The sets do not overlap; C-2291 is a separate run, not one of the 40.
Orange: 9 answers from C-2291; teal: 416 answers from 40 completers. C-2291 is not one of the 40.
Left/right bar lengths encode share within each named answer set. Descriptive 1–5 effectiveness input, no hiring inference.

Candidate comparison evidence / 4

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Fourteen bank-scenario means and third shifts



Top/bottom thirds form on the other answers of each person who answered that scenario; shown only where n is 12 or more. This is item review, not candidate grading.

Source: scenario means and n, S01-S14, 40 completers

Reading guide: each scenario has a mean answer rating and a shift between the top and bottom thirds, formed on the other answers per person ($n \geq 12$ only). This is item review from the same synthetic 40-person sample, not candidate grading.

ANSWER QUALITY · FOUR TRAITS

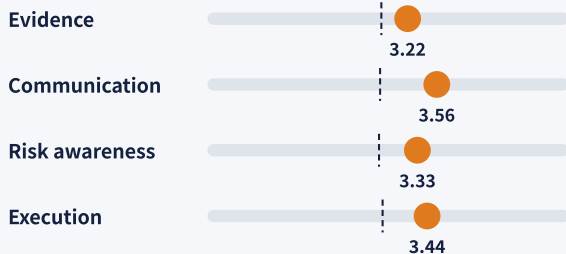
Ratings vary by scenario. The lowest one sits in the workload scenario (⑧).

Scenario × four answer-quality traits (1–5)

	Evidence	Communication	Risk awareness	Execution	Overall
① Escalated renewal at risk	4	3	4	4	4
② Conflicting asks from sales and product	4	5	4	4	4
③ Three urgent tickets, one hour	not administered in this adaptive run				–
④ Churn dashboard disagrees with field notes	2	3	3	3	3
⑤ A follow-up you promised slipped	not administered in this adaptive run				–
⑥ Customer demands an unsupported feature	3	4	3	2	3
⑦ Bad news for an executive sponsor	not administered in this adaptive run				–
⑧ Quarter-end workload spike	2	2	1	2	2
⑨ Handover from a departing colleague	4	4	3	4	4
⑩ Onboarding stalled at a key account	not administered in this adaptive run				–
⑪ Team disagreement on an account plan	3	4	5	5	4
⑫ Deadline vs quality trade-off	3	4	2	3	3
⑬ Pricing exception request	not administered in this adaptive run				–
⑭ Owning the post-incident review	4	3	5	4	4

Left colour bar = competency the scenario measured · green 4–5 · yellow 3 · red 1–2 · overall = effectiveness for that scenario

Trait means vs cohort mean

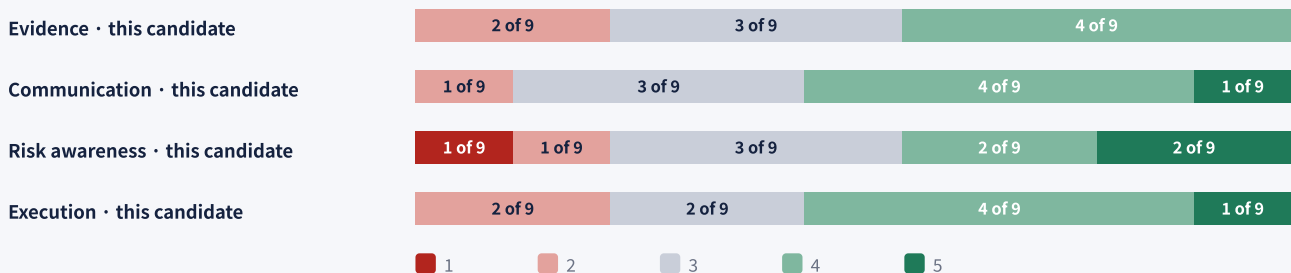


Orange = candidate mean · dashed = cohort mean (40 completers, all answered scenarios)

Pattern

- Evidence: highest 4 in ①②⑨⑭; lowest 2 in ④⑧.
- Communication: highest 5 in ②; lowest 2 in ⑧.
- Risk awareness: highest 5 in ⑪⑭; lowest 1 in ⑧.
- Execution: highest 5 in ⑪; lowest 2 in ⑥⑧.

Trait score distribution: this candidate, 9 administered scenarios (1–5)



Answer quality traits / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Source ledger · one row per scenario

The four generated 1–5 answer-quality inputs per scenario. The next sheet presents the same 36 inputs (9 administered scenarios x 4) as a heatmap; this is not a workplace trait assessment.

ID	Scenario	Group	Evid.	Comm.	Risk	Exec.	Eff.
S01	Escalated renewal at risk	C1	4	3	4	4	4
S02	Conflicting asks from sales and product	C2	4	5	4	4	4
S03	Three urgent tickets, one hour (not asked)	C3	–	–	–	–	–
S04	Churn dashboard disagrees with field notes	C4	2	3	3	3	3
S05	A follow-up you promised slipped (not asked)	C5	–	–	–	–	–
S06	Customer demands an unsupported feature	C1	3	4	3	2	3
S07	Bad news for an executive sponsor (not asked)	C2	–	–	–	–	–
S08	Quarter-end workload spike	C3	2	2	1	2	2
S09	Handover from a departing colleague	C5	4	4	3	4	4
S10	Onboarding stalled at a key account (not asked)	C1	–	–	–	–	–
S11	Team disagreement on an account plan	C2	3	4	5	5	4
S12	Deadline vs quality trade-off	C3	3	4	2	3	3
S13	Pricing exception request (not asked)	C3	–	–	–	–	–
S14	Owning the post-incident review	C5	4	3	5	4	4

Group code (assigned, not inferred): C1 Customer problem solving, C2 Stakeholder communication, C3 Prioritizing under pressure, C4 Data-informed judgment, C5 Ownership & follow-through. Eff. is a separate generated 1–5 effectiveness input; it is not the mean of four trait inputs.

Two answers worth comparing

S11 (team disagreement on an account plan) has answer-quality inputs 3/4/5/5 and effectiveness 4. S08 (quarter-end workload spike) has 2/2/1/2 and effectiveness 2. These are different prompts, so the contrast is not evidence that workload caused a drop.

Interview check: Ask for a real example of how they prioritized competing requests under time pressure and caught a risk before committing to a response.

Answer quality traits / 2

Scenario trait heatmap

	1	2	3	4
S01	4	3	4	4
S02	4	5	4	4
S03	–	–	–	–
S04	2	3	3	3
S05	–	–	–	–
S06	3	4	3	2
S07	–	–	–	–
S08	2	2	1	2
S09	4	4	3	4
S10	–	–	–	–
S11	3	4	5	5
S12	3	4	2	3
S13	–	–	–	–
S14	4	3	5	4

Generated rating (1-5) 1 2 3 4 5

Columns: 1 Evidence; 2 Communication; 3 Risk awareness; 4 Execution. Rows S01–S14 follow scenario order; a dash means the scenario was not asked in this adaptive run. Cells are generated answer-quality inputs on 1–5, not personality traits.

Answer quality traits / 3

Answer-quality inputs · means on a 1-5 scale

- 1 evidence: 3.22
- 2 communication: 3.56
- 3 risk awareness: 3.33
- 4 execution: 3.44



Synthetic answer data; no personality or calibrated competency inference.

Synthetic adaptive source, trait radar. Masked employer view; figures describe generated answers, not job performance.

Trait means · candidate and 40-person median

Input	C-2291 mean	40-person median	Δ
evidence	3.22	3.00	+0.22
communication	3.56	3.08	+0.48
risk awareness	3.33	3.00	+0.33
execution	3.44	2.95	+0.49

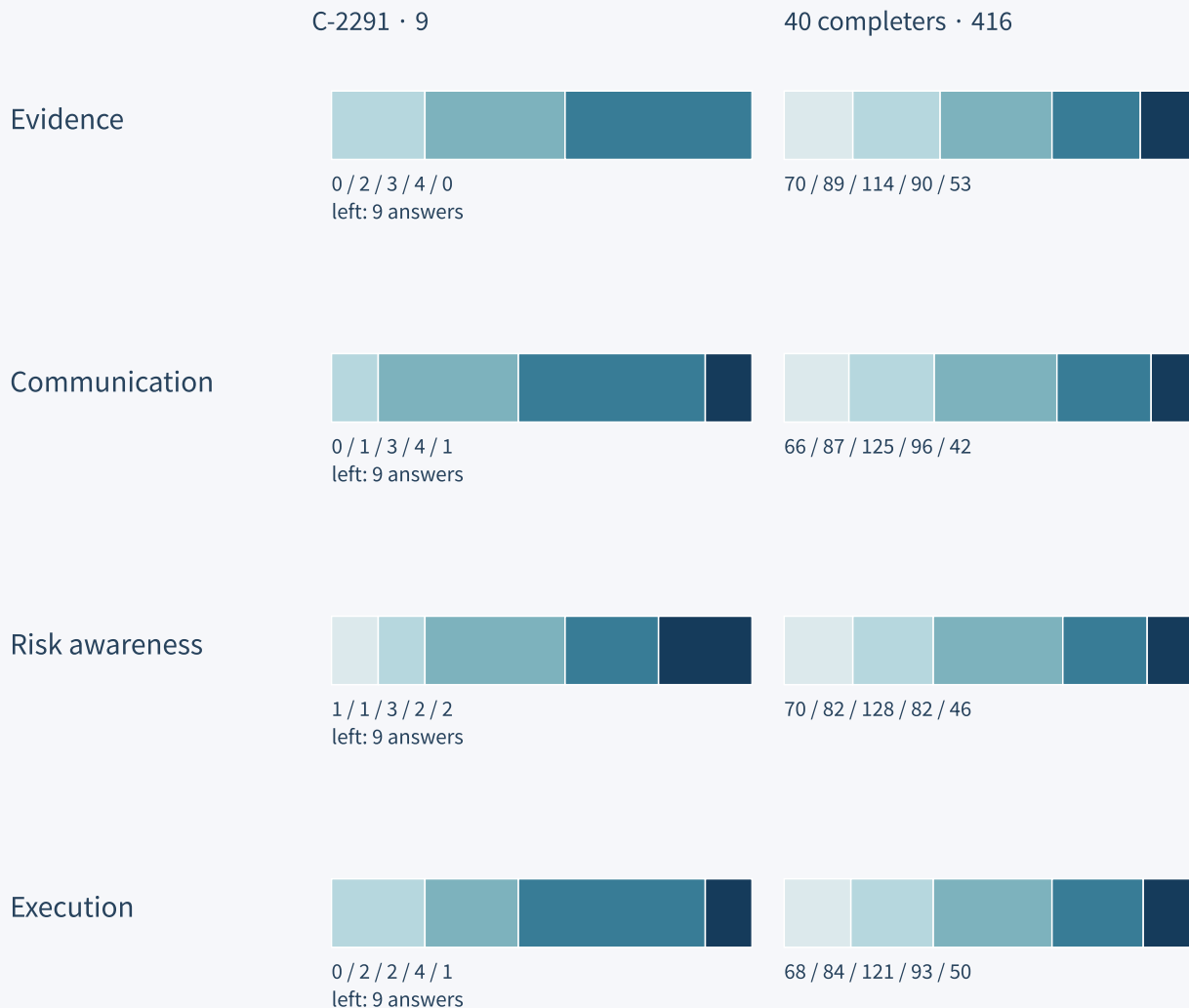
The comparator is the median of the 40 completers' own trait means (C-2291 is not one of the 40); separate sets, descriptive only, no normative rank.

What the reference changes

The execution input has the largest displayed mean-minus-median difference (+0.49); the other three are within 0.48 of their medians. This is a comparison of synthetic answer-quality inputs, not evidence of a stable strength or workplace performance.

Answer quality traits / 4

Four answer-quality traits · full five-level distributions



Rating 1 2 3 4 5 · Each row encodes share within the named answer set.



Each trait has 9 generated respondent ratings and 416 ratings from all 40 completers. Each bar sums to 100%; counts appear in rating order 1/2/3/4/5. The sets do not overlap (C-2291 is a separate run); this is not a hiring comparison.

Reading guide: each bar is one trait's distribution across the five rating levels, shown for C-2291's 9 ratings and for all 40 completers' 416 ratings.

