

PLAYBOOK

Consulting in Monitoring & Evaluation with AI

● SEVEN STAGES ● TWO EXAMPLES ● ONE METHOD

Win the right assignments, contract them well, deliver evidence that holds up, and see it used. Then do it faster with AI, without spending your clients' trust.



Start where you are.

For anyone who takes on monitoring and evaluation work for clients: solo consultants, small firms and in-house teams who work like consultants. Pick the path that fits today.

1

Completely new

Chapter 01, page 3

The job, ten words, a 30-day plan.

2

Deciding whether to bid

Chapter 02, page 6

Spot the risks in a ToR.

3

Running an assignment

Chapter 04, page 10

Seven stages, two examples.

4

Top-firm toolkit

Chapter 07, page 22

21 tools; templates in the pack.

Contents 18 CHAPTERS IN FIVE PARTS

PART 1 GET STARTED		
01 New to M&E consulting? Start here.	3	
What an M&E consultant does, how an assignment flows, the words you will meet first, and a 30-day plan to get from zero to your first practice assignment.		
02 Spot the risk in this ToR.	6	
Eight lines from realistic terms of reference. Each hides a risk that sinks assignments.		
03 What kind of consulting is this?	8	
The client asks for a report. What they need may be a decision, a system or a skill. Know which role you are playing.		
PART 2 THE METHOD		
04 Seven stages. One assignment.	10	
The CONSULT method takes any M&E assignment from the first ToR to findings that get used.		
05 Why it works.	17	
It joins two traditions: how the best consultants solve problems and build trust, and the standards evaluators are held to.		
06 Judgement and methods.	19	
How top evaluators get from evidence to a verdict they can defend, choose a causal design that fits the question, and evaluate programmes that keep changing.		
PART 3 TOOLS AND TEMPLATES		
07 The top-firm toolkit.	22	
Twenty-one tools that strategy firms, proposal teams and evaluators rely on, explained from scratch, with a template, a worked example and an AI prompt for each.		
08 The engagement studio.	26	
Six working tools from top-firm practice. Fill them in for a live assignment and copy the result into a proposal, an inception report or a status note.		
09 The field guide.	29	
The templates and rules to keep open on every assignment.		
10 Consulting with AI.	32	
AI makes a good consultant faster and a careless one dangerous. Client trust is the asset you protect.		
11 The bid desk.	36	
Four tools for your next bid: scan a ToR, price it, set your day rate and build the evaluation matrix.		
12 The kit: calculators, templates and concepts.	38	
Everything you need to work on your own: six calculators for the numbers clients ask about, nine document templates you can download and fill, and the evaluation concepts explained in plain words.		
13 Build your name: be found, trusted and remembered.	41	
Six tools to turn good work into a name clients can find and describe: your positioning, your profile, a case study, a content plan, your relationships and the events you attend.		
PART 4 YOUR CAREER		
14 Build a practice, not a gig.	44	
Most consultants are good at the work and weak at the business around it. Both decide whether you last.		
15 Leading and growing.	45	
Top consultants lead teams, pass due diligence, negotiate well, run a room, get evidence used, and become known for one thing.		
16 From first assignment to global consultant.	48	
A five-level path from new evaluator to leading international evaluations: what you do at each level, what proves it, and the move that gets you to the next one.		
PART 5 REFERENCE		
17 Ready?	51	
Ten checks before you bid, and ten before you submit the final report.		
18 The library of ideas.	52	
The research, standards and practices behind every stage, a glossary of every term, and a one-page cheat sheet.		



PART 1 GET STARTED

1 START HERE

New to M&E consulting? Start here.

What an M&E consultant does, how an assignment flows, the words you will meet first, and a 30-day plan to get from zero to your first practice assignment.

- 1 The decision is the product; the report is the vehicle
- 2 Every assignment follows the same seven stages
- 3 Learn by doing: a practice assignment is waiting

An M&E consultant helps clients make better decisions with evidence. THE JOB IN PLAIN WORDS

1 Who hires you NGOs, foundations, company CSR teams, governments, UN agencies and development banks, and the consulting firms that serve them.	2 What they buy An answer to a question they must decide on, a monitoring system that works, or skills their team keeps after you leave.	3 What you deliver Plans, tools, data, analysis, and a report or workshop that changes a decision. The decision is the product; the report is only the vehicle.
---	--	---

Every assignment follows the same seven stages. THE JOURNEY

C Choose the work Decide whether the job is right for you.	O Offer Write a proposal that shows you understand the problem.	N Negotiate the contract Agree the contract: scope, money and the rules.	S Set up the inception Plan the work in detail with the client.	U Uncover the evidence Collect and analyse the evidence.	L Land the findings Share the findings so people accept and act on them.	T Transfer and track use Hand over, follow up, and learn for next time.
--	---	--	---	--	--	---

Ten words you will meet first. LEARN THESE AND MOST TORS MAKE SENSE

ToR Terms of reference: the client's description of the job, its questions, deliverables, timeline and budget.	Proposal Your written offer: how you understand the problem, your approach, team, workplan and price.
Inception report The detailed plan you agree with the client before collecting data. It becomes the working contract.	Evaluation question A question the work must answer, such as 'Did incomes change, and for whom?'
Theory of change How and why a programme is expected to lead to change, step by step.	Indicator Something you measure to see whether change is happening, such as average monthly income.
Baseline The measurement at the start, so you can see change later.	Evaluation matrix A table that links each question to indicators, data sources, methods and analysis.
OECD DAC criteria Six common lenses for judging a programme: relevance, coherence, effectiveness, efficiency, impact and sustainability.	Management response The client's formal reply to each recommendation, with owners and dates.

1 START HERE

Ten M&E essentials, and a 30-day plan.

The evaluation basics underneath consulting, then four weeks from zero to your first practice assignment.

01 Monitoring and evaluation

Monitoring tracks what is happening, continuously. Evaluation judges whether it worked and why, at set moments.

You will use it: Every ToR: knowing which one is being asked for changes the method, time and price.

02 Theory of change

A picture of how activities lead to outputs, outcomes and impact, and the assumptions in between.

You will use it: Designing questions, and when a client has no clear logic for the programme.

03 Results framework and logframe

A table of the programme's goals, indicators, targets and sources at each level.

You will use it: MEL system set-ups and checking what data already exists.

04 Indicators

Specific, measurable signs of change, broken down by group where it matters (for example by gender or district).

You will use it: Building monitoring systems and evaluation matrices.

05 Evaluation designs

Ways to judge cause and effect: comparison groups (experimental or quasi-experimental), theory-based designs such as contribution analysis, and participatory approaches.

You will use it: Whenever a ToR uses the word 'impact'.

06 Sampling

Choosing who to ask so the answers represent the people you care about, with a size you can afford.

You will use it: Every survey, and every budget conversation.

07 Qualitative methods

Interviews, focus groups and observation that explain why and how things happened.

You will use it: Most evaluations: numbers show what changed; people explain why.

08 Mixed methods and triangulation

Combining numbers and stories, and checking one source against another before you call something a finding.

You will use it: Analysis, and rating the strength of evidence.

09 Data quality

Checking that data is complete, accurate, timely and consistent before anyone uses it.

You will use it: Monitoring reviews and every fieldwork week.

10 Ethics and consent

Doing no harm: informed consent, privacy, safety, and reporting honestly.

You will use it: From the bid to the final report.

Four weeks from zero to your first practice assignment. A 30-DAY STARTER PLAN

WEEK 1

Learn the language

- ☐ Read this Start here chapter and the ten words
- ☐ Take the test: spot the risk in eight ToR lines
- ☐ Read what kind of consulting this is, and meet the two examples
- ☐ Read the M&E essentials and open one link a day

WEEK 2

Learn the method

- ☐ Walk through the seven stages with both examples
- ☐ Read why it works: trust, problem solving, standards
- ☐ Study three tools: bid gate, problem statement, issue tree
- ☐ Fill in a problem statement in the studio

WEEK 3

Practise on a real-looking ToR

- ☐ Do practice tasks 1 to 4 below
- ☐ Compare each answer with the model answer
- ☐ Use the ToR scanner on the practice ToR
- ☐ Write down three things you would ask the client

WEEK 4

Get ready for real work

- ☐ Do practice tasks 5 to 7
- ☐ Rate yourself on the eight competencies
- ☐ Write your one-sentence niche
- ☐ Start a master CV in the donor format

1 START HERE

Your practice assignment.

Seven tasks on the practice ToR (the mid-term evaluation ToR in the bid desk). Try each one, then compare with the model answer.

1 Decide whether to bid

20 minutes

Read the practice ToR. Score it on the six bid criteria and decide: go, go with conditions, or no-go.

Model answer: About 11 or 12 of 18: go with conditions. The conditions: fewer questions, a design that fits 'impact', milestone payments, independent enumerators, and a written change process instead of 'any other tasks'.

2 List the risks

15 minutes

Paste the practice ToR into the ToR scanner. Then find any risk it missed.

Model answer: Open-ended scope ('any other tasks'), an impact claim with no comparison or baseline, payment of 100% on approval, approval of findings by the programme team, a vague deliverable, staff collecting data, all districts and beneficiaries, 30 days, and 10 questions.

3 Write three clarifying questions

20 minutes

Write three polite questions to send before the bid deadline. Each should change your design or price.

Model answer: 1. Which decisions will the findings inform, and when? 2. Is there baseline or monitoring data on household income, and can we access it? 3. Could payment follow milestones (inception and final report submission), as is common practice?

4 Fix the key question

20 minutes

Turn the ToR's purpose into one specific key question for a problem statement.

Model answer: Is the livelihoods programme on track to raise household income across the four districts, what explains differences between them, and what should change in the second half?

5 Group the 10 questions

30 minutes

Group the ToR's 10 questions into no more than five branches with no overlaps, and write a hypothesis for each.

Model answer: Relevance (Q1); Income and savings change (Q2, Q3, Q7); What explains differences: groups, federations, training, markets (Q4, Q5, Q6, Q8); Sustainability (Q10); Efficiency (Q9). Hypothesis for branch 3: strong federations explain most of the difference.

6 Price it and propose payments

30 minutes

Use the days and budget estimator with the mid-term preset. Adjust days for 30 calendar days, then set milestone payments.

Model answer: About 38 lead days and 29 team days do not fit into 30 calendar days with reviews: propose 14 weeks, or a smaller scope. Payments: 30% on signing, 40% on the inception report, 30% on submission of the final report.

7 Draft the first page of the report

30 minutes

Write a governing message and three supporting action titles, as if your hypotheses were confirmed.

Model answer: Results follow strong federations; invest in them before expanding. Supporting: incomes rose in two districts and were flat in two; the gains track federation strength; federated groups are likelier to last.

PART 1 GET STARTED

2 THE TEST



Spot the risk in this ToR.

Eight lines from realistic terms of reference. Tick the risk you think each one hides, then turn the page.

1 Most failed assignments were set up to fail before signing

2 The risks hide in ordinary sentences

3 Almost all of them can be fixed before you sign

1 · SCOPE

The consultant will also undertake any other tasks assigned by the programme team.

- ☐ Open-ended scope
- ☐ Vague deliverable
- ☐ Too many questions for the time

2 · DESIGN

The evaluation will assess the impact of the programme on household income across all 12 districts.

25 days, no baseline

- ☐ Impact claim with no design
- ☐ Unrealistic timeline
- ☐ Open-ended scope

3 · PAYMENT

Payment: 100% of the fee on approval of the final report by the programme manager.

- ☐ Payment risk
- ☐ A threat to independence
- ☐ Vague deliverable

4 · QUESTIONS

The evaluation will answer the 14 questions listed in Annex 2.

20 working days

- ☐ Too many questions for the time
- ☐ Unrealistic timeline
- ☐ Impact claim with no design

5 · INDEPENDENCE

Findings will be finalised after review and approval by the programme team.

- ☐ A threat to independence
- ☐ Payment risk
- ☐ Conflict of interest

6 · DELIVERABLES

Deliverable: a high-quality evaluation report.

- ☐ Vague deliverable
- ☐ Open-ended scope
- ☐ Too many questions for the time

7 · TIMELINE

Proposals due in five days. Fieldwork to begin the following week and the final report in one month.

- ☐ Unrealistic timeline
- ☐ Too many questions for the time
- ☐ Payment risk

8 · ELIGIBILITY

Preference will be given to consultants who helped design the programme.

- ☐ Conflict of interest
- ☐ A threat to independence
- ☐ Impact claim with no design

2 THE ANSWERS

The risks, and what to ask for.

Each one is fixable before you sign. After you sign, each one costs you days, money or trust.

1 Open-ended scope · Scope

Every unpriced task you agree to comes out of your margin or your quality. "Any other tasks" means unlimited work for a fixed fee.

Ask for: "Additional tasks will be agreed in writing as a change request, with revised days and fees."

2 Impact claim with no design · Design

Impact means change caused by the programme. Without a baseline or comparison, and with 25 days, nobody can measure it honestly.

Propose: "Assess contribution to income change in four districts, using monitoring data and case studies, and state the limits."

3 Payment risk · Payment

You fund the whole assignment, and approval can be delayed or tied to findings the client likes.

Ask for milestones: 30% on signing, 40% on the inception report, 30% on submission of the final report.

4 Too many questions for the time · Questions

Fourteen questions in 20 days is about a day and a half each, including fieldwork. All get shallow answers.

Propose: "We will group these into five key questions, agreed with you at inception."

5 A threat to independence · Independence

Approval of findings by the people being evaluated invites pressure to soften them.

Ask for: "The client reviews drafts for factual accuracy. The evaluator decides the findings."

6 Vague deliverable · Deliverables

Nobody can tell when "high quality" is done, so review rounds never end.

Specify: "A report of up to 30 pages with a 3-page summary, one round of written comments, and a findings workshop."

7 Unrealistic timeline · Timeline

No time for inception, tool testing, ethics approval or recruiting a field team. Quality is lost before you start.

Propose a realistic plan: two weeks of inception, three of fieldwork, three of analysis and reporting.

8 Conflict of interest · Eligibility

People who designed a programme are judging their own work. Funders and boards will question the findings.

Declare any past involvement. If you designed it, offer a learning review, not an independent evaluation.

1

Scoped by hope

Too many questions, too few days, and no change process. The work grows; the fee does not.

2

Built for the shelf

Nobody asked who will use the findings, for which decision. The report is filed and forgotten.

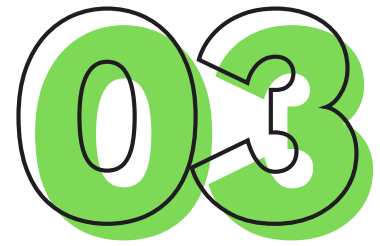
3

Trust spent early

A missed deadline, a surprise finding in the final meeting, an unpaid invoice. Relationships end on small things.

PART 1 GET STARTED

3 KINDS OF WORK



What kind of consulting is this?

The client asks for a report. What they need may be a decision, a system or a skill. Know which role you are playing.

1 Three roles: expert, doctor, process partner

2 Eight common M&E assignments

3 Two examples, followed throughout

Three ways to help a client EDGAR SCHEIN, HUMBLE CONSULTING

1

Expert

The client knows the problem and buys your skill to solve it.

"Design a sampling plan for our survey."

Risk: you solve the wrong problem very well.

2

Doctor

The client knows the symptoms and wants a diagnosis and a prescription.

"Our data is always late. Fix our M&E."

Risk: the client never owns the cure.

3

Process partner

You help the client find and solve the problem themselves.

"Help our team agree what success looks like."

Risk: it feels slow to clients who wanted an answer.

What clients buy, and the usual size

ASSIGNMENT	WHAT THE CLIENT GETS	TYPICAL SIZE	USUAL ROLE
MEL system set-up	Results framework, indicators, forms, dashboards and routines	10 to 30 days	Process partner
Baseline study	The starting values for key indicators, often with a survey	20 to 60 days	Expert
Monitoring or data review	Checks data quality and how well the system works	5 to 20 days	Doctor
Mid-term evaluation	Is it on track, and what should change?	30 to 80 days	Doctor
Final evaluation	Did it work, for whom, and was it worth it?	40 to 120 days	Expert
Impact evaluation	What changed because of the programme, against a counterfactual	Months, with a team	Expert
Learning facilitation	Reflection workshops, after-action reviews, learning agendas	3 to 15 days	Process partner
Capacity strengthening	Training and coaching staff to run M&E themselves	5 to 30 days	Process partner

Days vary widely with scale and team. Treat these as a starting point for scoping, not a price list.

