

Titan AI™

# Cohort Insight Report

Adaptive conversational assessment · Cohort

Where this campaign's candidates are strong and where they wobble, whether the evidence is enough, and what to raise first in interviews and onboarding.

**Sample campaign · Customer Success Manager · 42 candidates**

68 pages · Sample v3.0

Free-text + follow-up probes

Descriptive score per competency

No ranking · no composite

Thin evidence stays visible

**01** Executive summary**02** Summary**03** Competency distributions**04** Competency deep dive**05** Answer traits**06** Scenario results**07** Difficulty · separation**08** Probe analysis**09** Evidence quality**10** Consistency**11** All candidates**12** How competencies relate**13** Interviews · onboarding · note**14** Interview kit · by profile**15** Onboarding 12 weeks**Verify this report**  
TV-XXXX-XXXX-C



# Titan AI™ · Cohort report

Adaptive scenarios · synthetic 40-person source, 8 to 12 of 14 scenarios each

## Synthetic cohort source

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

### Synthetic session population

|                      |    |
|----------------------|----|
| Completed candidates | 40 |
| Started              | 42 |

### First read: a median can hide a different tail

Medians of the four eligible competencies sit between 48.66–50.73, yet scores at or above 75 occur for 4 people on Customer problem solving versus 2 on Prioritizing under pressure. This within-sample difference does not establish job performance. Data-informed judgment has one observation each and is excluded from comparison.

Source: synthetic adaptive 40-person cohort (8 to 12 of 14 scenarios per person)

## Reading path

|    |                                  |       |
|----|----------------------------------|-------|
| 01 | Summary and campaign coverage    | p. 4  |
| 02 | Competency distributions         | p. 9  |
| 03 | Scenario source and comparisons  | p. 26 |
| 04 | Evidence flow and consistency    | p. 40 |
| 05 | Linked candidate profiles        | p. 49 |
| 06 | Interview, onboarding and limits | p. 60 |

## EXECUTIVE SUMMARY

OBSERVED

DESCRIPTIVE

# This campaign on one page: where it's strong, where it's weak, and who to look at first.

Customer problem solving

50.00

median · middle 50% 35.51–

58.44

Stakeholder communication

50.30

median · middle 50% 36.08–

56.85

Prioritizing under pressure

50.73

median · middle 50% 36.03–

62.23

Data-informed judgment

51.69

median · middle 50% 39.96–

64.17

Ownership &amp; follow-through

48.66

median · middle 50% 36.47–

61.11

Candidates

42

40 completed

Completed as-is

37

no QC selection

Check first

5

QC-selected or timed out

Scenarios per person

10

8–12 · 1.6 probes each

Median final SE

0.29

0.30 met by 36 of 40

## Four things to know

**Data-informed judgment** has the highest median (51.69, reference only); **prioritizing under pressure** and **ownership & follow-through** bracket the eligible medians (50.73 / 48.66).

**Risk awareness** is the lowest trait mean (2.89 of 5); traits differ by 0.04.

Lowest scenario means (n ≥ 30): 'Handover from a departing colleague' 2.83 and 'Customer demands an unsupported feature' 2.85; thin S03, S05 not ranked.

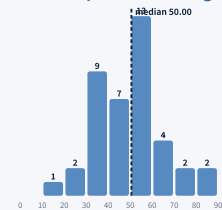
3 completed through QC selection and 2 timed out; check these first.

## Next steps

- 1 Review transcripts for the QC-selected and timed-out group
- 2 Interview on ownership & follow-through first (lowest eligible median; differences are small)
- 3 Onboarding practice on risk awareness (lowest trait mean, small gap)
- 4 Review the lowest-rated scenarios before the next campaign

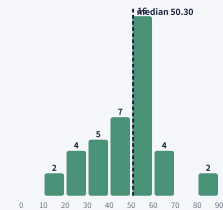
## Distribution by competency

Customer problem solving

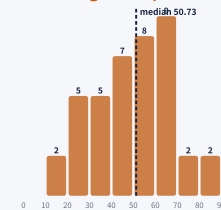


40 completers · dashed line = median

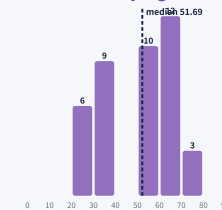
Stakeholder communication



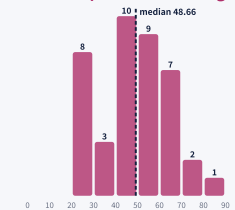
Prioritizing under pressure



Data-informed judgment



Ownership &amp; follow-through



## Completion · 42 candidates

Completed as-is 37

3

2

QC-selected or timed out = check first.

Not a verdict

No composite · no rank

Thin evidence stays visible

No personality or sentiment inference

## Four traits · mean over all scenarios

|                |             |      |
|----------------|-------------|------|
| Evidence       | <div></div> | 2.92 |
| Communication  | <div></div> | 2.91 |
| Risk awareness | <div></div> | 2.89 |
| Execution      | <div></div> | 2.94 |

416 answers · 1–5 scale.

# Executive summary / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## What this generated campaign can show

Of four comparison-eligible competencies, Prioritizing under pressure has the highest descriptive median (50.73/100) and Ownership & follow-through the lowest (48.66/100). The spread is small and this is not a validated hiring ranking.

Data-informed judgment has one observation per person and is reference-only, even though its descriptive median is 51.69/100.

Started / completed / timed out                      42 / 40 / 2

Evidence status                                              37 completers as-is; three sampled for QC within completed records

## Next steps

Use independent structured interviews and a versioned, human-rated calibration study before making applicant comparisons or inferring workplace performance.

Interpretation                      Read distributions and individual evidence counts, not just mean scores.

Access                                      Keep generated answer dialogue masked in the employer report.

## Descriptive medians and middle-half ranges

|                                       |       |
|---------------------------------------|-------|
| Customer problem solving median       | 50.00 |
| Customer problem solving IQR lower    | 35.51 |
| Customer problem solving IQR upper    | 58.44 |
| Stakeholder communication median      | 50.30 |
| Stakeholder communication IQR lower   | 36.08 |
| Stakeholder communication IQR upper   | 56.85 |
| Prioritizing under pressure median    | 50.73 |
| Prioritizing under pressure IQR lower | 36.03 |
| Prioritizing under pressure IQR upper | 62.23 |
| Data-informed judgment median         | 51.69 |
| Data-informed judgment IQR lower      | 39.96 |
| Data-informed judgment IQR upper      | 64.17 |

SAMPLE · SYNTHETIC DATA · NOT REAL COHORT DATA

TITAN AI™ · COHORT

## Executive summary / 2

### Read the five distributions once

The source-backed five competency histograms, including Data-informed judgment (one observation per person, reference only), appear once in page 14. This summary keeps the medians and IQRs. Its cohort comparison remains descriptive, not a hiring rank.

### Ownership median and session totals

|                                   |       |                                      |       |                                      |       |
|-----------------------------------|-------|--------------------------------------|-------|--------------------------------------|-------|
| Ownership & follow-through median | 48.66 | Ownership & follow-through IQR lower | 36.47 | Ownership & follow-through IQR upper | 61.11 |
| Started                           | 42    | Completed                            | 40    | Timed out                            | 2     |
| Completed as-is                   | 37    | Total probes                         | 682   | Median theta SE                      | 0.29  |
| Completed · QC-selected           | 3     |                                      |       |                                      |       |

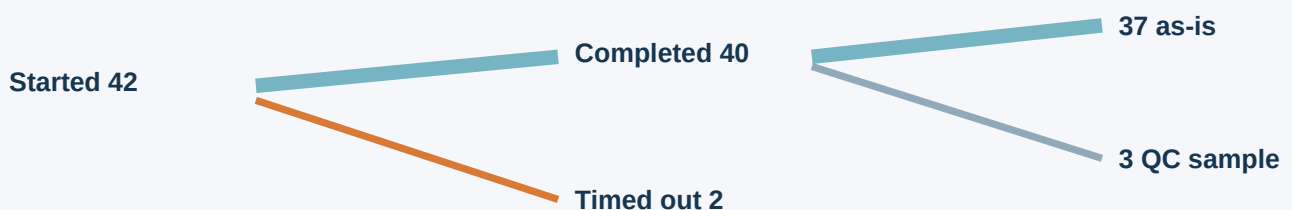
### Four traits across 416 generated answers

Means on the separate behavioral 1-5 scale, not transformed into a 0-100 competency score. Risk awareness is lowest descriptively, but the four means differ by only 0.04; this does not establish an onboarding priority or predict job performance.

|          |               |                |           |
|----------|---------------|----------------|-----------|
| Evidence | Communication | Risk awareness | Execution |
| 2.92 / 5 | 2.91 / 5      | 2.89 / 5       | 2.94 / 5  |

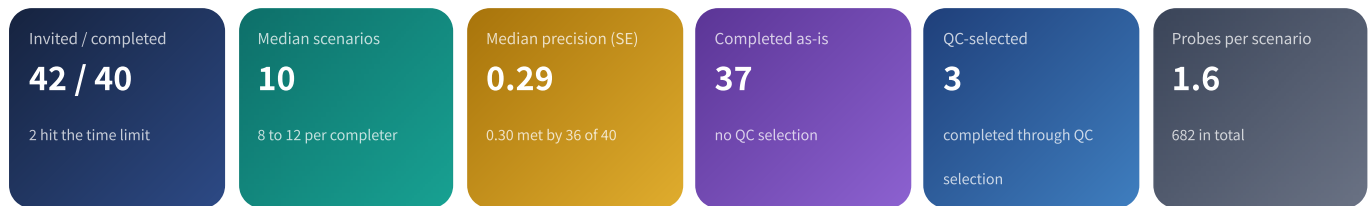
### Started, completed and timed out

Recorded status counts: 42 started = 40 completed + 2 timed out; 40 completed = 37 as-is + 3 preselected QC cases. No evidence grades or automatic exclusion decision.

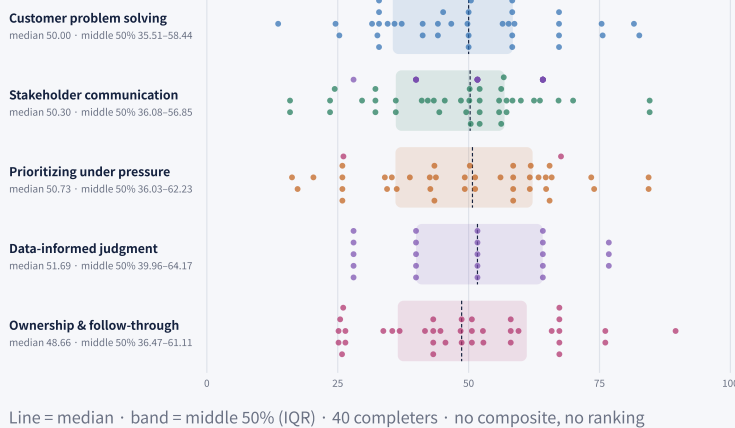


## SUMMARY · THE CAMPAIGN AT A GLANCE

# Competency medians sit close together (48.66–50.73) with wide spread around them; no competency stands out.



## Distribution by competency (one dot = one candidate)



## Key findings

- 1 Prioritizing under pressure and Ownership & follow-through bracket the four eligible medians (50.73 / 48.66); Data-informed judgment (51.69) is reference only. Differences are small and not a ranking.
- 2 Of the four answer traits, Risk awareness is the lowest (2.89 of 5) and Execution the highest (2.94); the gap is 0.04.
- 3 Lowest scenario means ( $n \geq 30$ ): 'Handover from a departing colleague' 2.83 and 'Customer demands an unsupported feature' 2.85; thin S03, S05 not ranked.
- 4 37 of 40 completers finished as-is and 3 completed through QC selection. 2 candidates timed out and have no result.
- 5 10 of 40 completers have 30 or more points between their highest and lowest competency (competencies with 2+ observations). Read results by competency, not by average.

### Look at first in interviews

## Ownership & follow-through

Median 48.66 (lowest eligible); 21 of 40 completers score below 50.

### Cover in onboarding

## Risk awareness

Lowest trait mean by a small margin.

### Check before using results

## 5

QC-selected or timed out. The list is on page 5.

No composite score

No ranking

QC selection ≠ misconduct

People make the decision

## What to do with this report

### 1 · Pick what matters

Choose the 1–2 competencies that matter most for the role and read only those columns on the candidate pages.

### 2 · Check first

Before using results, read page 5 for the 3 QC-selected cases and 2 time-outs.

### 3 · Prepare interviews

The page 8 questions come from this campaign's shared gaps. Use them with the questions in each individual report.

### 4 · Carry into onboarding

Put risk awareness practice into the first 30 days for everyone hired.

## Why sessions ended



# Campaign at a glance

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## Five-competency descriptive medians

|                                    |       |                                  |       |
|------------------------------------|-------|----------------------------------|-------|
| Customer problem solving median    | 50.00 | Stakeholder communication median | 50.30 |
| Prioritizing under pressure median | 50.73 | Data-informed judgment median    | 51.69 |
| Ownership & follow-through median  | 48.66 |                                  |       |

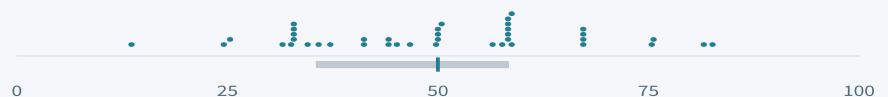
Started **42** → Completed **40** + Timed out **2**  
 Completed as-is **37** + Completed · QC-selected **3**  
 Total probes **682** · Median theta SE **0.29**

## Read these patterns, then verify

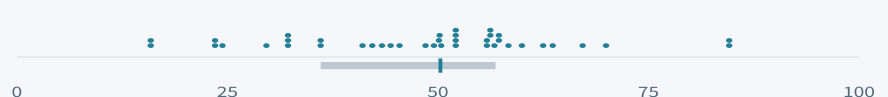
Risk awareness is the lowest of four generated-answer trait means (2.89/5; spread 0.04), not the lowest within every competency. Among scenarios with 30 or more answers, S09, Handover from a departing colleague, has the lowest mean (2.831/5). The 3 review cases were declared QC samples; 2 other sessions timed out. 36 of 40 sessions stopped when the 0.30 SE target was met and 4 at the 12-item maximum; no evidence grade is inferred.

## 40 person scores with medians and middle 50%

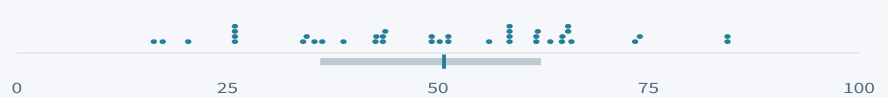
### Customer problem solving



### Stakeholder communication

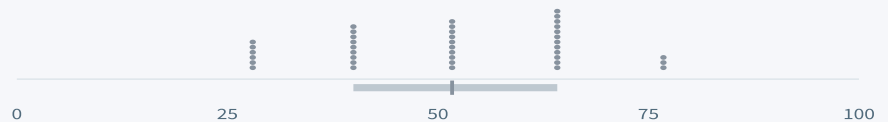


### Prioritizing under pressure

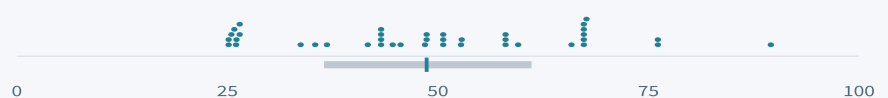


### Data-informed judgment

one observation · ref



### Ownership & follow-through

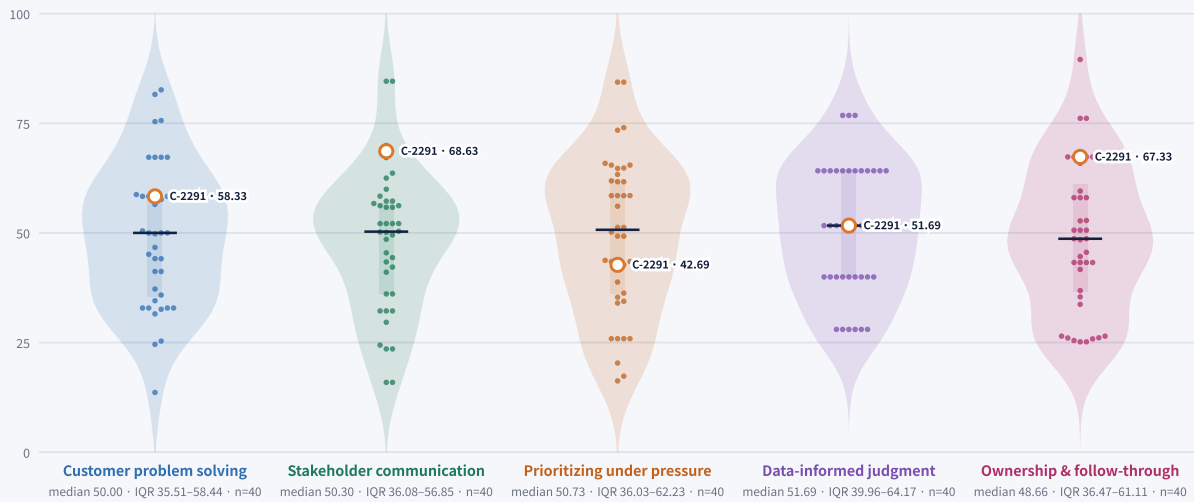


One dot per generated person, 40 per row. Repeated scores stack vertically; dots retain person ID in the source. Gray horizontal segments show p25–p75 (linear interpolation), vertical marks show medians. Descriptive 0–100 scores, not calibrated confidence intervals. Data-informed judgment: one answer/person, reference only.

COMPETENCY DISTRIBUTIONS · V3 DESCRIPTIVE

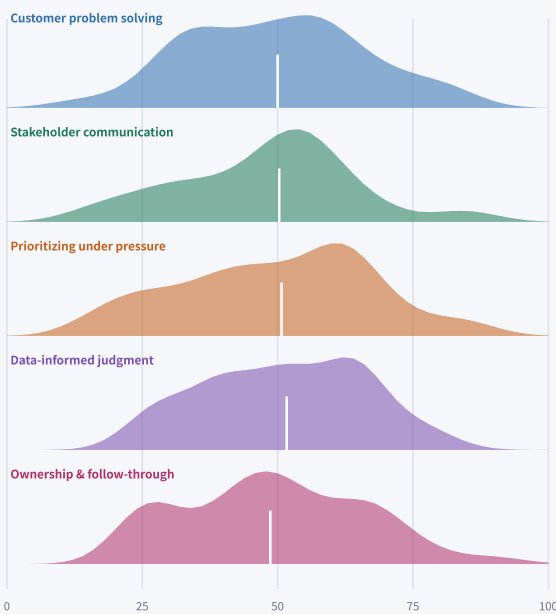
40 completers. Medians sit between 48.66 and 50.73, and scores spread widely around them.

## Violin + beeswarm



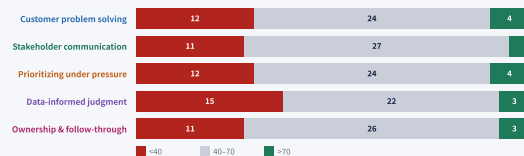
One dot = one person. White ring marks the sample individual (C-2291), a separate run that is not one of the 40. Black line = cohort median · pale bar = middle 50%.

## Ridgeline



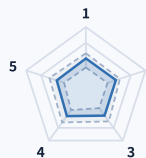
White line = median.

## People per score band



Cells = number of the 40 completers per band (score <40, 40–70, >70).

## Medians and spread on one radar



- 1 Customer problem solving · median **50.0** · IQR 35.5–58.4
- 2 Stakeholder communication · median **50.3** · IQR 36.1–56.8
- 3 Prioritizing under pressure · median **50.7** · IQR 36.0–62.2
- 4 Data-informed judgment · median **51.7** · IQR 40.0–64.2
- 5 Ownership & follow-through · median **48.7** · IQR 36.5–61.1

How to read. Solid = cohort median (40 completers); dashed = 25th and 75th percentile; rings 25/50/75/100.

# Competency distributions / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## Descriptive medians and middle-half ranges

| Competency                                      | Person n | Median | Middle 50%     |
|-------------------------------------------------|----------|--------|----------------|
| Customer problem solving                        | 40       | 50.00  | 35.51 to 58.44 |
| Stakeholder communication                       | 40       | 50.30  | 36.08 to 56.85 |
| Prioritizing under pressure                     | 40       | 50.73  | 36.03 to 62.23 |
| Data-informed judgment (one answer/person; ref) | 40       | 51.69  | 39.96 to 64.17 |
| Ownership & follow-through                      | 40       | 48.66  | 36.47 to 61.11 |

## Every person score with 5-point-bin silhouettes

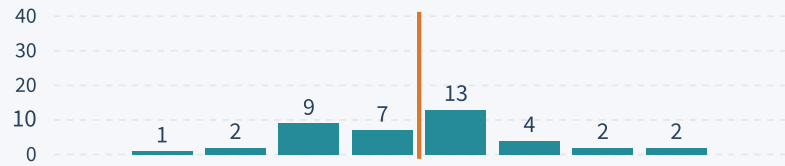


One dot per person, 40 per row. Pale shapes show counts per 5-point score bin (thickness 0.8 SVG units per person on each side), not a smoothed or validated density. Gray horizontal segment: p25–p75 (linear); navy tick: median; orange ring: first generated person (C-0001). Descriptive 0–100. Data-informed judgment is reference-only; its comparison markers are withheld.

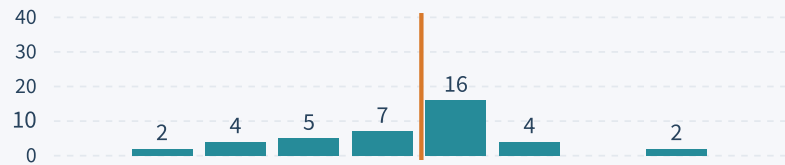
## Competency distributions / 2

### Source score frequencies · 10-point bins

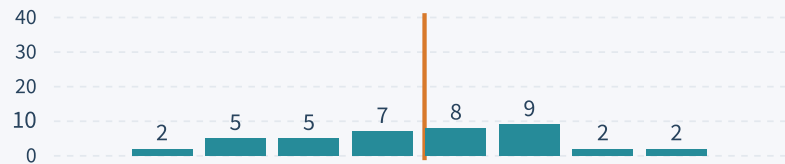
Customer problem solving  
n=40



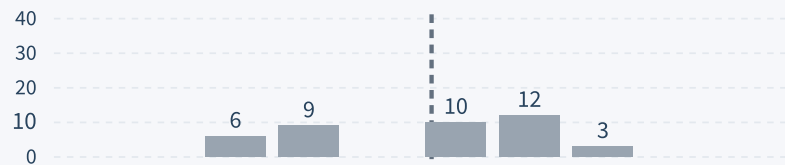
Stakeholder communication  
n=40



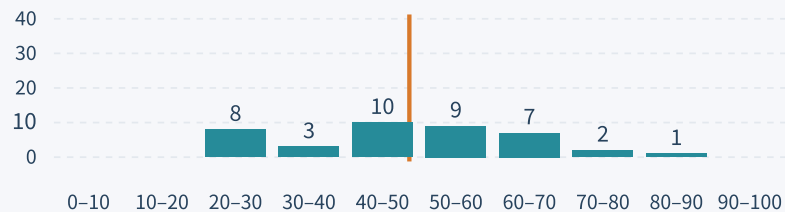
Prioritizing under pressure  
n=40



Data-informed judgment \*  
n=40; reference only



Ownership & follow-through  
n=40



Horizontal: descriptive score 0–100; 10-point bins

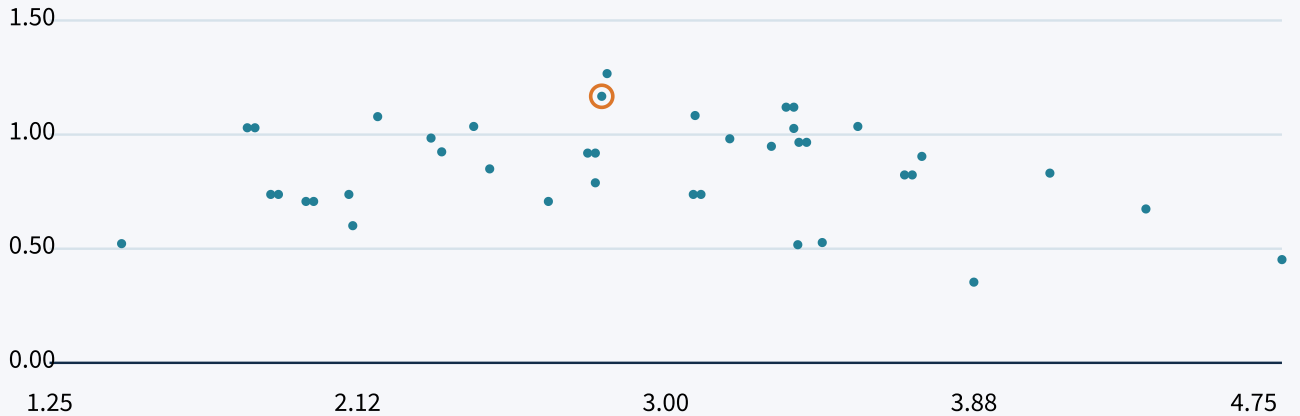
Vertical: people 0–40; orange: median; gray dashed: reference median

Every row uses the same zero-based count scale. Zero-count bins have zero height.