Candidate comparison evidence / 5

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Scenario by four answer traits

	Evidence	Communication	Risk	Execution
S01	4	3	4	4
S02	4	5	4	4
S03	–	–	–	–
S04	2	3	3	3
S05	–	–	–	–
S06	3	4	3	2
S07	–	–	–	–
S08	2	2	1	2
S09	4	4	3	4
S10	–	–	–	–
S11	3	4	5	5
S12	3	4	2	3
S13	–	–	–	–
S14	4	3	5	4

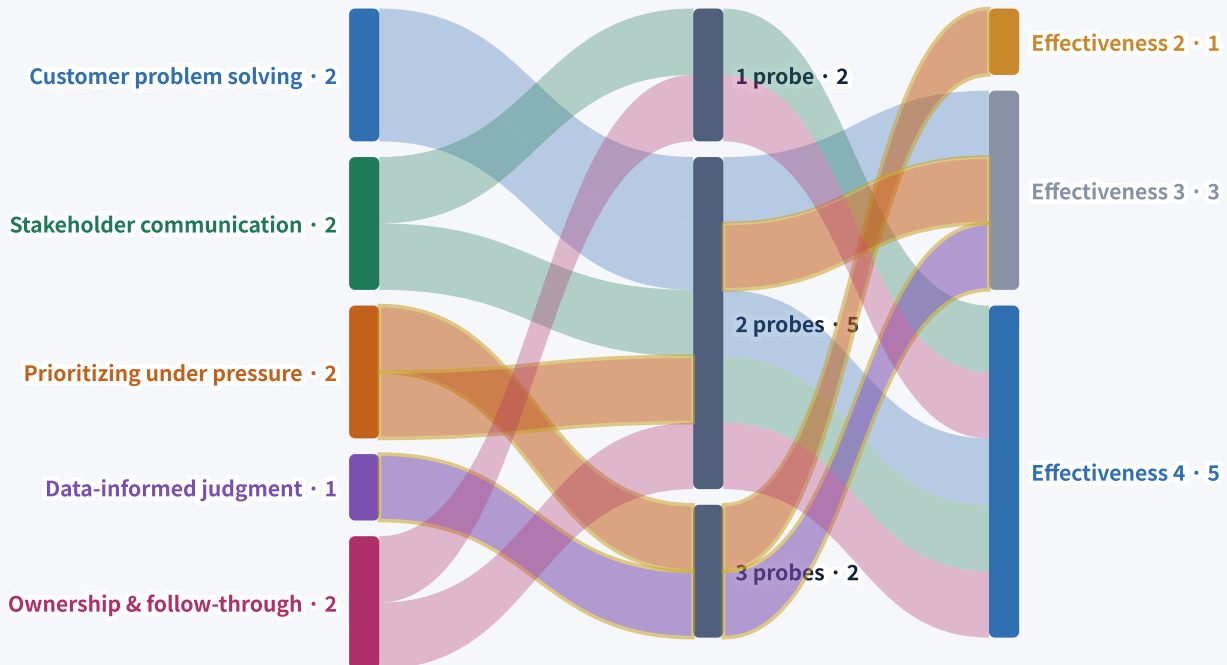
Generated rating (1-5)  1  2  3  4  5

Each cell is one generated 1–5 answer-quality rating, not a personality trait.

Source: answer-quality trait ratings for C-2291, 9 administered of 14 bank scenarios (dash = not asked)

The 2 administered scenarios with three probes averaged 2.5; the overall mean of the 9 is 3.44.

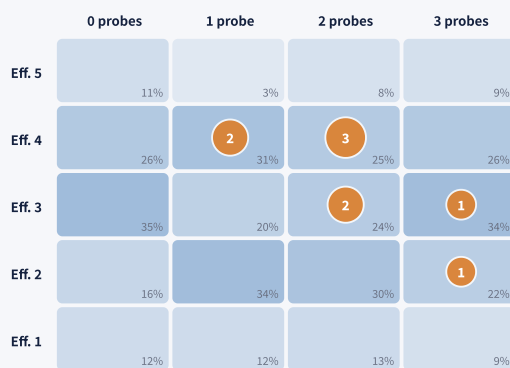
Sankey · competency → probes → effectiveness



Gold outline = scored with an anchor check

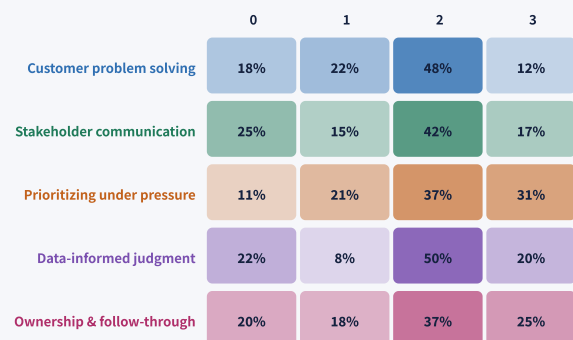
One strand per scenario. Gold outline = anchor check.

Bubble matrix · probes × effectiveness



Orange = this candidate's scenario count · background % = share of all administered cohort answers in that cell (per column).

Cohort · probe share by competency



416 administered answers from 40 completers. Association, not cause.

Follow-up probes / 1

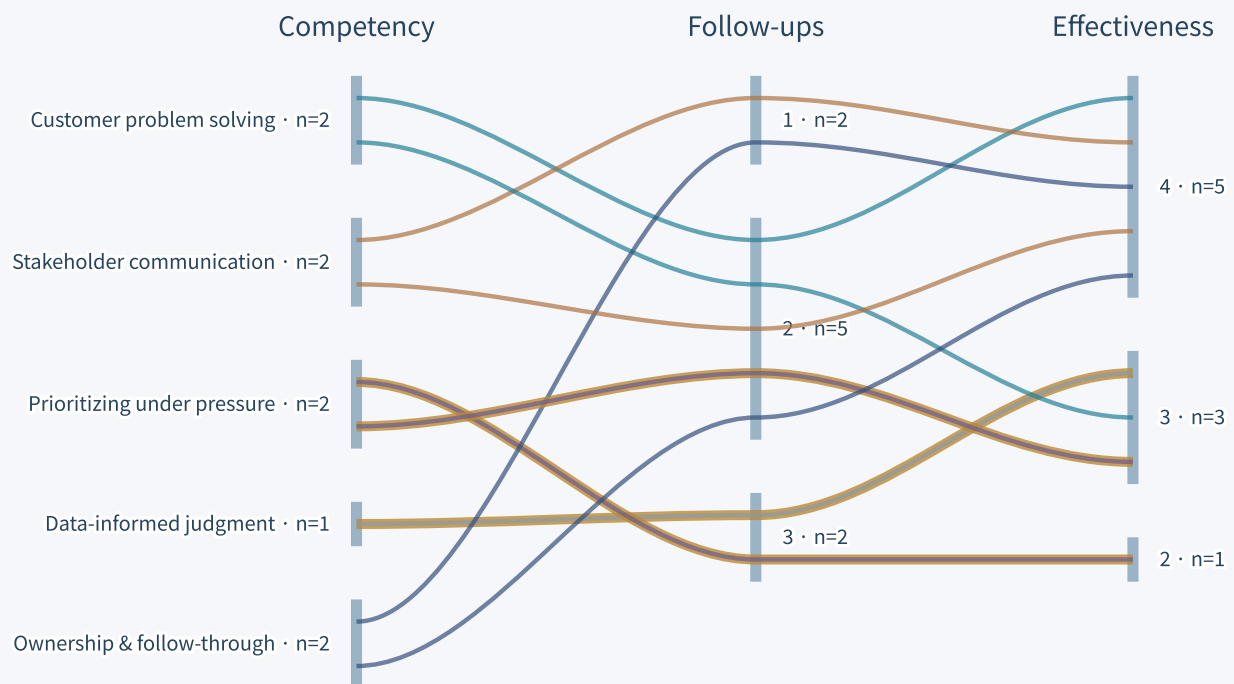
Follow-up and anchor counts

Follow-up probes	18
Anchor checks	3
Overall theta SE	0.30

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Nine administered scenario strands · competency → probes → rating



One strand = one generated answer · gold outline = anchor check

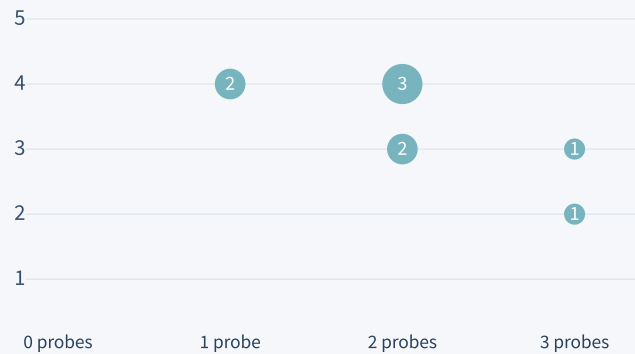
Each administered scenario stays linked across all three columns. Colors identify competency; gold outlines mark three anchor checks. Category counts refer to answers, not total probes. Crossings do not imply cause, ability or job performance. This is the synthetic C-2291 adaptive record, replayed from fixed ratings; it is not a live process.

Reading guide: each of the nine strands runs from competency to probes to rating, and gold outlines mark the three anchor checks.

Follow-up probes / 2

Probe effectiveness bubble

Answer effectiveness 1–5 (vertical), bubble = administered answer count



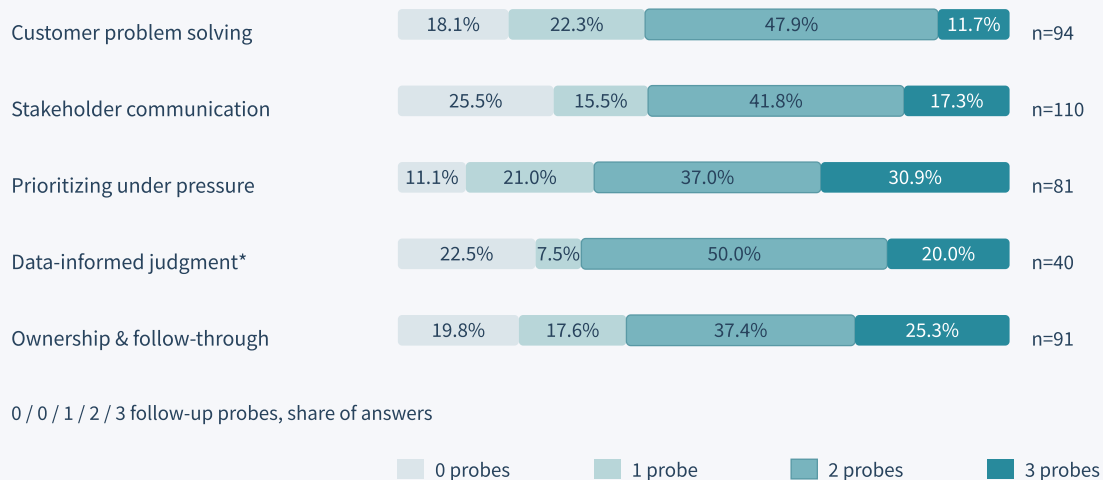
5 occupied cells total 9 administered answers; empty cells mean zero. Radius = $3 + 2.5 \times \text{count}$ in SVG units, so area is not proportional to count. Read the printed numerals for exact frequency; the table below gives a separate descriptive cohort conditional reference.

Synthetic C-2291 adaptive record, probe effectiveness bubble. Masked employer view; figures describe generated answers, not job performance.

Probe count	Cohort answers	Rating 1	Rating 2	Rating 3	Rating 4	Rating 5
0	81	10 (12.3%)	13 (16.0%)	28 (34.6%)	21 (25.9%)	9 (11.1%)
1	74	9 (12.2%)	25 (33.8%)	15 (20.3%)	23 (31.1%)	2 (2.7%)
2	175	22 (12.6%)	53 (30.3%)	42 (24.0%)	44 (25.1%)	14 (8.0%)
3	86	8 (9.3%)	19 (22.1%)	29 (33.7%)	22 (25.6%)	8 (9.3%)

Same-sample conditional reference: percentages divide by all generated answers with that follow-up count (across 40 people). Counts total 416, not 416 independent people. Not causal, calibrated, or a reason to change probe policy.