3 KINDS OF WORK

Two examples, followed throughout.

Both are fictional and followed through every stage. No sector knowledge needed.

SMALL ASSIGNMENT

The MEL system set-up

A 25-person NGO runs 40 after-school learning centres in one district. Its new funder asks for "a proper M&E system". The director emails a one-paragraph request to a solo consultant.

AT THE START

The request: "Set up our M&E system and train staff. Please share your quote." No budget, no timeline, no definition of done.

AT THE END

15 consultant days over eight weeks. A results framework, 12 indicators, three short forms, a monthly dashboard and two staff who can run it. Paid on time; referred to two other NGOs.

LARGE ASSIGNMENT

The mid-term evaluation

A foundation funds a three-year livelihoods programme for 6,000 women in self-help groups across four districts. Halfway through, it issues a ToR for a mid-term evaluation.

AT THE START

The ToR: 14 questions, "impact on income", 30 days, payment on approval, and a budget that covers about 60% of the work it describes.

AT THE END

Five agreed questions, mixed methods, a team of three plus eight field investigators. 9 of 11 recommendations accepted in the management response; invited back for the final evaluation.

Which role are you playing?

You are the expert

Agree the question tightly, deliver the answer well, and check it is still the right question at inception.

You are a process partner

Your job is to build the team's ownership. Plan workshops, not just deliverables, and leave tools they can run.

Expert with a handover

Deliver the technical work, and build in coaching so the client can keep it going.

You are the doctor

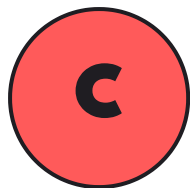
Start with a short diagnosis before you prescribe. Price the diagnosis separately if the scope is unclear.

Ask two questions. Does the client already know exactly what they need? If yes, lean expert. Will the client run this themselves after you leave? If yes, lean process partner.

PART 2 THE METHOD

04

4 · The CONSULT method. The CONSULT method takes any M&E assignment from the first ToR to findings that get used.



THE CONSULT METHOD · STAGE 1 OF 7

Choose the work

Should we bid, and on what terms?

Choose before you chase.

Read the ToR for its real purpose, users and risks. Score the fit: your skills, the client's readiness, the budget against the work, and any conflict of interest. Ask clarifying questions before you bid. Saying no to a badly framed assignment protects your reputation.

THE MEL SET-UP

A 30-minute discovery call showed the real need: the board wants to know whether children read better. The consultant scoped that, and declined to take on "and any other tasks".

Discovery call notes

GO

ASKED	"Set up our M&E system"
REAL NEED	Board wants to know if reading improves
READINESS	Data on paper; one keen staff member
DECISION	Bid, with a tight scope

THE MID-TERM EVALUATION

The ToR scored 6 of 10. The consultant sent four questions in the pre-bid window and learned the budget could stretch by 20%. They decided to bid with a team.

Go or no-go score

CHECK

FIT WITH SKILLS	3 of 3
BUDGET VS WORK	1 of 3
CLIENT READINESS	2 of 3
CONFLICTS	None declared

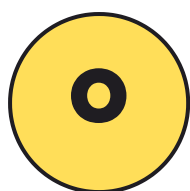
AI can help: Summarise the ToR, list its risks and draft clarifying questions. Check every risk yourself.

The ideas behind this: [Flawless consulting](#) · [The trust equation](#) · [Terms of reference](#) · [Conflict of interest](#)

Your turn

on your next assignment

- ☐ Find the real decision behind the ToR
- ☐ Score fit, readiness, budget and conflicts
- ☐ Ask clarifying questions before you bid
- ☐ Walk away from work set up to fail



THE CONSULT METHOD · STAGE 2 OF 7

Offer

Does the proposal show we understand the problem?

**Lead with the answer you expect,
then show how you will test it.**

A strong proposal restates the client's problem better than they did, offers a first hypothesis, breaks the questions into a tree, and shows a workplan and budget that match. Price the work, not your hope of winning it.

THE MEL SET-UP

A two-page proposal: the problem in one sentence, three deliverables, 15 days, a fixed price and a payment schedule.

Proposal, page 1

PROBLEM	No way to see learning progress monthly
DELIVERABLES	Framework · forms · dashboard · training
DAYS	15 over 8 weeks
PRICE	Fixed, in three payments

THE MID-TERM EVALUATION

The technical proposal grouped 14 questions into five, with a first hypothesis: "Results are strongest where group federations are active." The budget showed days by role and stage.

Technical proposal

QUESTIONS	14 grouped into 5
HYPOTHESIS	Stronger where federations are active
DESIGN	Mixed methods, 4 districts
TEAM	Lead, researcher, 8 investigators

AI can help: Draft an issue tree and a first proposal structure. Write the hypothesis and the price yourself.

The ideas behind this: [Seven-step problem solving](#) · [The pyramid principle](#) · [Hypothesis-driven work](#) · [No overlaps, no gaps](#)

Your turn

on your next assignment

- | | |
|--|--|
| ☐ Restate the problem in one sentence | ☐ Offer a first hypothesis |
| ☐ Break questions into a tree | ☐ Show days by role and stage |



THE CONSULT METHOD · STAGE 3 OF 7

Negotiate the contract

Have we agreed what done looks like, and what happens when things change?

Contract the expectations, not just the price.

Agree the questions, deliverables, review rounds, payment milestones, data ownership, independence of findings and a change process, in writing. Say what you need from the client too: documents, access and timely feedback.

THE MEL SET-UP

Three payment milestones (30, 40, 30), one review round per deliverable, and a one-line change process. The director agreed to name one staff member as the focal point.

Contract terms

PAYMENTS	30% · 40% · 30%
REVIEWS	One round per deliverable
CHANGES	In writing, with days and fees
CLIENT GIVES	A focal point and records

THE MID-TERM EVALUATION

The scope came down to five questions. The contract separated factual comments from findings, set payments on submission rather than approval, and agreed how AI may be used with data.

Contract terms

SCOPE	5 questions, agreed
FINDINGS	The evaluator decides
PAYMENTS	On submission, not approval
AI USE	Only on de-identified data

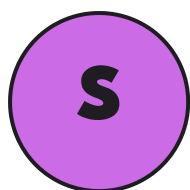
AI can help: Compare the draft contract with your proposal and list gaps. A person reads every clause before signing.

The ideas behind this: [Flawless consulting](#) · [Change requests](#) · [Independence of findings](#) · [AI terms in contracts](#) · [Principled negotiation](#)

Your turn

on your next assignment

- | | |
|--|---|
| ☐ Deliverables with page limits and review rounds | ☐ Payments on milestones, not approval |
| ☐ A written change process | ☐ What you need from the client |



THE CONSULT METHOD · STAGE 4 OF 7

Set up the inception

Does everyone agree on exactly what will be done?

Design for use from day one.

Meet the people who will use the findings. Turn questions into an evaluation matrix, sketch the final report before you collect anything, and settle sampling, tools, ethics, risks and the workplan. The signed inception report becomes your working contract.

THE MEL SET-UP

A half-day workshop with staff agreed 12 indicators. The consultant drew the monthly dashboard on paper before any form was built.

Inception note

USERS	Director, 4 coordinators, board
INDICATORS	12, from 31 proposed
GHOST DASHBOARD	Drawn before any form
SIGNED	Week 2

THE MID-TERM EVALUATION

The inception report had an evaluation matrix, a sample of 400 households, 12 focus groups and 20 interviews, ethics approval, a risk log and a "ghost" report outline. Signed in week three.

Evaluation matrix, row 2

QUESTION	Did incomes change, for whom?
INDICATOR	Monthly income, self-reported
SOURCES	Survey · FGDs · records
ANALYSIS	By district and group age

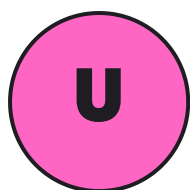
AI can help: Draft the evaluation matrix and a first report outline from the agreed questions. Check every method against the budget.

The ideas behind this: [Utilisation-focused evaluation](#) · [Evaluation matrix](#) · [Ghost report](#) · [RealWorld evaluation](#) · [Evaluative reasoning](#) · [Gender-responsive evaluation](#)

Your turn

on your next assignment

- ☐ Meet the people who will use the findings
- ☐ Sketch the report before collecting data
- ☐ An evaluation matrix for every question
- ☐ Sign off the inception report



THE CONSULT METHOD · STAGE 5 OF 7

Uncover the evidence

Will the evidence stand up when someone challenges it?

Collect only what answers a question, and check it every day.

Pilot tools, train the team, run daily quality checks and triangulate across sources. Keep a decision log. Share emerging findings with the client before the final meeting, so nobody is surprised.

THE MEL SET-UP

Forms were piloted in four centres for two weeks. Two fields were dropped; the dashboard filled itself from the first month of data.

Pilot log

CENTRES	4 for two weeks
FIELDS DROPPED	2 nobody used
DATA CHECKS	Weekly completeness
STATUS	Dashboard live

THE MID-TERM EVALUATION

Daily checks caught one investigator rushing interviews. A mid-fieldwork call shared three emerging findings, so the foundation was ready for the hard one.

Fieldwork tracker

HOUSEHOLDS	400 of 400
FOCUS GROUPS	12 of 12
INTERVIEWS	20 of 20
FLAGS	1 investigator retrained

AI can help: Flag odd patterns in de-identified data and suggest first-pass codes. Humans adjudicate every code and finding.

The ideas behind this: [Triangulation](#) · [Pre-wiring](#) · [AI for qualitative coding](#) · [Fieldwork quality checks](#) · [Contribution analysis](#) · [Realist evaluation](#)

Your turn

on your next assignment

☐

Pilot every tool

☐

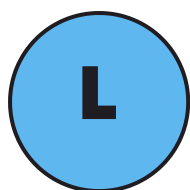
Triangulate across sources

☐

Daily quality checks

☐

Share emerging findings early



THE CONSULT METHOD · STAGE 6 OF 7

Land the findings

Will the people who decide understand and accept the findings?

Answer first, then the evidence, then the detail.

Synthesise by question: what we found, so what, and what to do. Write in a pyramid, with a one-page summary on top. Validate findings with stakeholders, accept factual corrections, and keep the findings your own.

THE MEL SET-UP

A four-page learning note and a working dashboard, walked through with the whole team in one meeting.

Learning note

HEADLINE	Reading levels visible monthly
PAGES	4, with the dashboard
MEETING	Whole team, 90 minutes
COMMENTS	One round, factual only

THE MID-TERM EVALUATION

A 1:3:25 report and a validation workshop. One finding was challenged; the evidence held, and it stayed, with the programme's view noted.

Report, 1:3:25

1 PAGE	Key messages for the board
3 PAGES	Summary by question
25 PAGES	Findings and evidence
WORKSHOP	Findings validated

AI can help: Draft summaries from your approved findings and tighten the language. Never let AI invent a finding or a quote.

The ideas behind this: [The pyramid principle](#) · [1:3:25 reports](#) · [The obligation to dissent](#) · [OECD DAC evaluation criteria](#) · [Evaluation influence](#) · [Liberating Structures](#)

Your turn

on your next assignment

- ☐ Answer first, then evidence
- ☐ Validate with stakeholders

- ☐ A one-page summary on top
- ☐ Keep factual and findings comments apart



THE CONSULT METHOD · STAGE 7 OF 7

Transfer and track use

Did anything change because of our work?

The assignment ends when the findings are used, not when the invoice is paid.

Help the client write a management response with owners and dates. Hand over data, tools and a short guide. Close the contract cleanly, ask for feedback, and check back in a few months.

THE MEL SET-UP

A three-month check-in found staff running the dashboard alone. The director referred the consultant to two other NGOs.

Close-out

GO

HANDOVER	Forms, dashboard, guide
CHECK-IN	Month three: running alone
FEEDBACK	9 of 10
NEXT	Two referrals

THE MID-TERM EVALUATION

The management response accepted 9 of 11 recommendations, each with an owner. The foundation invited the team to bid for the final evaluation.

Management response

GO

ACCEPTED	9 of 11
OWNERS	Named, with dates
FOLLOW-UP	Six-month review
NEXT	Invited to bid again

AI can help: Draft the management response template and a handover guide from the final report.

The ideas behind this: [Management response](#) · [Flawless consulting](#) · [The trust equation](#) · [Tracking use](#) · [Collaborating, learning and adapting](#)

Your turn

on your next assignment

- ☐ A management response with owners and dates
- ☐ Hand over data and tools
- ☐ Ask for feedback and a reference
- ☐ Check back in three months

PART 2 THE METHOD

5 WHY IT WORKS

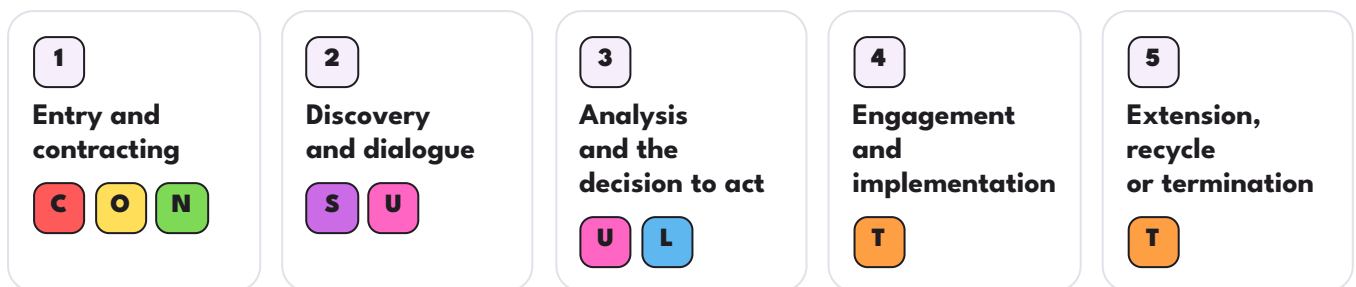
Why it works.

It joins two traditions: how the best consultants solve problems and build trust, and the standards evaluators are held to.



- 1 Consulting is a relationship with phases
- 2 Hypothesis first, then evidence
- 3 Standards protect independence and quality

Consulting is a relationship in five phases PETER BLOCK, FLAWLESS CONSULTING



The trust equation MAISTER, GREEN AND GALFORD, 2000

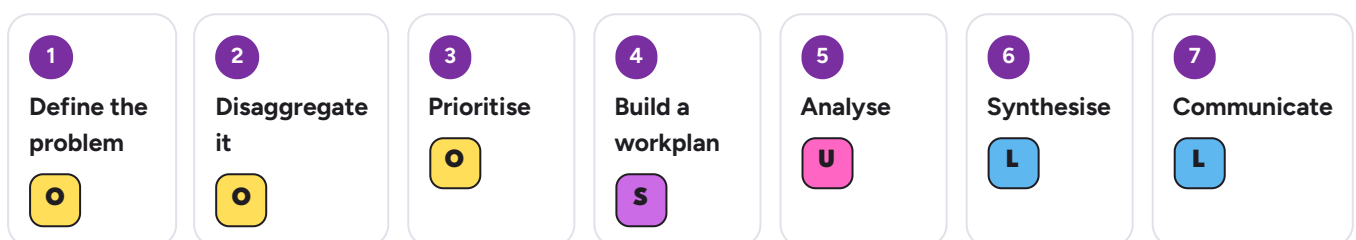
- **Credibility:** you know what you are talking about.
- **Reliability:** you do what you said, when you said.
- **Intimacy:** the client can tell you the hard things.
- **Self-orientation:** how much you focus on yourself, your fee or looking clever. It divides everything else.

