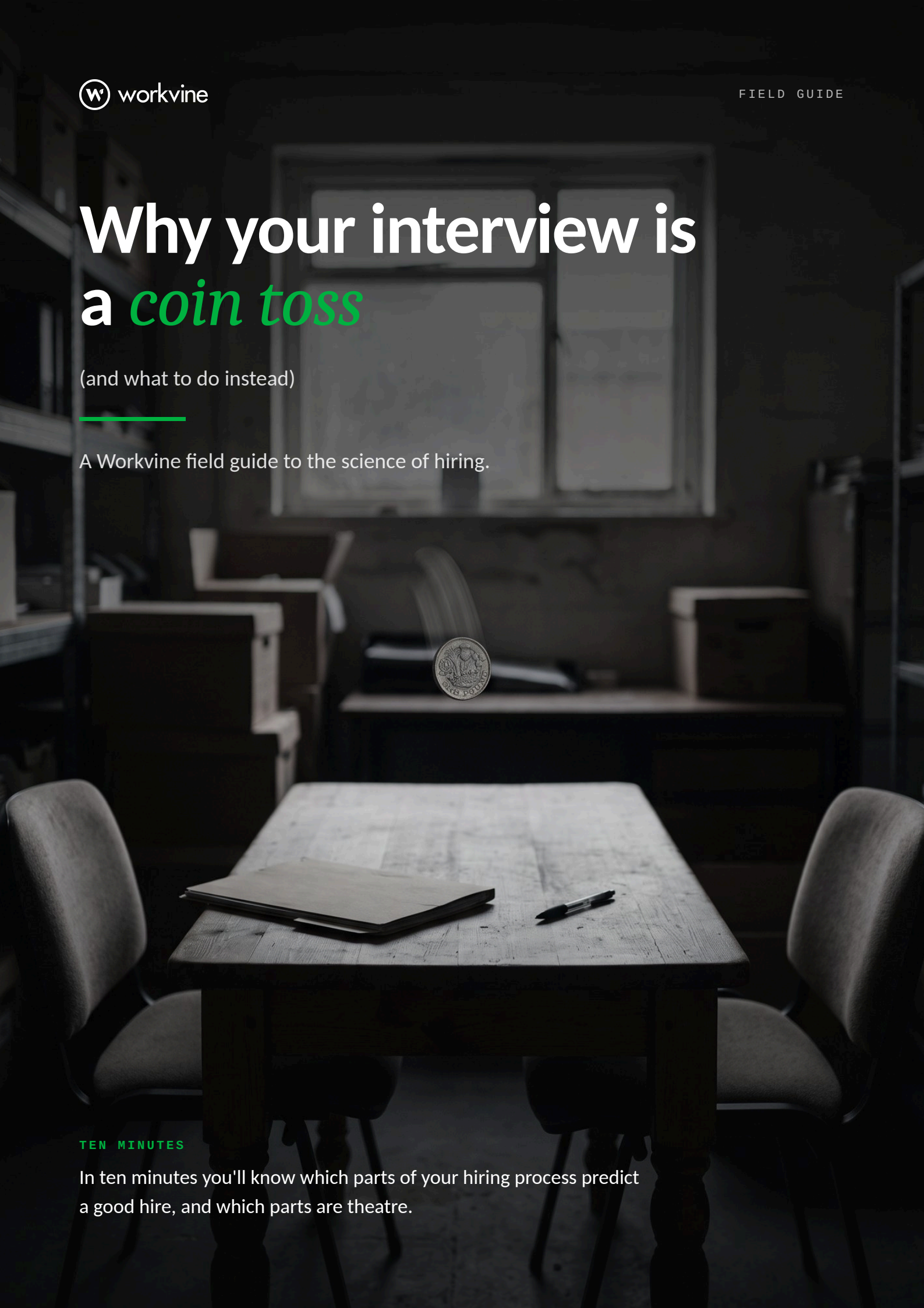
workvine

FIELD GUIDE

Why your interview is a *coin toss*

(and what to do instead)

A Workvine field guide to the science of hiring.



TEN MINUTES

In ten minutes you'll know which parts of your hiring process predict a good hire, and which parts are theatre.