Counts sum to 40 generated completers per row; orange lines show eligible row medians; gray dashed line shows the reference-only Data median. Bins are lower-inclusive and upper-exclusive, except the last bin includes 100. Data-informed judgment\* has one observation/person and is gray reference-only. Scores are descriptive, not calibrated job performance.

## Competency distributions / 3

### Forty matched person means and within-person variation

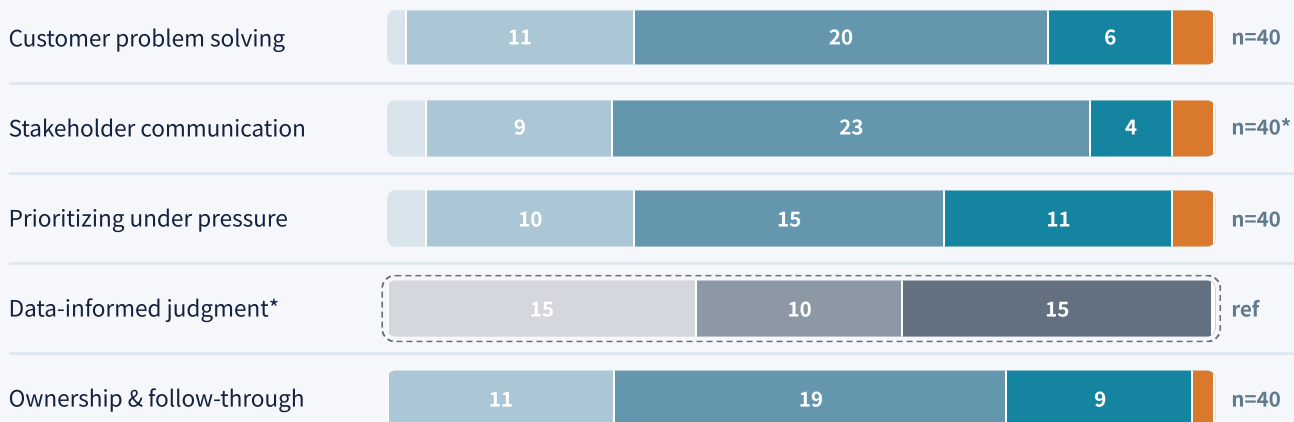


Sample SD across 8–12 answers (1–5)

Mean generated answer effectiveness (1–5)

40 dots, one per person (C-0001 has the orange ring), source IDs retained in each mark. Horizontal: mean generated effectiveness over that person's administered answers; vertical: sample SD over the same answers (8 to 12 per person). Dots that would overlap are nudged horizontally; descriptive, no smooth density or workplace inference.

### Five score bands · exact counts /40



0–19 20–39 40–59 60–79 80–100

Narrow segments (fewer than 3 people) carry no printed count; the ledger below lists every band count, including zeros.

Customer problem solving: 1 / 11 / 20 / 6 / 2

Stakeholder communication: 2 / 9 / 23 / 4 / 2

Prioritizing under pressure: 2 / 10 / 15 / 11 / 2

Data-informed judgment: 0 / 15 / 10 / 15 / 0

Ownership & follow-through: 0 / 11 / 19 / 9 / 1

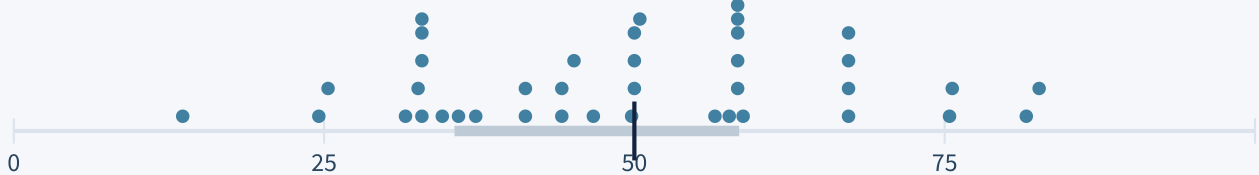
Segments and five-number ledger count people in 0–19 / 20–39 / 40–59 / 60–79 / 80–100 descriptive score bands (40 completed synthetic people). Data-informed judgment\* is one generated answer per person: show its counts for reference only, not as a peer comparison or hiring rank. No validated norm or competency coverage.

## Cohort comparison evidence / 1 of 9

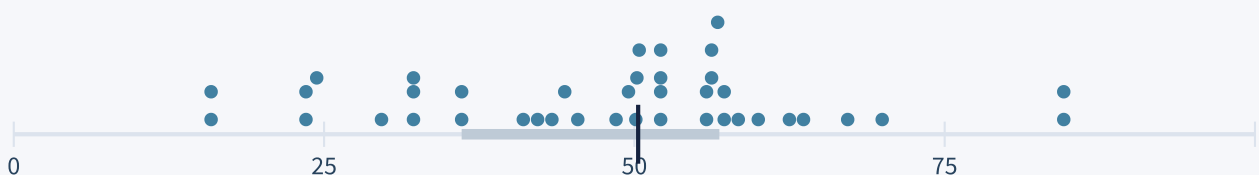
Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

### Score spread of the 40 people

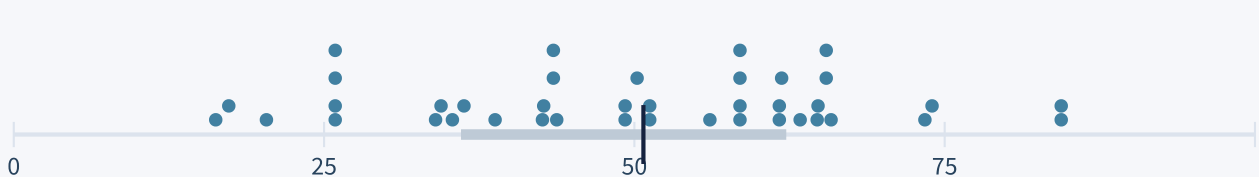
Customer problem solving median 50.00 · middle 50% 35.51–58.44



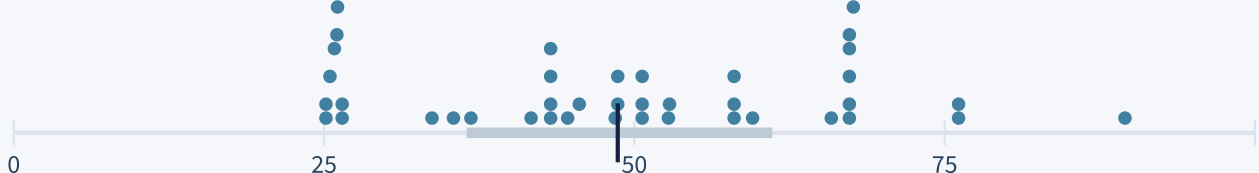
Stakeholder communication median 50.30 · middle 50% 36.08–56.85



Prioritizing under pressure median 50.73 · middle 50% 36.03–62.23



Ownership & follow-through median 48.66 · middle 50% 36.47–61.11



Data-informed judgment has one observation/person and stays reference-only, without a comparison mark.

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

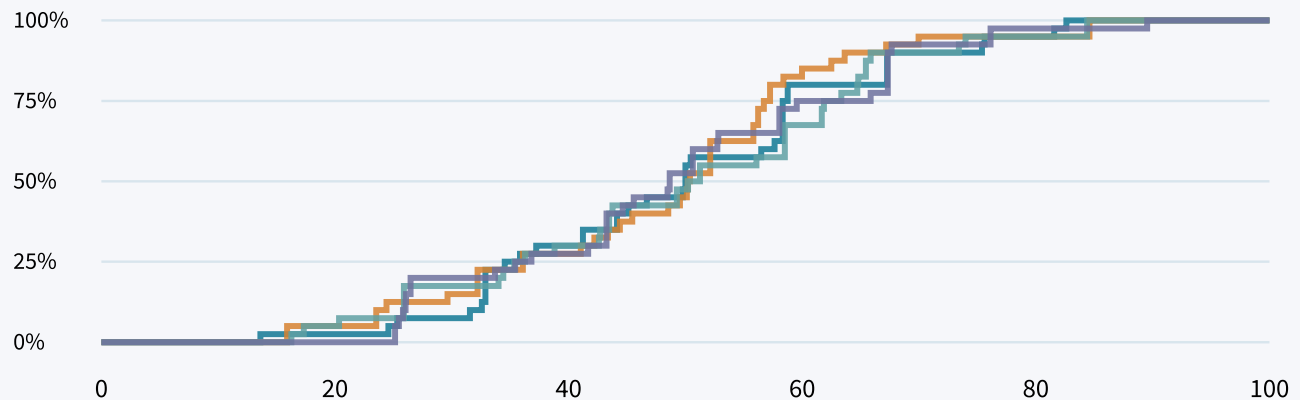
Reading guide: each row is one competency and each dot is one person, so every row has 40 dots. Dots that share a score stack vertically; the gray band is the middle 50% and the navy tick the median.

## Cohort comparison evidence / 2 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. 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. Ties cause steps. 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     |
| Stakeholder communication   | 0   | 2     | 4     | 5     | 7     | 16    | 4     | 0     | 2     | 0     |
| Prioritizing under pressure | 0   | 2     | 5     | 5     | 7     | 8     | 9     | 2     | 2     | 0     |
| Ownership & follow-through  | 0   | 0     | 8     | 3     | 10    | 9     | 7     | 2     | 1     | 0     |

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

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

Reading guide: each curve shows, for one competency, the share of the 40 generated people at or below a score on the 0-100 descriptive scale. Ties cause steps, and there is no external percentile.

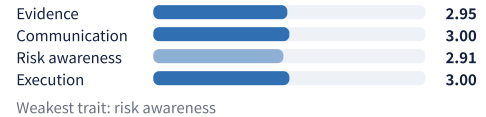
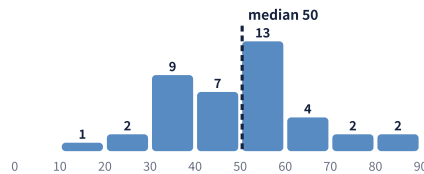
COMPETENCY DEEP DIVE OBSERVED DESCRIPTIVE

Each competency differs in shape, amount of evidence and lowest trait mean, though the gaps are small. Even at the same score, the interview question should differ.

## Customer problem solving

**50.00** median

middle 50% 35.51–58.44  
below 50: 19 · above 70: 4  
obs per person 2–3

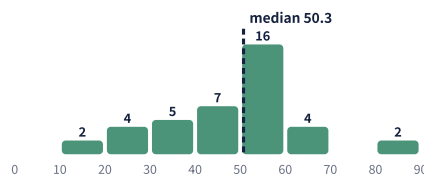


Weakest trait: risk awareness

## Stakeholder communication

**50.30** median

middle 50% 36.08–56.85  
below 50: 18 · above 70: 2  
obs per person 2–3

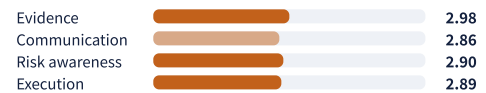
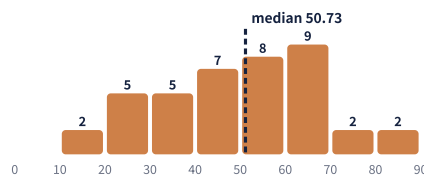


Weakest trait: communication

## Prioritizing under pressure

**50.73** median

middle 50% 36.03–62.23  
below 50: 19 · above 70: 2  
obs per person 1–3

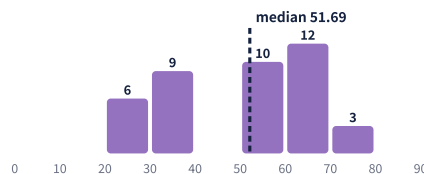


Weakest trait: communication

## Data-informed judgment

**51.69** median

middle 50% 39.96–64.17  
below 50: 15 · above 70: 3  
obs per person 1 · reference only

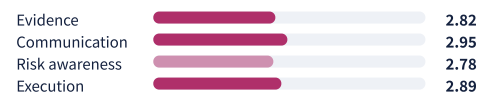
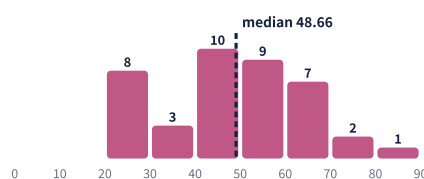


Weakest trait: communication

## Ownership &amp; follow-through

**48.66** median

middle 50% 36.47–61.11  
below 50: 21 · above 70: 3  
obs per person 2–3



Weakest trait: risk awareness

Histogram = each completer's competency score (0–100). Trait bars = mean of the four traits over every administered scenario for that competency (1–5). Thin = fewer than 2 observations (reference only).

## Observations per competency (people)

| Competency                  | 1 (ref. only) | 2  | 3  | 4+ |
|-----------------------------|---------------|----|----|----|
| Customer problem solving    | 0             | 26 | 14 | 0  |
| Stakeholder communication   | 0             | 10 | 30 | 0  |
| Prioritizing under pressure | 3             | 33 | 4  | 0  |
| Data-informed judgment      | 40            | 0  | 0  | 0  |
| Ownership & follow-through  | 0             | 29 | 11 | 0  |

40 completers.

# Competency deep dive / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

The five-competency median and IQR source tables remain on pages 4–5. Below is observation coverage for the same completed-person set.

## Observation coverage per competency

The 40 generated completers each answered 8 to 12 of the 14 scenarios, chosen adaptively, so observation counts vary by person. The Data-informed judgment lens has one observation per person and is not eligible for a relative rank.

| Competency                  | Observations / person | People | Use                    |
|-----------------------------|-----------------------|--------|------------------------|
| Customer problem solving    | 2.35 (2–3)            | 40     | descriptive comparison |
| Stakeholder communication   | 2.75 (2–3)            | 40     | descriptive comparison |
| Prioritizing under pressure | 2.02 (1–3)            | 40     | descriptive comparison |
| Data-informed judgment      | 1.00 (1–1)            | 40     | reference only         |
| Ownership & follow-through  | 2.27 (2–3)            | 40     | descriptive comparison |

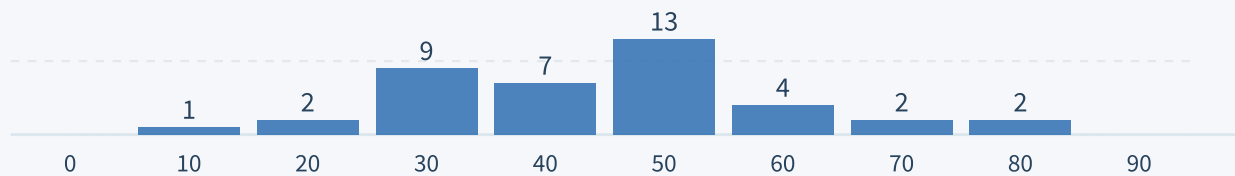
## Competency deep dive / 2

### Computed evidence · continued

|                                      |       |
|--------------------------------------|-------|
| Ownership & follow-through median    | 48.66 |
| Ownership & follow-through IQR lower | 36.47 |
| Ownership & follow-through IQR upper | 61.11 |

### Competency histogram (1 of 2)

#### Customer problem solving



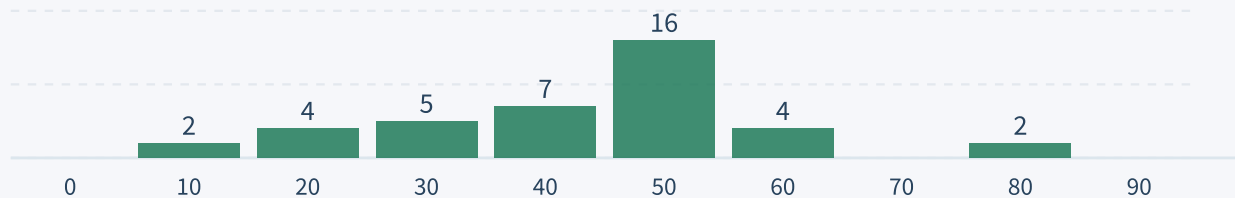
n=40; median 50.00/100; 10-point bins; last bin includes 100

Vertical: people per bin (0-30 scale). Horizontal: bin lower edge, 0-100 descriptive score; final bin includes 100.

Synthetic adaptive source, n=40 completed people. Figure shows competency histogram on its labeled scale; descriptive, not a hiring decision.

### Competency histogram (2 of 2)

#### Stakeholder communication



n=40; median 50.30/100; 10-point bins; last bin includes 100

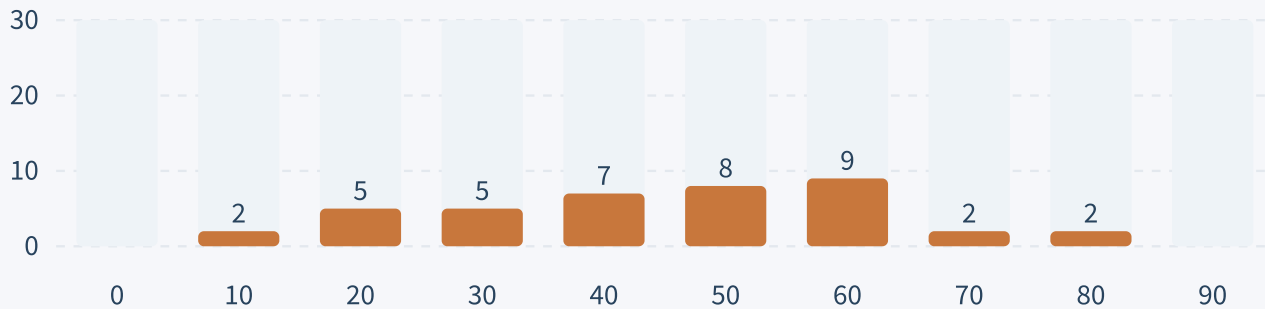
Vertical: people per bin (0-30 scale). Horizontal: bin lower edge, 0-100 descriptive score; final bin includes 100.

Synthetic adaptive source, n=40 completed people. Figure shows competency histogram on its labeled scale; descriptive, not a hiring decision.

## Competency deep dive / 3

### Competency histogram

#### Prioritizing under pressure



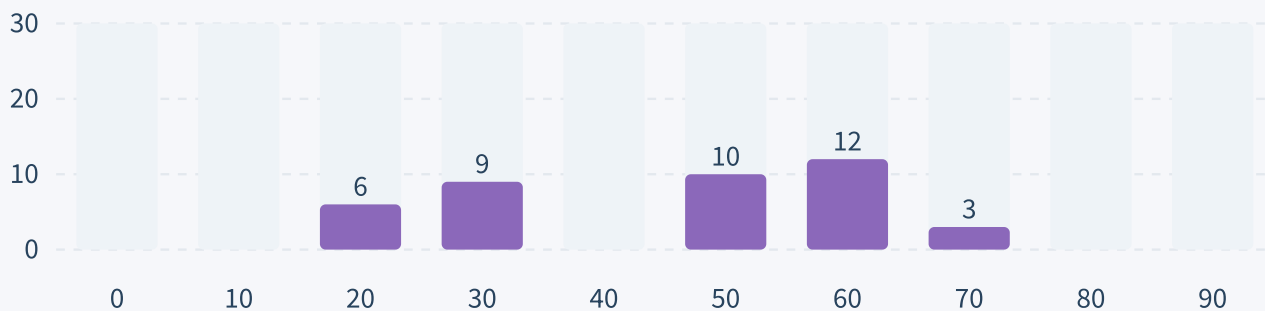
**n=40; median 50.73/100; 10-point bins; last bin includes 100**

Vertical: people per bin (0-30 scale). Horizontal: bin lower edge, 0-100 descriptive score; final bin includes 100.

Synthetic adaptive source, n=40 completed people. Figure shows competency histogram on its labeled scale; descriptive, not a hiring decision.

### Competency histogram · reference only\*

#### Data-informed judgment



**n=40; median 51.69/100; 10-point bins; last bin includes 100**

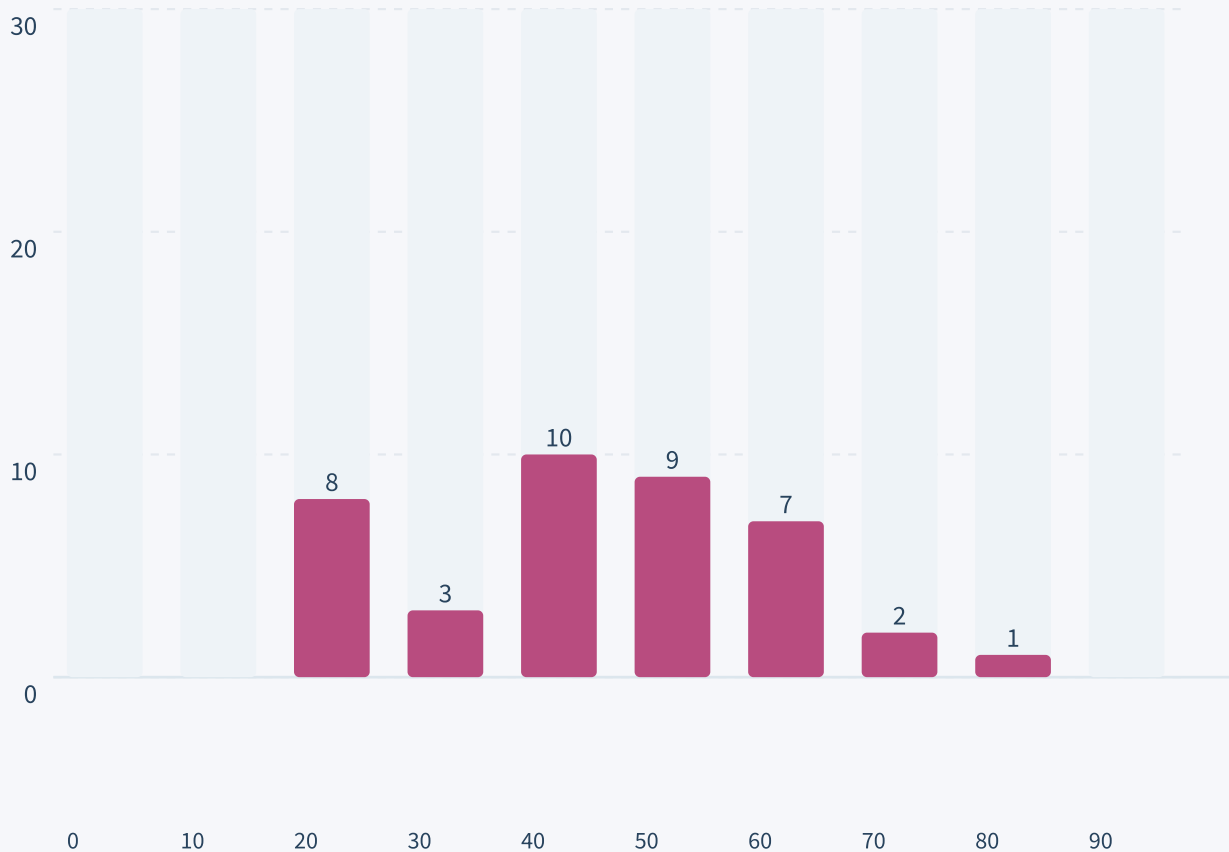
Vertical: people per bin (0-30 scale). Horizontal: bin lower edge, 0-100 descriptive score; final bin includes 100.