**Credibility + Reliability
+ Intimacy**

Self-orientation

At 7, 7 and 6 over 3, trust is 6.7. Raise self-orientation to 7 and it falls to 2.9.

How top consultants solve problems CONN AND MCLEAN, BULLETPROOF PROBLEM SOLVING



Start with a hypothesis, then look for the evidence that would prove it wrong. Break questions into groups with no overlaps and no gaps. Write the answer first, then the reasons, then the evidence.

5 WHY IT WORKS

Evaluation has its own rules.

Consulting craft wins and keeps clients. Evaluation standards keep the evidence honest. You need both.

OECD DAC criteria

Relevance, coherence, effectiveness, efficiency, impact and sustainability, revised in 2019. Use them as a menu, not a checklist.

[OECD DAC evaluation criteria](#)

UNEG ethical guidelines

Integrity, accountability, respect and beneficence, set out in 2020 for anyone evaluating with the UN.

[UNEG ethical guidelines](#)

AEA guiding principles

Systematic inquiry, competence, integrity, respect for people, and common good and equity, updated in 2018.

[AEA guiding principles](#)

Utilisation-focused evaluation

Judge an evaluation by its use by intended users. Find them first and design with them.

[Utilisation-focused evaluation](#)

Four findings that shape the method

Contract the wants on both sides

Say what you need from the client as clearly as what you will deliver. Responsibility is shared.

[Peter Block, Flawless Consulting](#)

Trust is the product

Credibility, reliability and intimacy build trust; putting yourself first destroys it.

[Maister, Green and Galford, 2000](#)

Answer first

Start with the governing thought, then the reasons, then the evidence.

[Barbara Minto, The Pyramid Principle](#)

Design within real constraints

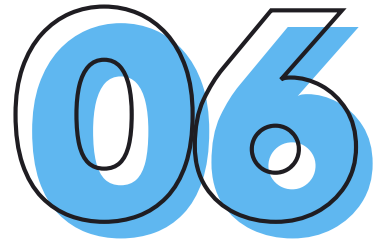
Budget, time, data and politics always limit an evaluation. Name them and design around them.

[Bamberger and Mabry, 2019](#)

Where the tensions sit. The client pays you, and your findings must still be independent. Clients check facts; evaluators decide findings. Put that in the contract before the first draft, not after it.

PART 2 THE METHOD

6 JUDGEMENT AND METHODS



Judgement and methods.

How top evaluators get from evidence to a verdict they can defend, choose a causal design that fits the question, and evaluate programmes that keep changing.

1 A judgement needs criteria and standards, not just data

2 Match the causal design to the question and the budget

3 Complex programmes need evaluation that adapts with them

A judgement needs criteria, standards and evidence, combined in the open. EVALUATIVE REASONING

1

Criteria

Decide what matters: the aspects of performance you will judge.

Income change, sustainability of groups, value for money.

2

Standards

Describe what excellent, good, adequate and poor look like for each criterion, before you see the data.

A rubric agreed with the client at inception.

3

Evidence

Gather and weigh evidence against each standard, from more than one source.

Survey, records, focus groups, cost data.

4

Synthesis

Combine the ratings into an overall judgement, and show your reasoning.

'Good overall: strong where federations exist, weak elsewhere.'

Agree what good looks like before you see the data. A RUBRIC FROM THE MID-TERM EVALUATION

CRITERION	EXCELLENT	GOOD	ADEQUATE	POOR
Income change About 22% and 19% in two districts; flat in two; the poorest members gained least.	Average incomes rose by more than 20% in all districts, including the poorest members.	Rose by more than 15% in at least half the districts, with no group worse off.	Rose in some districts; some groups unchanged.	No clear rise, or some groups worse off.
Sustainability of groups Federated groups meet regularly; groups without federations decline when visits stop.	Most groups meet, save and lend without programme staff.	Most groups linked to federations do so; others still need support.	Only a few groups show signs of independence.	Groups depend on staff visits to function.
Value for money Close to comparable programmes; higher in two remote districts that reach poorer women.	Cost per active member below comparable programmes, and falling.	Comparable to similar programmes, and stable.	Higher, but justified by reaching poorer women.	Higher, with no clear reason.

Highlighted: where the evidence landed.

Judge value for money across four Es, not a single ratio. VALUE FOR MONEY

E

Economy

Buying the right inputs at the right price.

Were trainers and materials procured at fair cost?

E

Efficiency

Turning inputs into outputs well.

Cost per group formed, per training delivered.

E

Effectiveness

Turning outputs into outcomes.

Did trained groups raise incomes?

E

Equity

Reaching those who need it most.

Did the poorest women benefit, even if it cost more?

6 JUDGEMENT AND METHODS

Choose a causal design that fits.

Match the design to the question, the data and the budget. Strong evaluations often combine two.

DESIGN	ANSWERS	NEEDS	WATCH OUT
Randomised controlled trial Experimental	How much difference did the programme make, on average?	Random assignment before the programme starts; large samples; time and budget.	Rarely possible mid-programme; says little about why.
Difference-in-differences Quasi-experimental	How did change compare with a similar group over the same time?	Baseline and follow-up data for both groups.	Assumes both groups would have changed in parallel.
Matching Quasi-experimental	How do participants compare with similar non-participants?	Rich data on both groups to find good matches.	Cannot match on what you did not measure.
Contribution analysis Theory-based	Is it reasonable to say the programme contributed, and how?	A theory of change, evidence for each link, and rival explanations.	Needs honest testing of other explanations.
Process tracing Theory-based	Did the programme cause this specific outcome, through which steps?	A detailed causal story and tests of how strongly evidence confirms each step.	Demanding; best for a few important cases.
Realist evaluation Theory-based	What works, for whom, in what circumstances, and why?	Programme theories of context, mechanism and outcome; mixed data.	Takes skill; findings are patterns, not averages.
Outcome harvesting Participatory and emergent	What changed, including the unexpected, and how did the programme contribute?	People who can describe and verify changes; time to substantiate.	Finds outcomes after the fact; less suited to measuring size.
Most Significant Change Participatory and emergent	Which changes matter most to the people involved, and why?	Stories collected and selected through structured discussion.	A complement to other evidence, not a measure of average effect.

When the programme keeps changing, evaluate alongside it. DEVELOPMENTAL EVALUATION; CLA

	TRADITIONAL EVALUATION	DEVELOPMENTAL EVALUATION
Purpose	Judge results at set points	Support innovation and adaptation as it happens
Timing	Mid-term and end	Continuous, alongside the team
Evaluator's role	Independent, at arm's length	Part of the team, asking evaluative questions in real time
Questions	Did it work?	What is emerging, and what should we try next?
Outputs	Reports and recommendations	Rapid feedback, learning sessions, updated theories of change

COLLABORATING, LEARNING AND ADAPTING

Collaborating. Working with the right partners and sharing what is learned.

Learning. Testing the theory of change and filling knowledge gaps on purpose.

Adapting. Using what is learned to change plans, budgets and activities.

THREE SYSTEMS QUESTIONS

Boundaries. What is inside and outside the evaluation, and who decided?

Perspectives. Whose view of success counts, and whose is missing?

Relationships. Which connections between actors make change happen or block it?

6 JUDGEMENT AND METHODS

Equity, power and quality.

Ask who benefits and who decides. Then check your report the way the commissioner will.

Gender-responsive evaluation

Looks at how a programme affects women, men and gender-diverse people differently, and at the power relations behind it.

UN Women Evaluation Handbook, 2022

Equity-focused evaluation

Asks whether the worst-off groups were reached and benefited, not only the average.

UNICEF and partners

Participatory evaluation

Involves the people affected in setting questions, collecting data and making sense of it.

BetterEvaluation

Locally led evaluation

Puts evaluators and knowledge from the region at the centre, as in the Made in Africa Evaluation movement.

African Evaluation Association and others

FIVE QUESTIONS ABOUT POWER TO ASK AT INCEPTION

1. Who chose the evaluation questions?
2. Whose data is collected, and who owns it?
3. Are respondents paid for their time, or only asked to give it?
4. Who will see the findings first, and who last?
5. Is the team from the region it evaluates?

Check your report against ten quality standards. **OECD DAC; UNEG**

- | | |
|---|---|
| ☐ The purpose, users and use are stated | ☐ Questions match the ToR and the inception report |
| ☐ The design and methods fit the questions, with limits stated | ☐ Every finding traces to evidence |
| ☐ Conclusions follow from findings | ☐ Recommendations are specific, feasible and prioritised |
| ☐ Gender, equity and human rights are addressed | ☐ Ethics, consent and data protection are described |
| ☐ Stakeholder views are represented fairly | ☐ The executive summary stands alone |

Based on the OECD DAC Quality Standards for Development Evaluation (2010) and UNEG guidance on report quality.

PART 3 TOOLS AND TEMPLATES

7 TOP-FIRM TOOLKIT



The top-firm toolkit.

Twenty-one tools that strategy firms, proposal teams and evaluators rely on, explained from scratch, with a template, a worked example and an AI prompt for each.

- 1 Eight habits separate top firms from the rest
- 2 Three tools for every stage of an assignment
- 3 Each tool: what it is, a template, an example, how to use it

Eight habits separate the top firms from the rest. STRATEGY CONSULTING PRACTICE

01 Answer first Lead with the conclusion, then the reasons, then the evidence. Busy clients read the first line. Minto; every top firm	02 Hypothesis-driven Write a best guess on day one and design the work to prove it wrong. It focuses scarce days. McKinsey, BCG	03 No overlaps, no gaps Break every question into parts that are mutually exclusive and together complete. McKinsey (MECE)	04 80/20 Find the few analyses that carry most of the answer. Do those brilliantly; drop the rest. Pareto; all firms
05 So what? Every chart and finding must say what it means for the client's decision. Facts alone are not findings. McKinsey	06 Pre-wire No surprises in the big meeting. Walk key people through hard news one to one first. McKinsey	07 Obligation to dissent Say when you disagree, with the client or your own team lead. Silence is a quality risk. Marvin Bower, McKinsey	08 Results, not reports Judge the work by the decisions it changes, not the pages delivered. Bain; utilisation-focused evaluation

Three proven tools for every stage of an assignment. MAPPED TO CONSULT

C CHOOSE	O OFFER	N NEGOTIATE	S SET UP	U UNCOVER	L LAND	T TRANSFER
Bid or no-bid gate	Issue tree and day-one hypothesis	Statement of work and change control	Kick-off and workplan	80/20 analysis plan	Storyline and action titles	Management response tracker
Problem statement worksheet	Three-option proposal	RAPID decision roles	RAID log	Weekly status note	Pre-wiring and validation	After-action review
Stakeholder power and interest map	Colour team reviews	Engagement charter	Ghost deck and report	Evidence strength rating	Impact and ease prioritisation	Client feedback score

7 TOP-FIRM TOOLKIT

Tools 1 to 7.

Choose and offer: frame the problem, map the people, design the bid.



SHIPLEY CAPTURE MANAGEMENT

Bid or no-bid gate

A short scored review before anyone writes a proposal.

In the examples: The mid-term evaluation scored 12 of 18: go, with conditions. The conditions became the clarifying questions.

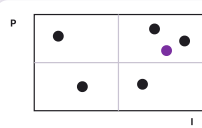


MCKINSEY

Problem statement worksheet

One page that fixes the key question, the decision it serves and what success looks like, before any method is chosen.

In the examples: The MEL set-up's request was 'set up our M&E'. The worksheet turned it into: 'How can the board see each month whether children's reading improves, within 15 days and existing staff?'

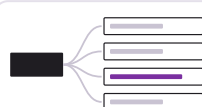


MENDELOW, 1991

Stakeholder power and interest map

A two-by-two that places people by their power over the work and their interest in it.

In the examples: In the mid-term evaluation, the foundation's board had high power and low interest: keep satisfied with a one-page brief. Federation leaders had high interest: involve them in validation.

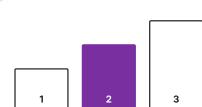


MCKINSEY AND BCG

Issue tree and day-one hypothesis

Break the key question into sub-questions with no overlaps or gaps, and write your best early guess at the answer.

In the examples: The 14 ToR questions became a tree of five branches. Day-one hypothesis: 'Income gains follow strong federations.' The design was built to test it.

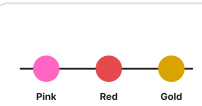


ALAN WEISS, VALUE-BASED FEES

Three-option proposal

Offer the client a choice of yes: three scopes at three prices, each tied to the value it delivers.

In the examples: The MEL set-up offered: essentials (framework and forms), standard (plus dashboard and training) and full (plus three months of coaching). The client chose standard.



SHIPLEY ASSOCIATES

Colour team reviews

Planned reviews of a proposal at set points: pink for the plan, red for a full draft read as the client, gold for final sign-off.

In the examples: For the mid-term evaluation bid, a colleague did a 45-minute red team as the foundation's panel and caught that the budget page did not match the workplan.



PROJECT MANAGEMENT PRACTICE

Statement of work and change control

A written scope with deliverables, acceptance rules and a change process, so that extra work becomes a decision, not a favour.

In the examples: When the foundation asked for two extra districts after inception, a one-page change request added 6 days and a revised fee. It was signed in a week.

Step-by-step guides and AI prompts for every tool are in the web version.

7 TOP-FIRM TOOLKIT

Tools 8 to 14.

Negotiate and set up: contract the work, decide who decides, plan it.



BAIN & COMPANY

RAPID decision roles

Name who recommends, who must agree, who performs, who gives input and who decides, for each key decision.

In the examples: For the management response, the programme head recommends, finance must agree, district teams perform, federations give input and the foundation's director decides.



PROJECT MANAGEMENT PRACTICE

Engagement charter

One page that both sides sign: objective, scope, deliverables, roles, timeline, risks and how you will work together.

In the examples: The MEL set-up's charter fitted on one page and named the two staff who would own the system after handover.



PROJECT MANAGEMENT PRACTICE

Kick-off and workplan

A first meeting that confirms the plan, and a workplan that shows stages, gates and client review time.

In the examples: The mid-term evaluation's kick-off agreed the five questions, the sample, review rounds of 10 working days and a steering group meeting at each gate.



PROJECT MANAGEMENT PRACTICE

RAID log

One living list of risks, assumptions, issues and dependencies, each with an owner and a date.

In the examples: Top risk: monsoon access to two districts. Assumption: programme records are complete. Dependency: ethics approval before fieldwork.

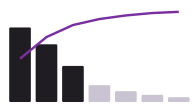


MCKINSEY

Ghost deck and report

Sketch the final report's headlines and charts before you collect data, then fill them in.

In the examples: Before fieldwork, the team wrote 12 headline placeholders, such as 'Incomes rose most where federations are strong', and designed tools to confirm or kill each one.



PARETO PRINCIPLE; ALL FIRMS

80/20 analysis plan

Rank your analyses by how much each could change the answer, and do the top few first.

In the examples: Three analyses carried most of the mid-term evaluation: income by district, income by federation strength, and cost per active member.



CONSULTING PRACTICE

Weekly status note

A five-part note each week: overall status in red, amber or green, progress, next steps, risks, and decisions needed.

In the examples: Week 6 of the mid-term evaluation was amber: two districts delayed by rain. The note asked the client to decide between a later report and a smaller sample.

Step-by-step guides and AI prompts for every tool are in the web version.

7 TOP-FIRM TOOLKIT

Tools 15 to 21.

Uncover, land and transfer: focus the analysis, tell the story, learn.

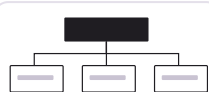


U EVALUATIVE RUBRICS (DAVIDSON); CONSULTING HARVEY BALLS

Evidence strength rating

Rate each finding by how strong its evidence is, so readers can see what is solid and what is tentative.

In the examples: 'Incomes rose in two districts' was rated strong: survey plus records plus focus groups. 'Federations cause the gains' was rated moderate.



L BARBARA MINTO; MCKINSEY

Storyline and action titles

A storyline where every page title is a full-sentence finding. Read only the titles and the whole argument still makes sense.

In the examples: The mid-term evaluation's title read: 'Results follow strong federations; invest in them before expanding.' Each supporting title proved one part of it.