Probe competency bars

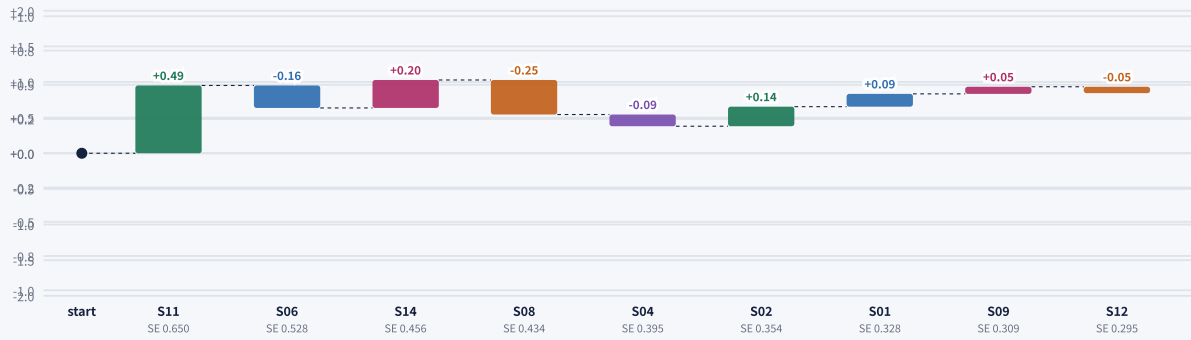


* Data-informed judgment has one scenario in the bank, so at most 40 answers, reference-only for competency-level interpretation. Bar colors identify follow-up count categories, not competency or ability.

Synthetic cohort, n=40 completed people. Figure shows probe competency bars on its labeled scale; descriptive, not a hiring decision. Shares use generated responses in each competency as the denominator. Color shows probe count, not ability.

Overall SE fell from 0.650 after the first scenario to 0.295 after the ninth. The 0.30 target was met, so the run stopped.

Estimate waterfall · overall θ



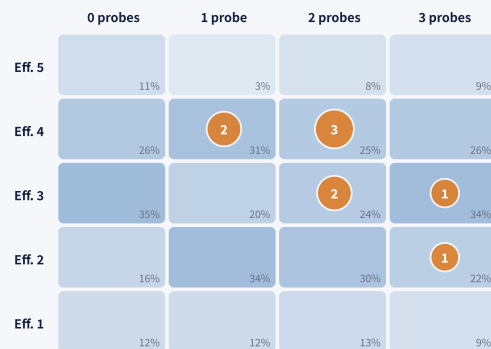
Bars = change in θ estimate per scenario (competency colour), in administered order · gold dot = anchor-checked · SE = overall θ standard error after each scenario
 θ is the engine's internal scale (mean 0), not the 0–100 competency unit.

Observations per competency



Dot = one observation. Fewer than 2 is reference only.

Bubble · this candidate's probes



The same data seen from the path side.

Scenario strip · administered order



Competency · effectiveness (1–5) · probes (dots) · anchor check (gold).

Adaptive scenario path / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Separate the sequence from the estimates

The competency source tables remain on page 7. This section separates the administered sequence, overall-theta updates and competency band widths.

Administered sequence

The engine chose the next scenario after each answer, so the order and the count differ by person. This candidate answered 9 of the 14 and stopped when the overall SE reached 0.30. The path does not by itself show an efficiency gain.

Overall-theta updates

Page 55 shows stepwise EAP changes and cumulative overall-theta intervals for the 9 administered scenarios. These are not intervals for the step changes or individual competencies.

Competency band widths

Page 56 shows the width of each 90% competency band. Their scale differs from overall-theta SE, so do not compare them directly.

Scenario source

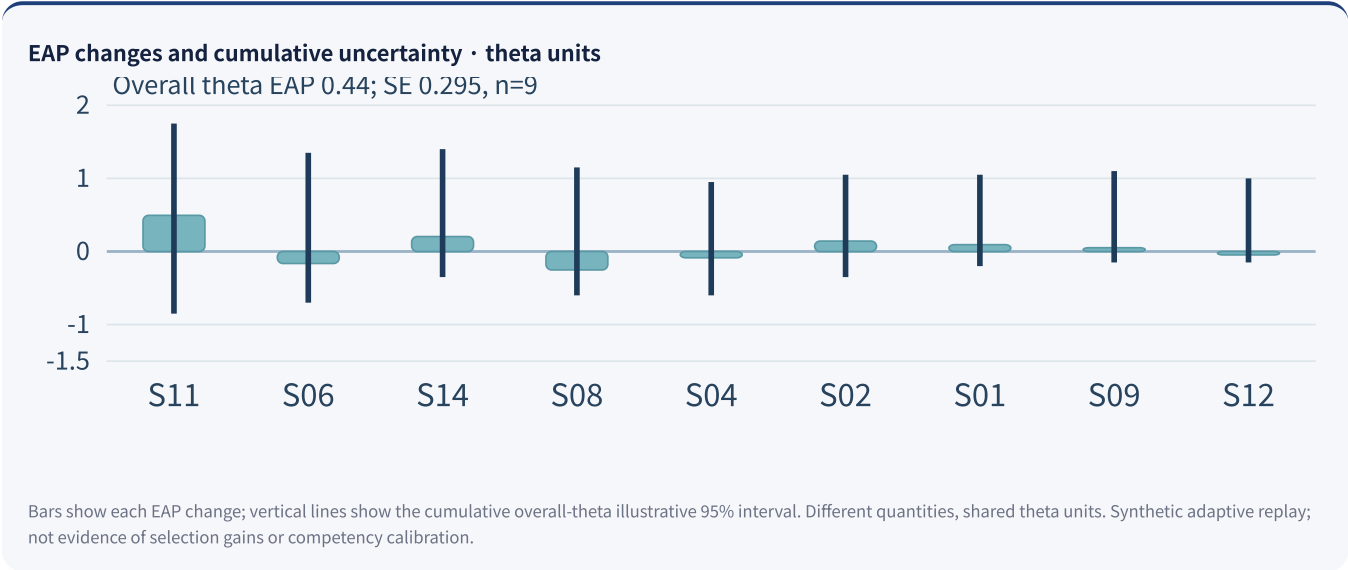
Page 57 checks scenario ratings, probe counts and anchor administration. No answer text is included.

Adaptive theta updates / 2

0.535 TOP CUM. EAP · S14

0.197 LOW CUM. EAP · S04

9 ROWS



9-step source check			
Step	Δ EAP	Cumulative EAP	95% interval (illustrative)
S11	+0.494	0.494	-0.850 to 1.750
S06	-0.164	0.331	-0.700 to 1.350
S14	+0.204	0.535	-0.350 to 1.400
S08	-0.252	0.283	-0.600 to 1.150
S04	-0.086	0.197	-0.600 to 0.950
S02	+0.143	0.340	-0.350 to 1.050
S01	+0.094	0.433	-0.200 to 1.050
S09	+0.052	0.485	-0.150 to 1.100
S12	-0.047	0.438	-0.150 to 1.000

Reading limit

The interval is about cumulative overall theta, not an interval for the step change. The final EAP and SE are descriptive synthetic engine outputs. They do not establish workplace performance, calibrated competency uncertainty, or benefits from adaptive item selection.

Adaptive scenario path / 3

Competency interval width



Synthetic descriptive chart; verify the underlying data before use.

Width of each 90% band on the 0–100 descriptive score scale, not calibrated competency coverage. Data-informed judgment has one answer, so it has no band and no width.

0 20 40 60 70

Interval width: band upper bound minus lower bound on the descriptive 0-100 scale. Rail spans 0-70; not calibrated competency uncertainty.

Read interval width separately from the sequence

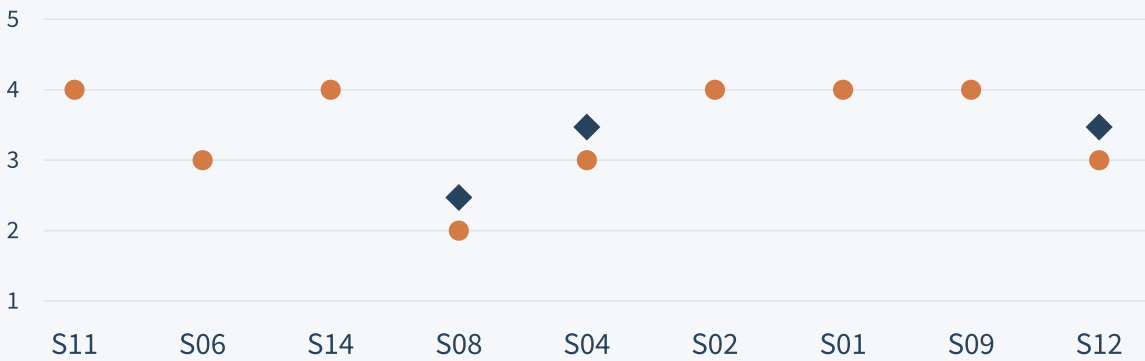
The chart above shows the width of each 90% band. It does not measure the sequence or the effect of follow-up questions.

Interval width	Each bar is the upper bound minus the lower bound on a descriptive score scale. A shorter bar does not establish more accurate calibrated competency measurement.
Administered sequence	Compare answer effectiveness, probe counts and anchor administration against the 9-step source check on page 57.
Follow-up distribution	The probe-count and answer-effectiveness grid is on page 52. It describes the same generated answers, not a causal improvement from probing.
Before use	Calibrated competency intervals need separate validation against human ratings. Do not use these widths for hiring decisions or claims of adaptive-selection gains.

Adaptive scenario path / 4

Scenario sequence strip

9 administered responses in the order asked; anchor check shown by ◆



Vertical: answer effectiveness 1–5; color is not a competency or rank

Synthetic adaptive replay, scenario sequence strip. Masked employer view; figures describe generated answers, not job performance.

9-step source check

Step	Answer effectiveness 1-5	Probe count	Anchor check
S11	4	2	No
S06	3	2	No
S14	4	2	No
S08	2	3	Yes
S04	3	3	Yes
S02	4	1	No
S01	4	2	No
S09	4	1	No
S12	3	2	Yes

Ratings, probe counts and anchor administration for 9 administered synthetic responses, in the order asked. No answer text is included. These are not job-performance measures or evidence of gains from adaptive item selection.

Overall SE fell from 0.650 to 0.395 within five scenarios and ended at 0.295. The 0.30 target was met after nine.

After 1

0.650

first scenario

After 5

0.395

39% below the first

After 7

0.328

49% below the first

Final

0.295

target 0.30 met · stopped after 9

Cohort median SE

0.29

40 completers

Adaptive path · SE falling



Observations per competency

Customer problem

2 obs

Stakeholder co

2 obs

Prioritizing u

2 obs

Data-informed

1 obs · reference only

Ownership & fo

2 obs

How the run stopped

Rule	This candidate
Target $SE \leq 0.30$	met · final SE 0.295
Maximum of 12 scenarios per person	not reached · stopped after 9
Competency with fewer than 2 observations	1 · data-informed judgment, reference only

Stopped because the 0.30 SE target was met, before the 12-scenario maximum.

Reading this page

- Data-informed judgment has 1 observation; 51.69 is reference only
- Cover it in the interview (see the interview kit)
- The step-to-step drops come from this candidate's adaptive path; one run does not show a selection gain

Error drop per scenario

#	Scenario	Competency	SE	Drop	#	Scenario	Competency	SE	Drop
1	Team disagreement on an account plan	Stakeholder communication	0.650	—	8	Handover from a departing colleague	Ownership & follow-through	0.309	+0.020
2	Customer demands an unsupported feature	Customer problem solving	0.528	+0.122	9	Deadline vs quality trade-off	Prioritizing under pressure	0.295	+0.013
3	Owning the post-incident review	Ownership & follow-through	0.456	+0.072					
4	Quarter-end workload spike	Prioritizing under pressure	0.434	+0.022					
5	Churn dashboard disagrees with field notes	Data-informed judgment	0.395	+0.039					
6	Conflicting asks from sales and product	Stakeholder communication	0.354	+0.041					
7	Escalated renewal at risk	Customer problem solving	0.328	+0.026					

The first scenarios cut error the most; each later one adds less.