One generated observation per person: scores run from 28.0 to 76.8; 15 of 40 below 50 and 3 at 70 or above. This histogram describes the synthetic sample only. \* Reference-only lens, not eligible for relative competency comparison, calibrated skill inference, or hiring decisions.

## Competency deep dive / 4

### Competency histogram

#### Ownership & follow-through



n=40; median 48.66/100; 10-point bins; last bin includes 100

Vertical: people per bin (0-30 scale). Horizontal: bin lower edge, 0-100 descriptive score; final bin includes 100.

Synthetic adaptive source, n=40 completed people. Figure shows competency histogram on its labeled scale; descriptive, not a hiring decision.

Reading guide: each bar counts the people in one 10-point score bin out of the 40 completed people. The horizontal axis is the 0-100 descriptive score and the last bin includes 100. This is descriptive, not a hiring decision.

## Cohort comparison evidence / 3 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

### Paired cohort rows 1–20, four eligible lenses



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

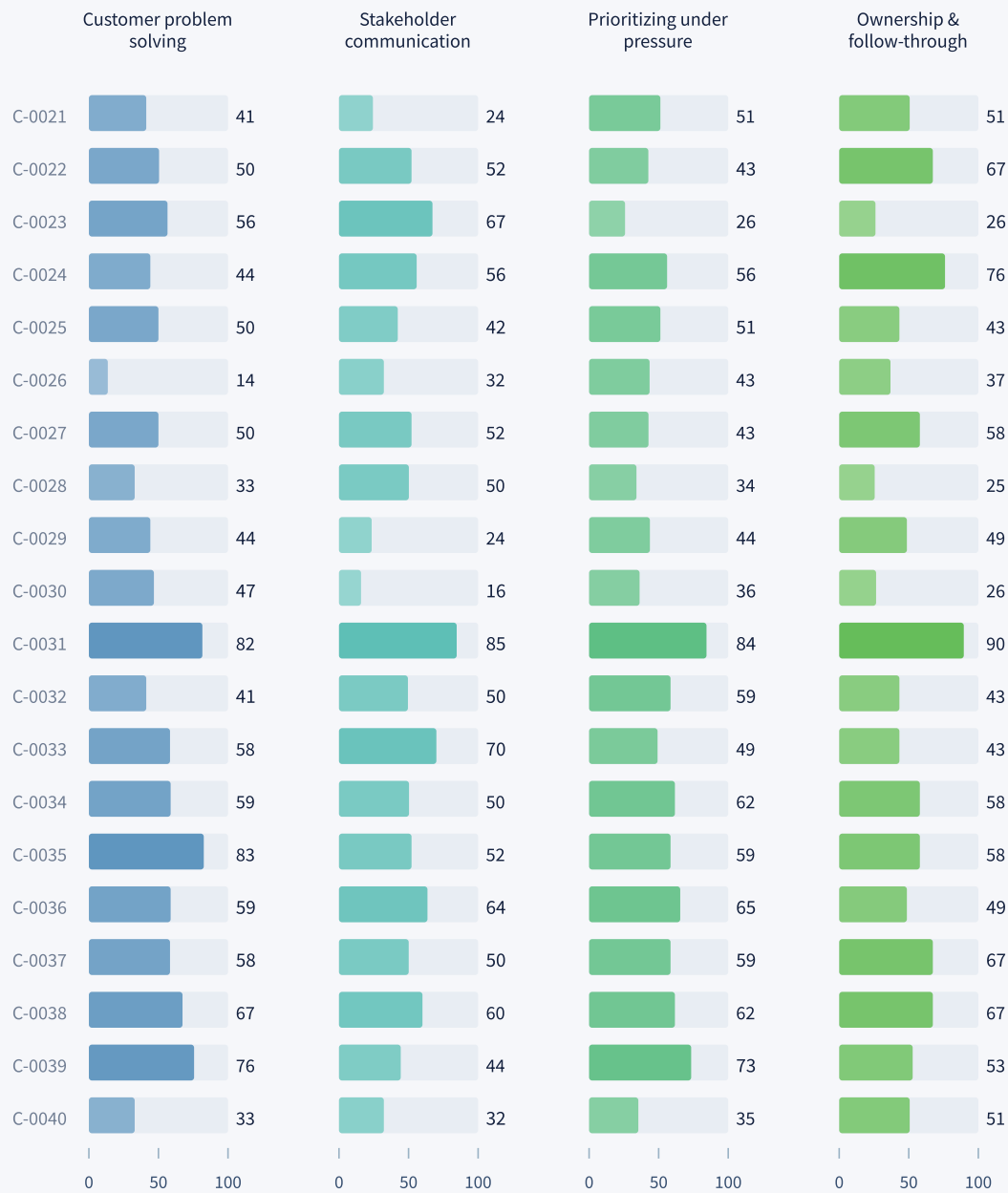
The fifth competency remains reference-only for all 40 people. Values are shown rounded to whole points. Row order is candidate ID, not ability order or hiring rank.

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

## Cohort comparison evidence / 4 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. 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. Values are shown rounded to whole points. Row order is candidate ID, not ability order or hiring rank.

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

ANSWER TRAITS

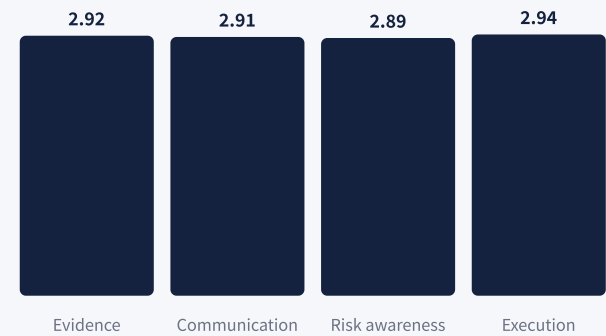
Risk awareness is the lowest trait mean overall, but the four traits differ by only 0.04 points. Treat any onboarding focus as a hypothesis.

Competency × answer trait mean (1–5, all scenario answers)

|                             | Evidence | Communication | Risk awareness | Execution | Mean |
|-----------------------------|----------|---------------|----------------|-----------|------|
| Customer problem solving    | 2.95     | 3.00          | 2.91           | 3.00      | 2.97 |
| Stakeholder communication   | 2.92     | 2.85          | 2.98           | 2.95      | 2.93 |
| Prioritizing under pressure | 2.98     | 2.86          | 2.90           | 2.89      | 2.91 |
| Data-informed judgment      | 2.98     | 2.83          | 2.85           | 2.92      | 2.89 |
| Ownership & follow-through  | 2.82     | 2.95          | 2.78           | 2.89      | 2.86 |
| All scenarios               | 2.92     | 2.91          | 2.89           | 2.94      |      |

Based on 416 answers · darker = higher (scale 2.7 to 3.3)

Mean by trait



Execution is highest (2.94) and Risk awareness is lowest (2.89).

Patterns

- Risk awareness is the lowest trait overall (2.89), a gap of only 0.04 from the highest.
- The widest gap between traits is in ownership & follow-through (2.95 communication vs 2.78 risk awareness, 0.16 points).
- Data-informed judgment rests on one scenario per person, so its row is reference only.

Evidence

2.92

Backs the answer with facts, numbers or examples

Communication

2.91

Clear and pitched to the listener

Risk awareness

2.89

Names risks, fallbacks and how to check

# Answer traits / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## Four answer-quality traits by competency

### Competency trait heatmap

|                             | Evidence | Communication | Risk awareness | Execution |
|-----------------------------|----------|---------------|----------------|-----------|
| Customer problem solving    | 2.95     | 3.00          | 2.91           | 3.00      |
| Stakeholder communication   | 2.92     | 2.85          | 2.98           | 2.95      |
| Prioritizing under pressure | 2.98     | 2.86          | 2.90           | 2.89      |
| Data-informed judgment*     | 2.98     | 2.83          | 2.85           | 2.92      |
| Ownership & follow-through  | 2.82     | 2.95          | 2.78           | 2.89      |

Generated mean (scale 2.7–3.3) 2.7 2.85 3.0 3.15 3.3

Color intensity scales from 2.7 (lightest) to 3.3 (darkest). Read the numeric means for exact values. Gray marks are reference-only, not this color scale.

Cells are 1–5 generated-answer trait means. \* Data-informed judgment has one generated answer per person (40 answers total); its four displayed cohort means are reference-only, not comparable evidence for competency-level skill or hiring decisions. The other rows are descriptive, not calibrated skill measures.

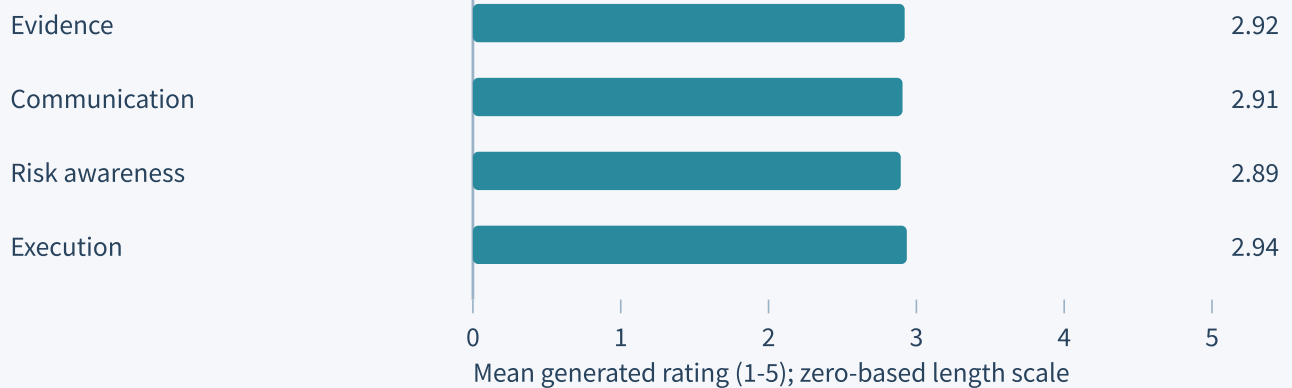
## Answer traits / 2

**2.94** HIGHEST MEAN ·  
EXECUTION

**2.89** LOWEST MEAN ·  
RISK AWARENESS

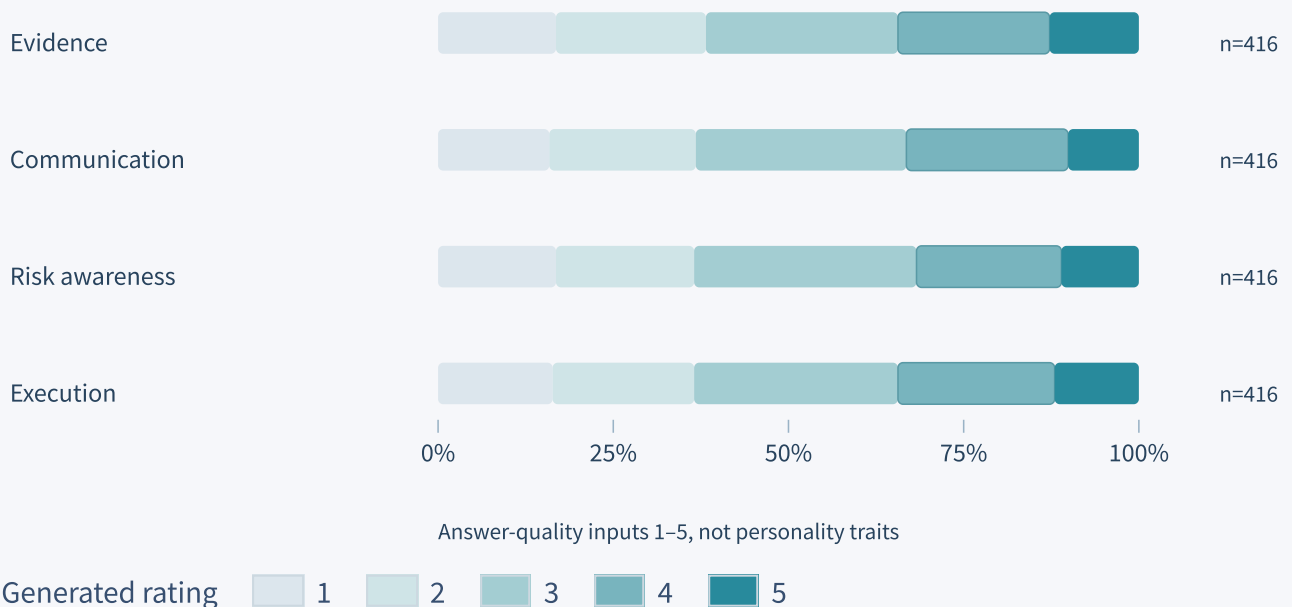
**416** RATINGS PER TRAIT

### Trait mean bars



Synthetic descriptive chart; verify the underlying data before use.

### Trait rating distribution



Share of 416 generated answers per trait (40 people × 8 to 12 administered scenarios). The mean chart above shows numeric means. In this distribution, n=416 is the answer count per trait; colors show rating categories, not evidence grades.

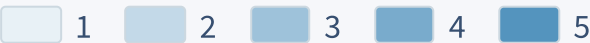
Synthetic adaptive source, n=40 completed people. Figure shows trait rating distribution on its labeled scale; descriptive, not a hiring decision.

## Cohort comparison evidence / 5 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

### Scenario by four answer traits

|            | Evidence | Communication | Risk | Execution |
|------------|----------|---------------|------|-----------|
| S01 · n=38 | 2.97     | 2.97          | 2.76 | 3.08      |
| S02 · n=38 | 2.87     | 2.82          | 3.13 | 3.03      |
| S03 · n=11 | 2.82     | 2.73          | 2.45 | 2.64      |
| S04 · n=40 | 2.98     | 2.83          | 2.85 | 2.92      |
| S05 · n=12 | 2.67     | 2.83          | 2.67 | 2.58      |
| S06 · n=39 | 2.85     | 2.90          | 2.85 | 2.82      |
| S07 · n=35 | 2.94     | 2.97          | 3.00 | 2.86      |
| S08 · n=40 | 3.02     | 2.80          | 2.88 | 2.88      |
| S09 · n=40 | 2.65     | 2.92          | 2.75 | 3.00      |
| S10 · n=17 | 3.12     | 3.29          | 3.41 | 3.24      |
| S11 · n=37 | 2.95     | 2.78          | 2.81 | 2.97      |
| S12 · n=29 | 2.93     | 3.00          | 3.07 | 2.97      |
| S13 · n=1  | 4.00     | 3.00          | 4.00 | 4.00      |
| S14 · n=39 | 3.05     | 3.00          | 2.85 | 2.87      |

Generated mean (1-5) 

Each cell is the mean of the generated 1–5 answer-quality ratings for that scenario (n varies by scenario because scenarios were chosen adaptively; S13 has a single answer); not a personality trait.

Source: answer-quality trait ratings, 40 people, 416 administered answers

## SCENARIO RESULTS

With 30+ answers, scenario means sit between 2.85 and 2.97; 'Handover from a departing colleague' is lowest. S05, S03, S13 and S10 have fewer answers.

## Results by scenario (lowest mean first)

| Situation                                  | Competency                  | Answers | Mean (1–5) | Rated 1–2 | Probes / answer | Rated 4–5 |
|--------------------------------------------|-----------------------------|---------|------------|-----------|-----------------|-----------|
| Three urgent tickets, one hour             | Prioritizing under pressure | 11†     | 2.64       | 5         | 1.73            | 2         |
| A follow-up you promised slipped           | Ownership & follow-through  | 12†     | 2.75       | 6         | 0.83            | 4         |
| Handover from a departing colleague        | Ownership & follow-through  | 40      | 2.85       | 19        | 1.93            | 15        |
| Bad news for an executive sponsor          | Stakeholder communication   | 35      | 2.86       | 12        | 1.63            | 9         |
| Customer demands an unsupported feature    | Customer problem solving    | 39      | 2.87       | 14        | 1.51            | 11        |
| Team disagreement on an account plan       | Stakeholder communication   | 37      | 2.89       | 15        | 1.73            | 13        |
| Quarter-end workload spike                 | Prioritizing under pressure | 40      | 2.90       | 16        | 1.77            | 17        |
| Churn dashboard disagrees with field notes | Data-informed judgment      | 40      | 2.92       | 15        | 1.68            | 15        |
| Escalated renewal at risk                  | Customer problem solving    | 38      | 2.95       | 15        | 1.76            | 15        |
| Conflicting asks from sales and product    | Stakeholder communication   | 38      | 2.95       | 14        | 1.18            | 11        |
| Owning the post-incident review            | Ownership & follow-through  | 39      | 2.97       | 12        | 1.69            | 12        |
| Deadline vs quality trade-off              | Prioritizing under pressure | 29†     | 3.07       | 12        | 2.03            | 11        |
| Onboarding stalled at a key account        | Customer problem solving    | 17†     | 3.23       | 4         | 1.06            | 7         |
| Pricing exception request                  | Prioritizing under pressure | 1†      | 4.00       | 0         | 3.00            | 1         |

† = fewer than 30 answers; scenarios were chosen adaptively, so each was asked of a different subset of the 40 completers. Rated 1–2 and Rated 4–5 are counts of answers. Means are generated answer ratings, descriptive only.

## Lowest mean (30+ answers)

## Handover from a departing colleague

mean 2.85 · 19 rated 1–2 · 40 answers

## Highest mean (30+ answers)

## Owning the post-incident review

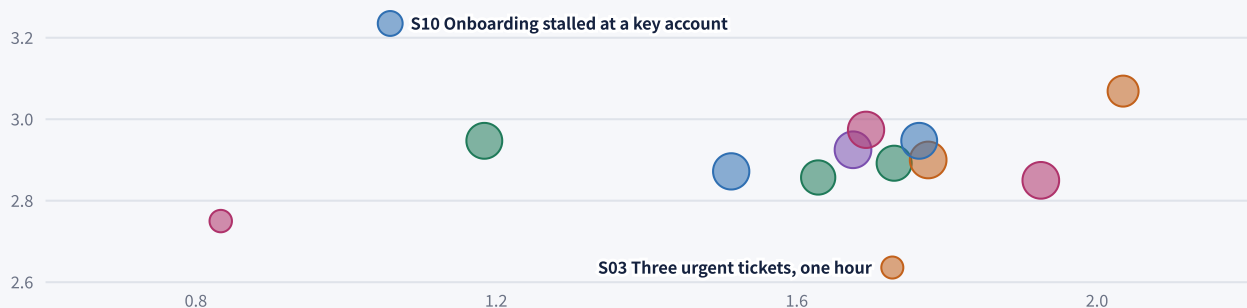
mean 2.97 · 1.69 probes per answer

## Most probing

## Handover from a departing colleague

1.93 probes per answer · cohort average 1.64

## Probes × mean score (circle size = number of answers)



## Scenarios sorted by generated answer mean

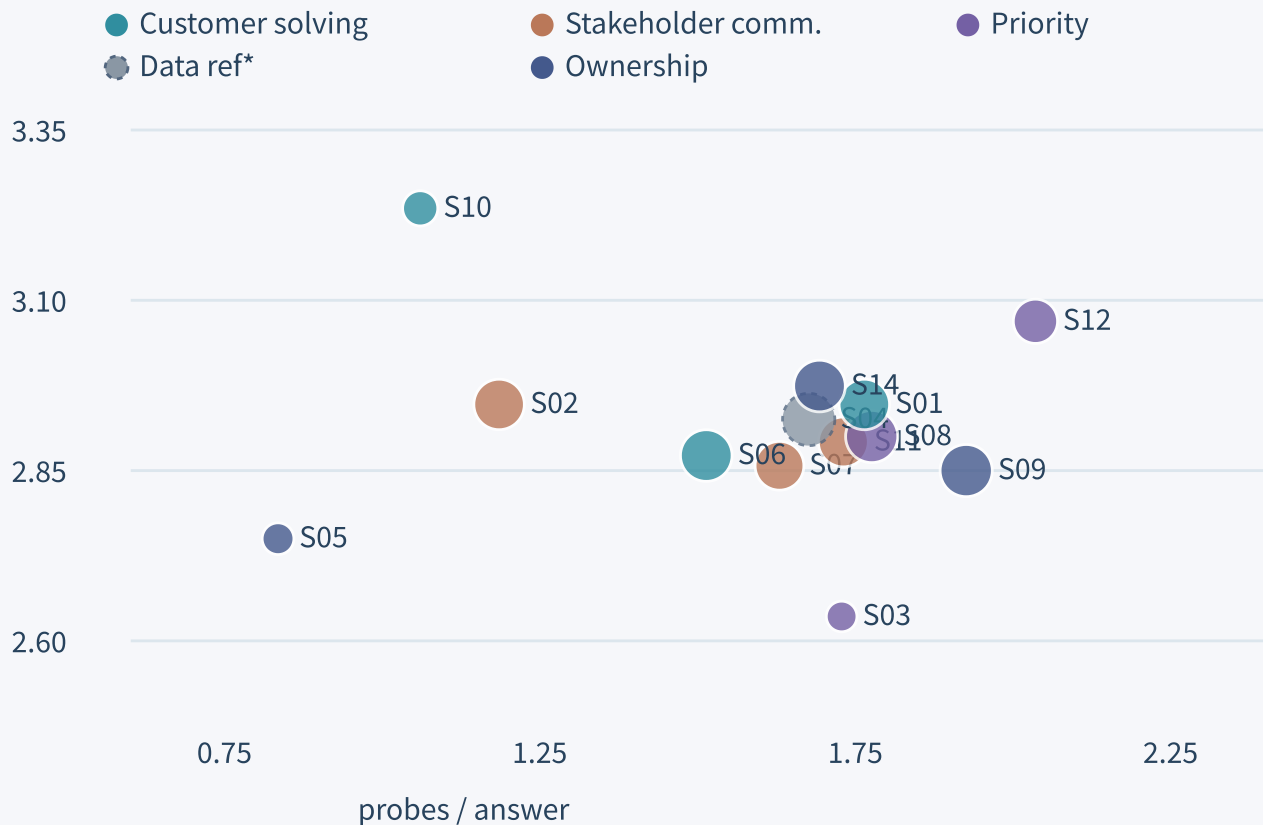
Among scenarios with 30+ answers, S09 has the lowest generated answer mean at 2.850/5 (40 answers), followed by S07 at 2.857/5 and S06 at 2.872/5. Of 416 answers, 159 are rated 1–2. Do not call a scenario a real-world weakness. S03, S05, S10, S12 and S13 have fewer than 30 answers († = thin). Probe counts are operational observations, not a causal explanation of score.

| ID  | Scenario                                   | Competency                  | Answers | Mean /5 | Rated 1–2 | Probes /answer | Rated 4–5 |
|-----|--------------------------------------------|-----------------------------|---------|---------|-----------|----------------|-----------|
| S03 | Three urgent tickets, one hour             | Prioritizing under pressure | 11†     | 2.636   | 5/11      | 1.73           | 2/11      |
| S05 | A follow-up you promised slipped           | Ownership & follow-through  | 12†     | 2.750   | 6/12      | 0.83           | 4/12      |
| S09 | Handover from a departing colleague        | Ownership & follow-through  | 40      | 2.850   | 19/40     | 1.93           | 15/40     |
| S07 | Bad news for an executive sponsor          | Stakeholder communication   | 35      | 2.857   | 12/35     | 1.63           | 9/35      |
| S06 | Customer demands an unsupported feature    | Customer problem solving    | 39      | 2.872   | 14/39     | 1.51           | 11/39     |
| S11 | Team disagreement on an account plan       | Stakeholder communication   | 37      | 2.892   | 15/37     | 1.73           | 13/37     |
| S08 | Quarter-end workload spike                 | Prioritizing under pressure | 40      | 2.900   | 16/40     | 1.77           | 17/40     |
| S04 | Churn dashboard disagrees with field notes | Data-informed judgment      | 40      | 2.925   | 15/40     | 1.68           | 15/40     |
| S01 | Escalated renewal at risk                  | Customer problem solving    | 38      | 2.947   | 15/38     | 1.76           | 15/38     |
| S02 | Conflicting asks from sales and product    | Stakeholder communication   | 38      | 2.947   | 14/38     | 1.18           | 11/38     |
| S14 | Owning the post-incident review            | Ownership & follow-through  | 39      | 2.974   | 12/39     | 1.69           | 12/39     |
| S12 | Deadline vs quality trade-off              | Prioritizing under pressure | 29†     | 3.069   | 12/29     | 2.03           | 11/29     |
| S10 | Onboarding stalled at a key account        | Customer problem solving    | 17†     | 3.235   | 4/17      | 1.06           | 7/17      |
| S13 | Pricing exception request                  | Prioritizing under pressure | 1†      | 4.000   | 0/1       | 3.00           | 1/1       |

\* Data-informed judgment (S04) was asked of all 40 people and is a reference-only competency lens. This table compares each scenario's answers descriptively, not people. Scenarios were chosen adaptively, so each person answered 8–12 of the 14 scenarios.