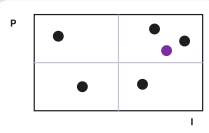


L MCKINSEY

Pre-wiring and validation

Walk the people most affected through hard findings one to one, then hold a validation workshop focused on decisions.

In the examples: Before the findings workshop, the evaluator met the district manager whose area showed no gains. In the workshop, she proposed the fix herself.



L CONSULTING PRACTICE

Impact and ease prioritisation

A two-by-two that sorts recommendations by their impact and how easy they are to do.

In the examples: Of 11 recommendations, three were quick wins, four were major projects and two were dropped as low impact.



T UNEG; EVALUATION PRACTICE

Management response tracker

A table where the client accepts, partly accepts or rejects each recommendation, with an owner, action and date.

In the examples: The foundation accepted 9 of 11 recommendations. At the three-month check-in, six actions were on track.

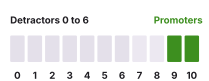


T US ARMY

After-action review

A short team review of four questions: what was planned, what happened, why, and what to do next time.

In the examples: The MEL set-up's review found training took twice as long as planned. The next proposal priced it at four days, not two.



T BAIN (NET PROMOTER, REICHHELD 2003)

Client feedback score

Ask one question at the end of each assignment: how likely is the client to recommend you, from 0 to 10, and why?

In the examples: The MEL set-up's director scored 9 and named two NGOs to refer. A 6 from another client led to a call that saved the relationship.

Step-by-step guides and AI prompts for every tool are in the web version.

PART 3 TOOLS AND TEMPLATES

8 ENGAGEMENT STUDIO



The engagement studio.

Six working tools from top-firm practice. The web version fills and checks them for you; here they are worksheets to print.

1 Frame the problem before you price it

2 Make the logic visible: tree, map, storyline

3 Report status the way steering committees read it

Problem statement worksheet MCKINSEY

Key question

One specific, time-bound question. What must be answered?

Example: How can the board see each month whether children's reading improves, within 15 days of consultant time?

Context

What is happening now that makes this matter?

New funder wants evidence; data is on paper; one keen staff member.

Criteria for success

How will the client know it worked?

Board reviews a reading dashboard every month; staff run it without help.

Scope

What is in, and what is out?

*In: framework, forms, dashboard, training.
Out: an evaluation of impact.*

Decision-makers and stakeholders

Who decides, who uses, who could block?

Director decides; board uses; centre staff collect.

Constraints

Time, money, data, politics.

15 days, eight weeks, two staff, low-end phones.

Sources of insight

Who and what already knows something?

Funder's template, staff, two similar NGOs, current registers.

RAPID decision roles BAIN & COMPANY

Decision	Programme director	Finance head	District managers	Federation leaders	Foundation director
Which recommendations to accept	R	A	P	I	D
Whether to expand to new districts	R	A	P	I	D

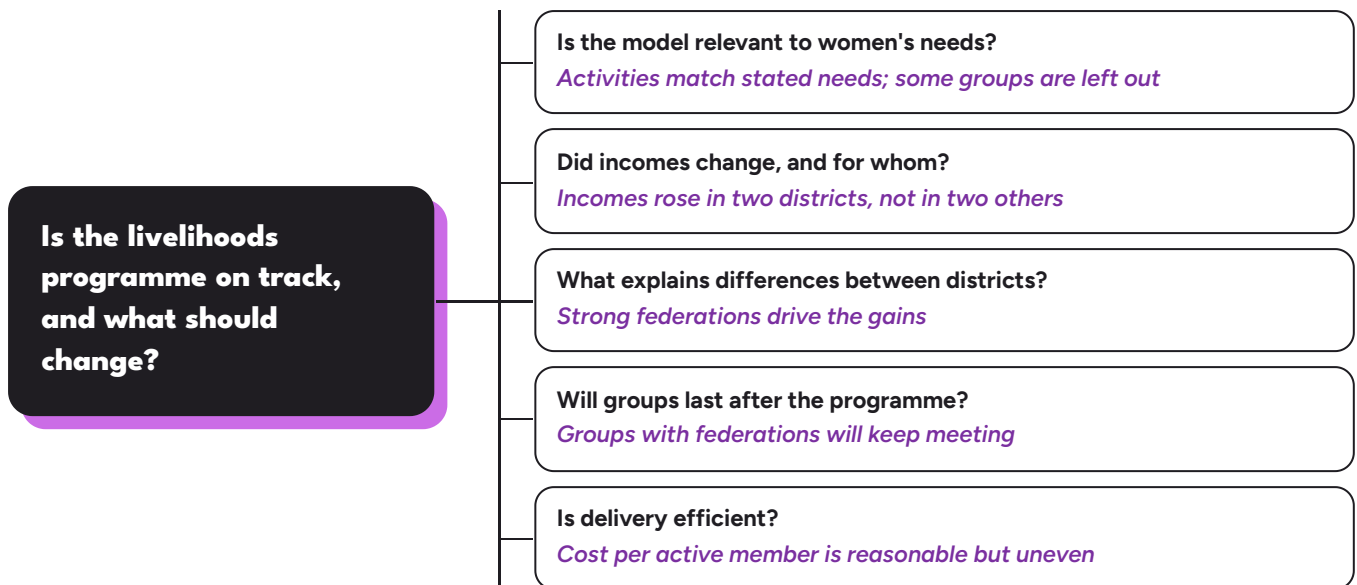
R Recommend: Makes the proposal, gathers input · A Agree: Must sign off; can block · P Perform: Carries out the decision · I Input: Consulted for facts and views · D Decide: Makes the final call, one person. One D per decision.

8 ENGAGEMENT STUDIO

Make the logic visible.

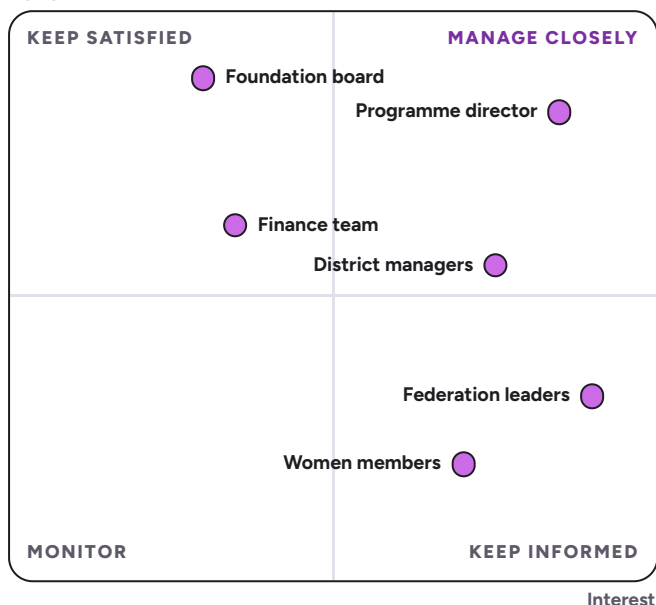
An issue tree for the questions, a map for the people, and action titles for the story.

Issue tree, with a hypothesis on every branch THE MID-TERM EVALUATION



Stakeholder power and interest MENDELOW, 1991

Power



A storyline the titles alone can tell MINTO; MCKINSEY

GOVERNING MESSAGE

Results follow strong federations; invest in them before expanding.

PAGE 2

Incomes rose by about a fifth in two districts and were flat in two others

PAGE 3

The gains track federation strength more closely than any other factor

PAGE 4

Groups linked to strong federations are the most likely to last

PAGE 5

Strengthening federations costs less per member than opening new districts

Test: read only the titles aloud. Topics such as "Background" or "Findings" fail.

8 ENGAGEMENT STUDIO

Report status the way steering committees read it.

One colour, five lines, the same day every week. Keep the RAID log beside it.

Weekly status note WEEK 6 OF THE MID-TERM EVALUATION

AMBER

Mid-term evaluation • weekly status At risk: action under way

WHY THIS COLOUR	Fieldwork in two districts is delayed by rain; analysis is on schedule elsewhere.
DONE THIS WEEK	Household survey complete in two districts (210 of 400); 8 of 12 focus groups held.
NEXT	Finish two districts by 14 November; first analysis readout on 21 November.
RISKS AND ISSUES	If the delay passes two weeks, the draft report moves by one week.
DECISION NEEDED	Choose: accept a one-week later draft, or a smaller sample in the delayed districts.

RAID log RISKS, ASSUMPTIONS, ISSUES, DEPENDENCIES

TYPE	ENTRY	OWNER	NEXT ACTION
Risk	Monsoon cuts access to two districts	Team lead	Plan fieldwork there first; keep a phone survey fallback
Assumption	Programme records are complete for 2024 and 2025	Client M&E officer	Check a 5% sample in week 2
Issue	Federation list is out of date	District managers	Updated list by 20 October
Dependency	Ethics approval before fieldwork	Evaluator	Submitted; decision due 25 October

Red off track: needs a decision now

Amber at risk: action under way

Green on track



The field guide.

The templates and rules to keep open on every assignment.

1 Read a ToR in ten minutes

2 Proposal, budget and contract anatomy

3 Inception, matrix and report structures

Read a ToR: twelve questions to ask first

1. Who will use the findings, for which decision?
2. Which questions matter most, and can they be answered in the time?
3. Is it monitoring, review, evaluation or impact?
4. What data already exists: baseline, monitoring, reports?
5. Is the budget stated, and does it match the work?
6. What are the deliverables, page limits and review rounds?
7. How is payment scheduled?
8. Who owns the data and the report?
9. Who approves what, and is independence protected?
10. Are ethics approval or government permission needed?
11. Any past involvement that could be a conflict?
12. What happens if scope changes?

A proposal that wins

Understanding	The problem in your words, the decision it serves, and what you heard between the lines.
Approach	A first hypothesis, the question tree, design and methods, and their limits.
Workplan	Stages, deliverables and dates, with client review time built in.
Team	Who does what, with days, and why they fit.
Budget	Days by role and stage, expenses, contingency and taxes, and the payment schedule.
Risks	What could go wrong and what you will do about it.
Credentials	Two or three similar assignments, with what changed because of them.

9 THE FIELD GUIDE

Contract it, then plan it.

Nine clauses protect the work. The inception report turns them into a plan everyone signs.

Nine clauses to contract

Scope	The agreed questions and what is excluded.
Deliverables	Each one with format, length and review rounds.
Payments	Milestones tied to submission, with an advance.
Change process	Extra work is agreed in writing with days and fees.
Independence	The client checks facts; the evaluator decides findings.
Data	Who owns, stores and deletes data, and how it is protected.
AI use	What tools may be used, on which data, and how it is disclosed.
Client duties	Documents, access, a focal point and feedback times.
Exit	How either side can end the contract, and what is paid.

Ten parts of an inception report

1. Purpose, users and use
2. Refined evaluation questions
3. Evaluation matrix
4. Design, methods and sampling
5. Tools, as annexes
6. Ethics and data protection
7. Workplan and deliverables
8. Risks and how you will handle them
9. Report outline
10. What you need from the client

The evaluation matrix, from the mid-term evaluation

EVALUATION QUESTION	SUB-QUESTIONS	INDICATORS OR JUDGEMENT CRITERIA	DATA SOURCES	METHODS	ANALYSIS
Is the model relevant to women's needs and priorities?	Do activities match what women say they need? Who is left out?	Share of women rating activities as useful; groups excluded	Survey; FGDs; programme records	Household survey; 12 FGDs	Compare by district and group age
Did incomes change, and for whom?	How much, in which districts, for which groups?	Change in monthly income, self-reported, against baseline	Baseline data; survey; FGDs; records	Survey of 400 households; case studies	Before and after by district; triangulate with FGDs
What explains differences between districts?	Role of federations, markets and staff?	Federation strength score; market access; staff turnover	KIIs; records; FGDs	20 KIIs; document review	Compare strong and weak federation areas
Are groups likely to last after the programme?	Can groups run without staff? What would they need?	Groups meeting and saving without staff support	Group records; FGDs; KIIs	Group record review; FGDs	Score groups on a sustainability rubric
Is delivery efficient?	Cost per active member; delays?	Cost per active member; share of activities on time	Budget and expenditure data; KIIs	Cost analysis; KIIs with finance staff	Compare costs across districts

9 THE FIELD GUIDE

Write it so it gets used.

Answer first, in layers, and deliver hard news early and in private.

One page, three pages, twenty-five pages THE 1:3:25 RULE

1 page

Key messages:
the answer to
each question,
and what to do.

3 pages

An executive summary
by question, with the
main evidence.

25 pages

Findings by question, methods, limits and
recommendations. Annexes hold the rest.

Situation, complication, question, answer BARBARA MINTO, THE PYRAMID PRINCIPLE

SITUATION

The programme
reached 6,000 women
in two years.

COMPLICATION

Incomes rose in two
districts but not in the
other two.

QUESTION

What explains the
difference, and what
should change?

ANSWER

Results follow strong
group federations.
Invest in federations
before expanding.

Delivering hard findings

Tell them early	No surprises in the final meeting. Pre-wire hard findings one to one.
Separate fact from finding	Accept factual corrections; keep the judgement yours.
Name the reaction	If someone resists, say what you notice, calmly, and listen.
Offer a way forward	A finding with a feasible next step is easier to accept.

Six ethics habits

1. Declare conflicts of interest before you bid
2. Get consent, and protect respondents' identities
3. Report what you found, including what nobody wanted to hear
4. Do not promise results you cannot measure
5. Say no to work you are not competent to do
6. Credit others' work, and disclose AI use



PART 3 TOOLS AND TEMPLATES

10 WITH AI

Consulting with AI.

AI makes a good consultant faster and a careless one dangerous.
Client trust is the asset you protect.

1 Where AI saves real days

2 Six ways it can cost you a client

3 A pact you can put in a contract

0.7+

A 2025 study tested generative models on coding news text. Without training examples, they matched human coders about as well as humans match each other (F1 around 0.70 to 0.83), but results varied by task and needed human validation.

Use AI as a second coder and a fast drafter. Keep humans responsible for findings. Than and colleagues, 2025

AI is good at

- Summarising a long ToR and listing its risks
- Drafting an issue tree and proposal outline
- Turning questions into a first evaluation matrix
- First-pass coding of de-identified transcripts
- Tightening a draft report and summaries
- Drafting templates: contracts, trackers, guides
- Checking a draft against the ToR and the matrix

It cannot do

- Knowing the politics behind the ToR
- Choosing the hypothesis worth testing
- Judging what the budget can really buy
- Deciding what a pattern means for the client
- Owning a finding in a difficult meeting
- Replacing a lawyer or tax adviser
- Seeing client data safely, unless the tool and contract allow it

The rhythm

You frame

The question, the client and the limits

AI drafts

Trees, matrices, outlines or summaries

You judge

Against the evidence and the client's reality

AI revises

And lists its assumptions

Client reviews

And you decide the findings

Brief it in six lines

1. The client, sector and assignment type
2. The decision the findings must serve
3. The ToR, or the questions, pasted in
4. Your constraints: days, budget, data, politics
5. What you want back, and in what format
6. What the AI must not do: invent sources, quotes or numbers

10 WITH AI

Watch two sessions, avoid six mistakes.

In both, the consultant frames, the AI drafts and the consultant judges.

THE MEL SET-UP · SCOPING

YOU

Here is a one-paragraph request from an NGO: "Set up our M&E system and train staff." List what I need to know before I can scope this, as questions for a 30-minute discovery call.

AI

Ten questions, grouped: who will use the data and for which decisions; what is collected now and how; who will run the system after you leave; budget and timeline; funder requirements. Ask first: "What would be different in a year if this worked?"

YOU

Good. Which three answers would make me decline or rescope the work?

THE MID-TERM EVALUATION · QUESTIONS

YOU

Here are 14 evaluation questions from a mid-term evaluation ToR (pasted, no names). Group them into no more than five key questions, mutually exclusive, and show which originals each covers.

AI

Five groups: relevance of the model; changes in income and assets; what drives differences between districts; sustainability of groups; efficiency of delivery. Two originals overlap, and one asks for impact, which needs a comparison the ToR does not fund.