BEFORE WE START

Is there a Heston Blumenthal way to *recruit*?

Heston's trick was never the liquid nitrogen. It was refusing to accept "that's how it's done". Why do chips go soggy? What does the sound of the sea do to a plate of fish? He pulled every step apart, tested each one on its own, and kept only what worked.

Recruitment runs the way his kitchens used to. Tradition and gut feel. The advert gets written the way the last one was written. The interview is a chat with whoever's free. The decision comes down to who "felt right". And the odd thing is, the science has been sitting there for years. Not new research. Old research, in journals nobody in a factory or a warehouse has any reason to read.

This guide takes three parts of hiring and treats each one like a kitchen experiment. Change one thing. Measure it. Keep it or bin it.

01	**The advert.** The part you can test on your next vacancy.
02	**The decision.** The interview, the screening, and why most of it doesn't work.
03	**The notebook.** How to find out, a year later, what actually worked.

None of it needs a budget. All of it needs you to write things down.

Three pull-out pages are built to be printed and used: an advert worksheet, a structured interview block, and a hiring log.



01 THE ADVERT

Start with the *advert*

THE PROBLEM

Most adverts are written once, posted, and never looked at again. If it didn't work, the conclusion is "there's nobody out there", and the same advert goes up again next time. Nobody knows whether the title, the salary line or the first paragraph made any difference, because nothing was changed and nothing was measured.

Heston's most interesting work was on perception: how what you hear, see and expect changes what you taste. With Professor Charles Spence at Oxford he found that oysters taste stronger and saltier when you eat them listening to the sea. That became the Fat Duck's "Sound of the Sea", and the point of it was never the gimmick. The point was that the dish was the same. Only the expectation changed.

A job advert is the same problem. The candidate isn't reading your advert. They're reading what it says about working for you. Same job, same money, same site: a different first line changes who applies.

27%

more applications for adverts that show a salary than for those that don't.

SOURCE: REED.CO.UK SITE DATA

THE FIVE VARIABLES THAT MOVE THE NUMBERS

01 **The job title.** What candidates search for, not what the org chart says. "Maintenance Engineer" gets searched. "Asset Reliability Technician (Band 4)" does not.

02 **Whether the salary is on it.** Reed.co.uk's own site data shows adverts that display a salary receive 27% more applications than those that don't. Put the salary on. If there's a band, put the band on.

03 **Whether the first line talks about them or about you.** "Established in 1987, we are a leading supplier of…" is about you. "You'll run the night shift on two CNC lines, with a team of four" is about them.

04 **How long the list of requirements is.** A requirements list reads as a filter, and people treat it as one. Tara Mohr's 2014 survey of over a thousand people found the main reason they didn't apply wasn't confidence; they believed the listed requirements were hard rules for being considered at all. Every "nice to have" you leave in the must-have list costs you applicants who could do the job. Cut the list to what someone needs on day one.

05 **Which boards it runs on, and when it goes live.** Covered below.

HOW TO RUN THE EXPERIMENT

You don't need to guess at any of this. Post two versions. Change one thing between them. Watch where the applications come from, how many make the shortlist, and who drops out. After enough vacancies, you know rather than believe which headlines bring the right people in your sector.

The discipline is the hard part: one variable per test, and the same measure every time. Change the title and the salary line at once and you've learned nothing. The worksheet on the next page is built for one change per vacancy.

TIMING AND BOARDS

Two things about distribution that cost nothing to get right.

- **Go live early in the week.** An advert that goes up on a Monday or Tuesday gets a full working week of searches before it drops down the listings. A Friday afternoon advert lands when fewer people are looking and is a week old by the time they are. Test this too: the worksheet has a column for it.
- **Post everywhere at once.** Candidates for the same role search on different boards, and most only use one or two. One advert on one board reaches that board's searchers. Workvine's DIY Recruitment sends one advert to around ten UK boards at once (Totaljobs, CV Library, Indeed, Google for Jobs and their onward networks) and runs it for 28 days. The same advert, one post, the whole market in one go.
- **Let it run the full 28 days, then stop and compare.** Don't repost halfway through. A fresh post resets the clock on your measurement, and you lose the one number the experiment was for.