## Scenario results / 2

### Follow-up probes versus answer rating · 6.1



Answers (n)    ● 10    ● 25    ● 40

Marker radius =  $4 + 5.2 \times (\text{answers}/40)$ . Marker area is not proportional to count.

Each point is one scenario with 10 or more answers (S13, n=1, not plotted). Horizontal: mean probes/answer; vertical: generated answer mean. Marker size increases with the number of answers; colors group competencies. Descriptive association is not evidence that probing caused a score.

Gray dashed S04 is the Data-informed judgment reference scenario (all 40 people). Other colors identify competencies, not performance levels.

SCENARIO OVERVIEW DESCRIPTIVE

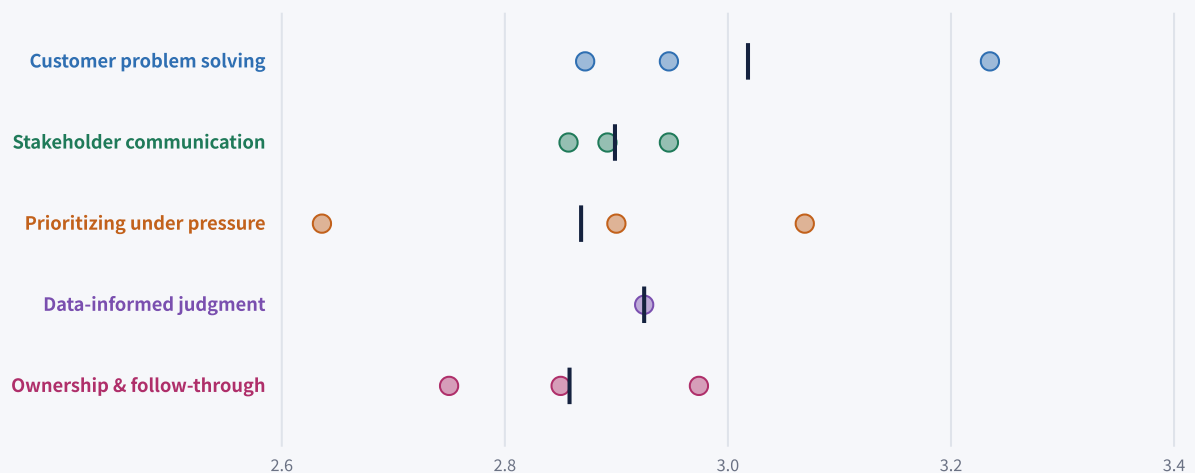
# All 14 scenarios at a glance: means span 2.64 to 4.00; S13 has one answer.

## 14 scenarios · metric heatmap

|                                            | Mean (1–5) | Rated 1–2 | Probes / answer | Rated 4–5 |
|--------------------------------------------|------------|-----------|-----------------|-----------|
| Three urgent tickets, one hour †           | 2.64       | 5         | 1.73            | 2         |
| A follow-up you promised slipped †         | 2.75       | 6         | 0.83            | 4         |
| Handover from a departing colleague        | 2.85       | 19        | 1.93            | 15        |
| Bad news for an executive sponsor          | 2.86       | 12        | 1.63            | 9         |
| Customer demands an unsupported feature    | 2.87       | 14        | 1.51            | 11        |
| Team disagreement on an account plan       | 2.89       | 15        | 1.73            | 13        |
| Quarter-end workload spike                 | 2.90       | 16        | 1.77            | 17        |
| Churn dashboard disagrees with field notes | 2.92       | 15        | 1.68            | 15        |
| Escalated renewal at risk                  | 2.95       | 15        | 1.76            | 15        |
| Conflicting asks from sales and product    | 2.95       | 14        | 1.18            | 11        |
| Owning the post-incident review            | 2.97       | 12        | 1.69            | 12        |
| Deadline vs quality trade-off †            | 3.07       | 12        | 2.03            | 11        |
| Onboarding stalled at a key account †      | 3.23       | 4         | 1.06            | 7         |
| Pricing exception request †                | 4.00       | 0         | 3.00            | 1         |

Colour depth is relative within each column. Sorted by mean, lowest first. † = fewer than 30 answers. Scenarios were chosen adaptively; n per scenario is on pages 25–26.

## Scenario mean by competency



One dot per scenario with 10+ answers (S13 not plotted); black tick = average of that competency's plotted scenarios. Data-informed judgment has one scenario.

## Difficulty and separation / 1

Pages 25–26 list the 14 scenarios with mean answer rating, answer count, count rated 1–2, mean follow-up probes and count rated 4–5. Other-scenario thirds are examined here. S13 (1 answer) is not shown; scenarios with fewer than 30 answers use smaller thirds (n/3 people each).

### Other-scenario thirds: focal answer difference · 7.1



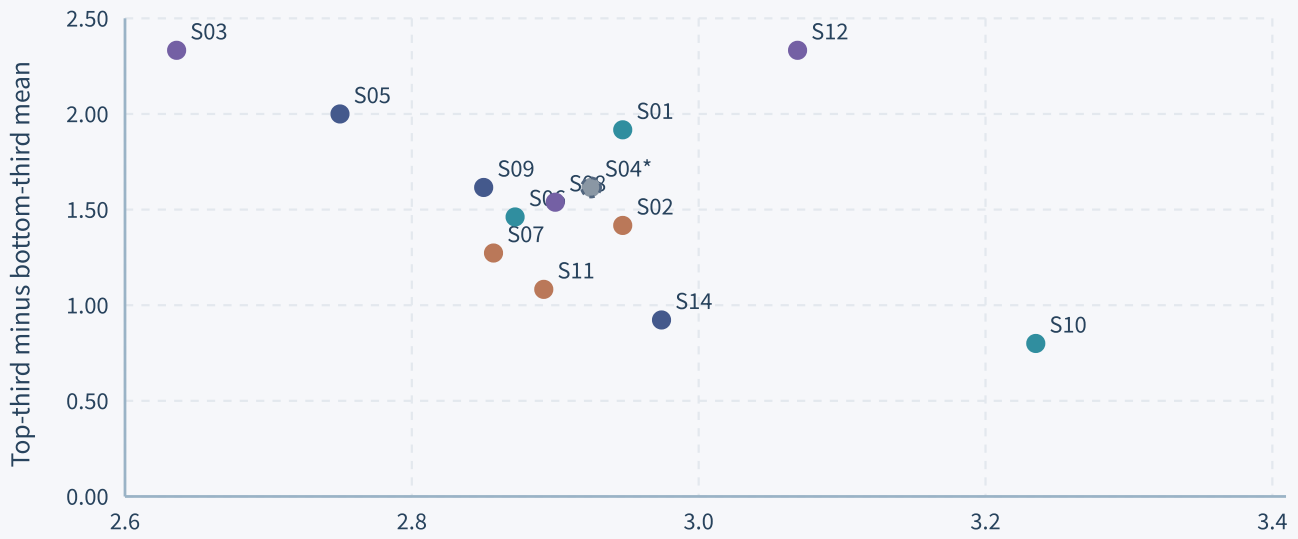
Scenario key: S02 Conflicting asks from sales and product · S04 Churn dashboard disagrees with field notes · S10 Onboarding stalled at a key account · S05 A follow-up you promised slipped · S06 Customer demands an unsupported feature · S09 Handover from a departing colleague · S13 Pricing exception request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

For each scenario, people who answered it are ranked on their mean across their other answered scenarios (ties by synthetic ID); the focal answer mean of the lower and upper third (n/3 people each) is compared. Adaptive item sets differ per person, so this is item review, not candidate grading.

\* S04 Data-informed judgment is a one-answer-per-person competency lens. Its thirds are a reference-only focal-scenario comparison, not eligible competency-level skill separation.

## Difficulty and separation / 2

Difficulty and separation scatter · 7.2



Scenario mean answer rating (1–5)

Marker size is constant. Colors: teal customer solving; orange stakeholder communication; purple priority; navy ownership; gray dashed S04\* Data reference. Colors are not performance levels.

Each equal-size dot is one scenario with 10+ answers (S13 not plotted). Differences are rating points, not a validity score; thirds are n/3 people each, so thin scenarios (S03, S05, S10, S12) are noisy.

| ID  | Scenario                                   | Mean 1–5 | Other-scenario thirds gap |
|-----|--------------------------------------------|----------|---------------------------|
| S01 | Escalated renewal at risk                  | 2.95     | +1.92                     |
| S02 | Conflicting asks from sales and product    | 2.95     | +1.42                     |
| S03 | Three urgent tickets, one hour             | 2.64     | +2.33                     |
| S04 | Churn dashboard disagrees with field notes | 2.92     | +1.62                     |
| S05 | A follow-up you promised slipped           | 2.75     | +2.00                     |
| S06 | Customer demands an unsupported feature    | 2.87     | +1.46                     |
| S07 | Bad news for an executive sponsor          | 2.86     | +1.27                     |
| S08 | Quarter-end workload spike                 | 2.90     | +1.54                     |
| S09 | Handover from a departing colleague        | 2.85     | +1.62                     |
| S10 | Onboarding stalled at a key account        | 3.23     | +0.80                     |
| S11 | Team disagreement on an account plan       | 2.89     | +1.08                     |
| S12 | Deadline vs quality trade-off              | 3.07     | +2.33                     |
| S13 | Pricing exception request                  | 4.00     | — (1 answer)              |
| S14 | Owning the post-incident review            | 2.97     | +0.92                     |

Synthetic descriptive chart; verify the underlying data before use.

## Difficulty and separation / 3

### Scenario metric matrix · 7.3

| Scenario | Mean rating<br>1-5 | Corrected r<br>-1 to +1 | Mean probes<br>per answer | Rated 4–5 share<br>0-1 fraction |
|----------|--------------------|-------------------------|---------------------------|---------------------------------|
| S01      | 2.95               | 0.66                    | 1.76                      | 0.39                            |
| S02      | 2.95               | 0.71                    | 1.18                      | 0.29                            |
| S03†     | 2.64               | 0.67                    | 1.73                      | 0.18                            |
| S04      | 2.92               | 0.61                    | 1.68                      | 0.38                            |
| S05†     | 2.75               | 0.69                    | 0.83                      | 0.33                            |
| S06      | 2.87               | 0.59                    | 1.51                      | 0.28                            |
| S07      | 2.86               | 0.54                    | 1.63                      | 0.26                            |
| S08      | 2.90               | 0.49                    | 1.77                      | 0.42                            |
| S09      | 2.85               | 0.62                    | 1.93                      | 0.38                            |
| S10†     | 3.23               | 0.46                    | 1.06                      | 0.41                            |
| S11      | 2.89               | 0.47                    | 1.73                      | 0.35                            |
| S12†     | 3.07               | 0.79                    | 2.03                      | 0.38                            |
| S13†     | 4.00               | —                       | 3.00                      | 1.00                            |
| S14      | 2.97               | 0.51                    | 1.69                      | 0.31                            |

Cells: generated answer mean (1-5); corrected item-versus-other-answers correlation  $r$  (-1 to +1, only where 10+ answers); mean follow-up probes per answer; share of answers rated 4–5 (0-1). † = fewer than 30 answers. Shade encoding was removed because these metrics use different units.

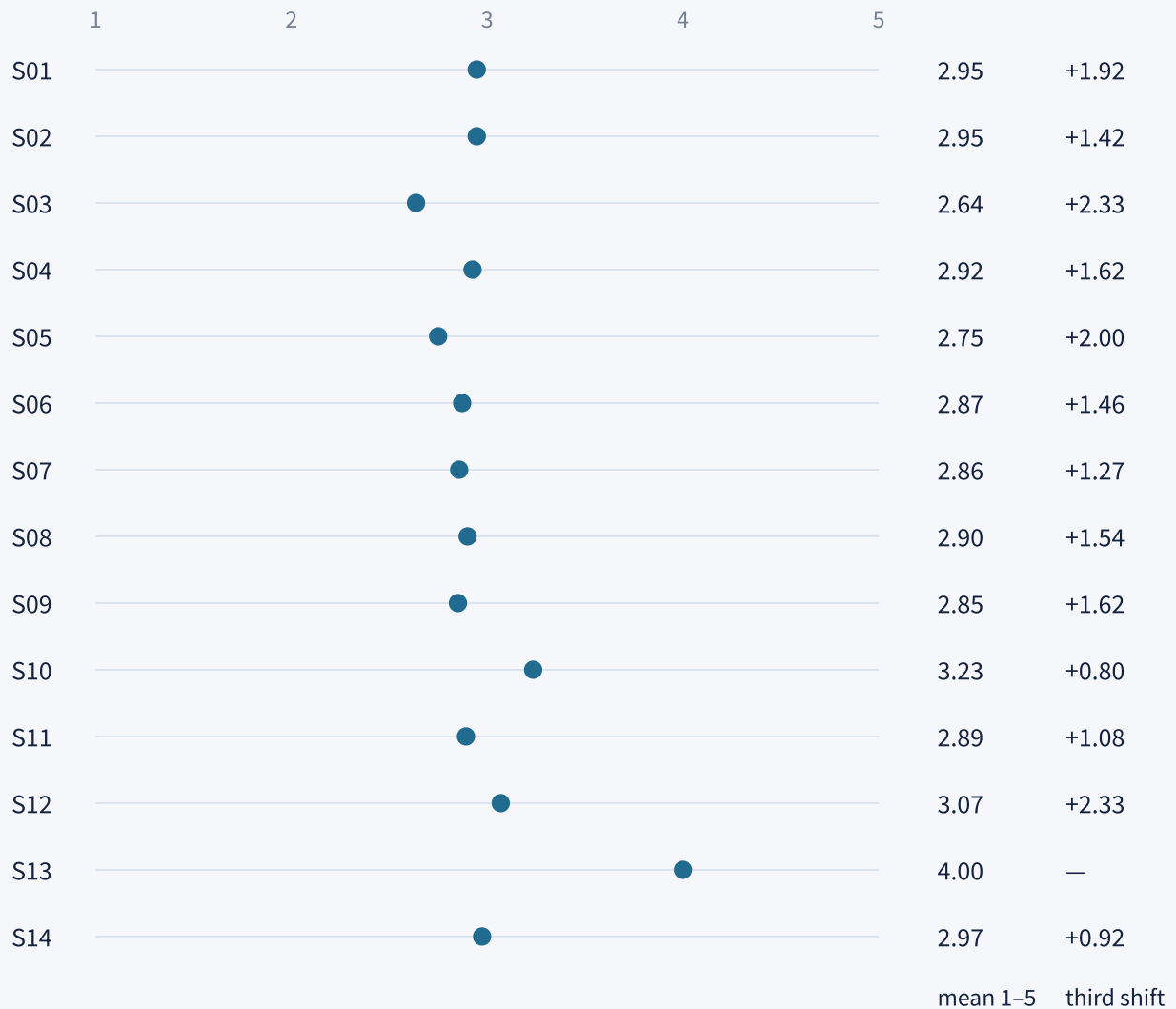
Source-derived scenario metric matrix; synthetic adaptive 40-person cohort. Descriptive, not a hiring rank or calibrated competency interval.

Reading guide: each column is one metric on its own scale. Read each column separately.  $r$  is computed against each person's other answered scenarios, so it is unstable where few people answered.

## Cohort comparison evidence / 6 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

### Fourteen scenario means and corrected third shifts



Rows remain in scenario order S01–S14, not ranked. All dots use the same color and show the mean answer rating of the people who answered. Third shift is the upper-third minus lower-third mean on this scenario, ordered by each person's other answered scenarios (n/3 each); S13 has one answer, so no shift.

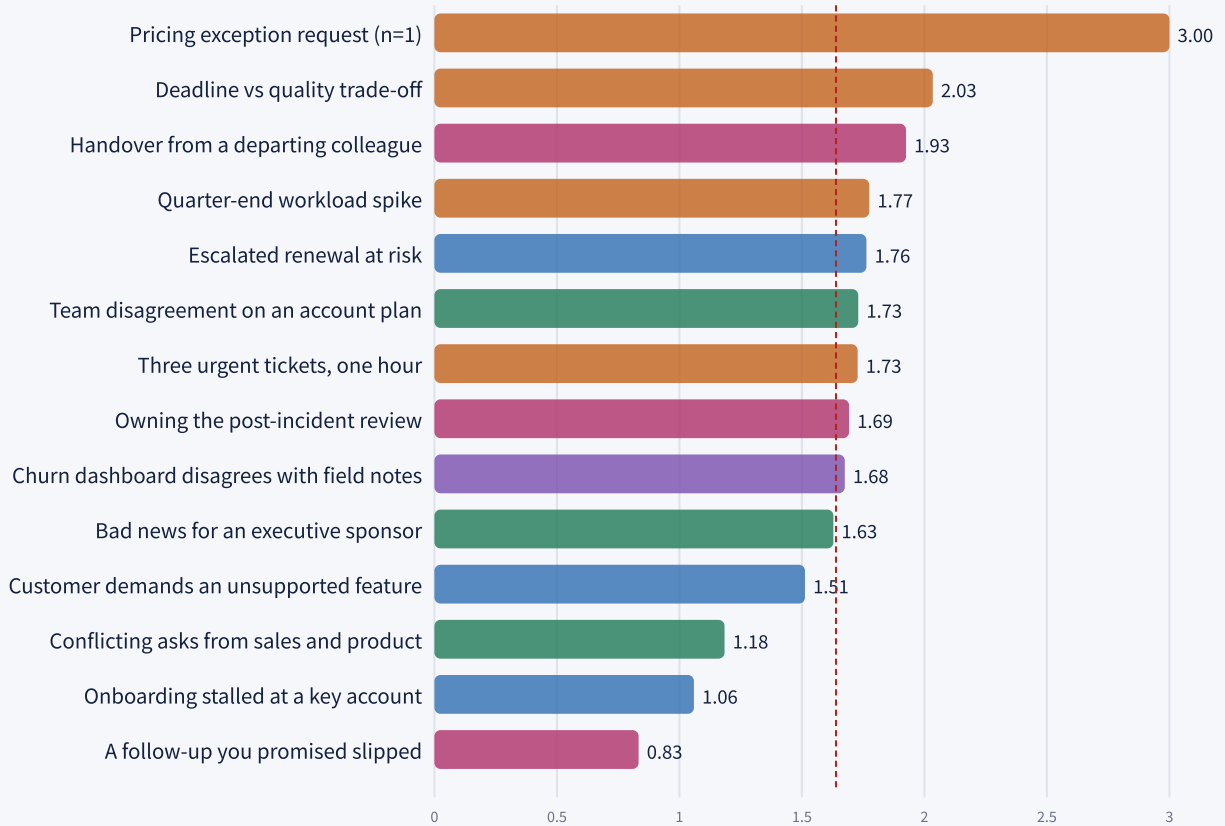
Thirds form on each person's other answered scenarios. This is item review, not candidate grading.

Source: scenario means, S01-S14, 416 answers from 40 people

PROBES DESCRIPTIVE

# Probes: 682 follow-up probes across 40 completers, 17.1 per person and 1.64 per answer.

## Probes per answer by scenario



Bar colour = competency. Dashed red line = cohort average (1.64 per answer). S13 has a single answer.

## Probes per answer by competency



Mean over all answers to that competency's scenarios. Descriptive only.

## Follow-up probes per answer at a glance



682 probes over 416 answers  
Mean 1.639 per answer across all positions.

How to read. Arc = 0 to 3 follow-up probes per answer; filled part = cohort mean. Describes follow-up volume, not answer quality.