YOU

Draft the clarifying question I should send about that impact question, politely.

1

Client data in the wrong tool

Pasting transcripts, names or unpublished reports into a tool the contract does not allow.

Agree AI use in the contract. Use approved tools and de-identified data only.

2

Invented sources

Confident references to reports, standards or statistics that do not exist.

Open every source yourself. No citation goes in unchecked.

3

Proposals that read like everyone's

Generic, polished text that shows no understanding of this client.

Write the problem statement and hypothesis yourself, from the discovery call.

4

Over-trusted coding

AI themes accepted without reading the transcripts.

Code a sample by hand, compare, and adjudicate every disagreement.

5

Undisclosed use

The client finds out later that AI drafted parts of the report.

Say in the method note what AI did and who checked it.

6

Hours saved, value lost

Racing to deliver faster and skipping the thinking the client paid for.

Spend the time AI saves on more stakeholder contact and sharper synthesis.

10 WITH AI

Eight roles, six promises.

Give AI a job with an edge. Then make promises your client can hold you to.

1 ToR reader

Summarises and lists risks and gaps. *You still own: deciding whether to bid.*

2 Tree builder

Drafts issue trees and question groupings. *You still own: the hypothesis.*

3 Matrix drafter

Turns questions into a first matrix. *You still own: matching methods to budget.*

4 Tool drafter

First drafts of interview guides and surveys. *You still own: testing with people.*

5 Second coder

First-pass codes on de-identified text. *You still own: adjudicating every code.*

6 Editor

Tightens drafts and summaries. *You still own: every finding and number.*

7 Template maker

Contracts, trackers, guides, invoices. *You still own: legal and tax advice.*

8 Checker

Compares drafts against the ToR and matrix. *You still own: the final sign-off.*

The six promises PUT THEM IN YOUR PROPOSAL

1 Client data stays where agreed

Only approved tools, only de-identified data, as the contract says.

2 Every source is checked

No reference, quote or number goes in unopened.

3 AI use is disclosed

The method note says what AI did and who checked it.

4 Humans own the findings

AI drafts and suggests; the evaluator decides and signs.

5 AI terms are in the contract

Permitted uses, data rules and disclosure, agreed up front.

6 Saved time goes to the client

More listening, sharper synthesis, or a lower price.

Guidance: [Ethical AI in UN evaluations](#) · [AI terms in contracts](#) · [AI for qualitative coding](#)

10 WITH AI

Prompt starters, by stage.

Remove names and confidential details first. Two per stage here, three in the web version. Each ends with a question to push back on the answer.

C Choose

Read the ToR

Summarise this ToR in ten lines: purpose, users, questions, deliverables, timeline, budget, payment terms. Then list every risk and gap: [paste ToR, no names].

Then ask: Which of these risks would make an experienced evaluator decline?

Clarifying questions

Draft five polite clarifying questions for the pre-bid window on these gaps: [gaps].

Then ask: Which question is most likely to change the budget?

N Negotiate

Gap check

Compare this draft contract with our proposal and list differences in scope, deliverables, payment and review rounds: [both, no names].

Then ask: Which difference costs us the most days?

Change clause

Draft a short, plain change-request clause for a consulting contract.

Then ask: How could the client read this clause against us?

U Uncover

Interview guide

Draft a 45-minute key informant interview guide for [role], covering these questions: [questions].

Then ask: Which question leads the respondent?

Daily checks

List daily quality checks for a household survey of [n] interviews with [k] investigators.

Then ask: Which check catches invented interviews?

T Transfer

Management response

Draft a management response table: recommendation, response (accept, adapt, defer, decline), action, owner, date: [recommendations].

Then ask: Which recommendation is likely to be quietly dropped?

Handover guide

Draft a two-page guide for staff to keep this monitoring system running: [system summary].

Then ask: What will break first when I leave?

O Offer

Issue tree

Turn these evaluation questions into an issue tree with no overlaps and no gaps, at most five branches: [questions].

Then ask: Which branch would a sceptical funder push on first?

First hypothesis

Given this programme summary, list three plausible hypotheses about what the evaluation will find, and the evidence that would disprove each: [summary].

Then ask: Which hypothesis would be most useful to the client if true?

S Set up

Evaluation matrix

Build an evaluation matrix for these questions with columns: question, sub-questions, indicators or judgement criteria, sources, methods, analysis: [questions].

Then ask: Which row is weakest given [days] days of fieldwork?

Ghost report

Draft a report outline with a headline finding placeholder for each question, before any data: [questions].

Then ask: Which placeholder can our data not fill?

L Land

Pyramid

Rewrite these findings as a pyramid: one governing message, three supporting points, evidence under each: [approved findings].

Then ask: Which supporting point is weakest?

One-page summary

Draft a one-page summary for a board from these approved findings and recommendations: [text].

Then ask: What would a board member ask first?



PART 3 TOOLS AND TEMPLATES

11 BID DESK

The bid desk.

The web version runs these as live tools. On paper, use them as worksheets for your next bid.

1 Flag risks in any ToR

2 Estimate days and a budget

3 Export an evaluation matrix

ToR scanner: phrases that signal risk

RISK	WHAT IT LOOKS LIKE	WHAT TO PROPOSE
Open-ended scope	"any other tasks assigned"	Ask for extra tasks to be agreed in writing with days and fees.
Impact claim	"assess the impact of"	Check there is a baseline or comparison, and budget for it. Otherwise propose contribution.
Payment on approval	"100% on approval of the final report"	Propose milestones: advance, inception, final submission.
Approval of findings	"findings subject to approval"	Separate factual review from the evaluator's judgement.
Vague deliverable	"a comprehensive, high-quality report"	Ask for page limits, structure and review rounds.
Staff collect the data	"project staff will collect data"	Independent enumerators, or a clear plan to limit bias.
Possible conflict of interest	"you designed this programme"	Declare any past involvement.
Very wide coverage	"all districts and all beneficiaries"	Check days and budget allow it, or propose a sample.
Tight timeline	"within two weeks of signing"	Show a realistic plan with inception and review time.
Ownership terms	"all data and rights vest in the client"	Agree data ownership, storage and deletion; keep the right to use anonymised learning.

Go or no-go SCORE 1 TO 3

Fit with your skills and sector	1	2	3
Budget matches the work	1	2	3
Client knows how findings will be used	1	2	3
Timeline allows a proper inception	1	2	3
No conflict of interest	1	2	3
Payment terms you can live with	1	2	3

15 or more: go. 11 to 14: go, with conditions. Below 11, or a conflict of interest: walk away.

Day-rate floor

$$(\text{income} + \text{yearly costs}) \times (1 + \text{buffer}) \div \text{billable days}$$

Example: $(\text{₹}12,00,000 + \text{₹}2,00,000) \times 1.15 \div (230 \text{ days} \times 50\%) = \text{₹}14,000 \text{ a day}$, rounded up to the nearest ₹500.

Billing more than about two thirds of your days leaves little time to find the next client.

11 THE WORKSHOP

Price the proposal.

Worked example: the mid-term evaluation. Rates are illustrative; set your own.

STAGE	LEAD DAYS	TEAM DAYS	FEES
Desk review and inception	6	3	₹114,000
Tools and ethics	4	3	₹84,000
Fieldwork	12	12	₹276,000
Analysis	8	8	₹184,000
Report and validation workshop	7	3	₹129,000
Management response support	1	0	₹15,000
Total days and fees	38	29	₹802,000
Travel and field costs			₹200,000
Contingency, 10%			₹100,200
Tax, 18%, if registered			₹198,396
Total			₹1,300,596

Lead at ₹15,000 a day, team at ₹8,000 a day. Fieldwork takes about a third of the fees, so it is the first place a short budget hurts.

A payment schedule to propose

On signing

₹390,179

30% · covers set-up and early travel

Inception report

₹520,238

40% · paid on submission

Final report

₹390,179

30% · paid on submission,
not approval

Your own estimate

STAGE	LEAD DAYS	TEAM DAYS	FEES



The kit: calculators, templates and concepts.

Everything you need to work on your own: six calculators for the numbers clients ask about, nine document templates you can download and fill, and the evaluation concepts explained in plain words.

1 Calculate samples, fieldwork, timelines and invoices yourself

2 Start every document from a tested template

3 Understand the concepts before a client asks

Six calculators for the numbers clients ask about. [LIVE IN THE WEB VERSION](#)

CALCULATOR	HOW IT WORKS	WORKED EXAMPLE
Sample size How many people to survey	$n = z^2 \times p(1 - p) \div e^2$. Then multiply by the design effect, adjust for small populations, and divide by $(1 - \text{non-response})$. Two groups: n per group = $(z + z \text{ power})^2 \times [p1(1 - p1) + p2(1 - p2)] \div (p1 - p2)^2$.	$p = 50\%$, $e = 5$ points, 95% ($z = 1.96$): 384. Design effect 1.5 and 10% non-response: 641.
Fieldwork Enumerators, days and cost	Interviews per enumerator per day = (working minutes – travel) $\div$ (interview + gap). Field days = total interviews $\div$ (per day $\times$ enumerators). Add 1 supervisor per 5 enumerators and 10% for back-checks.	400 interviews of 45 minutes, 7-hour days: 5 a day each; 8 enumerators need 10 field days.
Timeline Work back from the deadline	Start from the final report date and subtract, in working days: each review round plus revision, drafting, analysis, fieldwork, tools and ethics, and inception.	Report due 26 January 2027, one review round: start by 26 October 2026.
Qualitative plan Interviews, groups and hours	Focus groups = groups $\times 2$ or 3 (to saturation). Audio hours $\times 4$ to 6 for transcription, $\times 3$ for coding and analysis.	8 focus groups and 16 interviews = 24 audio hours: about 120 hours to transcribe.
Invoice Fee, GST and TDS	Invoice = fee + expenses + GST (if registered). Paid to you = invoice – TDS on the fee. Not tax advice.	₹3,00,000 fee + 18% GST = ₹3,54,000; less 10% TDS = ₹3,24,000 received.
Cost per outcome Value for money in numbers	Cost per person reached = cost $\div$ reach. Cost per outcome = cost $\div$ people achieving the outcome. Compare with a benchmark, then judge with a rubric.	₹3 crore, 2,400 outcomes: ₹12,500 each; 0.83 $\times$ a ₹15,000 benchmark.

Nine document templates. [IN THE TEMPLATE PACK, PART 2](#)

Technical proposal

Responding to a ToR or a client request

Cover letter, Understanding of the assignment, Approach and methodology, Ethics and data protection, ...

Inception report

Two to four weeks after signing, before data collection

Introduction and purpose, Programme background and theory of change, Evaluation questions, Evaluation matrix, ...

Evaluation report

At the end of the evaluation

Key messages (1 page), Executive summary (3 pages), Introduction, Programme and context, ...

Invoice

When a payment milestone is reached

Your details, Client details, Invoice number and date, Description, ...

Informed consent script

Before every interview, survey or focus group

Introduction, Purpose, What taking part involves, Voluntary, ...

Key informant interview guide

Interviewing staff, officials, partners or experts

Opening, Topic 1, linked to an evaluation question, Topic 2, Topic 3, ...

Focus group discussion guide

Group discussions with 6 to 10 people from similar backgrounds

Before the group, Opening and ground rules, Warm-up question, Core questions, ...

Data management plan

At inception, for every evaluation that collects data

Data to be collected, Consent and legal basis, Storage and security, De-identification, ...

Key-expert CV (donor format)

Bidding for donor, UN or firm-led assignments

Proposed position, Personal details, Education, Key qualifications, ...

Concepts explained.

The evaluation ideas clients, reviewers and ToRs assume you already know. Each with an example and the common mistake.

Outputs, outcomes and impact

Outputs are what a programme delivers (trainings held). Outcomes are the changes that follow (women start saving). Impact is the longer-term change (household income rises).

Example. A training programme reports 500 women trained (output), 300 now save monthly (outcome), and incomes rise after two years (impact).

Common mistake. Reporting outputs as if they were outcomes.

Attribution and contribution

Attribution asks how much of a change was caused by the programme. Contribution asks whether the programme plausibly helped cause it, alongside other factors.

Example. Incomes rose, but so did crop prices. Attribution needs a comparison group; contribution builds a case from evidence and rules out other explanations.

Common mistake. Claiming attribution without a comparison group.

Counterfactual

What would have happened without the programme. You cannot observe it directly, so you estimate it, usually with a comparison group.

Example. Similar villages without the programme saw incomes rise 8%; programme villages rose 20%. The counterfactual is about 8%.

Common mistake. Treating 'before and after' as if it were 'with and without'.

Baseline, midline and endline

Measurements at the start, middle and end of a programme, using the same indicators and methods so you can compare.

Example. Savings rates measured in 2024 (baseline), 2025 (midline) and 2027 (endline).

Common mistake. Changing the questions between rounds, so results cannot be compared.

SMART indicators

Indicators that are specific, measurable, achievable, relevant and time-bound.

Example. 'Share of group members saving at least monthly, by district, measured annually' instead of 'improved savings'.

Common mistake. Indicators no one can measure with the data available.

Disaggregation

Breaking results down by groups such as gender, age, district, disability or income.

Example. Average income rose 15%, but for the poorest fifth it rose only 3%.

Common mistake. Reporting only the average, which can hide who was left out.

Sampling

Choosing who to ask. Random sampling lets you generalise; stratified sampling guarantees groups are covered; cluster sampling saves travel; purposive sampling picks information-rich cases for qualitative work.

Example. Randomly choose 20 villages, then 20 households in each (a two-stage cluster sample).

Common mistake. Generalising from a purposive or convenience sample.

More concepts explained.

The evaluation ideas clients, reviewers and ToRs assume you already know. Each with an example and the common mistake.

Margin of error and confidence

A survey result is an estimate. 'Plus or minus 5% at 95% confidence' means that if you repeated the survey many times, 95% of the results would fall within 5 points of the true value.

Example. 62% save monthly, plus or minus 5 points: the true value is likely between 57% and 67%.

Common mistake. Comparing two groups whose difference is smaller than the margin of error.

Statistical and practical significance

Statistically significant means a difference is unlikely to be chance. Practically significant means it is big enough to matter for decisions.

Example. A 1% income difference can be statistically significant in a huge sample but too small to matter.

Common mistake. Treating 'significant' as 'important'.

Validity and reliability

Validity: are you measuring what you intend to? Reliability: would you get the same answer if you measured again in the same way?

Example. A question about 'income' that respondents interpret as 'salary' misses farm income (a validity problem).

Common mistake. Testing tools only for reliability, not for meaning.

Bias

Systematic error. Common types: selection bias (who you reach), response bias (telling you what you want to hear), recall bias (forgetting), and interviewer bias.

Example. Programme staff collecting data may get more positive answers (courtesy bias).

Common mistake. Assuming a large sample removes bias. It does not.

Triangulation

Checking a finding from more than one source, method or person before you call it a finding.

Example. A survey shows higher savings; records confirm deposits; focus groups explain why.

Common mistake. Calling a single quote a finding.

Saturation

In qualitative work, the point where new interviews add no new themes. It guides how many interviews you need.

Example. After 14 interviews with group leaders, the last three added nothing new.

Common mistake. Stopping at a fixed number without checking for new themes.

Theory of change and assumptions

The steps from activities to impact, and the assumptions that must hold for each step to work.

Example. Training leads to saving only if women have some cash income. That is an assumption to test.

Common mistake. Drawing the boxes and never testing the arrows.

PART 3 TOOLS AND TEMPLATES

13 BUILD YOUR NAME



Build your name: be found, trusted and remembered.

Six tools to turn good work into a name clients can find and describe: your positioning, your profile, a case study, a content plan, your relationships and the events you attend.

1 Say in one sentence who you help and how

2 Show proof clients can check, with consent

3 Grow relationships, not just a following

Visible does not mean loud. AN HONEST BRAND

In evaluation, your reputation is your independence and your judgement. Being visible matters, but in this field trust comes first. These tools help you become known through writing, speaking, teaching and giving back, within the same ethics you apply to your evaluations.