PULL-OUT 1

The advert experiment worksheet

Print it and pin it up for the 28 days the advert runs. Change ONE row between version A and version B. Everything else stays identical. Fill the results boxes on the day the advert closes, not from memory a month later.

Vacancy:

Live from: __ / __ / ____ **Closes:** __ / __ / ____

The one thing I changed:

VARIABLE	VERSION A	VERSION B
Job title (as searched)		
Salary shown? (figure / band / none)		
First line about them or about us?		
Must-haves on the list (count)		
Boards and go-live day		

RESULT	VERSION A	VERSION B
Applications received		
Applications that met the must-haves		
Shortlisted		
Shortlist rate (shortlisted ÷ applications)		
Top source (which board)		
Dropped out before interview		

Keep or bin: the winning version becomes next time's version A. Write the losing variable on the back so you don't test it again by accident.

The same job, written *twice*

A fictional role at a fictional Midlands manufacturer. No client, no real company. The two adverts describe exactly the same job on exactly the same money. Only the writing changes.

BEFORE: AS COMMONLY WRITTEN

Multi-Skilled Maintenance Technician (Electrical Bias), Grade 5

Salary: competitive, dependent on experience.

Established for over 40 years, we are a leading manufacturer of precision pressed components supplying tier-one automotive customers across Europe. Due to continued growth we are now seeking a Multi-Skilled Maintenance Technician to join our busy engineering team.

The successful candidate will be responsible for reactive and planned maintenance across the site, ensuring maximum uptime of production equipment and compliance with all H&S requirements.

Requirements: time-served apprenticeship, HNC/HND in Electrical or Mechanical Engineering, 17th/18th Edition, minimum 5 years' experience in a fast-paced manufacturing environment, experience of press tooling, hydraulics, pneumatics, PLC fault-finding (Siemens/Allen Bradley), robotics (ABB/Kuka), TPM, root cause analysis, IOSH, FLT licence, ability to work under pressure, excellent communication skills, flexible attitude, team player.

Hours: shifts. Benefits: pension, 25 days' holiday.

AFTER: SEARCHABLE TITLE, SALARY ON, CANDIDATE FIRST, MUST-HAVES ONLY

Maintenance Engineer (electrical), Hinckley. £42,000 to £46,000 plus shift allowance

You'll keep two press lines and a robot cell running on a continental shift pattern (four on, four off, days and nights). You're one of six engineers across the site, so you'll get the breakdowns, the planned work and a say in what gets fixed properly rather than patched.

The kit: 400 to 800 tonne presses, Siemens PLCs, ABB robots, hydraulics and pneumatics. Spares budget is yours to call on without a form.

You'll need: a time-served electrical or multi-skilled background, 18th Edition, and PLC fault-finding you've done yourself, not watched. Press or robot experience helps; we'll train the rest.

Pay: £42,000 to £46,000 depending on what you bring, plus 18% shift allowance on top. 25 days' holiday plus bank holidays, pension. Site is five minutes from the A5 with free parking.

What changed, and what to *measure*

VARIABLE	BEFORE	AFTER
Title	Internal grade and jargon; nobody searches "Grade 5"	The two words candidates type, plus the location
Salary	"Competitive"	The band, and the shift allowance that makes it real
First line	About the company	About the shift, the kit and the team they'd join
Must-haves	Nineteen items, most of them wishes	Three, plus two "helps"
Length	Longer, says less	Shorter, says more

The before version isn't a parody. Search any job board for this title and you'll find a dozen like it today.

> The after version is the same role written by someone who has spoken to the engineer leaving. That conversation is where the kit list and the spares-budget line came from, and it's the single best source of advert copy you have.



02 THE DECISION

Then fix the *decision*

THE PROBLEM

The interview is a chat. The questions change with the candidate and the mood of the room. Two interviewers come out with two opinions and the louder one wins. A year later, the hire that "felt right" is the one who left in March, and nobody can say what they saw in the room.

Heston's rule for a recipe was that every step had to earn its place. If he couldn't show that a step changed the dish, it went. Hiring has steps like that, and the research on which ones earn their place has been building for most of a century.