## Probe analysis / 1

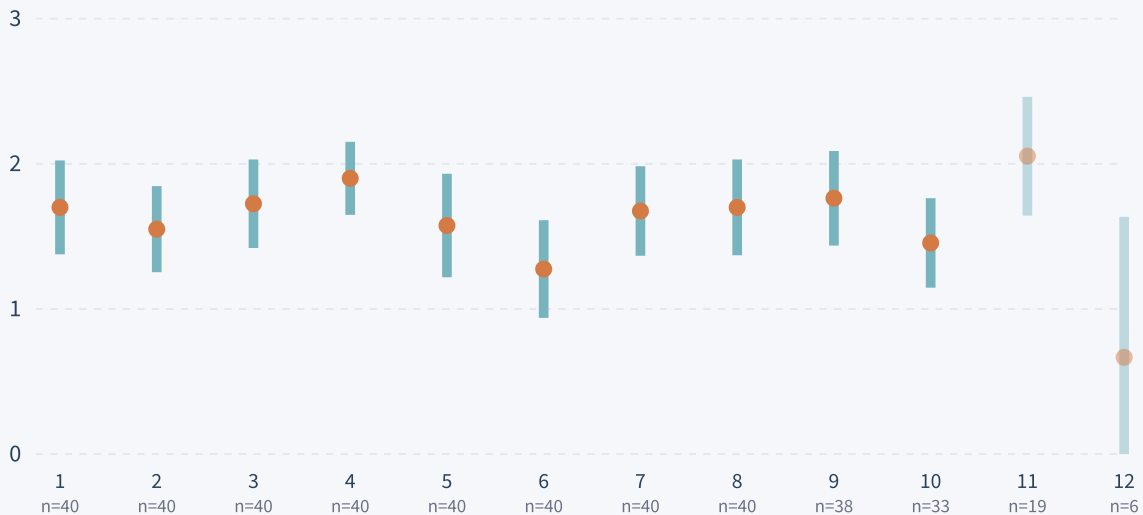
Pages 25–26 list mean follow-up probes for each scenario. This section examines probe patterns by answer position and competency. Positions 11–12 have fewer than 30 answers. The engine can ask 0–3 follow-up probes per answer.

**1.90** HIGHEST · POS. 4

**1.27** LOWEST · POS. 6

**1.64** AVERAGE · ANSWERS

### Probe position interval



Answer position (1–12); observed mean probes, descriptive normal 95% interval; faded = n under 30

Vertical: mean follow-up probes per answer (0–3 scale). Horizontal: position of the answer in the person's own adaptive sequence (1–12). Highest and lowest use positions with 30+ answers.

Synthetic adaptive source, n=40 completed people. Figure shows probe position interval on its labeled scale; descriptive, not a hiring decision.

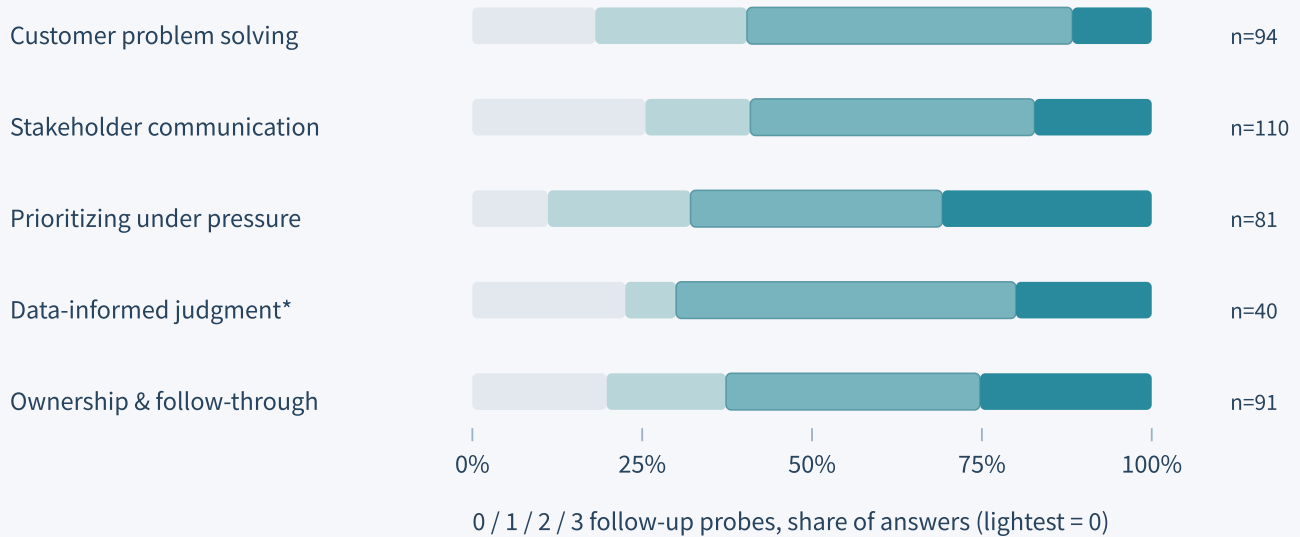
### Figures · mean follow-up probes by answer position

| Position    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   |
|-------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Answers     | 40   | 40   | 40   | 40   | 40   | 40   | 40   | 40   | 38   | 33   | 19   | 6    |
| Mean probes | 1.70 | 1.55 | 1.73 | 1.90 | 1.57 | 1.27 | 1.68 | 1.70 | 1.76 | 1.46 | 2.05 | 0.67 |

**How to read.** Mean = follow-up probes per answer (0–3 scale) at that position in each person's own adaptive sequence. Grey columns (positions 11–12, fewer than 30 answers) are reference only.

## Probe analysis / 2

### Competency probe bars



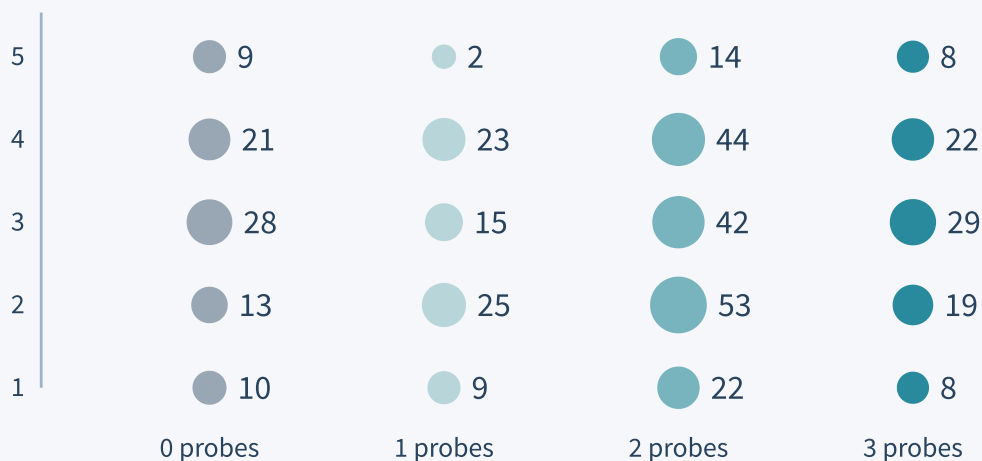
Follow-up count 0 probes 1 probe 2 probes 3 probes

Shares use each row's answer count, not 40 people. \* Data-informed judgment has one answer per person and is reference-only. Colors identify follow-up counts, not skill or evidence grades.

Synthetic adaptive source, n=40 completed people. Figure shows competency probe bars on its labeled scale; descriptive, not a hiring decision.

### Probe effectiveness bubble

Vertical: answer rating 1–5; circles count answers

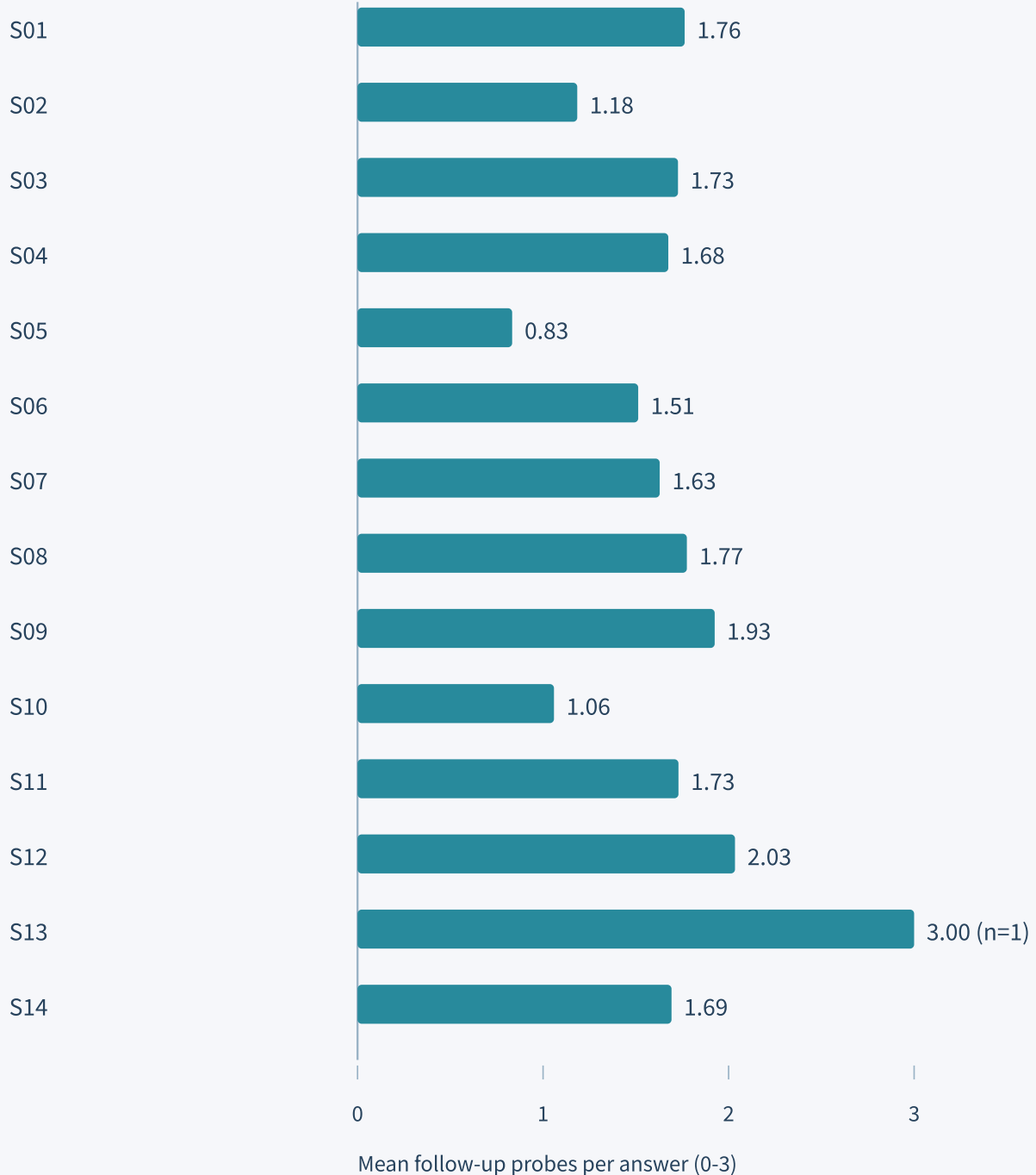


Printed counts sit beside circles. Total: 416 generated answers. Empty cells mean zero answers. Circle radius follows  $3 + \sqrt{\text{count}}$  SVG units; circle area is not proportional to count. Read the numerals for exact frequency.

Synthetic adaptive source, n=40 completed people. Figure shows probe effectiveness bubble on its labeled scale; descriptive, not a hiring decision.

## Probe analysis / 3

### Scenario probe bars

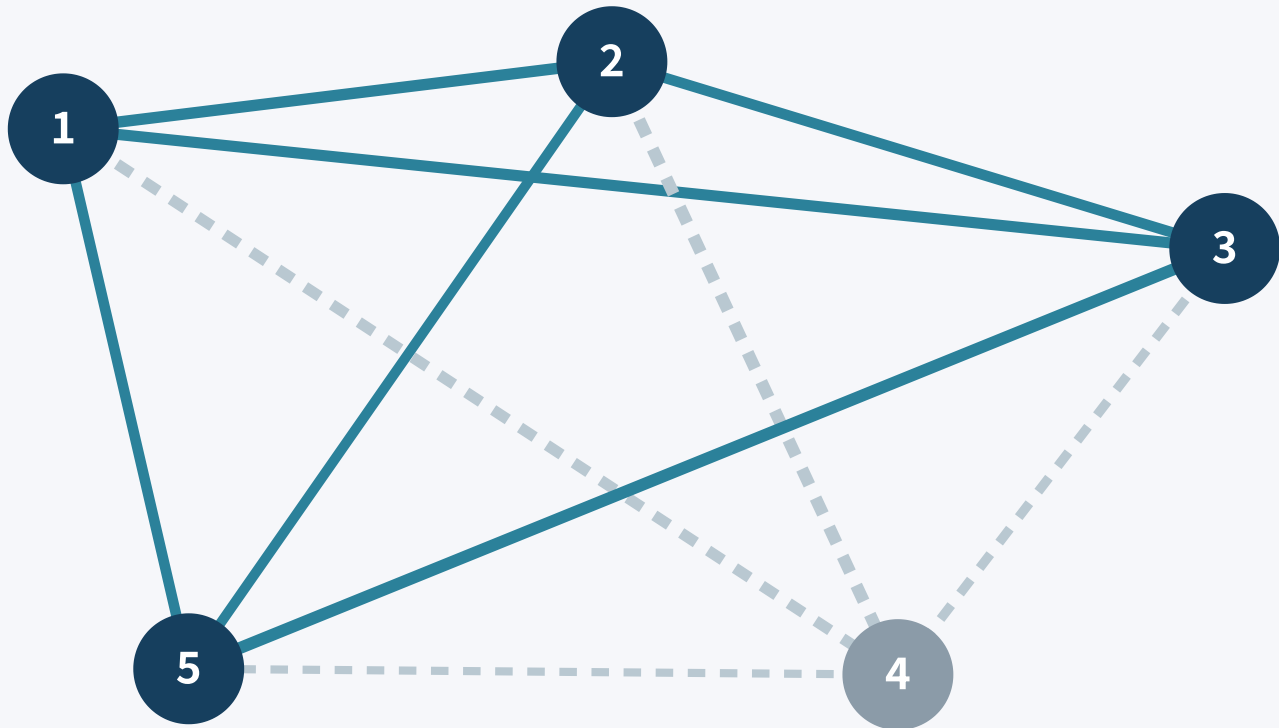


Answer counts differ by scenario (see pages 25–26); S13 has one answer. The same scenario IDs and titles appear in the scenario table and difficulty/separation section.

Synthetic descriptive chart; verify the underlying data before use.

## Probe analysis / 4

### Competency correlation network



Edge width: Pearson r

0.4

0.6

0.8

Width uses  $1 + 5 \times |r|$  in chart units, so it has a nonzero minimum. Gray dashed edges touch the reference-only Data-informed judgment node. Read the pair table for exact signed correlations; widths alone do not show sign.

Nodes 1–5 follow report competency order. Edge width = Pearson r; association only.

1 Customer problem solving

2 Stakeholder communication

3 Prioritizing under pressure

4 Data-informed judgment\*

5 Ownership & follow-through

### Pairwise Pearson r · synthetic n=40

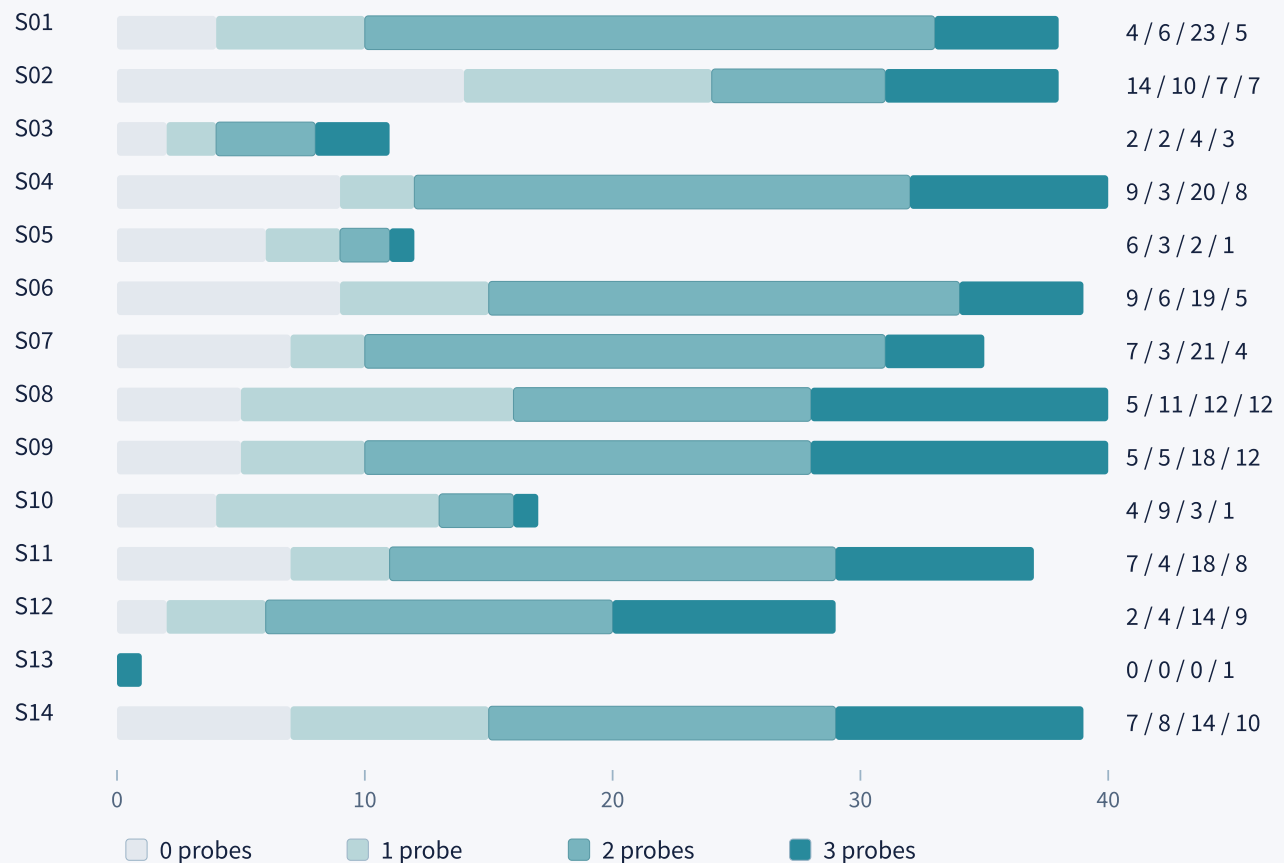
|       |      |       |      |       |      |       |      |       |      |
|-------|------|-------|------|-------|------|-------|------|-------|------|
| 1 / 2 | 0.58 | 1 / 3 | 0.56 | 1 / 4 | 0.56 | 1 / 5 | 0.53 | 2 / 3 | 0.53 |
| 2 / 4 | 0.65 | 2 / 5 | 0.49 | 3 / 4 | 0.45 | 3 / 5 | 0.63 | 4 / 5 | 0.36 |

Largest descriptive correlation in this adaptive sample: Stakeholder communication · Data-informed judgment (Pearson  $r=0.65$ ). Data-informed judgment comes mostly from a single generated answer per person (S04), so edges touching it are reference-only. Synthetic adaptive source, n=40 completed people. Descriptive association, not evidence for relative skill, causation, job performance or hiring.

## Cohort comparison evidence / 7 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

### Follow-up count in each of the fourteen scenarios



Horizontal: answer count 0–40, same scale for every scenario (scenarios were asked of different numbers of people, so bars have different lengths). Segments: 0 / 1 / 2 / 3 follow-up probes; printed quadruplets use that order. More follow-ups are not a higher skill grade.

Each stacked bar totals the answers given to that scenario. Counts are not question difficulty or score.

Source: follow-up probe counts, 416 answers from 40 people

Reading guide: each row is one scenario; its stacked bar splits the answers to that scenario by 0, 1, 2 or 3 follow-up probes.

Of 42 who started, 40 completed and 2 timed out. 37 finished as-is and 3 through QC selection.

Completion flow

Started · 42



Completed 40 · timed out 2



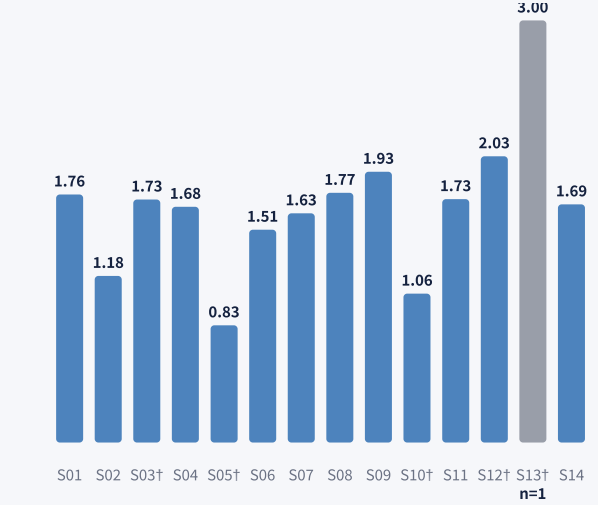
Each completer answered 8 to 12 of the 14 bank scenarios (adaptive).

Waterfall · 42 → 40



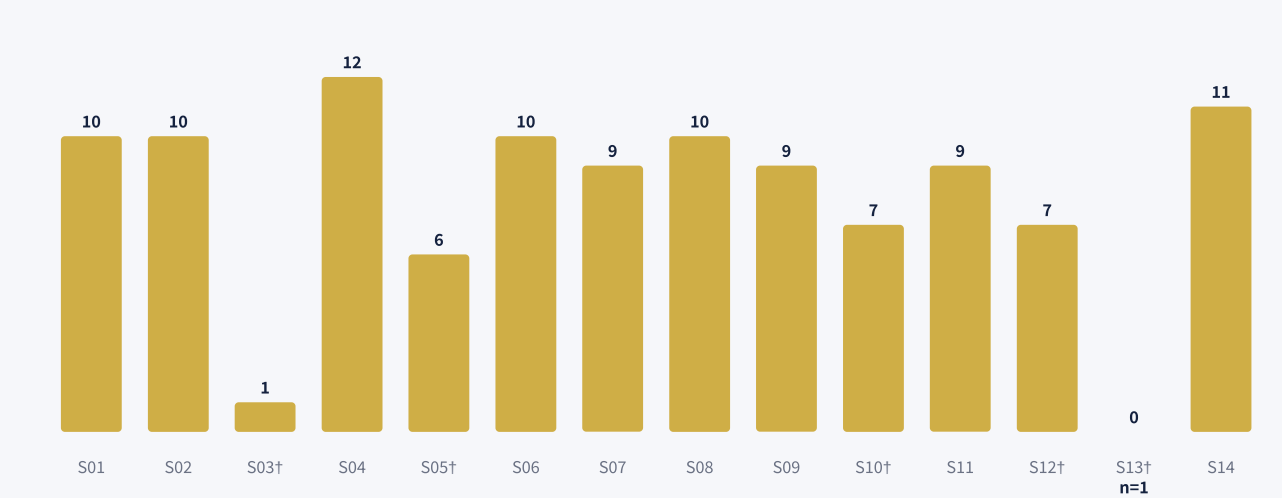
Timed-out sessions have no result.