Six promises to clients and peers

1 Claim only what you did
Name your role precisely. Credit the team, the client and the communities.

2 Ask before you name
Clients, results and quotes appear only with written permission.

3 Anonymise by default
Describe organisations and places by type unless the client agrees otherwise.

4 Share learning, not data
Method and lessons, never respondents' data or confidential findings.

5 Be specific, not superlative
Facts someone can check beat labels like 'leading' or 'top expert'.

6 Give before you ask
Share useful learning freely; ask for referrals only after good work.

Six habits of a trusted expert

☐ No inflated titles, results or numbers

☐ Testimonials in the client's approved words

☐ Keep consent records with every case study

☐ Disclose paid or sponsored content

☐ Refer on work outside your niche

☐ Correct a public claim the day you find it is wrong

Before you publish anything about an assignment: check the contract, ask the client, remove identifiers, and keep the written consent.

13 BUILD YOUR NAME

Six tools, one loop.

Be clear, be seen, be trusted. Start with your positioning, because the other five repeat it in different places.



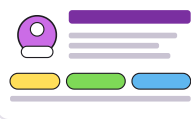
B POSITIONING AND NICHE SPECIALISATION

Positioning statement and brand canvas

One page that says who you help, with what problem, how, and what you will not do.

In three steps: fill the niche from four parts: sector, method, client and place; list proof points you can show, not adjectives; write the one-sentence statement last, and test it on three people.

In the examples: The solo consultant moved from 'M&E consultant' to 'MEL systems that small education NGOs can run themselves'.



B PROFESSIONAL-SERVICES MARKETING PRACTICE

Professional profile builder

A headline, about section, featured work and keywords for your LinkedIn and one-page web profile.

In three steps: write the headline from the formula: role, niche, result; structure the about section as problem, approach, proof, invitation; check every claim against your CV and your consent list.

In the examples: After the mid-term evaluation, the team leader rewrote the headline around livelihoods evaluation and mixed methods.



B SITUATION, APPROACH, RESULT FORMAT

Portfolio case study

A one-page, anonymised story of one assignment: context, question, approach, what changed and your role.

In three steps: anonymise first, then write; lead with what changed, and say what you cannot claim; send the draft to the client and get written consent before you publish.

In the examples: The MEL set-up became a one-page case study, approved by the director, with a quote used with written permission.



B VISIBLE EXPERT RESEARCH AND ASSOCIATION PROGRAMMES

Visibility and content plan

Twelve months of publishing, speaking, teaching and giving back, one quarter at a time.

In three steps: pick one theme from your positioning for the year; plan one item per column per quarter, no more; review each quarter: what was read, who called, what to drop.

In the examples: The team leader planned one practice note, one session proposal, one webinar and one reviewing role in a year.



B RELATIONSHIP MARKETING AND TRUSTED-ADVISOR PRACTICE

Relationship map and networking tracker

A short list of the people who matter to your work, by tier, with what you can give and when you will next be in touch.

In three steps: list people in six tiers, then keep the 30 that matter most; note one useful thing you can give each person; set a light cadence and follow up within 48 hours of every meeting.

In the examples: The solo consultant listed 30 people in six tiers, and a warm introduction from the director led to a second NGO client.



B PROFESSIONAL ASSOCIATION CONFERENCE PRACTICE

Event and conference plan

A one-page plan for before, during and after an event, so a conference turns into relationships and learning.

In three steps: before: set goals, list who to meet and submit a session if you can; during: have conversations, not pitches, and take notes the same day; after: follow up within 48 hours and share what you learned.

In the examples: Before a regional conference, the team leader set three goals, booked two meetings and followed up within 48 hours.

Each tool has a concept guide and a template in the Template Pack, Part 3 (pages 68 to 79).

13 BUILD YOUR NAME

Your name in action.

The solo consultant and the team leader use the tools on the two running examples.

Say it so a client can repeat it. POSITIONING AND PROFILE

POSITIONING

M&E-consultant

I help small education NGOs build MEL systems their own staff can run.

The proof: a 15-day set-up, and staff running the dashboard alone.

PROFILE HEADLINE

Independent-consultant

Evaluation team leader | Livelihoods and women's economic empowerment | Mixed methods

The about section opens with the problem programmes face mid-way.

One theme, four quarters. VISIBILITY PLAN: MAKING MID-TERM EVALUATIONS USEFUL

Q1

Publish

- ☐ A short practice note on grouping 14 questions into five, shared with the foundation's agreement.

Q2

Speak

- ☐ A session proposal for a regional evaluation conference.

Q3

Teach

- ☐ A free webinar for a national evaluation network.

Q4

Give back

- ☐ Reviewing conference abstracts and mentoring one emerging evaluator.

Each quarter took about two days.

Give first, follow up fast. RELATIONSHIPS AND EVENTS

1

After the MEL set-up, the director scored the consultant 9 out of 10 and mentioned two other NGOs.

2

The consultant asked whether the director would make an introduction, with a short note to forward.

3

One of the two NGOs became the next client.

4

The consultant later sent a useful dashboard template, asking nothing in return.

5

At the conference, each new contact had a message with one useful link within 48 hours.

Your turn

this quarter

- ☐ Write your positioning statement last, and test it on three people.
- ☐ Check every claim on your profile against your CV and your consent list.
- ☐ Get written consent before you publish a case study.
- ☐ Plan one item per column per quarter, no more.
- ☐ Follow up within 48 hours of every meeting or event.

PART 4 YOUR CAREER

14 YOUR PRACTICE



Build a practice, not a gig.

Most consultants are good at the work and weak at the business around it. Both decide whether you last.

1 Position on a narrow, credible niche

2 Keep a pipeline, not a hope

3 Get paid on time, every time

Be known for one specific thing

I help _____ (client) in _____ (place) answer questions about
_____ (sector) using _____ (method).

Where good work comes from

1

Repeat and referral

Your best source. Ask every satisfied client for a referral and keep in touch.

2

Networks

Evaluation associations, alumni and past colleagues. Be known for one thing.

3

Job boards and tenders

Useful, crowded and price-driven. Bid selectively.

4

Content

Short notes and tools that show how you think. They bring clients to you.

Three ways to charge

Day rate

Simple and fair for uncertain scope. Clients see every day.

Fixed price per deliverable

Rewards efficiency; needs a tight scope and a change process.

Retainer

A set number of days a month for ongoing support. Smooths income.

Six money habits

- | | |
|---|--|
| ☐ An advance on signing | ☐ Check registration, GST and tax deduction rules with an adviser |
| ☐ Payments on submission, not approval | ☐ Keep three months of costs in reserve |
| ☐ Invoice the day a milestone is met | ☐ Never start fieldwork unpaid |

The ideas behind this: [Positioning](#) · [A pipeline](#) · [Setting a day rate](#) · [Milestone payments](#) · [Win and loss reviews](#) · [Associates and teams](#)

PART 4 YOUR CAREER

15 LEADING AND GROWING



Leading and growing.

Top consultants lead teams, pass due diligence, negotiate well, run a room, get evidence used, and become known for one thing.

1 A firm is a promise of quality, not just more people

2 Negotiate interests, not positions

3 Plan for use, then build a body of work

Four ways to grow beyond your own days. GROWTH MODELS

Model 1

Solo specialist

You, with a sharp niche.

Gain Low overheads, full control.

Risk Income caps at your own days.

Model 2

Associate network

You lead; trusted freelancers join per assignment.

Gain Bid bigger without fixed costs.

Risk Quality varies unless you set standards.

Model 3

Boutique firm

A small core team with a shared method.

Gain Wins larger tenders; builds a brand.

Risk Payroll pressure between contracts.

Model 4

Productised practice

Repeatable services and tools, such as retainers, training or MEL as a service.

Gain Income that does not depend on every bid.

Risk Needs investment in products and marketing.

A team is only as good as its quality system. LEADING A TEAM

FIVE ROLES

Team leader. Owns design, quality, client and final report.

Specialists. Methods, sector or gender experts for parts of the work.

Field coordinator. Plans fieldwork, logistics, safety and daily checks.

Enumerators and facilitators. Collect data to the protocol, with training and supervision.

Quality reviewer. Reads drafts as the commissioner would, before the client does.

FOUR QUALITY CHECKPOINTS

Before fieldwork. Tools pre-tested; enumerators trained; consent and safety plans in place.

During fieldwork. Daily data checks by interviewer; back-checks on a sample; issues logged.

During analysis. Second coder on a sample; findings traced to evidence; peer review of the matrix.

Before submission. Red-team read; quality checklist; facts checked; sources opened.

DUTY OF CARE

Risk assessment for each field location

Safeguarding and sexual exploitation prevention policy, signed by every team member

Referral details for respondents in distress

Check-in routine for field teams

Insurance for team members where possible

Have due-diligence documents ready before anyone asks. DUE DILIGENCE

- ☐ Registration certificate of your firm or proof of self-employment
- ☐ Audited accounts or income statements for recent years
- ☐ Data protection policy
- ☐ Code of conduct
- ☐ CVs of the core team in donor format

- ☐ Tax registration (for example PAN and GST in India)
- ☐ Safeguarding policy
- ☐ Anti-fraud, anti-corruption and conflict of interest policies
- ☐ Professional indemnity insurance, if required
- ☐ Three references and two sample reports

15 LEADING AND GROWING

Negotiate, facilitate, plan for use.

Trade scope, not quality. Run the room so every voice shapes the findings. Decide who needs what, and when.

1

Separate the people from the problem

Be warm with the person and firm about the issue.

2

Focus on interests, not positions

Ask why the client wants 30 days: a board meeting? Then solve for that.

3

Invent options for mutual gain

Offer a short early briefing for the board and a full report later.

4

Use objective criteria

Refer to day rates in comparable tenders, or the days the questions really need.

5

Know your walk-away point

Decide before the call the minimum fee and scope you will accept, and your best alternative if there is no deal.

GIVE SOMETHING, GET SOMETHING

When: Budget is cut by 20% **Offer: Fewer questions or districts, not lower quality**

When: Deadline is moved earlier **Offer: An interim briefing now, the full report on the original date**

When: Client wants more questions **Offer: A change request with days and fee, or drop others**

When: Client wants payment on approval **Offer: Payment on submission, with a fixed review window**

When: Client wants to approve findings **Offer: Fact-checking by the client; findings stay with the evaluator**

A two-hour validation workshop. LIBERATING STRUCTURES

0:00	Welcome, purpose and how findings were reached	10 min
0:10	Findings walk: posters by question; people add notes	30 min
0:40	What, So What, Now What? in small groups	40 min
1:20	1-2-4-All on the three most important recommendations	30 min
1:50	15% Solutions: who will do what first	20 min
2:10	Close: what changes in the report, and next steps	10 min

Plan who needs what, and when. USE AND INFLUENCE

Foundation board	One-page brief and a 15-minute presentation	Before the board meeting
Programme team	Validation workshop and full report	Within two weeks
Federation leaders	Local-language summary and a meeting	Within a month
Women members	Short audio or poster summary	Within a month
Wider sector	Practice note or webinar, with permission	Within three months

15 LEADING AND GROWING

Build a body of work, and watch what is coming.

Six moves for reputation, five facilitation methods and four emerging themes.

1

Pick one niche and own it

Be the first name people think of for one thing, such as MEL for state education programmes.

2

Publish on a rhythm

A short practice note every month; one longer piece a year.

3

Speak where your clients learn

Association events, gLOCAL, sector conferences and webinars.

4

Turn methods into tools

Templates, checklists or platforms others can use, such as EvalCanvas.

5

Productise one service

A retainer, a training course or MEL as a service with a fixed scope and price.

6

Give back

Mentor, review, and contribute to your evaluation association.

Five facilitation methods for evaluation moments. LIBERATING STRUCTURES

1-2-4-All

For: Validation workshops

Everyone reflects alone, then in pairs, then fours, then the whole room. Quiet voices are heard.

What, So What, Now What?

For: Sense-making sessions

Look at the data together, agree what it means, then what to do.

15% Solutions

For: Recommendations

Each person names what they can do now, with the resources they control.

TRIZ

For: Risk and assumption reviews

Ask how to guarantee failure, then stop doing those things.

Troika Consulting

For: Peer advice for your team

One person shares a challenge; two others advise while the first listens.

Four themes clients are starting to ask about. WHAT IS COMING

Climate and environment

More ToRs ask how programmes affect, and are affected by, climate and the environment. Learn the basics of environmental sustainability as an evaluation lens.

Evaluating AI programmes

Clients increasingly fund AI tools in health, education and agriculture. Evaluate accuracy, fairness, safety and real use, not just adoption.

Data protection in India

The Digital Personal Data Protection Act, 2023 and its Rules, notified in November 2025 and being phased in, apply to personal data you collect. Plan consent, storage and deletion.

Remote and digital data

Phone surveys, digital monitoring and administrative data are cheaper but carry bias and privacy risks. State the limits.

PART 4 YOUR CAREER

16 YOUR PATHWAY

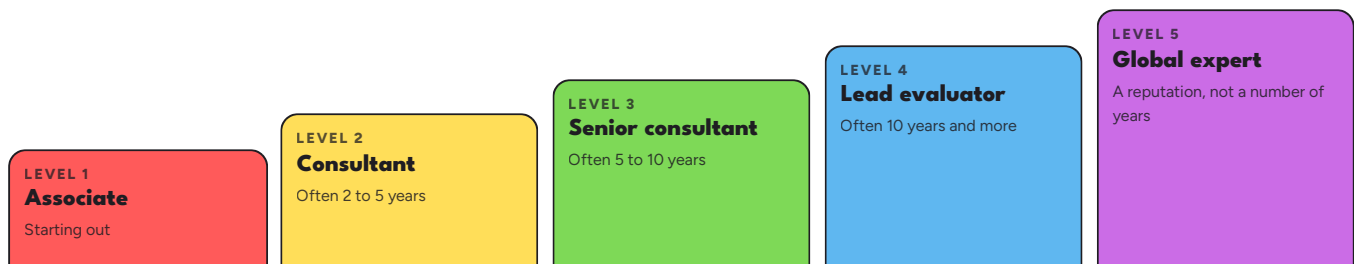


From first assignment to global consultant.

A five-level path from new evaluator to leading international evaluations: what you do at each level, what proves it, and the move that gets you to the next one.

- 1 A portfolio and referees prove competence, not a job title
- 2 Each level is earned with evidence you can show
- 3 Global work comes through rosters, firms and networks

Five levels take a new evaluator to global work. THE CAREER LADDER



	ASSOCIATE	CONSULTANT	SENIOR CONSULTANT	LEAD EVALUATOR	GLOBAL EXPERT
WHAT YOU DO	Collect and clean data, run interviews, review documents and draft sections under a lead.	Lead small assignments end to end: scope, tools, analysis, report and handover.	Lead mid-size evaluations and small teams. Own the method, the client relationship and quality.	Lead complex, multi-site or multi-country evaluations and quality-assure other people's work.	Shape methods and standards, advise evaluation offices and boards, and develop the next generation.
WHAT PROVES IT	Clean datasets, field notes, a desk review, and a lead who will vouch for you.	Two or three pieces you led, two client references, and a niche you can name in one sentence.	Evaluations you led as team leader, a management response that was acted on, and a method you are known for.	Team-leader roles on donor or UN evaluations, a peer review or credential, and work you have presented or published.	Invited roles, reference groups, journal or panel reviews, publications, keynotes, mentees who lead.
TYPICAL WORK	Data collection, desk reviews, transcription and coding, survey support.	MEL system set-ups, monitoring reviews, small evaluations and learning reviews.	Mid-term and final evaluations, baselines, theory-based and mixed-methods studies.	Country, thematic, portfolio and strategic evaluations; evaluation quality assurance.	Methods advice, meta-evaluations, evaluation policies, expert panels.
THE MOVE UP	Ask to draft one full report section and present it to the client.	Join a mid-size evaluation as deputy team leader.	Bid as team leader on a donor or UN tender, alone or with a firm.	Win long-term agreements and roster places; review for evaluation offices.	Give back: mentor, write, and build the evaluation ecosystem.