Measurement precision / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

What the uncertainty refers to

The competency values and band source tables remain on page 7. Overall-theta SE and the 90% competency bands are different quantities.

Overall-theta SE Page 84 traces overall-theta SE as the adaptive answers accumulate. Its units are theta, not competency score differences.

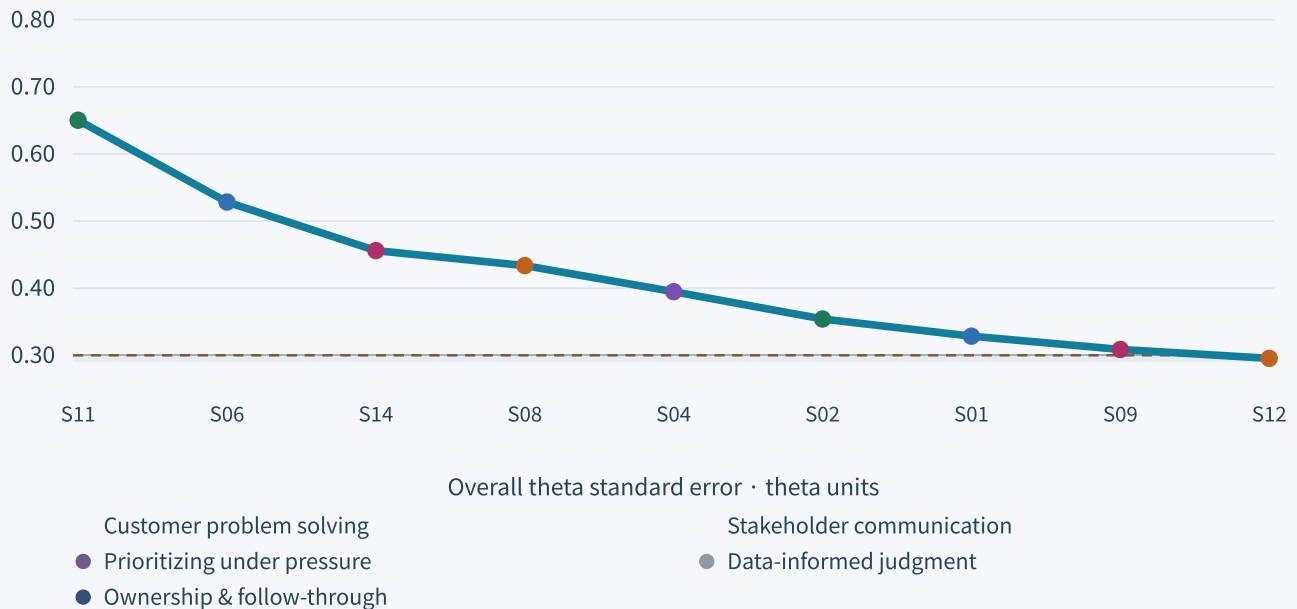
90% bands Competency band widths come from each competency's own posterior on the synthetic engine. They are not calibrated coverage or observed work evidence.

Calibration check Real competency intervals require independent human ratings and a coverage study. The current band widths do not replace that validation.

Comparison limit Do not interpret a sparsely observed axis from width alone. Data-informed judgment has one observation, so it is reference only and has no band.

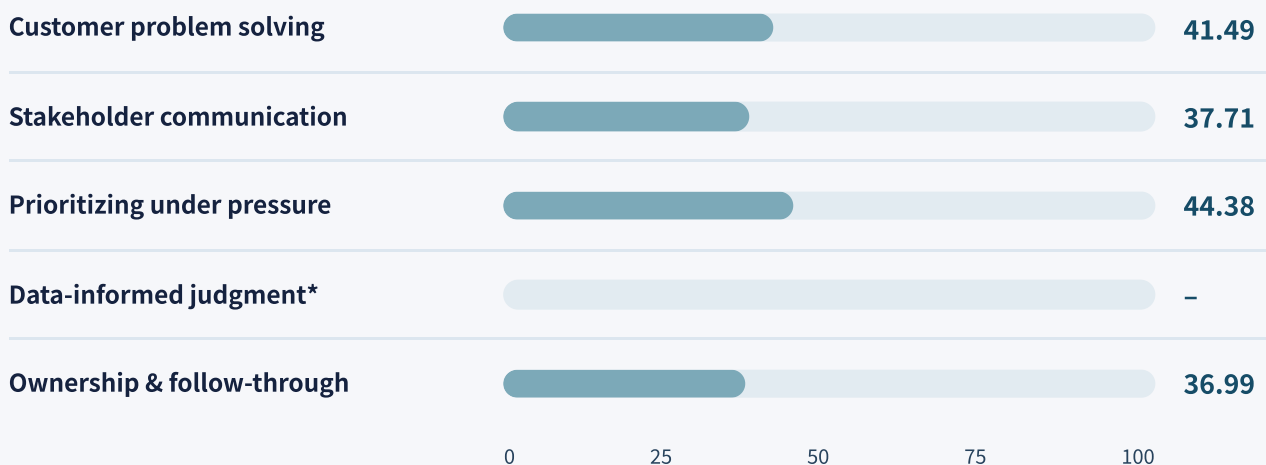
Measurement precision / 2

Overall theta standard error · 13.1



Synthetic adaptive replay. Dot color identifies competency. Pale band: final overall-theta SE p25–p75 of the 40 completers (0.290–0.298); it is not candidate uncertainty at each step or a confidence interval. Dashed brown line: 0.30 target; C-2291 finishes at 0.295, target met after 9 scenarios. Overall theta units, not per-competency 90% intervals.

Competency 90% band widths · 13.2



Band widths on a 0–100 descriptive scale, not calibrated per-competency coverage. * Data-informed judgment has one answer, so it has no band and must not be compared or ranked.

STRENGTHS · DEVELOPMENT AREAS

Strengths repeat across three behaviours. Both development areas link to judgment under pressure, in the 9 administered scenarios.

Strengths

Tailors the message to the listener

Puts two conflicting asks on one page and frames them on customer impact; in ⑪ keeps the team disagreement to the part that can be tested with data.

“I'd put both asks on one page side by side and ask the sales and product leads for 15 minutes together, framed on customer impact.”

Evidence: ②⑪ · communication 4–5

Owns commitments up front

On a handover the candidate lists the open commitments first and contacts each customer in the first week; the same ownership shows in ⑭ with a named owner and date for each fix.

“I'd build a list of open commitments first and send each customer an intro in my name within the first week.”

Evidence: ⑨⑭

Checks the cause before fixing

In the escalated renewal ① the candidate confirmed the cause with engineering before booking the call; in ⑥ asked what problem the customer wanted solved before offering a workaround.

“I'd book a call within 24 hours and lead with what changed and the fix timeline, not excuses.”

Evidence: ①⑥

Development areas · risks to check

Takes everything on when work piles up

Tried to stretch evenings rather than decide what to defer; talking to the manager came up only after a probe. With a bigger book of accounts this can turn into burnout or dropped balls.

“They're all important, so I'd put in evenings to hit every one.”

Evidence: ⑧⑫ · risk awareness 1–2 in ⑧⑫

Keeps the deadline before raising the trade-off

In ⑫ the candidate leaned toward keeping the deadline even if quality dropped, and thought of telling the customer what was dropped only afterwards.

“I'd keep the deadline even if quality drops a bit... I suppose I should tell the customer what's being dropped.”

Evidence: ⑫ · one scenario

Evidence gap

Data-informed judgment was observed in one scenario (④). Asking to check sample and time window when the dashboard disagreed with field notes is a good sign, but one instance can't be generalized. Ask for a real data example in the interview.

Trait profile per competency · small multiples



Dashed = this candidate's overall mean over the 9 administered answers · single-observation competency faded and dashed. Best/low = strongest and weakest scenario in that competency.

Selected generated-answer ratings

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Selected masked scenario ratings · administered scenarios only

S01 Escalated renewal at risk

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 4/5. Do not use unrelated personal information in a decision.

S08 Quarter-end workload spike

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 2/5. Do not use unrelated personal information in a decision.

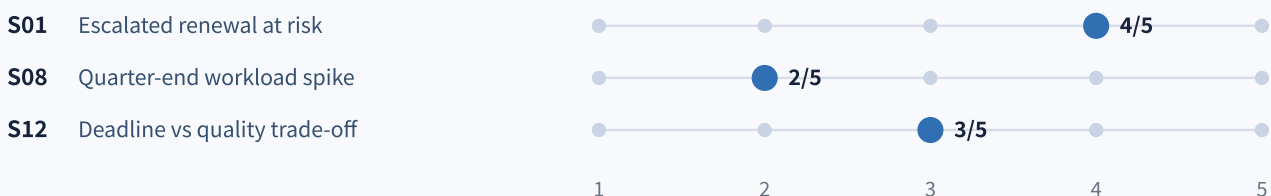
S12 Deadline vs quality trade-off

Restricted masked view. Synthetic scenario answer omitted from this employer report. Effectiveness: 3/5. Do not use unrelated personal information in a decision.

S05 · S07

Not administered in this adaptive run (NOT ASKED). No rating exists for these scenarios.

Ratings at a glance · 1–5 effectiveness scale



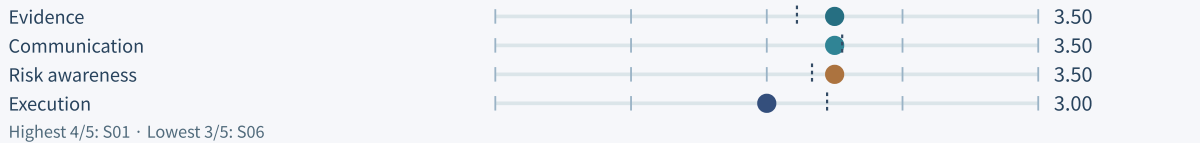
How to read. Large dot = the 1–5 effectiveness rating of that masked answer. A rating is not a competency score and has no band. S05 and S07 were NOT ASKED, so they have no rating.

Trait profile by competency / 2

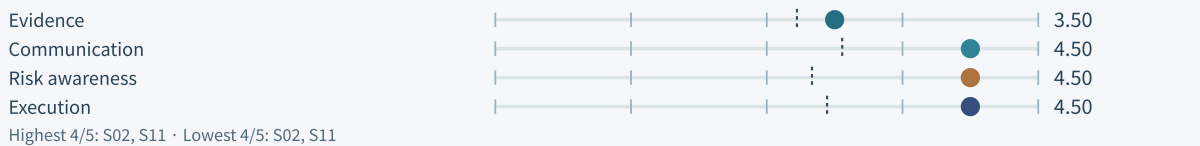
Generated 1–5 answer-quality traits, not an applicant diagnosis. The one-item Data-informed judgment row is reference-only.