Frank Schmidt and John Hunter pulled 85 years of it together in 1998, in Psychological Bulletin. In 2022 Paul Sackett and colleagues re-ran the maths with tighter corrections and published the revised figures in the Journal of Applied Psychology.

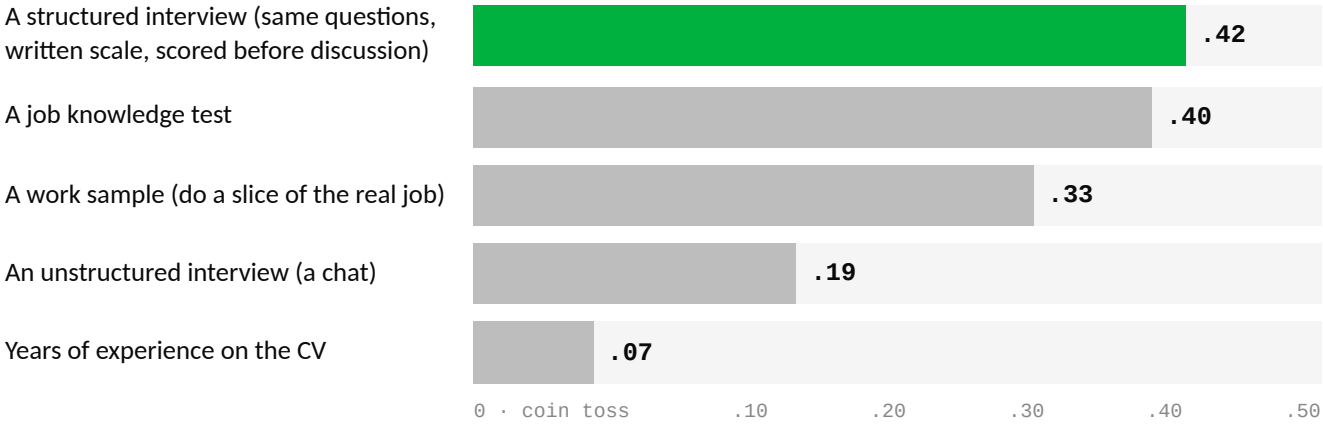
.19

is how well a chat-style interview predicts job performance. A coin toss is zero. A structured interview scores .42.

SOURCE: SACKETT AND COLLEAGUES, 2022

What each method predicts about job performance

Correlation with later job performance, 2022 figures. 1.0 would be perfect prediction. 0 would be a coin toss.



Source: Sackett, Zhang, Berry and Lievens, Journal of Applied Psychology, 2022. The 1998 figures most people quote are higher; the 2022 re-analysis lowered nearly every method.

Two things in that chart are uncomfortable. An unstructured interview predicts about as much as a personality questionnaire. Years of experience, the thing most CV sifts are built on, predicts almost nothing. Most hiring decisions are built on both.

What works is dull by comparison. Same questions for every candidate. Score each answer against a written scale before anyone talks. Give people a short task that looks like the real job. Keep interviewers apart until the scores are in.

Daniel Kahneman, Olivier Sibony and Cass Sunstein call this "decision hygiene" in Noise (2021). Their finding on interviews is blunt: interviewers of job candidates make widely different assessments of the same people. Their two fixes are structure and independence. Ask the same things in the same order, break the judgement into separate scores, and collect each interviewer's scores before the group discusses anyone. Kahneman put it this way in 2022: the most important recommendation of decision hygiene is structuring, and the deep principle behind it is independence.

None of this stops you using your judgement. It makes you use it on the same evidence for every candidate, and write it down while it's still yours rather than the room's.

The AI angle. Candidates now use AI to write applications and to coach themselves through phone and text screening. A phone screen with a script on the other end tells you about the script. The counter is a structured video screen: the same questions for everyone, a real person asking them, follow-ups the candidate can't have rehearsed, and a written score on each answer. Our line on it is smart tools, human decisions. Use AI to work faster behind the scenes, on the advert, the sift, the scheduling. Never let it talk to a candidate, and never let it decide who gets hired.

PULL-OUT 2

A structured interview block

Four questions for a maintenance engineer, each with a written 1 to 5 scale. Every candidate gets all four, in this order, with the same follow-up. Each interviewer scores alone, on paper, before anyone speaks. Swap the kit and the examples for your own role; keep the shape.