Probes · mean per answer, by scenario



0–3 follow-up probes per answer · overall mean 1.64  
n per scenario 1–40; † = under 30 answers; S13 reference only

Anchor checks · by scenario



Count of answers per scenario that used an anchor check (n per scenario varies)  
n per scenario 1–40; † = under 30 answers; S13 reference only

# Evidence flow / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## Session completion and protocol totals

|                         |      |
|-------------------------|------|
| Started                 | 42   |
| Completed               | 40   |
| Timed out               | 2    |
| Completed as-is         | 37   |
| Completed · QC-selected | 3    |
| Total probes            | 682  |
| Median theta SE         | 0.29 |

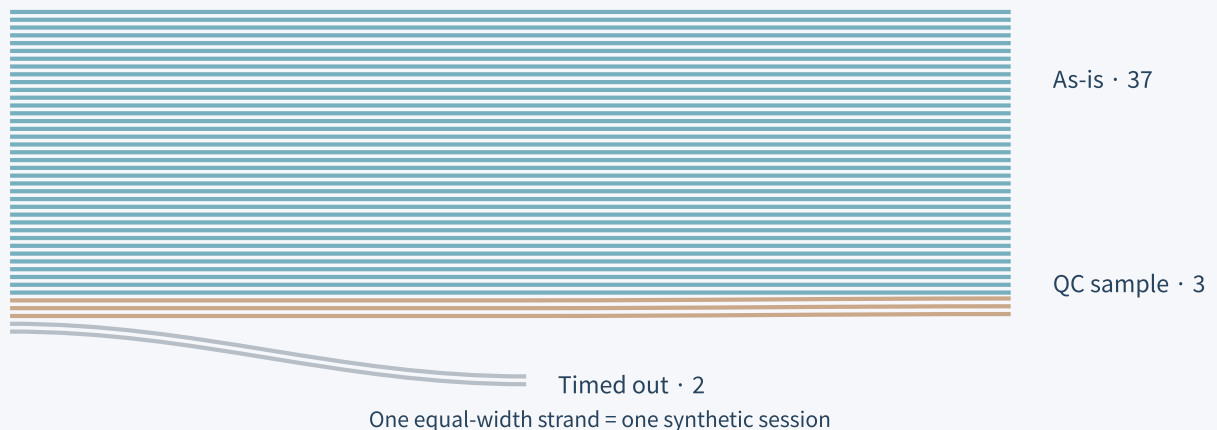
Median theta SE: median of the final recorded SE across 40 synthetic completers (0.2925, shown as 0.29). 36 stopped with the SE target met; 4 stopped at the 12-scenario maximum. These values come from the declared local illustrative posterior, not a calibrated Titan model or evidence grade. Total probes counts the 416 completed-person answers; timed-out answers are outside this total.

## Evidence flow / 2

### 42 linked sessions · completion and declared review

Started · 42

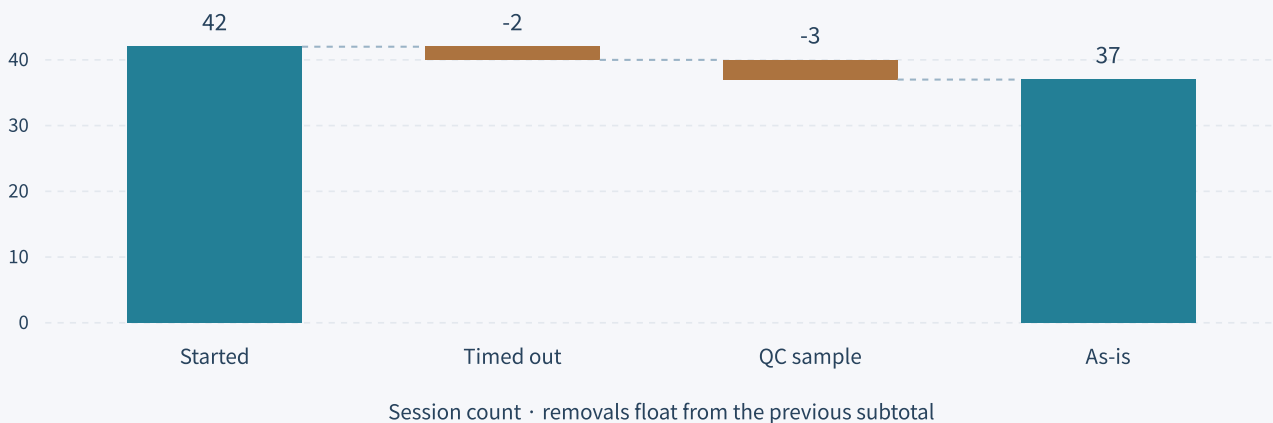
Completed · 40



Every equal-width strand represents one session. QC sample is a declared review selection, not an evidence grade. No target-count or early-stop categories are inferred.

Synthetic session log: 42 started = 40 completed + 2 timed out; 40 completed = 37 as-is + 3 declared QC review. The chart displays statuses across starters and completers, not a 40-person-only denominator or inferred grade.

### Session disposition waterfall · 42 → 37



42 started – 2 timed out = 40 completed; 40 completed – 3 declared QC sample = 37 completed as-is. QC-selected completers remain in descriptive statistics.

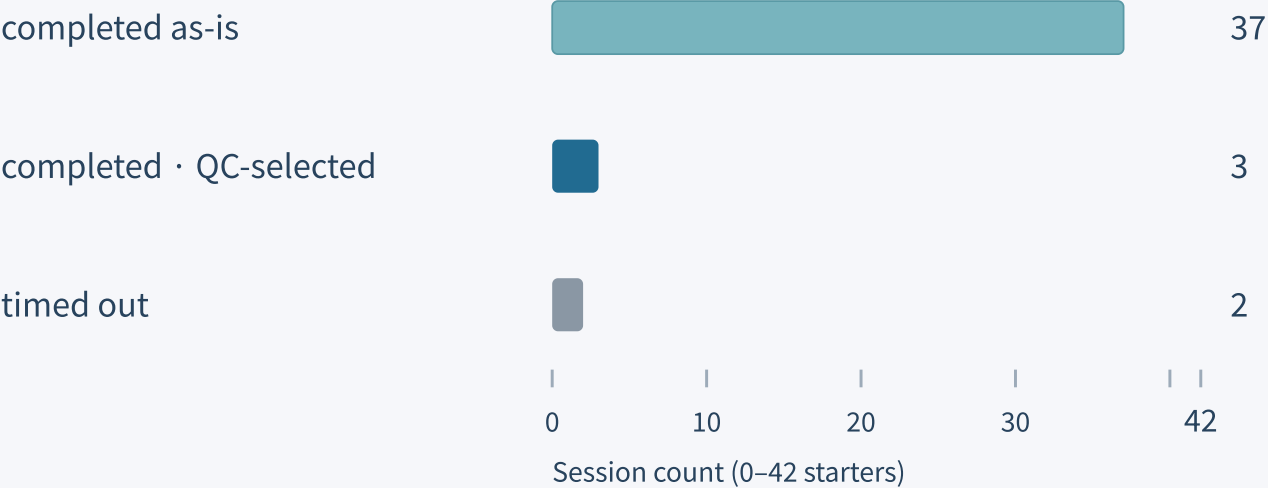
Synthetic session log: 42 started = 40 completed + 2 timed out; 40 completed = 37 as-is + 3 declared QC review. The chart displays statuses across starters and completers, not a 40-person-only denominator or inferred grade.

Reading guide: the first chart follows 42 started sessions, 40 completed and 2 timed out; the second shows 37 completed as-is and 3 declared QC review.

# Evidence flow / 3

Competency follow-up distribution: p. 35. This page separates started, completed and declared review status.

## Session completion and review status · 9.4



Status and declared QC sample, not inferred evidence grades.

Synthetic session log: 42 started = 40 completed + 2 timed out; 40 completed = 37 as-is + 3 declared QC review. The chart displays statuses across starters and completers, not a 40-person-only denominator or inferred grade.

## Review scope and denominators

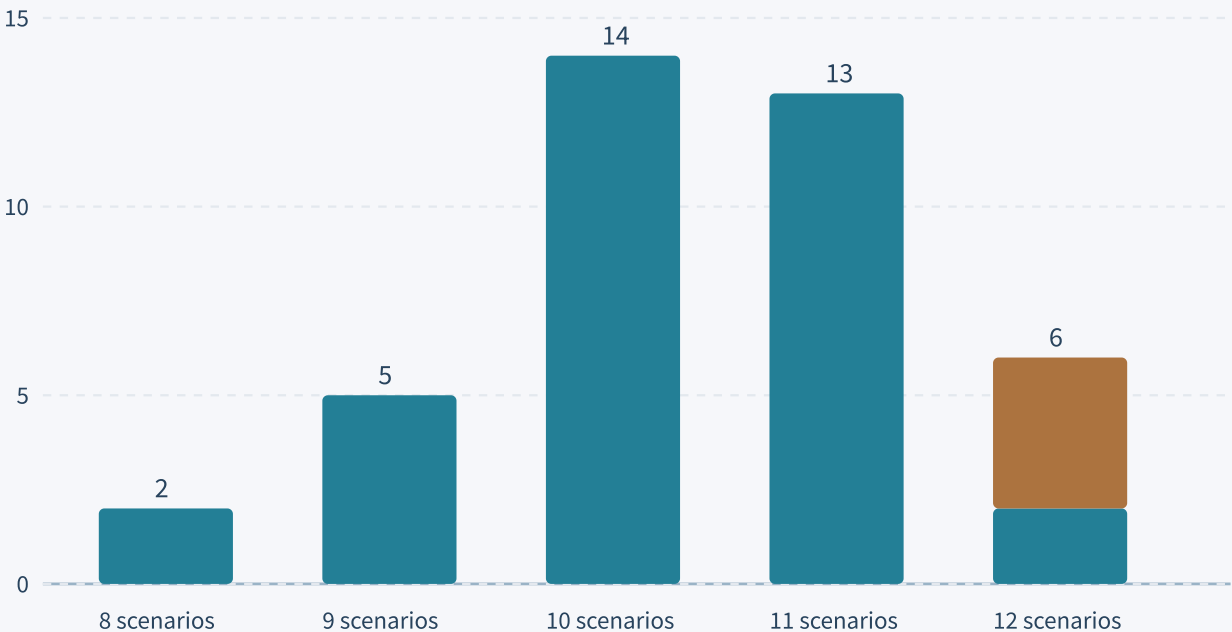
|                    |    |                                                    |
|--------------------|----|----------------------------------------------------|
| Started sessions   | 42 | Includes completed and timed-out sessions          |
| Completed sessions | 40 | Scope for individual and cohort descriptive values |
| Declared QC review | 3  | Included among completers, not additional people   |
| Timed-out sessions | 2  | Outside the 40 completed-person statistics         |

These are synthetic session-log states, not evidence grades or real applicant status.

Reading guide: the table separates started (42), completed (40), declared QC review (3) and timed-out (2) sessions; QC completers remain among the 40.

# Evidence flow / 4

Scenarios answered × stop reason · 40 completed



Vertical: number of completed people. Teal = stopped with the SE target met; brown = stopped at the 12-scenario maximum. Stacked bars start at zero; denominator = 40 completers. The 2 timed-out sessions have no recorded length and are not shown.

40 completed people: 36 stopped with the SE target met, 4 reached the 12-scenario maximum. Colors describe stop reasons, not evidence grades.

## Scenarios answered and stop reason

| Scenarios answered | SE target met | 12-scenario maximum | People | Share |
|--------------------|---------------|---------------------|--------|-------|
| 8                  | 2             | 0                   | 2      | 5%    |
| 9                  | 5             | 0                   | 5      | 12%   |
| 10                 | 14            | 0                   | 14     | 35%   |
| 11                 | 13            | 0                   | 13     | 32%   |
| 12                 | 2             | 4                   | 6      | 15%   |

These are engine stop reasons (the engine SE target and 12-scenario cap), not evidence grades. They describe why the adaptive sequence ended, not how well a person did.

Reading guide: each bar is one number of scenarios answered (8 to 12), stacked by stop reason; the table repeats the same counts.

DESCRIPTIVE

one completely

uneven (large spread) →

No composite · ordered by spread, not a rank.

# Answer consistency / 1

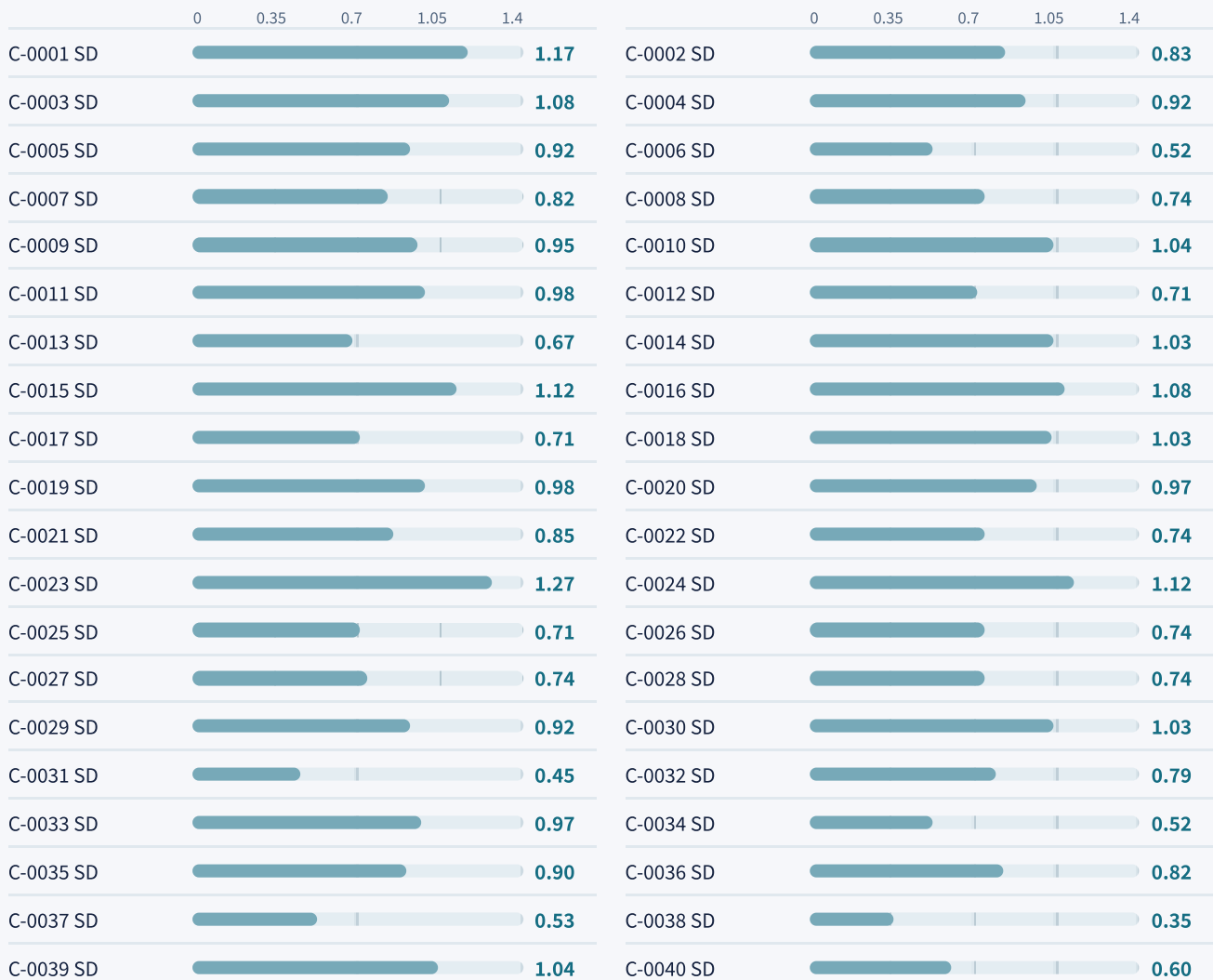
Synthetic template responses; not real candidate data or calibrated Titan AI™ output.

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## Answer-rating SD by person

Sample SD across each person's 8–12 scenario effectiveness ratings (1–5). It describes variation, not uncertainty or job fit.

Bar width: sample SD from 0 to 1.4 rating points, identical in both columns. The 1.00 split below is an illustrative descriptive cut, not a validated threshold or an automatic review trigger.



## Descriptive readout · synthetic adaptive 40-person sample

Median person-level sample SD of 8–12 answer ratings: 0.88 on the 1–5 input scale.

Under 1.00 SD: 29/40; at or above: 11/40. This is a descriptive split, not a reliability grade.

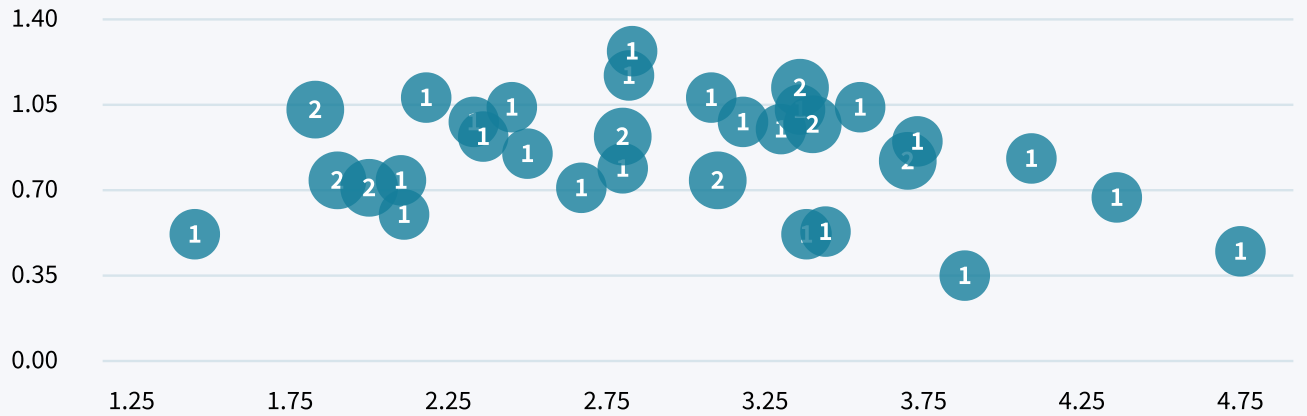
Median of each person's mean within-scenario trait ranges: 1.00 rating points.

Neither answer variation nor trait range predicts job fit or supports an automatic review rule.

## Answer consistency / 2

Same mean, different spread: C-0004 and C-0032 both average 2.80/5; their sample SDs are 0.92 and 0.79 (over 10 and 10 answers). That is variation across answers, not evidence that either person is more reliable.

### Mean rating × answer-rating SD · 40 people, 32 distinct points



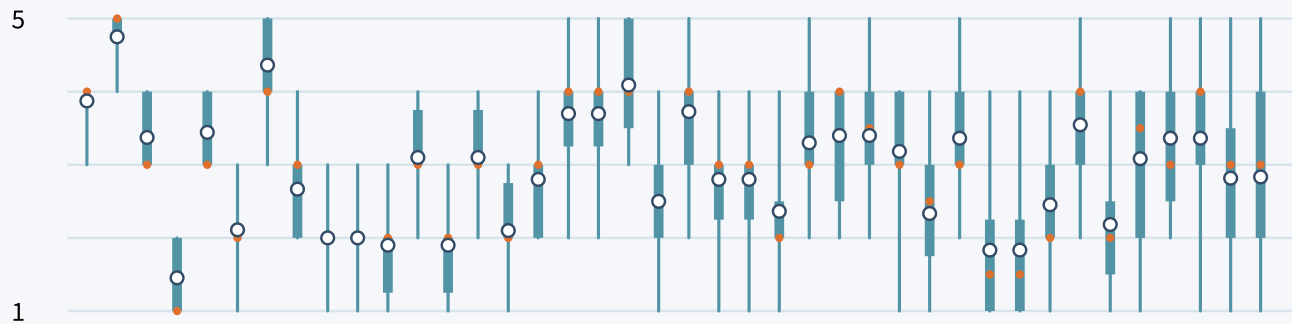
Sample SD across 8–12 answers (1–5)

Mean answer rating (1–5)

Each circle is one distinct mean/SD coordinate; its numeral is the number of people sharing it. Coincident points are counted, not silently hidden. Synthetic descriptive chart, not a hiring decision.

### Per-person answer-rating range, middle half, mean and median

40 completed people, ordered by answer-rating SD



Vertical: generated answer rating 1–5. Horizontal: columns 1–40 by ascending sample SD then source ID, not quality ranks. Thin teal min–max; thick teal p25–p75; hollow navy mean; orange median.

|            |            |            |            |            |            |            |            |
|------------|------------|------------|------------|------------|------------|------------|------------|
| 1: C-0038  | 2: C-0031  | 3: C-0034  | 4: C-0006  | 5: C-0037  | 6: C-0040  | 7: C-0013  | 8: C-0025  |
| 9: C-0012  | 10: C-0017 | 11: C-0008 | 12: C-0022 | 13: C-0026 | 14: C-0027 | 15: C-0028 | 16: C-0032 |
| 17: C-0007 | 18: C-0036 | 19: C-0002 | 20: C-0021 | 21: C-0035 | 22: C-0004 | 23: C-0005 | 24: C-0029 |
| 25: C-0009 | 26: C-0020 | 27: C-0033 | 28: C-0011 | 29: C-0019 | 30: C-0018 | 31: C-0014 | 32: C-0030 |
| 33: C-0010 | 34: C-0039 | 35: C-0016 | 36: C-0003 | 37: C-0024 | 38: C-0015 | 39: C-0001 | 40: C-0023 |

Lower answer-rating SD

Higher answer-rating SD

One column per person, in ascending sample SD then source ID (column numbers are not quality ranks). Thin teal line: min–max of that person's 8–12 generated answer ratings. Thick teal bar: p25–p75 with linear interpolation. Hollow navy circle: mean; small orange dot: median. All use rating 1–5, not competency intervals, evidence grades or workplace reliability.

# Cohort comparison evidence / 8 of 9

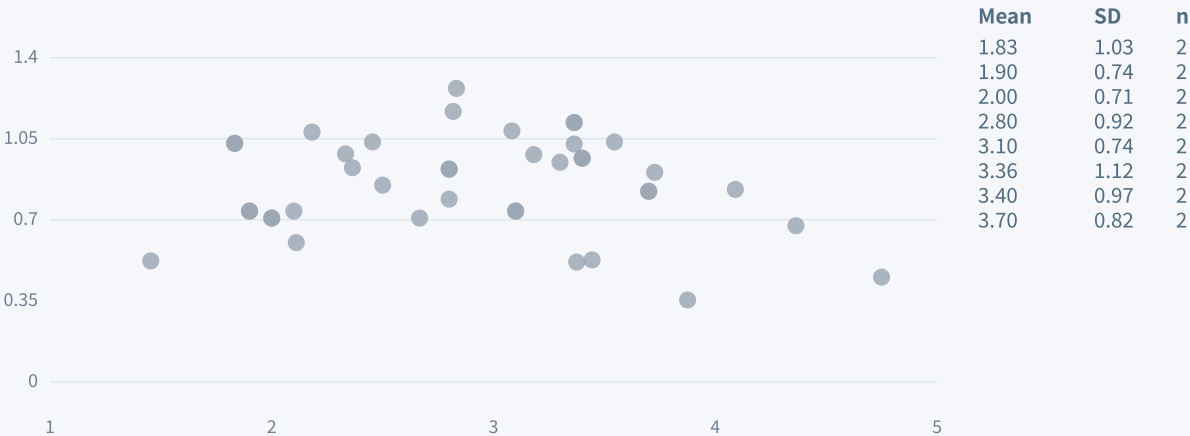
Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

40 PEOPLE IN LEDGER

32 DISTINCT MEAN-SD POSITIONS

1.27 HIGHEST SD · N=1

## Matched person mean versus response variation



Each point is one completed person from the same 8–12 scenario answers, not separately reconstructed marginals.

Source: generated scores and spread, 40 people

Gray marks are source people, not group classes. The right ledger lists the 8 positions shared by 2 people; the other 24 positions are single people, totaling 40. Same-position points overlap. No jitter or invented coordinates. SD describes 8–12-answer variation, not measurement uncertainty or reliability.

## Seven-point interpretation scorecard

Completed people

40

Comparable lenses

4 / 5

Data-informed observation per person

1

Observed scenarios per completer

8–12

Issued evidence grades

Not assigned

External percentile / hiring verdict

Unavailable

Relative source

Synthetic adaptive 40-person set

Descriptive sample only. None of these fields is a hiring decision or live product output.