Five competencies · four trait means

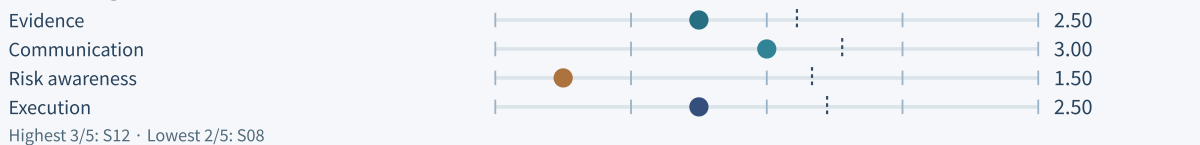
Customer problem solving



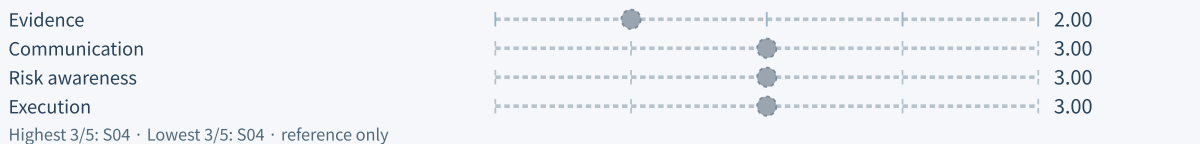
Stakeholder communication



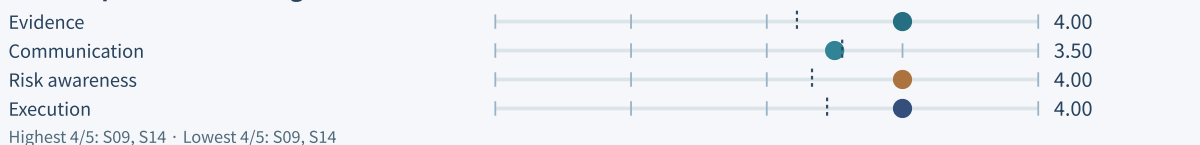
Prioritizing under pressure



Data-informed judgment*



Ownership & follow-through



1

generated trait rating

5

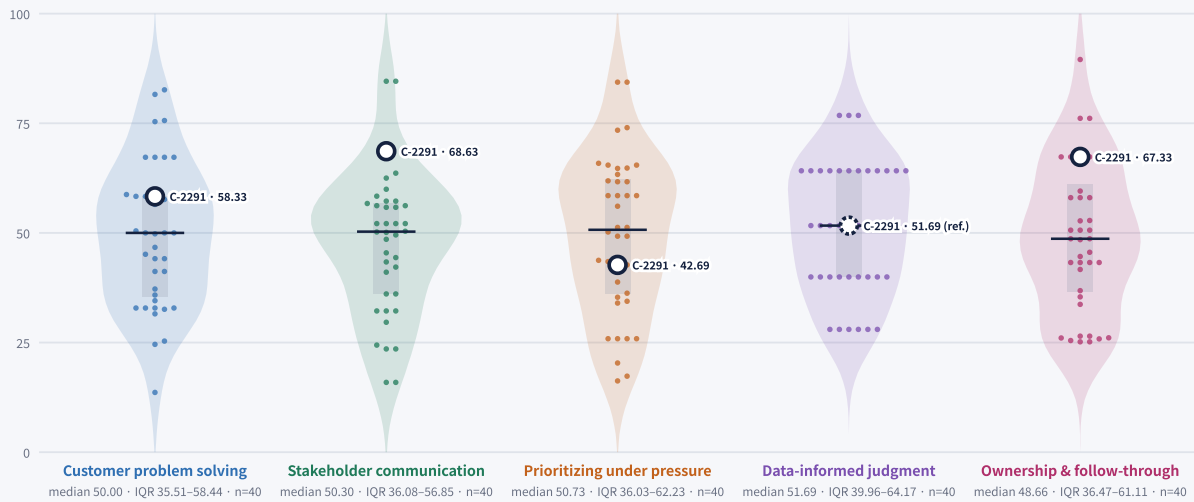
* Data-informed judgment has one generated answer. Gray dashed dots and rails retain the four answer-quality inputs for source traceability only, not competency comparison or ranking.

Dots: competency mean of four generated 1–5 answer-quality inputs. Dashed vertical ticks: this person's mean over the 9 administered answers, separately for each trait; not cohort norms or 0–100 competency scores. Highest/lowest list every tied scenario by effectiveness rating, not trait rating or job performance. Data-informed judgment has one answer; its extremes are the same source item and its comparison ticks are withheld.

POSITION IN THE COHORT · V3 DESCRIPTIVE

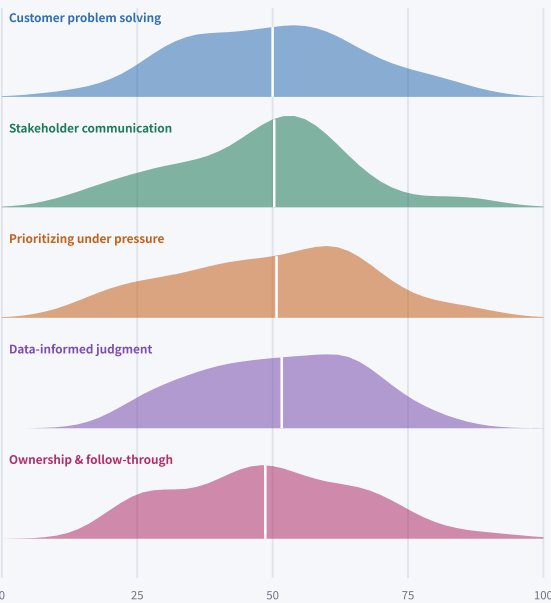
Violin = the 40 completers' spread; white ring = this candidate. No rank or percentile shown.

Violin + beeswarm



One dot = one completer. White ring = this candidate · black line = cohort median · pale bar = middle 50%.

Ridgeline · cohort distribution per competency



White line = median.

Dumbbell · score ↔ cohort median



Filled dot = this candidate · hollow = cohort median · grey bar = middle 50%.

Position in cohort / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Competency bands are 90% intervals from each competency's own posterior; a competency with one observation is reference only and has no band. Evidence grades were not assigned. Each person answers up to 12 of the 14 bank scenarios; the sample individual stopped after 9 because the 0.30 SE target was met.

Position in cohort · checks before comparison

The competency values, observation counts and band source tables remain on page 7. This section separates the chart scales and comparison scope rather than repeating those numbers.

Same comparison set	The distributions in this section use synthetic people who each completed an adaptive run drawn from the same 14-scenario bank (8 to 12 scenarios per person). They are not real applicant norms or a validated hiring population.
Competency distributions	Pages 66–67 show descriptive competency distributions and frequencies. Cohort medians and middle ranges describe the set, not an individual confidence interval.
Mean and variability	Page 68 shows the mean and sample standard deviation of the same answer ratings. Keep that scale separate from competency comparisons and do not order overlapping coordinates arbitrarily.
Ranks and ties	The added relative-comparison pages describe competency-specific tie intervals within the same source. Preserve joint positions rather than inventing an overall rank or hiring conclusion. Data-informed judgment is excluded from relative comparison.

Three scales, kept separate

Competency distributions
0–100 descriptive score
pages 66–67

Mean and variability
1–5 answer rating: mean and SD
page 68

Ranks and ties
competency-specific tie intervals
relative-comparison pages

How to read. Choose the question first, then read only the page for that scale. All three use the same synthetic 14-scenario set (8 to 12 scenarios per person), not real applicant norms.

Position in cohort / 2

Full competency source: page 7

All descriptive values, observation counts and bands remain on page 7. The chart uses competency-specific descriptive scales; Data-informed judgment has sparse observations and stays reference-only.

Every person score, spread and sample position



One dot per completed person (40 per row); x is the exact score and dots are spread vertically only so they stay visible. White orange-outlined circle: C-2291. Gray bar: descriptive p25–p75 (linear interpolation); navy line: median. These middle ranges are not individual confidence intervals. Data-informed judgment has one answer/person and is reference only.

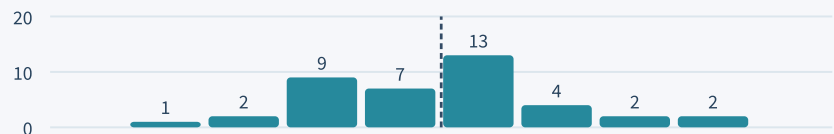
Reading guide: each row is one competency with one dot per person (40 per row); the orange-outlined circle is C-2291, the gray bar is p25–p75 and the navy tick is the median.

Position in cohort / 3

Competency score frequencies · 15.2

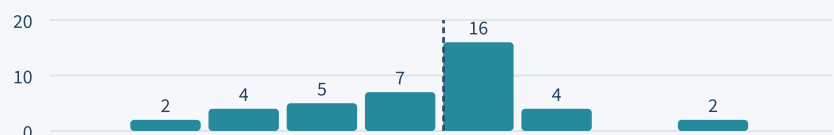
Customer problem solving

n=40 · median 50.00



Stakeholder communication

n=40 · median 50.30



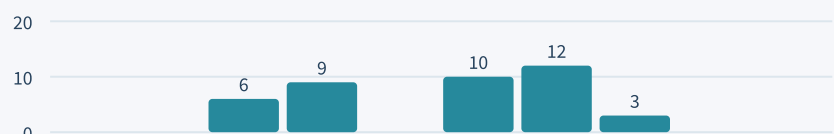
Prioritizing under pressure

n=40 · median 50.73



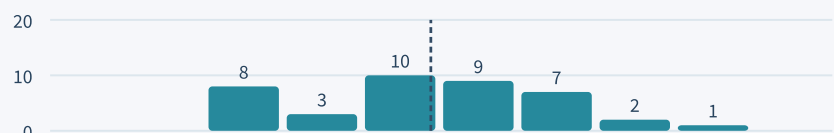
Data-informed judgment*

n=40 · reference only



Ownership & follow-through

n=40 · median 48.66

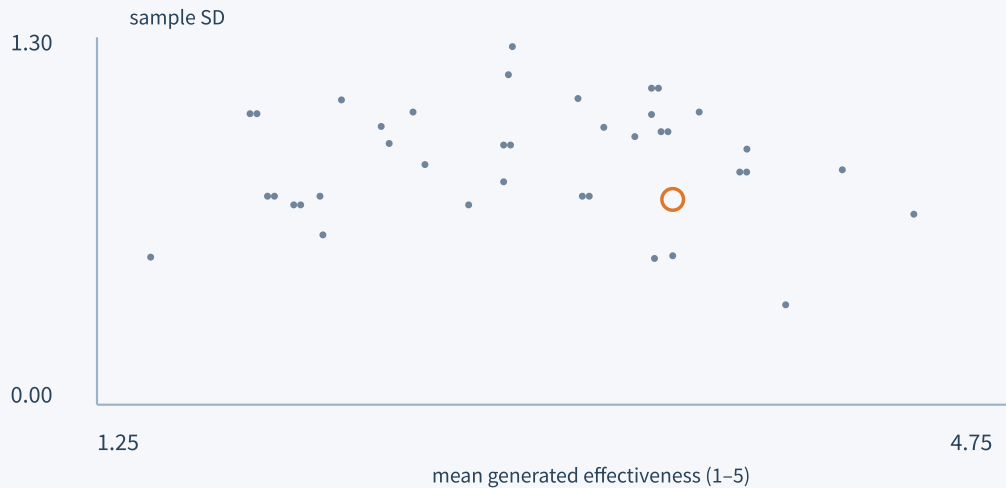


Descriptive score frequencies in ten-point bins (40 people). * Data-informed judgment has one item/person, reference-only, not a relative-comparison lens, hiring rank, or calibrated competency interval. Other rows are descriptive, not hiring ranks. Navy dashed line is the exact median of the unbinned scores, not the middle of a bin. Data-informed judgment has no median mark. The figure is a frequency histogram, not a smoothed density or calibrated interval. Scores of exactly 100 fall in the last bin.

Reading guide: each row is one competency's score frequencies across the 40 people; the navy dashed line marks that row's median, and Data-informed judgment has no median mark.

Position in cohort / 4

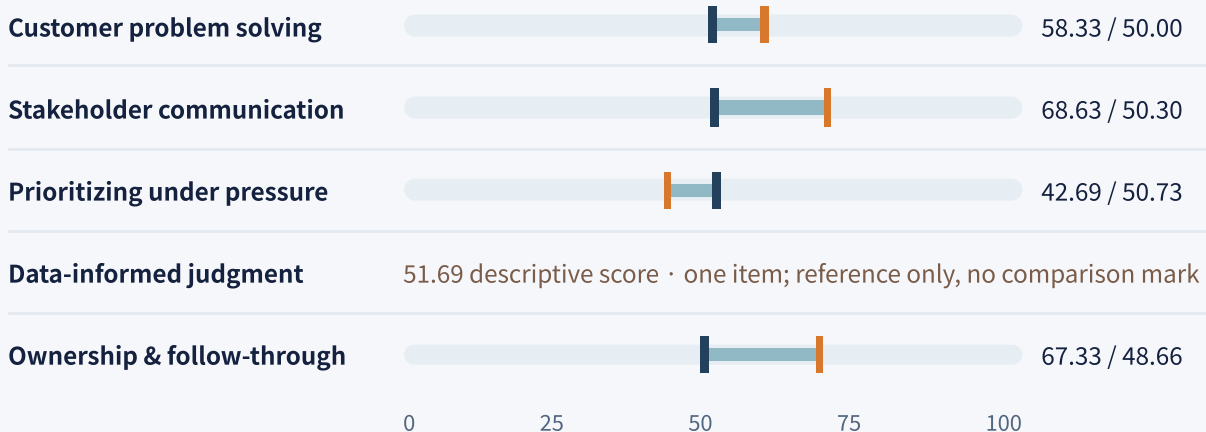
Generated answer mean and variation · 15.3



One dot per cohort person, 40 total, spread horizontally only for visibility. Orange ring: the sample individual (C-2291), not one of the 40. Horizontal: mean generated effectiveness (1-5); vertical: sample SD over that person's administered answers (8 to 12 per person; 9 for C-2291). Not smooth density, confidence or job-performance inference.