- Ask the question as written. Then the follow-up as written. Then stop talking.
- Score before you discuss. Write the number and one line of evidence (what they actually said).
- Scores are compared after all candidates are seen, not after each one.
- A 3 is a pass. A 2 is a concern you'd want a second source on. A 1 or a 5 should be rare, and you should be able to quote why.

Question 1: breakdown under pressure (behavioural)

"Tell me about the last time a line went down on your shift and you were the one who had to get it running. What did you do first?" **Follow-up:** "What did you find, and how long was it down?"

1	Can't give a specific example, or the example is someone else's fix
2	A real example, but vague on the fault-finding. "I checked it over and found the problem." No sequence, no evidence of isolating the fault
3	A specific fault, a sensible sequence (safe isolation, symptoms, likely causes, test, fix), a plausible time
4	As 3, plus they explain why they went where they went first, and what they ruled out and why
5	As 4, plus they changed something afterwards so it wouldn't recur, and can say what

Question 2: PLC fault-finding (technical)

"A conveyor stops. The HMI shows no alarm. Walk me through how you'd find the fault." **Follow-up:** "At what point would you go into the PLC, and what would you be looking at?"

1	Would call the OEM or wait for someone else
2	Jumps straight to swapping parts, or straight to the PLC, with no check of the obvious (power, e-stops, guards, sensors)
3	Checks the physical and safety circuit first, then inputs and outputs, then the logic. Knows what online monitoring shows them
4	As 3, and can describe forcing an output safely, or tracing a rung, in the kit they've used
5	As 4, and asks a clarifying question before answering (which PLC, is there a drive in the loop, has anything changed recently)

Question 3: planned work versus firefighting (behavioural)

"Tell me about a planned maintenance job you had to push back because of breakdowns. How did you decide, and what happened to the planned job?" **Follow-up:** "Who did you tell?"

1	"Planned work always gets pushed, that's maintenance." No example
2	An example, but the decision was made for them or they can't say why that order
3	A clear reason for the order, the planned job rescheduled, the right person told
4	As 3, and they weigh the risk of leaving the planned job (what fails if it slips again)
5	As 4, and they did something to stop the same clash next month

Question 4: the thing you fixed properly (behavioural)

"Tell me about a repeat fault you eventually fixed for good. What was the real cause?" **Follow-up:** "How many times had it been patched before that?"

1	No example. Doesn't distinguish between a repair and a fix
2	An example, but the "real cause" is a guess or a part swap that happened to hold
3	A specific repeat fault, a cause they found by looking rather than by luck, and a fix that matches the cause
4	As 3, and they can say what the repeat fault was costing (downtime, parts, overtime)
5	As 4, and they had to argue for the time or the money to do it properly, and did

Scoring sheet

One per candidate, per interviewer. Score alone, before anyone discusses.

Candidate: _____ **Interviewer initials:** _____ **Date:** _____

	Q1	Q2	Q3	Q4	TOTAL /20
Score					
One line of evidence					

Work samples for jobs that aren't at a *desk*

THE PROBLEM

Everyone assumes a "work sample" means a coding test or a presentation, so nobody runs one for a warehouse supervisor or a CNC setter. Yet those are the roles where twenty minutes of the real job tells you more than an hour of talking about it. The research puts a work sample at .33 against .19 for a chat. On the shop floor, that gap is wider than it looks, because the gap between "can describe it" and "can do it" is wider.

The rules: it must be a slice of the actual job, it must be the same slice for every candidate, it must be scored against a written sheet, and it must be short enough that a good candidate with another offer doesn't walk. Twenty minutes. Thirty at most.

WAREHOUSE SUPERVISOR, 20 MINUTES

Give them a printed sheet: tomorrow's shift. Nine pickers rostered, two have called in sick, one is on restricted duties. Three outbound trailers with cut-off times. A stock-check the site manager wants done by Friday. One pallet location that keeps coming up short.

Ask them to write the shift plan on the sheet. Who goes where, what gets dropped, what they'd escalate and to whom.

Score on: did they cover the trailer cut-offs first; did they say what they'd drop rather than pretend nothing has to give; did they put the restricted-duties person somewhere useful; did they flag the short location rather than ignore it; did they say who they'd tell. Five things, one point each.