ALL CANDIDATES

DESCRIPTIVE

SMALL SAMPLE

## 40 completers across five competencies. Order and labels carry no ranking.

Parallel coordinates · all 40 completers



Score grid · all 40 completers

|        | Cust. | Stake. | Prior. | Data | Own. |        | Cust. | Stake. | Prior. | Data | Own. |
|--------|-------|--------|--------|------|------|--------|-------|--------|--------|------|------|
| C-0001 | 66.7  | 83.3   | 43.8   | 50   | 75   | C-0021 | 75    | 75     | 68.8   | 75   | 75   |
| C-0002 | 50    | 50     | 50     | 50   | 50   | C-0022 | 41.7  | 50     | 50     | 50   | 50   |
| C-0003 | 58.3  | 50     | 50     | 50   | 58.3 | C-0023 | 50    | 50     | 50     | 50   | 50   |
| C-0004 | 58.3  | 58.3   | 50     | 50   | 66.7 | C-0024 | 58.3  | 50     | 50     | 50   | 66.7 |
| C-0005 | 58.3  | 50     | 68.8   | 75   | 75   | C-0025 | 58.3  | 50     | 50     | 75   | 75   |
| C-0006 | 75    | 75     | 75     | 75   | 75   | C-0026 | 66.7  | 66.7   | 75     | 75   | 75   |
| C-0007 | 75    | 58.3   | 75     | 75   | 75   | C-0027 | 75    | 75     | 62.5   | 75   | 75   |
| C-0008 | 50    | 41.7   | 50     | 50   | 50   | C-0028 | 75    | 66.7   | 75     | 75   | 75   |
| C-0009 | 50    | 50     | 50     | 50   | 58.3 | C-0029 | 50    | 41.7   | 50     | 50   | 58.3 |
| C-0010 | 50    | 50     | 50     | 50   | 58.3 | C-0030 | 50    | 50     | 50     | 50   | 58.3 |
| C-0011 | 58.3  | 50     | 50     | 75   | 66.7 | C-0031 | 58.3  | 50     | 50     | 50   | 66.7 |
| C-0012 | 66.7  | 50     | 68.8   | 75   | 75   | C-0032 | 50    | 50     | 50     | 75   | 75   |
| C-0013 | 75    | 66.7   | 75     | 75   | 75   | C-0033 | 58.3  | 50     | 68.8   | 75   | 66.7 |
| C-0014 | 75    | 66.7   | 75     | 75   | 75   | C-0034 | 58.3  | 75     | 75     | 75   | 75   |
| C-0015 | 50    | 50     | 50     | 50   | 50   | C-0035 | 75    | 75     | 75     | 75   | 75   |
| C-0016 | 50    | 50     | 56.2   | 50   | 50   | C-0036 | 50    | 50     | 50     | 50   | 50   |
| C-0017 | 58.3  | 58.3   | 50     | 50   | 58.3 | C-0037 | 50    | 50     | 50     | 50   | 50   |
| C-0018 | 50    | 50     | 50     | 75   | 75   | C-0038 | 58.3  | 58.3   | 56.2   | 75   | 66.7 |
| C-0019 | 58.3  | 50     | 75     | 75   | 66.7 | C-0039 | 58.3  | 58.3   | 50     | 75   | 75   |
| C-0020 | 75    | 75     | 62.5   | 75   | 75   | C-0040 | 58.3  | 50     | 68.8   | 75   | 58.3 |

Rows follow candidate ID, not score. Colour depth = score (0–100).

## Candidate matrix / 1 of 2

### Forty-person matrix · rows 1–20

|        | Customer | Stakeholder | Priority | Data* | Ownership |
|--------|----------|-------------|----------|-------|-----------|
| C-0001 | 58.3     | 43.4        | 63.3     | 40.0  | 25.2      |
| C-0002 | 67.3     | 56.2        | 84.4     | 64.2  | 76.1      |
| C-0003 | 37.2     | 57.2        | 61.9     | 64.2  | 45.6      |
| C-0004 | 58.3     | 56.2        | 25.9     | 40.0  | 50.6      |
| C-0005 | 57.6     | 48.5        | 34.4     | 51.7  | 43.3      |
| C-0006 | 25.3     | 15.9        | 16.3     | 40.0  | 25.8      |
| C-0007 | 58.3     | 56.7        | 74.0     | 51.7  | 67.3      |
| C-0008 | 32.9     | 41.0        | 20.4     | 51.7  | 25.2      |
| C-0009 | 50.0     | 52.1        | 64.8     | 64.2  | 48.5      |
| C-0010 | 32.6     | 36.1        | 25.9     | 64.2  | 59.5      |
| C-0011 | 67.3     | 45.4        | 58.5     | 64.2  | 44.6      |
| C-0012 | 24.6     | 36.1        | 50.2     | 28.0  | 33.7      |
| C-0013 | 75.4     | 84.6        | 65.9     | 76.8  | 67.6      |
| C-0014 | 45.1     | 29.6        | 17.3     | 28.0  | 26.5      |
| C-0015 | 34.5     | 62.5        | 65.5     | 40.0  | 65.9      |
| C-0016 | 35.8     | 23.5        | 43.5     | 28.0  | 52.7      |
| C-0017 | 32.9     | 32.2        | 25.9     | 51.7  | 35.4      |
| C-0018 | 49.8     | 55.8        | 64.7     | 76.8  | 41.7      |
| C-0019 | 31.6     | 58.4        | 38.8     | 51.7  | 26.1      |
| C-0020 | 67.3     | 57.2        | 49.3     | 40.0  | 67.3      |

Source ID order · no Ward clustering · data judgment\* is one observation/person

0–<45   45–<60   60–<75   75–100   Data\*: reference only

Shade bands are display bins on descriptive 0–100 scores, not calibrated performance levels. Gray dashed cells (all of Data\*, and any cell with only 1 observation) are reference-only and neutral. Printed values are retained for every cell.

Five descriptive 0–100 summaries, 20 source-ID rows on this sheet. Two sheets contain all 40 people and 200 cells. Assigned groups are not discovered clusters; reference-only cells (1 observation) are neutral.

## Candidate matrix / 2 of 2

### Forty-person matrix · rows 21–40

|        | Customer | Stakeholder | Priority | Data* | Ownership |
|--------|----------|-------------|----------|-------|-----------|
| C-0021 | 41.2     | 24.4        | 51.2     | 51.7  | 50.6      |
| C-0022 | 50.4     | 52.1        | 42.6     | 51.7  | 67.3      |
| C-0023 | 56.5     | 67.2        | 25.9     | 64.2  | 26.0      |
| C-0024 | 44.1     | 55.8        | 56.1     | 51.7  | 76.1      |
| C-0025 | 50.0     | 42.2        | 51.2     | 40.0  | 43.3      |
| C-0026 | 13.6     | 32.2        | 43.5     | 28.0  | 36.8      |
| C-0027 | 50.0     | 52.1        | 42.7     | 64.2  | 58.0      |
| C-0028 | 32.9     | 50.2        | 34.0     | 40.0  | 25.5      |
| C-0029 | 44.1     | 23.5        | 43.8     | 40.0  | 48.7      |
| C-0030 | 46.7     | 15.9        | 36.3     | 28.0  | 26.5      |
| C-0031 | 81.6     | 84.6        | 84.4     | 64.2  | 89.5      |
| C-0032 | 41.2     | 49.5        | 58.5     | 40.0  | 43.3      |
| C-0033 | 58.3     | 70.0        | 49.3     | 64.2  | 43.3      |
| C-0034 | 58.8     | 50.4        | 61.7     | 51.7  | 58.0      |
| C-0035 | 82.6     | 52.1        | 58.5     | 64.2  | 58.0      |
| C-0036 | 58.8     | 63.6        | 65.5     | 76.8  | 48.7      |
| C-0037 | 58.3     | 50.1        | 58.5     | 51.7  | 67.3      |
| C-0038 | 67.3     | 60.0        | 61.7     | 64.2  | 67.3      |
| C-0039 | 75.6     | 44.4        | 73.4     | 64.2  | 52.8      |
| C-0040 | 32.9     | 32.2        | 35.3     | 28.0  | 50.6      |

Source ID order · no Ward clustering · data judgment\* is one observation/person

0–<45
  45–<60
  60–<75
  75–100
  Data\*: reference only

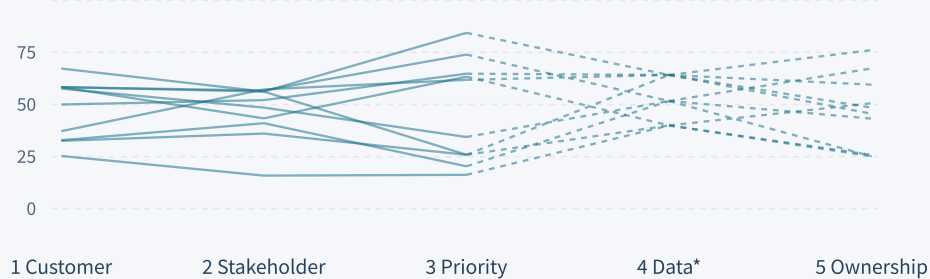
Shade bands are display bins on descriptive 0–100 scores, not calibrated performance levels. Gray dashed cells (all of Data\*, and any cell with only 1 observation) are reference-only and neutral. Printed values are retained for every cell.

Five descriptive 0–100 summaries, 20 source-ID rows on this sheet. Two sheets contain all 40 people and 200 cells. Assigned groups are not discovered clusters; reference-only cells (1 observation) are neutral.

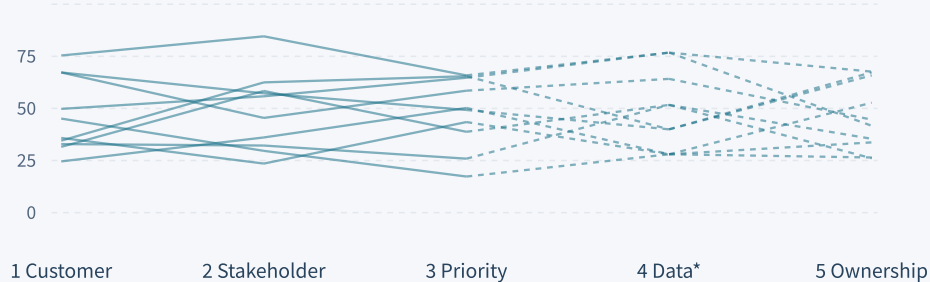
# Individual score paths

## 40 person-level paths · score shapes, not clusters

C-0001–C-0010



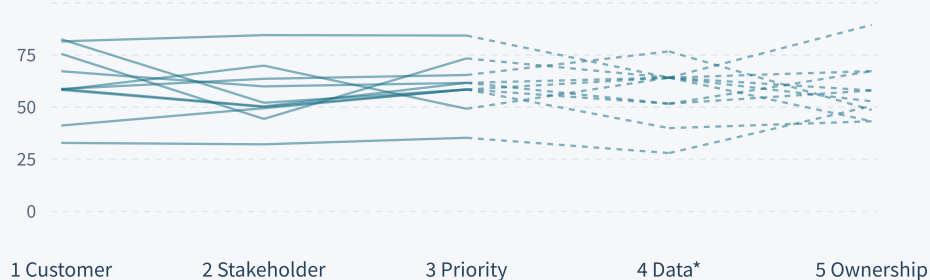
C-0011–C-0020



C-0021–C-0030



C-0031–C-0040



Vertical score scale: the full descriptive 0–100 range, equal in all four panels. Lines are not zero-based bars. Dashed segments touch Data\*, the one-item reference axis. A connected path is a visual guide, not a calibrated competency profile.

Vertical axis 0–100. Axes 1–5: customer / stakeholder / priority / data\* / ownership. Each thin path is one completed respondent on five descriptive 0–100 summaries, in source-ID batches to reduce overlap. Assigned groups are not clustering results. Data-informed judgment\* is reference-only; dashed segments touch it.

## Forty linked descriptive profiles

All five score columns are descriptive 0–100 generated-answer summaries, not validated ranking. Gray cells are reference-only (Data-informed judgment, and any cell with 1 observation). Each row is linked to the synthetic cohort candidate ID; the individual report's C-2291 is a separate run and not one of the 40.

| ID     | Customer | Stakeholder | Priority | Data ref* | Ownership |
|--------|----------|-------------|----------|-----------|-----------|
| C-0001 | 58.3     | 43.4        | 63.3     | 40.0      | 25.2      |
| C-0002 | 67.3     | 56.2        | 84.4     | 64.2      | 76.1      |
| C-0003 | 37.2     | 57.2        | 61.9     | 64.2      | 45.6      |
| C-0004 | 58.3     | 56.2        | 25.9     | 40.0      | 50.6      |
| C-0005 | 57.6     | 48.5        | 34.4     | 51.7      | 43.3      |
| C-0006 | 25.3     | 15.9        | 16.3     | 40.0      | 25.8      |
| C-0007 | 58.3     | 56.7        | 74.0     | 51.7      | 67.3      |
| C-0008 | 32.9     | 41.0        | 20.4     | 51.7      | 25.2      |
| C-0009 | 50.0     | 52.1        | 64.8     | 64.2      | 48.5      |
| C-0010 | 32.6     | 36.1        | 25.9     | 64.2      | 59.5      |
| C-0011 | 67.3     | 45.4        | 58.5     | 64.2      | 44.6      |
| C-0012 | 24.6     | 36.1        | 50.2     | 28.0      | 33.7      |
| C-0013 | 75.4     | 84.6        | 65.9     | 76.8      | 67.6      |
| C-0014 | 45.1     | 29.6        | 17.3     | 28.0      | 26.5      |
| C-0015 | 34.5     | 62.5        | 65.5     | 40.0      | 65.9      |
| C-0016 | 35.8     | 23.5        | 43.5     | 28.0      | 52.7      |
| C-0017 | 32.9     | 32.2        | 25.9     | 51.7      | 35.4      |
| C-0018 | 49.8     | 55.8        | 64.7     | 76.8      | 41.7      |
| C-0019 | 31.6     | 58.4        | 38.8     | 51.7      | 26.1      |
| C-0020 | 67.3     | 57.2        | 49.3     | 40.0      | 67.3      |

Columns are generated-answer scores on the same 0–100 descriptive scale: Customer problem solving, Stakeholder communication, Prioritizing under pressure, Data-informed judgment\*, Ownership & follow-through. Gray cells are reference-only (1 observation), retained for traceability only, not a comparison or ranking.

## Linked profiles / 21–40

| ID     | Customer | Stakeholder | Priority | Data ref* | Ownership |
|--------|----------|-------------|----------|-----------|-----------|
| C-0021 | 41.2     | 24.4        | 51.2     | 51.7      | 50.6      |
| C-0022 | 50.4     | 52.1        | 42.6     | 51.7      | 67.3      |
| C-0023 | 56.5     | 67.2        | 25.9     | 64.2      | 26.0      |
| C-0024 | 44.1     | 55.8        | 56.1     | 51.7      | 76.1      |
| C-0025 | 50.0     | 42.2        | 51.2     | 40.0      | 43.3      |
| C-0026 | 13.6     | 32.2        | 43.5     | 28.0      | 36.8      |
| C-0027 | 50.0     | 52.1        | 42.7     | 64.2      | 58.0      |
| C-0028 | 32.9     | 50.2        | 34.0     | 40.0      | 25.5      |
| C-0029 | 44.1     | 23.5        | 43.8     | 40.0      | 48.7      |
| C-0030 | 46.7     | 15.9        | 36.3     | 28.0      | 26.5      |
| C-0031 | 81.6     | 84.6        | 84.4     | 64.2      | 89.5      |
| C-0032 | 41.2     | 49.5        | 58.5     | 40.0      | 43.3      |
| C-0033 | 58.3     | 70.0        | 49.3     | 64.2      | 43.3      |
| C-0034 | 58.8     | 50.4        | 61.7     | 51.7      | 58.0      |
| C-0035 | 82.6     | 52.1        | 58.5     | 64.2      | 58.0      |
| C-0036 | 58.8     | 63.6        | 65.5     | 76.8      | 48.7      |
| C-0037 | 58.3     | 50.1        | 58.5     | 51.7      | 67.3      |
| C-0038 | 67.3     | 60.0        | 61.7     | 64.2      | 67.3      |
| C-0039 | 75.6     | 44.4        | 73.4     | 64.2      | 52.8      |
| C-0040 | 32.9     | 32.2        | 35.3     | 28.0      | 50.6      |

Columns are generated-answer scores on the same 0–100 descriptive scale: Customer problem solving, Stakeholder communication, Prioritizing under pressure, Data-informed judgment\*, Ownership & follow-through. Gray cells are reference-only (1 observation), retained for traceability only, not a comparison or ranking.

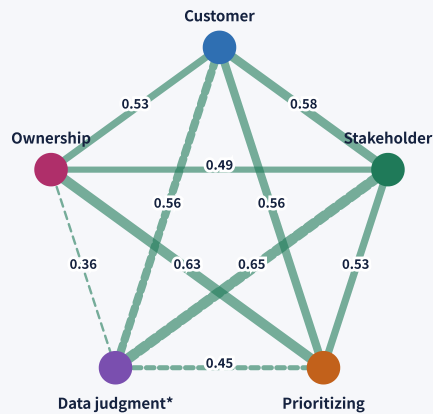
## HOW COMPETENCIES RELATE · NETWORK

ASSOCIATION · NOT CAUSAL

SMALL SAMPLE

Among the six pairs not involving Data-informed judgment, correlations range  $r$  0.49–0.63. With  $n=40$ , don't read much into the order.

## Competency network



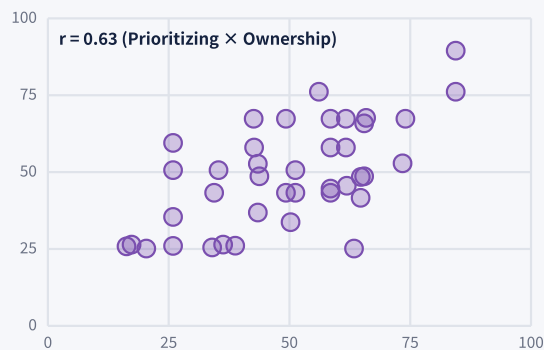
Line width = correlation  $r$  (40 completers). Dashed = reference only (pairs with Data-informed judgment). Association, not cause.

## Correlation matrix

|                | Cust. | Stake. | Prior. | Data* | Own.  |
|----------------|-------|--------|--------|-------|-------|
| Customer       | —     | 0.58   | 0.56   | 0.56* | 0.53  |
| Stakeholder    | 0.58  | —      | 0.53   | 0.65* | 0.49  |
| Prioritizing   | 0.56  | 0.53   | —      | 0.45* | 0.63  |
| Data judgment* | 0.56* | 0.65*  | 0.45*  | —     | 0.36* |
| Ownership      | 0.53  | 0.49   | 0.63   | 0.36* | —     |

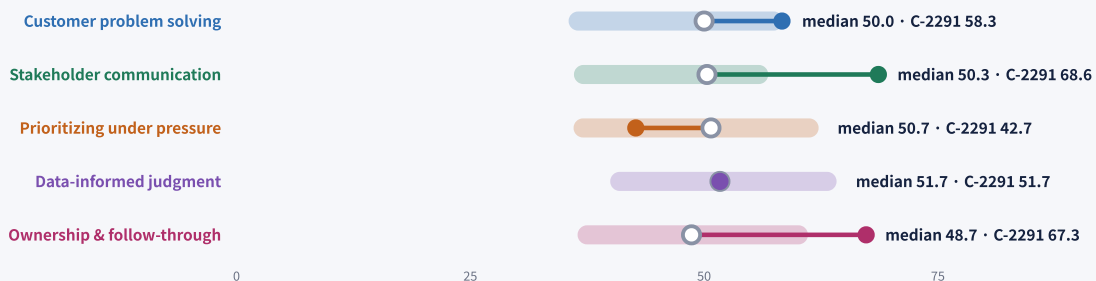
\* pair includes Data-informed judgment (one observation); reference only

## Strongest pair · scatter

All pairs ·  $r$ 

|                               |       |
|-------------------------------|-------|
| Stakeholder × Data judgment*  | 0.65* |
| Prioritizing × Ownership      | 0.63  |
| Customer × Stakeholder        | 0.58  |
| Customer × Data judgment*     | 0.56* |
| Customer × Prioritizing       | 0.56  |
| Stakeholder × Prioritizing    | 0.53  |
| Customer × Ownership          | 0.53  |
| Stakeholder × Ownership       | 0.49  |
| Prioritizing × Data judgment* | 0.45* |
| Data judgment* × Ownership    | 0.36* |

## Dumbbell · C-2291 (separate run) vs cohort median



Pale bar = cohort middle 50%; hollow dot = cohort median; filled dot = C-2291, a separate adaptive run that is not one of the 40. Shown here only for orientation; a live cohort report does not single anyone out.

# Competency relationships / 1

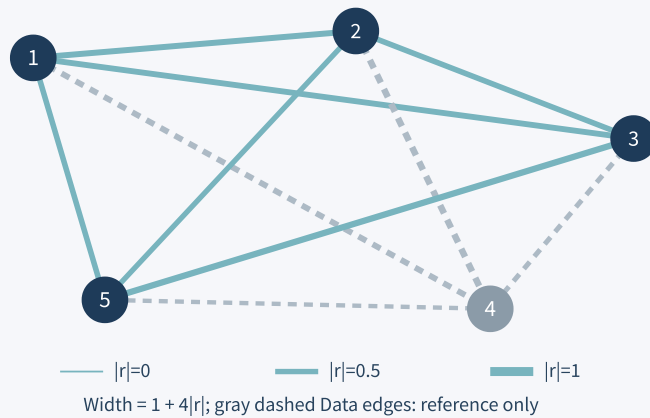
Synthetic template responses; not real candidate data or calibrated Titan output.

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

## Source medians, separate from correlations

|                                    |       |
|------------------------------------|-------|
| Customer problem solving median    | 50.00 |
| Stakeholder communication median   | 50.30 |
| Prioritizing under pressure median | 50.73 |
| Data-informed judgment median      | 51.69 |
| Ownership & follow-through median  | 48.66 |

## Competency correlation network



1 Customer problem solving  
3 Prioritizing under pressure  
5 Ownership & follow-through

2 Stakeholder communication  
4 Data-informed judgment

| Pair   | Pearson r | Pair   | Pearson r |
|--------|-----------|--------|-----------|
| 1 / 2  | 0.582     | 2 / 4* | 0.647     |
| 1 / 3  | 0.565     | 2 / 5  | 0.491     |
| 1 / 4* | 0.565     | 3 / 4* | 0.448     |
| 1 / 5  | 0.525     | 3 / 5  | 0.626     |
| 2 / 3  | 0.530     | 4 / 5* | 0.359     |

\* Includes one-item Data lens. Gray dashed edges are not comparison-eligible.

Among the four comparison-eligible lenses, the strongest descriptive pair is Prioritizing under pressure × Ownership & follow-through ( $r=0.63$ ). The thicker Data-informed judgment × Customer problem solving/Stakeholder edges include a lens with mostly one item per person and are reference-only, not eligible comparisons or recommendations.

Synthetic adaptive source,  $n=40$  completed people. Figure shows competency correlation network on its labeled scale; descriptive, not a hiring decision.

Data-informed judgment has mostly one observation per person. Correlations using it are descriptive artifacts of this generation, not a usable relative-comparison lens or hiring evidence.