Synthetic 40-person adaptive cohort: 31 lower, 1 tied, 8 higher generated answer means than C-2291 (3.44).

competency dumbbell 15.4



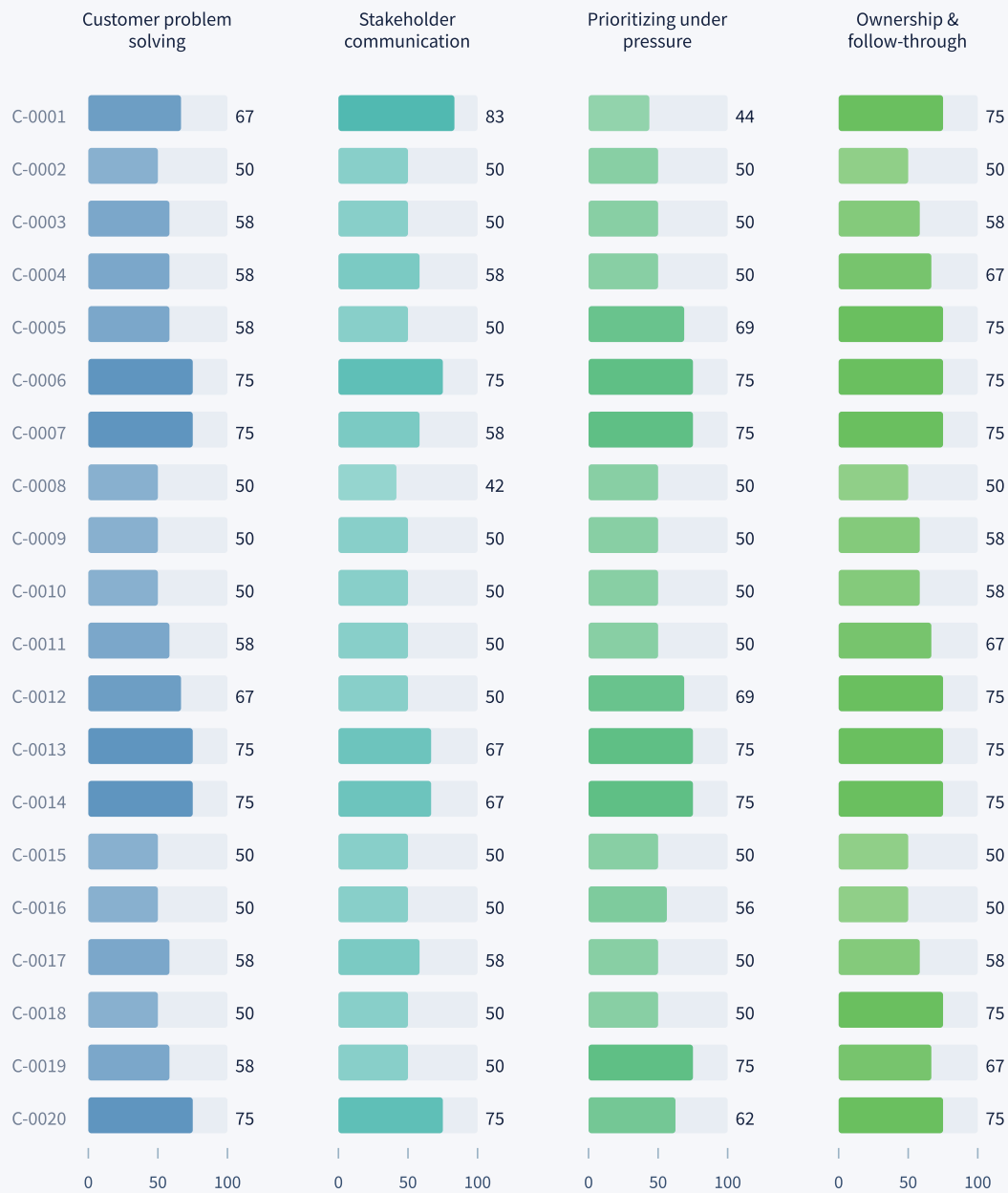
Orange: C-2291 (synthetic) / navy: synthetic cohort median; descriptive 0-100

Synthetic adaptive source. Four eligible comparison marks; Data-informed judgment has one item per person and its descriptive score is reference-only. Masked employer view; generated answers do not establish job performance.

Candidate comparison evidence / 7

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Paired cohort rows 1–20, four eligible lenses



The fifth competency remains reference-only for all 40 people. Row order is candidate ID, not ability order or hiring rank.

Source: generated 0-100 competency scores, people 1-20

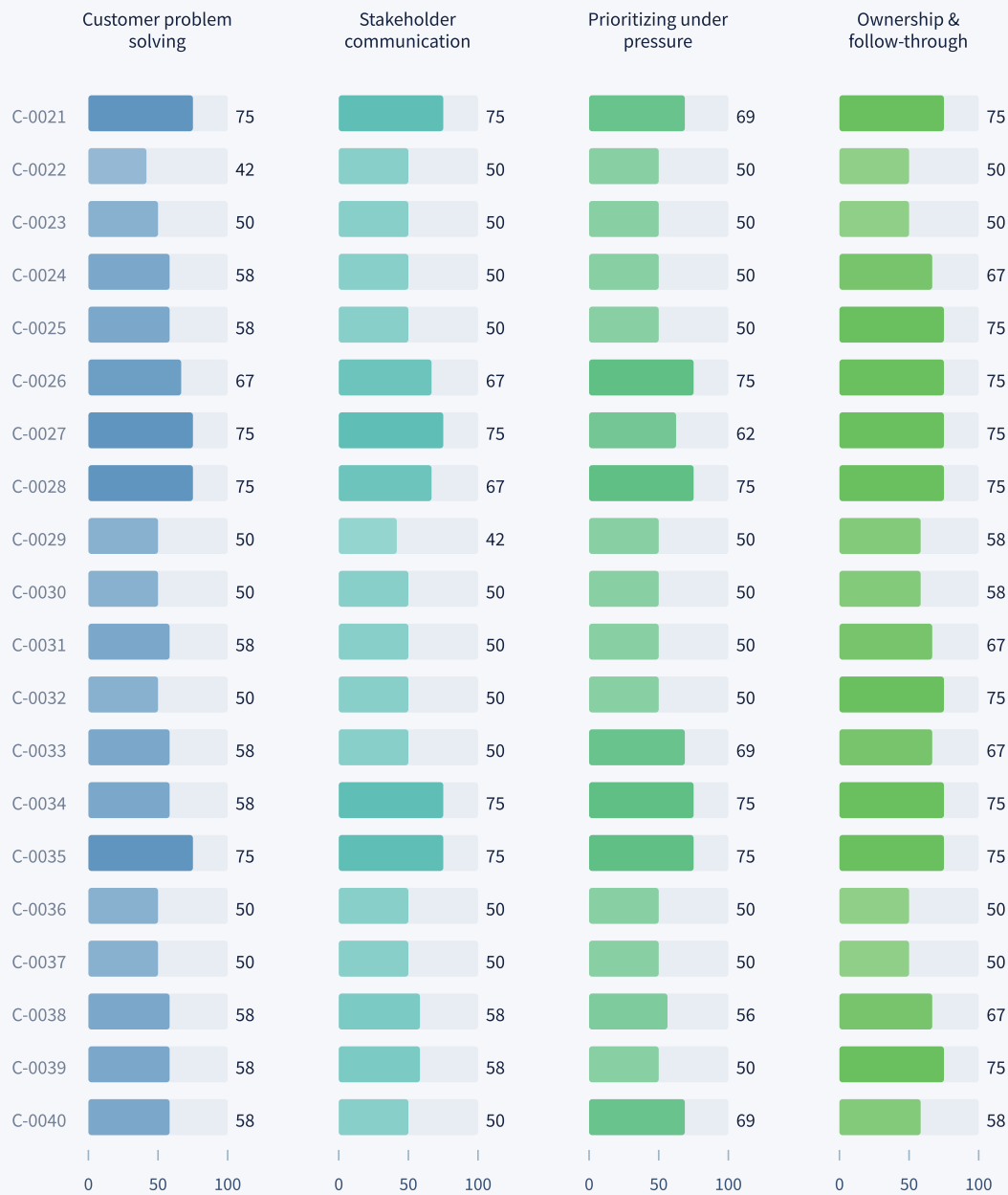
TITAN AI™ · INDIVIDUAL

EN · SYNTHETIC-ADAPTIVE-V350

Candidate comparison evidence / 8

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

Paired cohort rows 21–40, four eligible lenses



The same person across all columns. A bar spans 0–100 and is not independently sorted.

The fifth competency remains reference-only for all 40 people. Row order is candidate ID, not ability order or hiring rank.

Source: generated 0-100 competency scores, people 21-40

Candidate comparison evidence / 9

Additional descriptive views from the same synthetic 40-person, 14-scenario sample. Not a validated hiring norm or issued assessment.

3 tied

Customer 10–12

unique

Stakeholder 1–1

unique

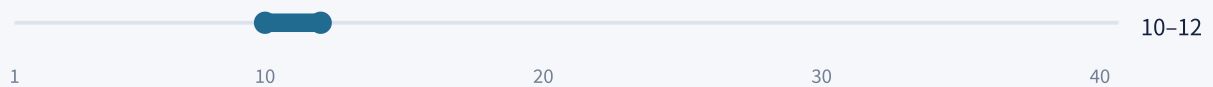
Prioritizing 40–40

18 tied

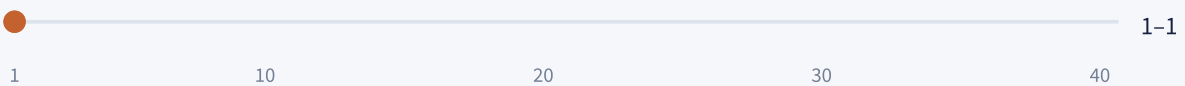
Ownership 1–18

Tie intervals, within each synthetic competency set

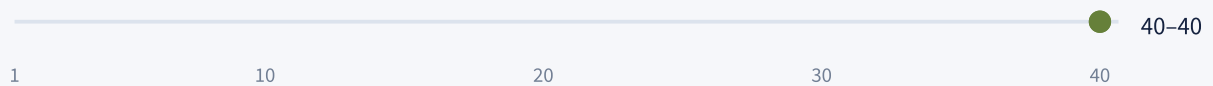
Customer problem solving



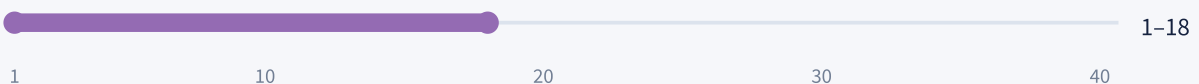
Stakeholder communication



Prioritizing under pressure



Ownership & follow-through



Position 1 is highest within a single competency. Ties occupy a range, not a forced ordering. Fifth lens and overall rank unavailable.

Source: C-0001 rank with tie interval, 40-person sample

Reading guide: the axis is rank position in the 40-person sample (1 = highest). Each bar is the tie range for that competency, and the label at the right end gives the range.

INTERVIEW · ONBOARDING GUIDE

Four questions: three fill evidence gaps, one confirms a strength.

① Prioritizing under pressure · Takes-it-all-on pattern in ③⑧⑫

“Tell me about a time in the last six months when work piled up beyond what you could handle. What did you defer, and who did you decide that with?”

Signals of a strong answer

Decides what to defer · talks to the manager early · names criteria (revenue impact, customer risk)

Signals to probe further

Ends with 'I got it all done' · overtime as the only fix · talked to no one

Follow-up

“If you did it again, what would you change?”

② Data-informed judgment · One observation · reference only

“Give me a real case where the data and what you heard in the field disagreed. What did you trust, and how did you check?”