CNC SETTER, 25 MINUTES

At the machine, with the drawing for a simple part they'll have seen the like of before. Ask them to talk you through the set-up: work-holding, tooling, offsets, first-off checks. Then hand them a first-off that's 0.05mm oversize on one dimension and ask what they'd do.

Score on: safe sequence; sensible work-holding for the part; knows which offset moves which dimension; checks the first-off against the drawing rather than against the last part; asks about tolerance or material before answering the oversize question. Five things, one point each.

DOING IT FAIRLY

Same brief, same time, same sheet, same scorer where you can. Tell candidates in advance that there's a practical and roughly what it covers. Make reasonable adjustments if anyone asks. The point is to see the job done, not to spring a surprise.



03 THE NOTEBOOK

Keep the *notebook*

THE PROBLEM

You find out whether a hire worked twelve months on, by which point nobody remembers what "felt right" at interview, which advert they came from, or what the other candidate scored. Every vacancy starts from scratch. The business that hires six people a year has done six experiments and kept the results of none of them.

The last piece of the Heston method is the least glamorous. He wrote everything down. Every version of a dish, what changed, what the tasters said. The lab notebook is what turned a hunch into a recipe someone else could follow.

Hiring has a slow feedback loop, which is exactly why it needs a notebook. The hire you make in October tells you in the following October whether the interview score meant anything. If the score, the source and the advert version aren't written down somewhere, that lesson arrives and leaves without anyone noticing.

1 row per vacancy. Ten columns. Five minutes at offer and two minutes at each check-in.

PULL-OUT 3, OVERLEAF

PULL-OUT 3

The hiring log

One row per vacancy. Log the runner-up too, on the row below, with "No" in the Hired column. You'll want their scores when you're asking whether the interview is any good.

VACANCY	ADVERT VERSION	SOURCE	SCREEN SCORE	INTERVIEW SCORE	HIRED	IN POST AT 3 MONTHS	6 MONTHS	12 MONTHS	NOTE
Maintenance Engineer, Oct 26	*B (salary on)*	*CV Library*	*14/20*	*16/20*	*Yes*				

HOW TO READ IT AFTER A YEAR

Sort by the 12-month column and look at what the stayers have in common: if the high interview scores stayed and the low ones left, the interview is working, and if there's no pattern, change the questions. Do the same with the source column, and you'll see which board sent the people who are still here rather than which board sent the most CVs. Then do it with the advert version, and that's the end of the "there's nobody out there" conversation.

TO FINISH

So yes, there is a scientific way to *recruit*

Tested adverts, structured questions, written scoring, a short slice of the real job, and a stubborn habit of writing down what happened. None of it is glamorous. All of it beats instinct, and all of it is in this guide.

Start with the worksheet. Your next vacancy is the experiment.

At Workvine we write every advert from a detailed brief, run it across around ten boards, and screen every candidate on a structured Teams video call with the same questions and the same scale. We log the lot. The advert data from dozens of campaigns a year is the part a business hiring twice a year never gets to see.

SOURCES

Interview and selection figures. Sackett, Zhang, Berry and Lievens, "Revisiting meta-analytic estimates of validity in personnel selection", Journal of Applied Psychology, 2022 (doi 10.1037/apl0000994). Schmidt and Hunter, Psychological Bulletin 124(2), 1998 (doi 10.1037/0033-2909.124.2.262).

Decision hygiene. Kahneman, Sibony and Sunstein, Noise (2021); Kahneman interview, Issues in Science and Technology, spring 2022.

Sound of the sea. Charles Spence, Gastrophysics (2017); Heston Blumenthal quoted in the Mail and Guardian, 16 April 2007.

Salary on adverts. Reed.co.uk site data, reported in HRreview and theHRDirector.

Requirements lists. Tara Mohr, "Why Women Don't Apply for Jobs Unless They're 100% Qualified", Harvard Business Review, 25 August 2014.

Sources checked October 2026. The advert example is fictional.

YOUR NEXT STEP

Want to see the advert data from our last campaigns in your *sector*?

Ask Kelly or Adrianna.

WORKVINE.CO.UK