## Strongest eligible relationship · $r=0.63$

Prioritizing under pressure × Ownership & follow-through:  $r=0.63$  among four multi-observation lenses. Relationships with the mostly one-item Data-informed judgment are reference-only.

### Five-lens correlation matrix · row/column 4 is reference-only

|   | 1    | 2    | 3    | 4    | 5    |
|---|------|------|------|------|------|
| 1 | 1    | 0.58 | 0.56 | 0.56 | 0.53 |
| 2 | 0.58 | 1    | 0.53 | 0.65 | 0.49 |
| 3 | 0.56 | 0.53 | 1    | 0.45 | 0.63 |
| 4 | 0.56 | 0.65 | 0.45 | 1    | 0.36 |
| 5 | 0.53 | 0.49 | 0.63 | 0.36 | 1    |

r=0
r=0.25
r=0.5
r=0.75
r=1
 Gray dashed: row/column 4, reference only

1 Customer problem solving

2 Stakeholder communication

3 Prioritizing under pressure

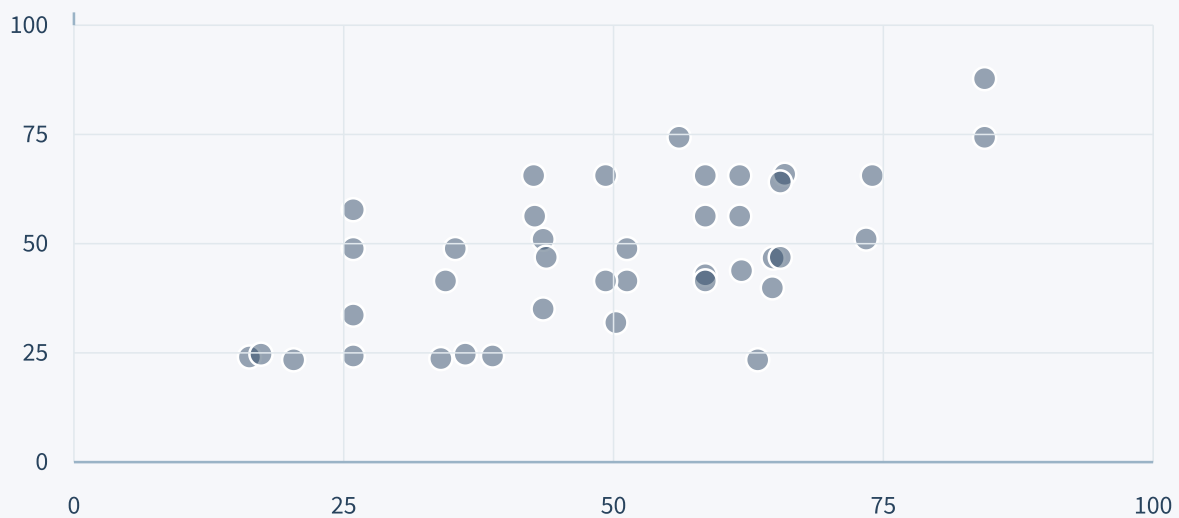
4 Data-informed judgment · reference only

5 Ownership & follow-through

Synthetic descriptive chart; verify the underlying data before use.

Rows and columns 1–5 follow the named legend above. These are generated-score associations only.

### Strongest eligible pair · each dot is one person



Prioritizing under pressure (horizontal) vs Ownership & follow-through (vertical), descriptive 0–100 synthetic scores. Forty completed people, one dot each. Pearson  $r=0.63$ . Scores come from 2–3 observations per person. This describes generated answers, not a validated hiring signal.

# Competency relationships / 3

## Generated-score distributions · 40 people per competency

### Customer problem solving

n=40 · median 50.00



### Stakeholder communication

n=40 · median 50.30



### Prioritizing under pressure

n=40 · median 50.73



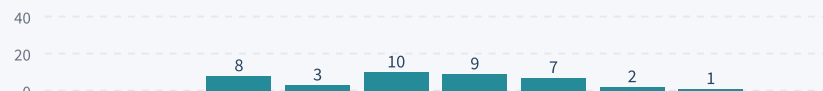
### Data-informed judgment

n=40 · median 51.69



### Ownership & follow-through

n=40 · median 48.66



Count: people per bin · same 0–40 scale · zero bins have zero height

Descriptive 0–100 ten-point bins; each row totals 40. The one-item data-judgment lens is reference-only, not a ranking. Navy dashed line is the exact median of the unbinned source scores, not the middle of a histogram bin. Data-informed judgment has one answer/person and no comparison median mark. The figure is a frequency histogram, not a smoothed density or calibrated interval.

## Generated case vs cohort median · descriptive 0–100 scores

### Customer problem solving

1

### Stakeholder communication

1

### Prioritizing under pressure

1

### Data-informed judgment\*

reference only

### Ownership & follow-through

1

0 25 50 75 100

Orange: C-2291 (separate run, not one of the 40); navy: synthetic cohort median; 0–100 illustrative display

Four candidate-versus-median connections use the four comparison-eligible generated-answer lenses. \* Data-informed judgment: candidate 50/100, cohort median 75/100, one observation/person. These values remain descriptive in the table, but no comparative mark is drawn; reference only, not a hiring rank.

The data-informed judgment row has mostly one observation per person and is reference-only, not an eligible relative-comparison lens. Orange is C-2291, a separate adaptive run that is not one of the 40; navy is the 40-person cohort median. These are illustrative scores, not calibrated job performance.

# Cohort comparison evidence / 9 of 9

Additional descriptive views from the same synthetic 40-person adaptive cohort. Not a validated hiring norm or issued assessment.

## Paired, same-person competency associations



Legend:   
|r| 0 (lightest blue)   
|r| 0.25 (light blue)   
|r| 0.5 (medium blue)   
|r| 0.75 (dark blue)   
|r| 1 (darkest blue)

Shade encodes absolute Pearson r, 0–1. Printed signed values give direction; all associations in this synthetic matrix are positive. Diagonal 1.00 is self-correlation, not independent evidence.

A Customer problem solving; B Stakeholder communication; C Prioritizing under pressure; D Ownership & follow-through. Pearson r across 40 matched rows; descriptive, not causal or a hiring signal.

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

## INTERVIEWS · ONBOARDING · MEASUREMENT NOTE

# Interview on prioritizing under pressure and data-informed judgment; start onboarding with risk awareness.

## Interview questions (from the campaign's shared gaps)

### Prioritizing under pressure · median 51

"Tell me about a time work piled up. What did you decide to push back, and how did you tell people?"

Signs of a strong answer  
Decides what to defer and explains why

Signs to probe further  
Takes it all on, or ends with 'I worked hard'

Follow-up  
What would you change if it happened again?

### Data-informed judgment · median 52 (reference only)

"When data and what people told you disagreed, what did you check first?"

Signs of a strong answer  
Checks sample, time window, definitions first

Signs to probe further  
Picks one side straight away

Follow-up  
What would you change if it happened again?

### Stakeholder communication · median 50

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

Signs of a strong answer  
Impact, options and the ask, briefly

Signs to probe further  
Long background, no ask

Follow-up  
What would you change if it happened again?

### Risk awareness · across competencies

"What was most likely to go wrong with that plan, and what was your fallback?"

Signs of a strong answer  
A specific risk and when to check it

Signs to probe further  
'I expected it to go fine'

When to use  
Add after any answer

## 90-day onboarding suggestions

**Days 1–30** · add one line on 'what could go wrong' to every plan (risk awareness)

**Days 31–60** · in heavy weeks, decide what to defer together with the manager (prioritizing under pressure)

**Days 61–90** · review one case where the dashboard and field notes disagree (data-informed judgment)

## Measurement note

- Candidates answer in their own words, and the engine asks follow-ups based on the answer.
- Each answer is scored on four traits (1–5): evidence, communication, risk awareness, execution. When free text isn't enough, a multiple-choice anchor confirms it.
- Each competency gets its own descriptive 0–100 score, with no interval; they are never added up.
- The evidence base is shown as observation counts and the final SE; there is no evidence grade.
- Every number in this sample is synthetic.

## Evidence trail

| Campaign            | Sample · CSM            |
|---------------------|-------------------------|
| Invited / completed | 42 / 40                 |
| Scored answers      | 416                     |
| Scenario bank       | 14 · 5 competencies     |
| Stop rule           | SE 0.30 or 12 items     |
| Version             | synthetic-adaptive-v350 |

## Scenario bank check before the next campaign

| Competency                  | Scenarios in bank | Avg. obs. per person | Median | Hardest scenario mean |
|-----------------------------|-------------------|----------------------|--------|-----------------------|
| Customer problem solving    | 3                 | 2.4                  | 50.00  | 2.85                  |
| Stakeholder communication   | 3                 | 2.8                  | 50.30  | 2.88                  |
| Prioritizing under pressure | 4                 | 2.0                  | 50.73  | 2.89                  |
| Data-informed judgment      | 1                 | 1                    | 51.69  | 2.89                  |
| Ownership & follow-through  | 3                 | 2.3                  | 48.66  | 2.83                  |

Data-informed judgment has a single scenario, so its score stays reference-only. Competencies with 3 scenarios may repeat situations on a retake. We suggest at least 4 per competency. † Hardest scenario mean uses scenarios with at least 30 answers.

## Measurement settings

|                               |                     |
|-------------------------------|---------------------|
| Target scenarios              | 14                  |
| Min. observations for a score | 2                   |
| Target SE                     | 0.30                |
| Scenario bank                 | 14 · 5 competencies |
| Score scale                   | 0–100 · descriptive |

## Before the next campaign

- Review the wording of the lowest-rated scenario, 'Handover from a departing colleague' (2.83 of 5)
- Competency with a single scenario (reference only): data-informed judgment
- 2 timed out: check the time-limit notice

# Interview, onboarding and measurement / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

Median source table: p. 55. The questions and work checks below are separate drafts, not selection or staffing decisions based on generated scores.

## Question bank and measurement conditions

Use these as prompts for a separate structured interview, not as a hiring recommendation. Each completed synthetic record has 8 to 12 scenario answers. Data-informed judgment has one observation per person, so its descriptive score is not an eligible relative lens.

|                    |                                                                                                                                 |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------|
| Priority trade-off | Ask about a real deadline decision; the generated prioritization median is 50.7/100 from about 2 items per person.              |
| Data versus field  | Ask for a real example and independent evidence; one synthetic item per person is too thin for a comparison.                    |
| Bank check         | Version the scenario bank, item counts, scoring rule, tenant scope and human-rated calibration before use with real applicants. |

## Interview questions · source-limited prompts

|                                       |                                                                                                                                                |
|---------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| Customer problem solving              | About 2.4 items/person, median 50.0/100. Ask how the person would restore customer trust after a failure; seek ownership, timing and evidence. |
| Stakeholder communication             | About 2.8 items/person, median 50.3/100. Ask for a real example of delivering a difficult update; check audience, options and a clear ask.     |
| Prioritizing under pressure           | About 2.0 items/person, median 50.7/100. Ask what was deferred when deadlines collided, why, and what changed.                                 |
| Data-informed judgment                | 1 item/person, median 51.7/100. This is reference-only; use independent job-related evidence, not a relative comparison.                       |
| Risk awareness (answer-quality input) | Ask about a concrete risk and fallback. This is a generated answer-quality input, not a validated person trait.                                |

## Interview, onboarding and measurement / 2

416 COMPLETED ANSWERS

682 PROBE FOLLOW-UPS

36 STOPPED AT SE TARGET

Synthetic, descriptive only. Do not use as a hiring decision.

### Optional onboarding exercises

|            |                                                                                                                                                                                  |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Days 1–30  | If the role calls for it, practice a written customer-risk check and manager review. This is a possible development exercise, not a finding that these applicants need training. |
| Days 31–60 | Use a priority tradeoff drill for the role, with a manager deciding what can be deferred.                                                                                        |
| Days 61–90 | Review one real case where an operational dashboard and field notes disagree; document what evidence settles it.                                                                 |

### Evidence trail · adaptive synthetic protocol

|                                   |                                                                                              |
|-----------------------------------|----------------------------------------------------------------------------------------------|
| Source                            | synthetic-adaptive-v350 seeded run; not a real applicant cohort                              |
| Starts/completions                | 42 / 40, plus 2 timed out after 3 and 4 scenarios                                            |
| Completed answers                 | 416 answers = 40 people × 8 to 12 scenarios (adaptive)                                       |
| Probe follow-ups (0–3 per answer) | 682                                                                                          |
| Stop reasons                      | 36 reached SE 0.30; 4 reached the 12-item maximum                                            |
| Review sample                     | 3 of 40 completions are declared QC samples, not evidence grades                             |
| Stop rule                         | median final theta SE 0.2925; 36 of 40 met the 0.30 target, 4 stopped at the 12-item maximum |
| Version                           | synthetic-adaptive-v350; separate from the v349 fixed-sample reports                         |

### Scenario-bank coverage

|                             |                                                                                                                    |
|-----------------------------|--------------------------------------------------------------------------------------------------------------------|
| Customer problem solving    | 3 scenarios in bank · about 2.4 items/person · 40 completers · median 50.0/100 · lowest scenario mean 2.85/5 (S06) |
| Stakeholder communication   | 3 scenarios in bank · about 2.8 items/person · 40 completers · median 50.3/100 · lowest scenario mean 2.88/5 (S11) |
| Prioritizing under pressure | 4 scenarios in bank · about 2.0 items/person · 40 completers · median 50.7/100 · lowest scenario mean 2.89/5 (S08) |
| Data-informed judgment      | 1 scenarios in bank · about 1.0 items/person · 40 completers · median 51.7/100 · lowest scenario mean 2.89/5 (S04) |
| Ownership & follow-through  | 3 scenarios in bank · about 2.3 items/person · 40 completers · median 48.7/100 · lowest scenario mean 2.83/5 (S09) |

# Interview, onboarding and measurement / 3

Synthetic, descriptive only. Do not use as a hiring decision.

## Measurement settings and validity limits

|                         |                                                                                                                          |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------|
| Scenario protocol       | Adaptive: 8 to 12 of a 14-scenario bank across five competencies (3/3/4/1/3 in the bank)                                 |
| Score interpretation    | Generated 1–5 effectiveness and four 1–5 answer-quality inputs; descriptive competency summary 0–100 is a separate scale |
| Competency intervals    | Illustrative local bands only; valid 90% coverage requires independent human-rated calibration                           |
| Overall theta precision | Median final theta SE 0.2925 across 40 completers; 36 met the 0.30 target; no separate competency coverage claim         |
| Dialogue access         | Employer report masks free-text answers; restricted tenant-bound access is a separate permissioned workflow              |

## Before another campaign

|                   |                                                                                                                                                                     |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Scenario bank     | Current bank has 14 scenarios; Data-informed judgment has one/person. Do not claim it has three. Consider new validated items before using that lens comparatively. |
| Lowest separation | Review the scenario with the lowest corrected item–total $r$ first (see p. 44-47); synthetic descriptive screening, not a real item-quality verdict.                |
| Timed out         | Two declared timeout records reached three and four scenarios; check real-world time limits and notices separately.                                                 |
| Validation        | Version fixed engine, tenant access, report model and human-rated calibration before any real-candidate use.                                                        |

Reading guide: this page lists the measurement settings and limits for this synthetic run, followed by items to review before another campaign.

## All five medians sit near 50; use the same questions for every competency.

### Stakeholder communication · median 50.3

at or below 50: 18 of 40

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

**Follow-up**  
What was your first sentence?

**Strong-answer signs**  
Bottom line first, then cause and next steps

**Weak-answer signs**  
Long lead-in; the bottom line comes late

### Prioritizing under pressure · median 50.7

at or below 50: 19 of 40

**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?

**Strong-answer signs**  
Chooses what to defer and flags it early

**Weak-answer signs**  
Tries to do it all by working longer; escalates late

### Customer problem solving · median 50.0

at or below 50: 19 of 40

**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?

**Strong-answer signs**  
Checks cause first and avoids promises that can't be kept

**Weak-answer signs**  
Jumps to fixes or credits without checking cause

### Ownership & follow-through · median 48.7

at or below 50: 21 of 40

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

**Follow-up**  
Did you tell them before they noticed?

**Strong-answer signs**  
Flags first, gives a new date and a checkpoint

**Weak-answer signs**  
Tells them late or spreads the blame

### Data-informed judgment · median 51.7

at or below 50: 15 of 40

**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?

**Strong-answer signs**  
Checks source, sample size and timing

**Weak-answer signs**  
Trusts one side as-is or decides on gut

### Interviewer sheet

| Competency                  | 1                        | 2                        | 3                        | 4                        | 5                        | Notes |
|-----------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|-------|
| Stakeholder communication   | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |       |
| Prioritizing under pressure | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |       |
| Customer problem solving    | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |       |

Record interview ratings separately; don't merge them with report scores.

# Shared interview questions / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

Median source table: p. 55. The questions and work checks below are separate drafts, not selection or staffing decisions based on generated scores.

## 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?

Trait means are close (2.89–2.94); risk awareness is lowest by a small margin. The plan below is a discussion draft that starts there.

12-week plan

Weeks 1–2

Risk awareness basics: write one line on 'what could go wrong' before each decision

Weeks 3–4

Prioritizing under pressure case workshop, replaying a low-rated scenario ('Quarter-end workload spike')

Weeks 5–8

Data-informed judgment pair coaching on two live cases

Weeks 9–12

Re-measure with new scenarios for the same competencies; no change threshold is defined

Cohort trait means

Evidence

2.92

Communication

2.91

Risk awareness

2.89

Execution

2.94

Mean over 416 answers (40 completers × 8 to 12 scenarios), 1–5.

Onboarding focus by competency

Customer problem solving

19 / 40

Stakeholder communication

18 / 40

Prioritizing under pressure

19 / 40

Data-informed judgment

15 / 40

Ownership & follow-through

21 / 40

People at or below 50 (scores shown to one decimal). 50 is the scale midpoint, not a pass mark.

Onboarding material by competency

| Competency                  | Median | At or below 50 | Weakest trait         | Lowest-rated scenario                                             |
|-----------------------------|--------|----------------|-----------------------|-------------------------------------------------------------------|
| Prioritizing under pressure | 50.7   | 19             | Communication (2.86)  | Quarter-end workload spike (2.89)                                 |
| Data-informed judgment      | 51.7   | 15             | Communication (2.83)  | Churn dashboard disagrees with field notes (2.89, reference only) |
| Stakeholder communication   | 50.3   | 18             | Communication (2.85)  | Team disagreement on an account plan (2.88)                       |
| Customer problem solving    | 50.0   | 19             | Risk awareness (2.91) | Customer demands an unsupported feature (2.85)                    |
| Ownership & follow-through  | 48.7   | 21             | Risk awareness (2.78) | Handover from a departing colleague (2.83)                        |

Lowest-rated scenario = the lowest cohort mean among that competency's scenarios. Use it as the workshop case.

What managers do

· In weekly 1:1s, replay one decision through 'risk awareness'

· Share escalation rules in month one

Use the strengths

· Ask strong ownership & follow-through people to present cases

· Let them coach in pairs

How to check progress

· Re-measure the same competencies at 12 weeks

· No change threshold is defined

· Individual results go only to the person and their manager

# Onboarding evidence / 1

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

These are descriptive synthetic answer metrics, not calibrated competency intervals or evidence grades. Each completed person answered 8 to 12 of the 14 scenarios, chosen adaptively.

Median source table: p. 55. The questions and work checks below are separate drafts, not selection or staffing decisions based on generated scores.

## Generated scores do not assign peer coaches

The synthetic cohort median for Ownership & follow-through is 48.66/100. If managers want to use an example as a discussion case, first verify the person’s real work and ask permission. Generated answer ratings do not identify capable coaches, prove strengths at work, or justify assigning anyone to mentor peers.

## Onboarding and progress checks

Risk awareness is the lowest generated input-axis mean in this dataset, not an observed workplace deficit. A 12-week plan requires manager agreement and actual work observations. Never use synthetic scores as the baseline for improvement.

|                 |                                                                                                                   |
|-----------------|-------------------------------------------------------------------------------------------------------------------|
| Generated input | Risk-awareness mean 2.89/5; uniquely lowest for 9 of 40 completers (ties excluded).                               |
| Weeks 1–4       | Record real decisions, context and risks in a manager-agreed work log.                                            |
| Weeks 5–8       | Review how evidence and communication affected actual decisions.                                                  |
| Weeks 9–12      | Review work evidence separately; a calibrated retest requires comparable items and independent coverage evidence. |

## Lowest generated answer-quality trait per person

|                |                                 |
|----------------|---------------------------------|
| evidence       | 6/40 unique minima              |
| communication  | 11/40 unique minima             |
| Risk awareness | 9/40 unique minima              |
| execution      | 6/40 unique minima              |
| Tied minima    | 8/40; not assigned to one trait |

## Onboarding material and limits / 2

Synthetic answers do not establish anyone's training need. The 50-point cut is descriptive, not a pass mark. A manager may use these as optional job cases only after real work evidence.

### Five-competency source inventory

|                             |                                                                                                                       |
|-----------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Customer problem solving    | Median 50.0/100 · 19/40 below 50 · lowest generated trait risk awareness (2.91/5) · lowest scenario mean S06 (2.85/5) |
| Stakeholder communication   | Median 50.3/100 · 18/40 below 50 · lowest generated trait communication (2.85/5) · lowest scenario mean S11 (2.88/5)  |
| Prioritizing under pressure | Median 50.7/100 · 19/40 below 50 · lowest generated trait communication (2.86/5) · lowest scenario mean S08 (2.89/5)  |
| Data-informed judgment      | Median 51.7/100 · 15/40 below 50 · lowest generated trait communication (2.83/5) · lowest scenario mean S04 (2.89/5)  |
| Ownership & follow-through  | Median 48.7/100 · 21/40 below 50 · lowest generated trait risk awareness (2.78/5) · lowest scenario mean S09 (2.83/5) |

Data-informed judgment has one item per person: its 51.7 median and 15/40 below the descriptive 50-point cut are reference-only, never a relative-comparison signal.

### Optional manager checks, not assignments

|            |                                                                                                                                                     |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| Weeks 1–2  | Agree what real work decisions and risks to document, if the role calls for it.                                                                     |
| Weeks 3–4  | Discuss a real deadline-versus-quality case and what could be deferred.                                                                             |
| Weeks 5–8  | Compare a real dashboard with field notes; evaluate the evidence, not synthetic scores.                                                             |
| Weeks 9–12 | Review work evidence separately. A retest needs comparable items and human-rated calibration; overlapping illustrative bands are not a change rule. |

## How to use this report

Synthetic demonstration, not real candidate data or calibrated selection evidence.

### Four checks before a decision

1

#### 1 · Set the lens

Choose 1–2 role-relevant lenses before reading the five descriptive competency distributions and 40-person matrix. Data-informed judgment has one item per person, so keep it reference-only and out of relative ranks.

2

#### 2 · Check completeness

Review the recorded session states (page 39, adaptive sample): 42 started, 40 completed, 2 timed out; 3 of the completions were preselected for quality review. These are not evidence grades or an automatic exclusion list.

3

#### 3 · Ask, then record

Use the shared interview questions and blank record (pages 59–63) to seek independent, job-related examples. The lower generated-answer means for S09, S06, S11 can suggest questions, not proven workplace weaknesses. Apply the same criteria to everyone; synthetic metrics are not a hiring decision.

4

#### 4 · Carry forward cautiously

The optional onboarding cases and source inventory (page 65) are discussion material only. Use real-work evidence before setting a plan. The current example intervals are not calibrated for change decisions.