Years are indicative. People move faster with deliberate practice, mentors and a clear niche, and slower without them.

Your first three assignments should build the proof, not just the income. FOR PEOPLE NEW TO CONSULTING

1. Second chair

Work under an experienced lead.

Ask for: one section to draft, a seat in client meetings, and feedback on your writing.

Keep: the lead as a referee, and an anonymised sample of your section.

2. Small and complete

Lead a small, tightly scoped piece from scope to handover.

Ask for: a written scope, a change process, and an advance on signing.

Keep: a short case study, a client reference and your first feedback score.

3. Stretch with support

Join a bigger evaluation as deputy team leader.

Ask for: ownership of one evaluation question and the matrix rows behind it.

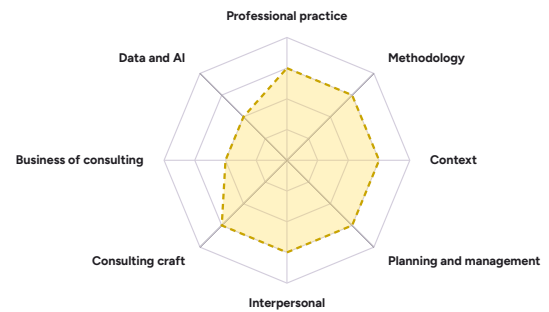
Keep: your role in the CV in the donor's words, and a second referee.

16 YOUR PATHWAY

Rate yourself, then grow the two biggest gaps.

Eight competencies: the five AEA domains plus three from consulting practice. Circle your level, mark the target, and plot yourself on the radar.

Competency	1 Aware	2 Practising	3 Leading	4 Teaching	Senior target
Professional practice Ethics, standards, independence, reflective practice.	☐	☐	☐	☐	3
Methodology Designs, methods, sampling, analysis and judging evidence.	☐	☐	☐	☐	3
Context Sector, culture, politics, power and the people affected.	☐	☐	☐	☐	3
Planning and management Workplans, budgets, teams, risks and quality control.	☐	☐	☐	☐	3
Interpersonal Communication, facilitation, negotiation and handling conflict.	☐	☐	☐	☐	3
Consulting craft Framing problems, issue trees, proposals, storylines, trust.	☐	☐	☐	☐	3
Business of consulting Niche, pipeline, pricing, contracts and getting paid.	☐	☐	☐	☐	2
Data and AI Data protection, digital tools and responsible AI use.	☐	☐	☐	☐	2



Dashed: the senior consultant target. Draw your profile over it. Growth actions for each gap are in the web version.

A 12-month plan from foundations to your first global bid. FOUR QUARTERS

MONTHS 1 TO 3 Foundations

- ☐ Take the competency self-assessment and pick two gaps
- ☐ Join your national evaluation association and EvalYouth if you are early career
- ☐ Read the UNEG norms, the OECD DAC criteria and one evaluation a week
- ☐ Write your one-sentence niche

MONTHS 4 TO 6 First proof

- ☐ Take a second-chair role with a lead you respect
- ☐ Lead one small, well-scoped piece end to end
- ☐ Write a one-page case study of it, with permission
- ☐ Ask for two references while the work is fresh

MONTHS 7 TO 9 Visible expertise

- ☐ Publish one short practice note in your niche
- ☐ Present at an association event or during gLOCAL Evaluation Week
- ☐ Build a master CV in the key-expert format donors use
- ☐ Find a mentor one or two levels above you

MONTHS 10 TO 12 Global entry

- ☐ Register on UNGM and the consultant systems of two agencies or banks
- ☐ Apply to two roster calls or long-term agreements in your niche
- ☐ Join a firm's bid as a named expert
- ☐ Retake the self-assessment and set next year's two gaps

16 YOUR PATHWAY

Go global through rosters, firms and networks.

Cold applications rarely win international work. Being known, being listed and being named in bids does.

1. UN rosters and long-term agreements

Evaluation offices of UN agencies such as UNDP and UNICEF regularly call for roster members and long-term agreements. Once you are on one, you are invited to bid for specific evaluations.

Watch UNEG's vacancies page and agency job sites; apply in your niche.

2. UN Global Marketplace

The common procurement portal of the UN system, where tenders are published and suppliers register.

Register and set alerts for evaluation and monitoring tenders.

3. Development bank systems

Multilateral development banks run their own consultant registration and procurement systems for firms and individuals.

Register with the banks that fund your sector and region.

4. Teaming with firms

Most large donor tenders are won by firms that must name key experts. Being on a firm's bench is the fastest route to team-leader roles.

Send three firms in your niche a short note and your donor-format CV.

5. Foundations and philanthropy

Many foundations hire through networks and referrals rather than open tenders.

Ask past clients for one introduction each, every year.

6. Job boards

Global boards such as ReliefWeb and Devex list consultancy calls from many organisations.

Use them to find calls, then apply selectively to the ones that fit.

Six rules for a donor CV that scores. THE KEY-EXPERT CV

Use their format	Many donors and firms ask for a key-expert CV in a set template. Keep a master CV and adapt it in 30 minutes.
Lead with relevant evaluations	For each: your role, client, country, dates, days worked, design and methods.
Quantify	Team size, countries, sample sizes, budget range, number of evaluation questions.
Mirror the ToR	Use the words in the expert requirements: 'team leader', 'mixed methods', 'theory-based'.
Be honest about languages and regions	Overstating them is easy to check and ends relationships.
Three referees who saw you lead	Ask before you list them, and tell them about each bid.

Credentials and communities.

PROVE AND GROW YOUR PRACTICE

AEA evaluator competencies American Evaluation Association, 2018

Five domains to assess yourself against: professional practice, methodology, context, planning and management, and interpersonal.

Credentialed Evaluator (CE) Canadian Evaluation Society

A professional designation based on education, experience and demonstrated competencies.

IPDET

International Program for Development Evaluation Training, Bern

An intensive training programme in development evaluation, with core and specialised courses.

Your national evaluation association

IOCE and EvalPartners

Voluntary organisations for professional evaluation exist in most countries. Start presenting and find mentors here.

UNEG evaluation competency framework United Nations Evaluation Group, 2016

The competencies UN evaluation offices expect of evaluators and evaluation managers.

Voluntary Evaluator Peer Review European Evaluation Society and UK Evaluation Society

A structured peer review of your practice that helps you plan your development.

gLOCAL Evaluation Week Global Evaluation Initiative

A yearly, worldwide week of local events: a low-cost way to present and connect.

EvalYouth Global Network EvalPartners

A network supporting young and emerging evaluators, including mentoring.



PART 5 REFERENCE

17 READY?

Ready?

Ten checks before you bid, and ten before you submit the final report.

Print this page and tick them off.

1 Before you bid

2 Before you submit

3 Print it for the team

BEFORE YOU BID

The proposal

- ☐ The decision behind the ToR is clear
- ☐ Your skills fit the sector and methods
- ☐ Budget and days match the questions
- ☐ Conflicts of interest declared or ruled out
- ☐ Clarifying questions asked
- ☐ A first hypothesis and question tree
- ☐ Days by role and stage, with contingency
- ☐ Payment milestones you can live with
- ☐ Change process and review rounds proposed
- ☐ AI and data use stated

Score: ____ / 10

BEFORE YOU SUBMIT

The final report

- ☐ Every question answered, in order
- ☐ Findings trace to evidence in the matrix
- ☐ Limits and constraints stated plainly
- ☐ A one-page summary on top
- ☐ Recommendations with owners and timeframes
- ☐ Factual comments addressed, findings kept
- ☐ Respondents cannot be identified
- ☐ Every source and quote checked
- ☐ AI use disclosed in the method note
- ☐ Data and tools ready to hand over

Score: ____ / 10



The library of ideas.

72 ideas, standards and practices this playbook stands on, each with something to do on Monday. Links are in the web version.

1 Ideas from strategy consulting and evaluation

2 A glossary in plain words

3 A printable cheat sheet

Flawless consulting

Consulting craft · Peter Block; 4th edition, 2023

Consulting runs in five phases, and contracting is the one most people skip.

On Monday: Before your next contract, write down what you need from the client and say it out loud.

Helping roles

Consulting craft · Edgar Schein, Process Consultation and Humble Consulting

You can help as an expert, a doctor or a process partner.

On Monday: Ask early: do you want an answer, a diagnosis, or help finding your own answer?

The trust equation

Consulting craft · David Maister, Charles Green and Robert Galford, The Trusted Advisor, 2000

Trust = (credibility + reliability + intimacy) ÷ self-orientation.

On Monday: In your next meeting, ask one more question before offering any solution.

Seven-step problem solving

Consulting craft · Charles Conn and Robert McLean, Bulletproof Problem Solving, 2019

Define, disaggregate, prioritise, plan, analyse, synthesise, communicate.

On Monday: Draw the tree for your evaluation questions before writing the methodology.

The pyramid principle

Consulting craft · Barbara Minto, The Pyramid Principle

Put the answer first, then the reasons, then the evidence.

On Monday: Rewrite your report summary so its first sentence is the answer.

Hypothesis-driven work

Consulting craft · Strategy consulting practice; Ethan Rasiel, The McKinsey Way, 1999

Start with a best guess at the answer, then try to prove it wrong.

On Monday: Write a one-line hypothesis for each evaluation question at inception.

No overlaps, no gaps

Consulting craft · Barbara Minto; consulting practice

Group ideas so they do not overlap and together cover everything.

On Monday: Check that your evaluation questions do not double-count and leave nothing important out.

Pre-wiring

Consulting craft · Consulting practice

Share hard findings one to one before the big meeting.

On Monday: Before a findings workshop, call the two people most affected by the hardest finding.

The obligation to dissent

Consulting craft · Consulting practice

You owe the client your honest view, especially when it is unwelcome.

On Monday: Write down in advance how you will respond if asked to soften a finding.

Problem statement worksheet

Consulting craft · McKinsey problem solving

Fix the question before you fix the method.

On Monday: Fill it in after your discovery call and send it to the client to confirm.

RAPID decision roles

Consulting craft · Bain & Company; Rogers and Blenko

Clear decision roles make findings easier to act on.

On Monday: Name the D for each decision your evaluation must inform, at inception.

Power and interest map

Consulting craft · Aubrey Mendelow, 1991

Engage people according to their power and interest.

On Monday: Map stakeholders at inception and plan one engagement for each quadrant.

RAID log

Consulting craft · Project management practice

Track risks, assumptions, issues and dependencies in one place.

On Monday: Start a RAID log in your inception report with at least five entries.

The 80/20 rule

Consulting craft · Vilfredo Pareto; consulting practice

A few analyses carry most of the answer.

On Monday: Before analysis, rank your planned analyses and share the top three results early.

Weekly status note

Consulting craft · Consulting practice

Short, regular, honest updates build reliability.

On Monday: Send one every Friday, even when nothing has changed.

Change requests

Contracts and money · Project management practice

Scope changes are agreed in writing with days, fees and dates.

On Monday: Add a one-paragraph change clause to every proposal.

Terms of reference

Contracts and money · Evaluation practice; EvalCanvas guidance

A good ToR states the decision, few answerable questions, deliverables and review points.

On Monday: Read every ToR first for who will use the findings, and for which decision.

● Setting a day rate

Contracts and money · Freelance and consulting practice

Your rate must cover unbilled days, overheads and gaps between contracts.

On Monday: Use the rate builder in the workshop with your own numbers.

● Milestone payments

Contracts and money · Consulting practice

Tie payments to submission of deliverables, with an advance.

On Monday: Never accept 100% on approval of the final report.

● Colour team reviews

Contracts and money · Shipley Associates

Review proposals at planned points, as the client would.

On Monday: Book a 45-minute red team with a colleague before every major bid.

● Value-based fees

Contracts and money · Alan Weiss, Value-Based Fees

Price on the value to the client, and offer a choice of yes.

On Monday: Put three options in your next proposal and see which one the client picks.

● Principled negotiation

Contracts and money · Roger Fisher and William Ury, Getting to Yes

Negotiate interests, not positions.

On Monday: Before your next fee discussion, write your walk-away point.

● Utilisation-focused evaluation

Evaluation design · Michael Quinn Patton

Judge an evaluation by its use by intended users.

On Monday: At inception, name three people who will use the findings and ask what they will decide.

● RealWorld evaluation

Evaluation design · Michael Bamberger and Linda Mabry, 3rd edition, 2019

Every evaluation works under budget, time, data and political constraints.

On Monday: List your four constraints in the inception report, with how you will work around each.

● Evaluation matrix

Evaluation design · Evaluation practice

One table that links every question to indicators, sources, methods and analysis.

On Monday: If a method appears in no row of the matrix, cut it.

● Ghost report

Evaluation design · Consulting practice, adapted

Sketch the final report before you collect any data.

On Monday: Draft the report table of contents at inception, with one placeholder finding per question.

● The Rainbow Framework

Evaluation design · BetterEvaluation

Seven clusters of evaluation tasks: manage, define, frame, describe, understand causes, synthesise, report and support use.

On Monday: Use it to check your workplan has a line for every cluster.

● Triangulation

Evaluation design · Evaluation practice

Check findings across different sources, methods and people.

On Monday: Mark in the matrix which findings rest on one source only.

● Fieldwork quality checks

Evaluation design · Survey practice

Daily checks by interviewer, back-checks and supervision.

On Monday: Agree the daily checks with your field team before day one.

● Evaluative rubrics

Evaluation design · E. Jane Davidson, Evaluation Methodology Basics

Make the standard for 'good' explicit before you judge.

On Monday: Co-write one rubric with the client at inception for the most contested question.

● Evaluability assessment

Evaluation design · Rick Davies, DFID Working Paper 40, 2013

Check whether a programme can be evaluated, and how, before you commit.

On Monday: When a ToR asks for impact with no baseline, propose a short evaluability assessment first.

● Evaluative reasoning

Evaluation design · Michael Scriven; E. Jane Davidson; Julian King

A judgement needs criteria, standards and evidence, combined transparently.

On Monday: For your next evaluation, agree a rubric with the client before data collection.

● Value for money with rubrics

Evaluation design · Julian King; Oxford Policy Management

Judge value for money across economy, efficiency, effectiveness and equity.

On Monday: Add a value-for-money question and rubric to your next inception report.

● Contribution analysis

Evaluation design · John Mayne

Build a credible case that a programme contributed to change.

On Monday: Use it when there is no comparison group but a clear theory of change.

● Process tracing

Evaluation design · Political science and evaluation practice

Test whether specific steps caused a specific outcome.

On Monday: Use it for one or two high-stakes outcomes.

● Realist evaluation

Evaluation design · Ray Pawson and Nick Tilley, 1997

What works, for whom, in what circumstances, and why?

On Monday: Ask 'for whom and in what circumstances' of every finding.

● Outcome harvesting

Evaluation design · Ricardo Wilson-Grau and colleagues
Collect changes first, then work back to the contribution.

On Monday: Use it for advocacy and networks where outcomes are hard to predict.

● Most Significant Change

Evaluation design · Rick Davies and Jess Dart
Let people say which changes matter most.

On Monday: Add it alongside a survey to explain what the numbers mean to people.

● Developmental evaluation

Evaluation design · Michael Quinn Patton, 2010
Evaluation for innovation in complex settings.

On Monday: Offer it when a client is designing something new, not judging something finished.

● Sample size for surveys

Evaluation design · Survey statistics
Precision depends on sample size, not on population size, for large populations.

On Monday: Calculate before you price fieldwork, and state the margin of error in the inception report.

● 1:3:25 reports

Reporting and use · Canadian Health Services Research Foundation; BetterEvaluation
One page of key messages, a three-page summary, 25 pages of detail.

On Monday: Write the one-page messages first, then build the rest under them.

● Management response

Reporting and use · Evaluation practice; EvalCanvas guidance
For each recommendation: accept, adapt, defer or decline, with an action, owner and date.

On Monday: Offer a management response template with your final report.

● Tracking use

Reporting and use · EvalCanvas guidance
Decide in advance how you will know whether findings were used.