Signals of a strong answer

Checks sample, time window and definitions · respects both sources · changed their mind once

Signals to probe further

Dismisses one side · just quotes the number

Follow-up

“How did you check how that decision turned out?”

③ Prioritizing under pressure · Weak criteria on the ⑧ pricing exception

“Tell me about a time a customer asked for an exception outside policy and you said no.”

Signals of a strong answer

Explains criteria and precedent · offers an alternative · says no and keeps the relationship

Signals to probe further

No example of saying no · only 'the relationship comes first'

Follow-up

“If you did it again, what would you change?”

④ Stakeholder communication · To confirm a strength (②⑦⑪)

“Tell me about delivering bad news to a senior leader yourself. What did you say first?”

Signals of a strong answer

Facts → impact → decision needed · short · has options ready

Signals to probe further

Long backstory · shifts blame

Follow-up

“How did you check how that decision turned out?”

First-90-days onboarding ideas

Days 1–30

Write down account priority rules (revenue, risk, renewal date) with the manager.

Days 31–60

In weekly 1:1s, name one thing deferred that week.

Days 61–90

Have them present in a data review so data-informed judgment can be seen first-hand.

Avoid in the interview

- Reading report scores to the candidate
- Inferring personality or values from this report
- Summarizing with a composite or a rank
- Calling a 1-observation competency a weakness

Interview time split

Rapport		5'
Prioritizing under pressure		10'
Data-informed judgment		8'
Stakeholder communication		5'
Questions		2'

After the interview

- Store interview notes separately from this report
- Note what you learned on the 1-observation competency
- Decide by interviewer consensus; use this report as evidence only

Where each question comes from

Prioritizing under pressure	Questions ① ③		42.69	obs 2
Data-informed judgment	Question ②		51.69	obs 1
Stakeholder communication	Question ④		68.63	obs 2

Bars show the 0 to 100 descriptive score and observation count of the competency each question targets. Data-informed judgment has one observation: reference only.

Interview questions / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Adaptive 9-scenario run (C-2291, synthetic): scores, 90% bands and observation counts follow the v350 rule; non-administered scenarios were not asked.

Individual descriptive-score source table: p. 27. The questions and work checks below are separate drafts, not selection or staffing decisions based on generated scores.

Interview and onboarding questions, not verdicts

Ask for a real example of a heavy-workload trade-off and another example of using data against field observations. Make a separate interview record; do not infer workplace ability from this generated case.

Onboarding prompts can be considered only after the person starts and the manager sees their actual work.

Interview	5 min context · 10 min priority trade-off · 8 min data check · 5 min communication · 2 min questions
Avoid	Do not label a one-observation competency a weakness or merge interview ratings with this score.

Four source-limited topics

Four draft job-evidence questions. They are not cues proven by C-2291's masked dialogue, validated scoring keys or a hiring recommendation. Data-informed judgment has only one synthetic observation.

Priority trade-off	Question: Tell me about a time urgent work exceeded your capacity. What did you defer, and who agreed? Follow-up: What was the criterion and when did you tell others?
Data versus field	Question: Tell me about data and field observations disagreeing. What did you check first? Follow-up: What source, time window and definition changed your view?
Policy exception	Question: Tell me about a customer request outside policy. How did you decide what you could offer? Follow-up: What approval and precedent did you check?
Senior update	Question: Tell me about personally giving a senior stakeholder difficult news. What did you say first? Follow-up: What decision did you ask them to make?

After the interview / 2

Store independent interview notes separately. Any onboarding plan starts only after real work evidence and manager agreement.

Optional 30/60/90 work checks

Days 1–30	Agree job-specific priority rules, if relevant.
Days 31–60	Review one actual deferral decision with the manager.
Days 61–90	Review a real data-versus-field case before making any development inference.

Safety and process

Do not read scores aloud	Do not read generated report scores aloud to the candidate as an interview script. Ask the same job-based questions and record real examples separately; these synthetic scores are not a hiring verdict.
Do not infer	No personality, values, composite rank or one-item weakness from this generated report.
Record separately	Document the question, job evidence and interviewer rating separately from generated scores.
Decision	Use a pre-agreed job-based rubric and interviewer review; this synthetic source is not decision evidence.

Independent interview record - blank worksheet

Capture the interviewee's real example and the interviewer's job-based observation. These blanks are not generated evidence, a rating, or a hiring recommendation. Store them apart from this synthetic report.

Priority trade-off

Real example _____

Job-based observation _____

Data versus field

Real example _____

Job-based observation _____

Policy exception

Real example _____

Job-based observation _____

Senior update

Real example _____

Job-based observation _____

Ask about the two thinnest competencies first, then confirm one strength. The signal cues come from this candidate's own answers.

Prioritizing under pressure · 42.69

lowest · 2 observations

Q. Tell me about a time several things were urgent and you deliberately let one slip.

Follow-up

Who did you agree the slip with, and when?

Signs of a strong answer

Chooses what to defer and flags it early

Signs of a weak answer

Tries to do it all by working longer; escalates late

Data-informed judgment · 51.69

1 observation · reference only

Q. Tell me about a time the data and what people on the ground said disagreed.

Follow-up

How did you decide which to trust?

Signs of a strong answer

Checks source, sample size and timing

Signs of a weak answer

Trusts one side as-is or decides on gut

Stakeholder communication · 68.63

confirm strength

Q. Tell me about delivering bad news to someone senior.

Follow-up

What was your first sentence?

Signs of a strong answer

Bottom line first, then cause and next steps

Signs of a weak answer

Long lead-in; the bottom line comes late

Customer problem solving · 58.33

confirm strength

Q. Tell me about a time what a customer wanted and what you could give were different.

Follow-up

What was the first thing you said to the customer?

Signs of a strong answer

Checks cause first and avoids promises that can't be kept

Signs of a weak answer

Jumps to fixes or credits without checking cause

Ownership & follow-through · 67.33

confirm strength

Q. Tell me what you did when something you promised slipped.

Follow-up

Did you tell them before they noticed?

Signs of a strong answer

Flags first, gives a new date and a checkpoint

Signs of a weak answer

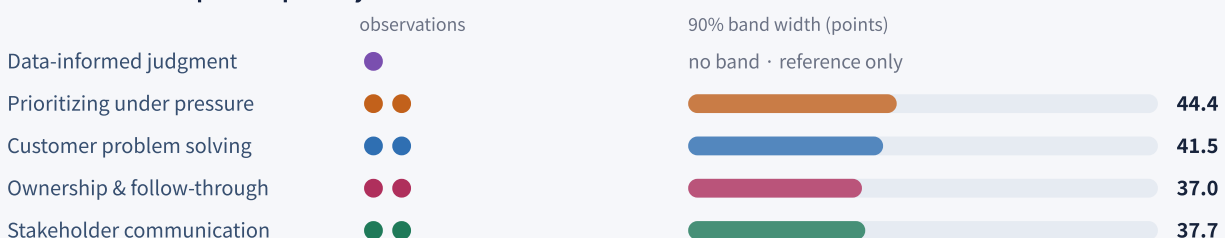
Tells them late or spreads the blame

Interviewer sheet

Competency	1	2	3	4	5	Notes
Prioritizing under pressure	☐	☐	☐	☐	☐	
Data-informed judgment	☐	☐	☐	☐	☐	
Stakeholder communication	☐	☐	☐	☐	☐	

Don't merge interview ratings with this report's scores. Record them separately.

Evidence amount per competency



Sorted by fewest observations first. A wider band means less information, not a lower score; the width scale is 0 to 100 points. Guidance, not a score.

Interview kit · five prompts / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Adaptive 9-scenario run (C-2291, synthetic): scores, 90% bands and observation counts follow the v350 rule; non-administered scenarios were not asked.

Individual descriptive-score source table: p. 27. The questions and work checks below are separate drafts, not selection or staffing decisions based on generated scores.

Interviewer record - separate from the synthetic report

Record observed interview evidence and the question asked. Leave ratings blank until an interviewer supplies them.

Prioritizing under pressure	Observed example: _____ Independent rating: ____ / 5
Data-informed judgment	Observed example: _____ Independent rating: ____ / 5
Stakeholder communication	Observed example: _____ Independent rating: ____ / 5

Five job-evidence prompts

Ask the same core questions and use a shared job-based rubric. These are draft prompts, not validated scoring keys. Data-informed judgment has one synthetic item per person and is not a relative-comparison lens.

Customer problem solving	Q: Tell me about a customer request you could not meet. What did you check before offering a fix? Follow-up: How did you agree on the next step and verify it happened?
Stakeholder communication	Q: Tell me about giving a senior stakeholder a difficult update. What was your first sentence? Follow-up: What decision or help did you ask for?
Prioritizing under pressure	Q: Tell me about several urgent tasks colliding. What did you defer, and why? Follow-up: Who did you tell, and when?
Data-informed judgment	Q: Tell me about data conflicting with what you heard from the field. What did you check? Follow-up: How did you test the source, time window and definitions?
Ownership & follow-through	Q: Tell me about a commitment that slipped. What did you do first? Follow-up: How did the other person confirm the new date or outcome?

Reading guide: the interviewer record stays blank until an interviewer fills it in; the five prompts are drafts, not selection decisions based on generated scores.

Independent interview record / 2

Blank worksheet. Interview ratings are not generated report scores and must never be averaged into them.

Blank interviewer worksheet

Record each job-related answer using the same 1–5 rubric defined by the employer before interviews. Blank cells are for the interviewer, not scores assigned by Titan.

Competency	1	2	3	4	5	Notes
Customer problem solving	☐	☐	☐	☐	☐	_____
Stakeholder communication	☐	☐	☐	☐	☐	_____
Prioritizing under pressure	☐	☐	☐	☐	☐	_____
Data-informed judgment	☐	☐	☐	☐	☐	_____
Ownership & follow-through	☐	☐	☐	☐	☐	_____

Evidence behind each independent rating

For each answer, write an observable action and the relevant job criterion. A blank row is not a zero. Do not copy or average generated report scores.

Customer problem solving

Observed action _____

Job criterion _____

Stakeholder communication

Observed action _____

Job criterion _____

Prioritizing under pressure

Observed action _____

Job criterion _____

Data-informed judgment

Observed action _____

Job criterion _____

Ownership & follow-through

Observed action _____

Job criterion _____

DEVELOPMENT PLAN · 30/60/90 GUIDANCE · NOT A SCORE

The first 90 days after hire. Invest first in prioritizing; settle the data-informed judgment plan once there is more evidence.

Competency	First 30 days	By 60 days	By 90 days
Prioritizing under pressure 42.69 · obs 2	Each morning, share with your manager the one thing you'll defer today	Build a quarter-end load plan and rebalance two weeks ahead	Re-measure: retake the prioritizing scenarios to check for change
Data-informed judgment 51.69 · obs 1	Thin evidence: check in the interview first	Use a source / sample / timing checklist when reading data	Settle the plan once there are two or more observations
Customer problem solving 58.33 · obs 2	Already a strength: take one hard renewal and repeat the approach	Share the decision-criteria question as a team template	Coach a new hire by joining an escalation call
Ownership & follow-through 67.33 · obs 2	Strength: write a handover checklist	Lead one post-incident review	Run the team's commitment tracker
Stakeholder communication 68.63 · obs 2	Keep it up: draft one executive update yourself	Mediate one sales-product conflict	Make bottom line / cause / action the team standard for quarterly reports

Tied to the evidence

- ⑧ Quarter-end workload spike · 2/5 → Stretches own time; doesn't choose what to defer
- ⑫ Deadline vs quality trade-off · 3/5 generated effectiveness rating
- ④ Churn dashboard disagrees with field notes · 3/5 → Checking sample and time window is a good sign; next step unclear · one observation
- ⑥ Customer demands an unsupported feature · 3/5 → Says no, then asks what the real problem is

The plan's first lines come from the four lowest scenarios.

Re-measure

Retaking new scenarios for the same competencies after 90 days lets you compare change on the same scale (0–100). A difference is descriptive; no threshold for “real change” is defined.

What the manager does

- For the first two weeks, agree one thing to defer in each weekly 1:1
- Check workload three weeks before quarter end

Use the strengths

- Stakeholder communication: give them executive update drafts
- Ownership & follow-through: let them lead handovers and post-incident reviews

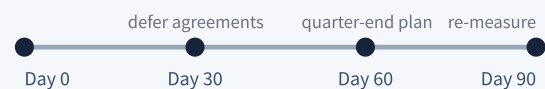
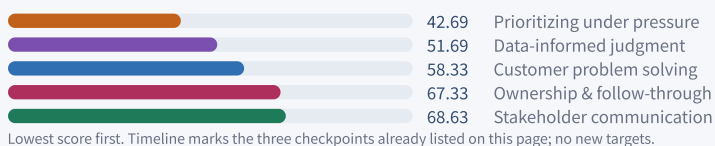
Checkpoints

- Day 30: 4+ recorded defer agreements
- Day 60: quarter-end plan submitted
- Day 90: re-measure the same competencies

90-day re-measure note

Competency	Now	On a retest
■ Prioritizing under pressure	42.69	compare on the same 0–100 scale
■ Data-informed judgment	51.69	reference only · get observations first
■ Customer problem solving	58.33	compare on the same 0–100 scale
■ Ownership & follow-through	67.33	compare on the same 0–100 scale
■ Stakeholder communication	68.63	compare on the same 0–100 scale

Scores behind the plan, and the 90-day timeline



Development plan · evidence / 1

Synthetic template responses; not real candidate data or calibrated Titan output.

Adaptive 9-scenario run (C-2291, synthetic): scores, 90% bands and observation counts follow the v350 rule; non-administered scenarios were not asked.

Individual descriptive-score source table: p. 27. The questions and work checks below are separate drafts, not selection or staffing decisions based on generated scores.

Generated high answer score - verify with real examples

Stakeholder communication has the highest eligible descriptive score in this generated case (68.63/100). Ask for a real example and a contrasting situation before using it in a development plan; these generated answers do not predict job performance.

Development and re-measurement gates

This generated case cannot establish job performance or a 90-day improvement. If the manager later agrees to a plan, record real decisions and check-ins rather than using this synthetic rating as a baseline.

A later retest requires a calibrated measure, matched item bank and human-rated coverage evidence before claiming meaningful change.

Day 30	Review real prioritization decisions with the manager.
Day 60	Check data-versus-field decisions in actual work.
Day 90	Assess work evidence separately; do not compare synthetic 90% bands.

Optional manager work checks / 2

Only after hire and manager agreement. These synthetic answers do not establish a development need or an employee action.

Four source-linked lower-answer examples

The four lowest of the nine administered generated ratings (2–3/5); S08 first, ties in administered order.

S12 · Deadline vs quality trade-off	3/5 generated effectiveness rating. A job-relevant interview or actual work example is needed before making a plan.
S04 · Churn dashboard disagrees with field notes	3/5 generated effectiveness rating. A job-relevant interview or actual work example is needed before making a plan.
S06 · Customer demands an unsupported feature	3/5 generated effectiveness rating. A job-relevant interview or actual work example is needed before making a plan.
S08 · Quarter-end workload spike	2/5 generated effectiveness rating. A job-relevant interview or actual work example is needed before making a plan.

Optional checks, never automatic assignments

First 2 weeks	If relevant, agree in writing on one real priority/defer decision and its owner.
Before a busy period	If relevant, review available capacity and who approves trade-offs three weeks before a known peak.
Day 30 / 60 / 90	Review independent interview/work notes, not generated score changes.
Later measurement	Use a validated, comparable measure with observed reliability/coverage; do not infer change from nonoverlap of 90% bands.

Generated score and 90% band

The band is a 90% interval from the synthetic adaptive model, not a hiring cutoff or change-detection rule. Data-informed judgment has one observation and is reference-only.

Competency	Observations	Descriptive score /100	90% band /100
Customer problem solving	2	58.33	37.33–78.82
Stakeholder communication	2	68.63	48.14–85.85
Prioritizing under pressure	2	42.69	22.22–66.59
Data-informed judgment	1	51.69	—
Ownership & follow-through	2	67.33	47.36–84.36

Treat these as display-only synthetic values. At retest, use matched items and human-rated calibration and coverage, then analyze change with an independently validated method.

Above the cohort median on three competencies, below it on prioritizing under pressure. Data-informed judgment is reference only.

Compared with others in the same campaign

Customer problem solving

cohort median 50.00 · middle 50% 35.51–58.44

Stakeholder communication

cohort median 50.30 · middle 50% 36.08–56.85

Prioritizing under pressure

cohort median 50.73 · middle 50% 36.03–62.23

Data-informed judgment

cohort median 51.69 · middle 50% 39.96–64.17

Ownership & follow-through

cohort median 48.66 · middle 50% 36.47–61.11

Small dots = the 40 completers · orange = candidate (hollow = reference only) · dashed = cohort median · a plain comparison within this campaign, not a percentile; no comparison to other campaigns or norms.

Measurement note

- Scoring: free-text and follow-up answers are scored against a competency rubric; an anchor check is added only when an answer is unclear.
- Scale: combined effectiveness per competency scaled 0–100; descriptive scores, with no interval.
- Evidence base: observation counts per competency are shown on the evidence trail. It is not pass/fail.
- Limits: a conversation shows how someone says they'd act, not actual behaviour. Confirm with interviews and references.

Evidence trail

Scenarios · stop	9 of 14 · Stopped when the 0.30 SE target was met
Final estimate · SE	+0.44 · 0.295 (0.30 target met)
Competencies under 2 obs.	1 (Data-informed judgment)
Fallback scoring	0/9

Position in the campaign by competency

Competency	Score	Cohort median	Middle 50%
Customer problem solving	58.33	50.00	35.51–58.44
Stakeholder communication	68.63	50.30	36.08–56.85
Prioritizing under pressure	42.69	50.73	36.03–62.23
Data-informed judgment	51.69	51.69	39.96–64.17
Ownership & follow-through	67.33	48.66	36.47–61.11

Position within each competency, not an overall rank. Cohort = the 40 completers (C-2291 is a separate run, not one of them).

Amount of evidence · candidate vs completers

Competency	Obs (candidate)	Obs (every completer)	Probes (candidate)	Probes mean
Customer problem solving	2	2–3	4	3.6
Stakeholder communication	2	2–3	3	4.2
Prioritizing under pressure	2	1–3	5	3.8
Data-informed judgment	1	1	3	1.7
Ownership & follow-through	2	2–3	3	3.8

Each completer answers 8 to 12 of the 14 scenarios, so observation counts differ by person. This compares amount of evidence, not scores.

Campaign comparison / 1

68.63 TOP · STAKEHOLDER

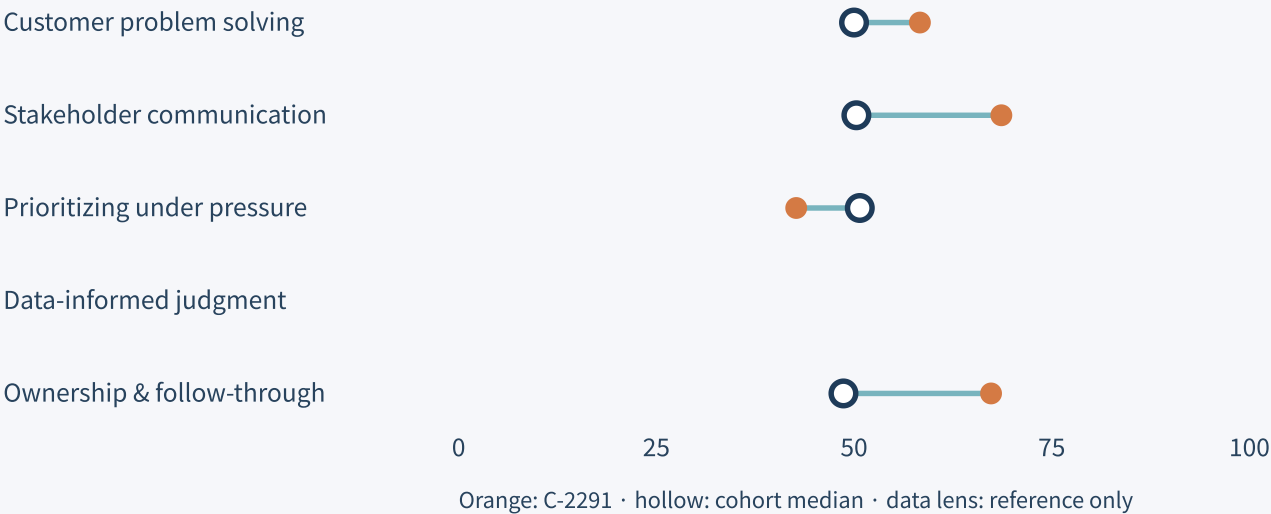
42.69 LOW · PRIORITIZING

3 of 4 ABOVE MEDIAN

Synthetic template responses; not real candidate data or calibrated Titan output.

Adaptive 9-scenario run (C-2291, synthetic): scores, 90% bands and observation counts follow the v350 rule; non-administered scenarios were not asked.

C-2291 versus cohort median · descriptive 0–100



Synthetic adaptive 40-person cohort; C-2291 is a separate run, not one of the 40. The one-observation Data-informed judgment lens has no comparative mark; other marks are descriptive, not hiring ranks or calibrated intervals.

How much evidence sits behind each position

"Lower / tied" compares C-2291 with the 40 cohort people (C-2291 is not one of them). Counts use descriptive generated-answer scores, not a validated hiring rank. Data-informed judgment has one item per person, so it is reference only.

Competency	Score	Cohort median	Observations	Follow-ups	Cohort 40: lower / tied
Customer problem solving	58.33	50.00	2	4	25 / 5
Stakeholder communication	68.63	50.30	2	3	37 / 0
Prioritizing under pressure	42.69	50.73	2	5	13 / 1
Data-informed judgment	51.69	51.69	1	3	reference only
Ownership & follow-through	67.33	48.66	2	3	31 / 5

Measurement and evidence trail / 2

Synthetic adaptive generation, not a production score API. No evidence grade, language/duration claim or calibrated competency interval is issued beyond the synthetic 90% band shown elsewhere.

Synthetic adaptive trace

Scenario completion	9 of 14 scenarios administered; stopped when the 0.30 SE target was met
Overall theta EAP / SE	+0.438 / 0.295; SE target 0.30 met
Competencies with fewer than two answers	1: Data-informed judgment; reference only
Generated follow-ups / anchors	18 / 3; neither count measures skill
Evidence grade	Not assigned; no earlier label or fallback percentage transfers
Answer text	Withheld in this employer-facing proof; restricted access requires tenant-bound authorization

Amount of evidence, not ability

Customer problem solving	C-2291: 2 obs, 4 probes · cohort: 2.4 obs (range 2–3), 3.60 probes/person
Stakeholder communication	C-2291: 2 obs, 3 probes · cohort: 2.8 obs (range 2–3), 4.15 probes/person
Prioritizing under pressure	C-2291: 2 obs, 5 probes · cohort: 2.0 obs (range 1–3), 3.80 probes/person
Data-informed judgment	C-2291: 1 obs, 3 probes · cohort: 1.0 obs (range 1–1), 1.68 probes/person
Ownership & follow-through	C-2291: 2 obs, 3 probes · cohort: 2.3 obs (range 2–3), 3.83 probes/person

Reading guide: the first table lists the adaptive trace for this one synthetic run; the second counts observations and probes, which describe the amount of evidence, not ability.

What each administered scenario changed in overall uncertainty

This is overall theta standard error over the 9 scenarios the engine administered in this synthetic run. Step-to-step differences are not item selection gains or calibrated competency-level uncertainty.

Order	Scenario	Competency	Overall SE	Drop from prior
1	S11	Stakeholder communication	0.650	—
2	S06	Customer problem solving	0.528	+0.122
3	S14	Ownership & follow-through	0.456	+0.072
4	S08	Prioritizing under pressure	0.434	+0.022
5	S04	Data-informed judgment	0.395	+0.039
6	S02	Stakeholder communication	0.354	+0.041
7	S01	Customer problem solving	0.328	+0.026
8	S09	Ownership & follow-through	0.309	+0.020
9	S12	Prioritizing under pressure	0.295	+0.013

After five scenarios SE was 0.395; after the ninth, 0.295. The 0.30 target was met, so the run stopped before the 12-scenario maximum.