On Monday: Put a three-month check-in in your close-out email.

● The shelf report

Reporting and use · EvalCanvas guidance
Technically excellent, submitted and never read.

On Monday: Ask at inception when the decision is, and deliver before it.

● Action titles and horizontal logic

Reporting and use · Consulting practice; Cole Nussbaumer Knaflic
If the titles alone tell the story, the report works.

On Monday: Read only the headings of your next draft aloud. Fix any that are topics, not findings.

● Collaborating, learning and adapting

Reporting and use · USAID Learning Lab
Build learning into how a programme is managed.

On Monday: Propose one pause-and-reflect session each quarter.

● Evaluation influence

Reporting and use · Karen Kirkhart, 2000
Evaluations influence people beyond the report.

On Monday: Plan products for each audience, and check back on influence later.

● Liberating Structures

Reporting and use · Henri Lipmanowicz and Keith McCandless
Simple ways to involve everyone in a room.

On Monday: Use one in your next validation workshop.

● Conflict of interest

Ethics and standards · Evaluation standards and practice
Past involvement with a programme can compromise an evaluation's credibility.

On Monday: Declare any involvement in your proposal and let the client decide.

● Independence of findings

Ethics and standards · UNEG norms and evaluation practice
Clients review for facts; evaluators decide findings.

On Monday: Put the factual-review rule in your contract.

● OECD DAC evaluation criteria

Ethics and standards · OECD DAC, revised 2019
Relevance, coherence, effectiveness, efficiency, impact and sustainability.

On Monday: Choose the two or three criteria that matter most for this client's decision.

● UNEG ethical guidelines

Ethics and standards · UNEG, 2020
Integrity, accountability, respect and beneficence.

On Monday: Use them as the ethics section of your inception report.

● AEA guiding principles

Ethics and standards · American Evaluation Association, 2018
Systematic inquiry, competence, integrity, respect for people, and common good and equity.

On Monday: Check your proposal against the competence principle: do you have the skills, or a team member who does?

● DAC Quality Standards

Ethics and standards · OECD DAC, 2010
The standards commissioners use to judge evaluation quality.

On Monday: Check your report against them before you submit.

● Gender-responsive evaluation

Ethics and standards · UN Women Evaluation Handbook, 2022
Evaluate how programmes affect gender equality and power.

On Monday: Add a gender question to your next evaluation matrix.

● India's data protection law

Ethics and standards · Digital Personal Data Protection Act, 2023;
Rules 2025
Personal data collected in India now has legal rules.

On Monday: Add consent, storage and deletion to your data management plan, and check with a legal adviser.

● AI terms in contracts

AI · UK Evaluation Society, AI in Evaluation guidelines, November 2025
Agree permitted AI uses, data rules and verification in the contract.

On Monday: Add a short AI clause to your standard proposal.

● Ethical AI in UN evaluations

AI · UNEG, 2025
Five ethical principles for using AI in UN evaluations.

On Monday: Read it before using AI on any UN-funded assignment.

● AI for qualitative coding

AI · Nga Than and colleagues, Sociological Methods and Research, 2025
Generative models can code text about as reliably as human coders, on some tasks.

On Monday: Hand-code a sample first, then use AI as a second coder and compare.

● Positioning

Your practice · Consulting practice
Be known for one specific thing.

On Monday: Write your niche as one sentence a past client could repeat.

● A pipeline

Your practice · Consulting practice
Track leads, proposals and decisions in one place.

On Monday: Keep a simple list: lead, stage, value, next step, date.

● Win and loss reviews

Your practice · Consulting practice
Ask why you won or lost every bid.

On Monday: After every decision, send one short question: what made the difference?

● Associates and teams

Your practice · Consulting practice
Grow capacity without fixed costs.

On Monday: Name three people you would bring onto a larger bid.

● After-action review

Your practice · US Army; widely adopted
Learn from every assignment in 30 minutes.

On Monday: Hold one with your team within a week of every final report.

● Client feedback score

Your practice · Fred Reichheld, Bain; Harvard Business Review, 2003
One question predicts referrals: would you recommend us?

On Monday: Ask every client at the end of the assignment, and call anyone who scores 6 or below.

● AEA evaluator competencies

Your practice · American Evaluation Association, 2018
Assess yourself against five domains.

On Monday: Rate yourself on each domain and pick two to grow this year.

● UNEG evaluation competency framework

Your practice · United Nations Evaluation Group, 2016
Know what UN evaluation offices expect.

On Monday: Before a UN roster call, map your evidence to the framework.

● Credentialed Evaluator

Your practice · Canadian Evaluation Society
A professional designation for evaluators.

On Monday: Check whether a credential would help in the markets you want to work in.

● Voluntary Evaluator Peer Review

Your practice · European Evaluation Society; UK Evaluation Society
Peers review your practice and your development plan.

On Monday: Use its framework as a checklist even before you apply.

● IPDET

Your practice · International Program for Development Evaluation Training
Intensive training in development evaluation.

On Monday: Plan one structured course in the next 12 months, then apply it at once.

● gLOCAL Evaluation Week

Your practice · Global Evaluation Initiative
Present to a global audience at low cost.

On Monday: Propose one session with a colleague or your association.

● Evaluation associations and EvalYouth

Your practice · IOCE; EvalPartners; EvalYouth Global Network
Your network is your first market.

On Monday: Join one, volunteer for one task, and present within a year.

18 GO DEEPER

Glossary.

69 terms used in this playbook, in plain words.

action title: A heading written as a full-sentence finding, not a topic.

after-action review: A short team review of what was planned, what happened, why, and what to change.

attribution: How much of a change can be credited to the programme alone.

baseline: The measurement at the start, used to see change later.

BATNA: Best alternative to a negotiated agreement: what you will do if there is no deal.

change request: A written agreement to change scope, with days, fees and dates.

colour team review: A planned proposal review: pink for the plan, red for the full draft, gold for sign-off.

conflict of interest: A personal or professional interest that could bias your work.

contingency: A small reserve in the budget for the unexpected, often around 10%.

contribution: How far a programme plausibly helped cause a change, alongside other factors.

contribution analysis: A theory-based way to build a credible case that a programme contributed to change.

counterfactual: What would have happened without the programme, often estimated with a comparison group.

day rate: The fee for one day of consulting work.

deliverable: Something you hand over: a report, a tool, a workshop, a dataset.

design effect: How much a cluster sample's precision is reduced compared with a simple random sample; multiply the sample by it.

developmental evaluation: Evaluation alongside a team, in real time, to support innovation in complex settings.

due diligence: Checks a client or firm makes on you before contracting: registration, accounts, policies, references.

engagement charter: A one-page agreement of objective, scope, roles, timeline and ways of working.

evaluability assessment: A short study of whether, and how, a programme can be evaluated.

evaluation matrix: A table linking each evaluation question to indicators, data sources, methods and analysis.

evaluation question: A question the evaluation must answer, agreed with the client.

focus group discussion: A guided discussion with a small group of people.

ghost report: An outline of the final report drafted before data collection.

GST: Goods and services tax: charged on your fee if you are registered.

hypothesis: A best guess at the answer, stated early so you can test it.

inception report: The detailed plan agreed with the client before fieldwork. It becomes the working contract.

independence: The evaluator's freedom to reach findings based on evidence, without pressure.

indicator: Something you measure to see whether change is happening.

issue tree: A diagram that breaks a big question into smaller ones with no overlaps or gaps.

issue tree: A diagram that breaks a key question into smaller questions with no overlaps or gaps.

key expert: A named person in a firm's bid whose CV is scored, such as the team leader.

key informant interview: An interview with someone who knows the programme or context well.

logframe: A table linking activities, outputs, outcomes and indicators.

logframe: A table of a programme's goals, indicators, targets and sources at each level.

long-term agreement: A framework contract that lets an agency hire you for specific tasks over several years without a new tender each time. Often called an LTA.

management response: The client's formal reply to each recommendation, with actions, owners and dates.

MECE: Mutually exclusive and collectively exhaustive: no overlaps, no gaps.

milestone payment: A payment tied to a deliverable or stage.

mixed methods: Using both numbers and qualitative evidence to answer questions.

niche: The specific combination of sector, method, client and place where you do your best work.

non-response: People who cannot be reached or refuse; plan extra sample to cover them.

OECD DAC criteria: Six common evaluation criteria: relevance, coherence, effectiveness, efficiency, impact and sustainability.

positioning: A clear statement of who you help, with what problem and how, so people know what to call you about.

pre-wiring: Sharing difficult findings privately before a group meeting.

proposal: Your written offer: understanding, approach, workplan, team and budget.

pyramid principle: Writing with the answer first, then reasons, then evidence.

RAG status: Red, amber or green: a one-word signal of whether work is on track.

RAID log: A list of risks, assumptions, issues and dependencies, each with an owner.

RAPID: Bain's decision roles: recommend, agree, perform, input, decide.

realist evaluation: An approach that asks what works, for whom, in what circumstances and why.

retainer: A fixed number of consulting days bought each month.

roster: A list of pre-approved consultants an organisation invites to bid for or take on assignments.

rubric: A description of what different levels of performance look like for a criterion.

rubric levels: Descriptions of excellent, good, adequate and poor performance for one criterion.

scope creep: Work that grows beyond what was agreed, without more time or fees.

statement of work: The part of a contract that defines scope, deliverables and acceptance.

steering committee: A small group of senior client staff who review progress and make decisions at key points.

TDS: Tax deducted at source: the share of your fee an Indian client deducts and pays to the tax department on your behalf.

team leader: The evaluator responsible for the design, team, quality and final report.

terms of reference: The client's description of the assignment: purpose, questions, deliverables, timeline and budget. Often called the ToR.

theory of change: A description of how and why a programme is expected to lead to change.

ToR: Short for terms of reference.

triangulation: Checking a finding across different sources, methods or people.

utilisation rate: The share of your working days that are billed to clients.

utilisation-focused evaluation: An approach that designs evaluations around their intended users and uses.

validation workshop: A meeting where stakeholders review draft findings before the report is final.

value for money: Whether a programme makes good use of resources, often judged on economy, efficiency, effectiveness and equity.

VOPE: A voluntary organisation for professional evaluation, such as a national evaluation association.

warm introduction: An introduction to a new contact made by someone who knows you both.

From Zero to World-Class

The Monitoring & Evaluation Consulting Toolkit

The hands-on companion to this playbook. Every tool and document opens as a two-page spread: a concept guide on the left and a template on the right, grouped by the seven CONSULT stages. Use the playbook to learn the method and the toolkit to work through it.

36

TEMPLATES

36

CONCEPT GUIDES

7

STAGES

C

CHOOSE

O

OFFER

N

NEGOTIATE

S

SET UP

U

UNCOVER

L

LAND

T

TRANSFER

PART 1

21 tools

Three tools for each of the seven stages, from deciding whether to bid to tracking how findings are used.

PART 2

9 documents

The documents every consultant writes, from the technical proposal to the key-expert CV.

PART 3

6 brand and network tools

Positioning, profile, case study, visibility plan, relationships and events, with an honest-brand guide.

Also inside

- A before-and-after ToR spread: seven risky clauses become written agreements
- Two worked examples that run through every concept guide
- Key terms in plain words, and a guide to ethics and AI use
- A first hour: three tools you can use today
- The CONSULT method at a glance, with a page reference for every tool
- Notes pages, and a clean A4 layout for double-sided printing



CURATED BY

Shashidhar S A

- Founder, HubForge Global · MERL & AI Innovator
- En-Intrapreneur
- Driving evaluation ecosystems through People, Process & Product

Ask for the toolkit
hubforgeglobal@gmail.com · www.hubforgeglobal.com

The CONSULT method

- C

Choose the work: Choose before you chase.
Find the real decision behind the ToR · Score fit, readiness, budget and conflicts · Ask clarifying questions before you bid
- O

Offer: Lead with the answer you expect, then show how you will test it.
Restate the problem in one sentence · Offer a first hypothesis · Break questions into a tree
- N

Negotiate the contract: Contract the expectations, not just the price.
Deliverables with page limits and review rounds · Payments on milestones, not approval · A written change process
- S

Set up the inception: Design for use from day one.
Meet the people who will use the findings · An evaluation matrix for every question · Sketch the report before collecting data
- U

Uncover the evidence: Collect only what answers a question, and check it every day.
Pilot every tool · Daily quality checks · Triangulate across sources
- L

Land the findings: Answer first, then the evidence, then the detail.
Answer first, then evidence · A one-page summary on top · Validate with stakeholders
- T

Transfer and track use: The assignment ends when the findings are used, not when the invoice is paid.
A management response with owners and dates · Hand over data and tools · Ask for feedback and a reference
- Trust** = (credibility + reliability + intimacy) ÷ self-orientation.
Maister, Green and Galford

Read a ToR: ask first

- Who will use the findings, for which decision?
- Which questions matter most, and can they be answered in the time?
- Is it monitoring, review, evaluation or impact?
- What data already exists: baseline, monitoring, reports?
- Is the budget stated, and does it match the work?
- What are the deliverables, page limits and review rounds?
- How is payment scheduled?
- Who owns the data and the report?

Contract the expectations

- Scope

The agreed questions and what is excluded.
- Deliverables

Each one with format, length and review rounds.
- Payments

Milestones tied to submission, with an advance.
- Change process

Extra work is agreed in writing with days and fees.
- Independence

The client checks facts; the evaluator decides findings.
- Data

Who owns, stores and deletes data, and how it is protected.
- AI use

What tools may be used, on which data, and how it is disclosed.

Write answer first: 1:3:25

- 1 page

Key messages: the answer to each question, and what to do.
- 3 pages

An executive summary by question, with the main evidence.
- 25 pages

Findings by question, methods, limits and recommendations. Annexes hold the rest.

Eight habits of top firms

Answer first · Hypothesis-driven · No overlaps, no gaps · 80/20 · So what? · Pre-wire · Obligation to dissent · Results, not reports.

21 top-firm tools, by stage

- C

Bid or no-bid gate, Problem statement worksheet, Stakeholder power and interest map
- O

Issue tree and day-one hypothesis, Three-option proposal, Colour team reviews
- N

Statement of work and change control, RAPID decision roles, Engagement charter
- S

Kick-off and workplan, RAID log, Ghost deck and report
- U

80/20 analysis plan, Weekly status note, Evidence strength rating
- L

Storyline and action titles, Pre-wiring and validation, Impact and ease prioritisation
- T

Management response tracker, After-action review, Client feedback score

With AI: six promises

1. **Client data stays where agreed.** Only approved tools, only de-identified data, as the contract says.
2. **Every source is checked.** No reference, quote or number goes in unopened.
3. **AI use is disclosed.** The method note says what AI did and who checked it.
4. **Humans own the findings.** AI drafts and suggests; the evaluator decides and signs.
5. **AI terms are in the contract.** Permitted uses, data rules and disclosure, agreed up front.
6. **Saved time goes to the client.** More listening, sharper synthesis, or a lower price.

Get paid

An advance on signing · Payments on submission, not approval · Invoice the day a milestone is met · Check registration, GST and tax deduction rules with an adviser.



written by

Shashidhar S A

Founder, HubForge Global · MERL & AI Innovator

● En-Intrapreneur

● Driving evaluation ecosystems through People, Process & Product

LinkedIn: [linkedin.com/in/shashidharsa](https://www.linkedin.com/in/shashidharsa)

Also from HubForge

HF.

PLAYBOOK

**Consulting in
Monitoring &
Evaluation
with AI**

YOU ARE READING

HF.

PLAYBOOK

**Designing
Surveys
and Forms with
AI**

HF.

PLAYBOOK

**Claude
Dashboard
Playbook**

HF.

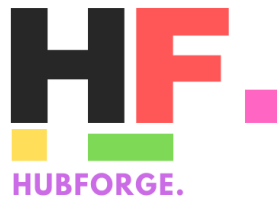
PLAYBOOK

**Designing
Dashboards
with AI**

Designing Surveys and Forms with AI Playbook: the SURVEY method for monitoring forms and evaluation surveys.

Claude Dashboard Playbook and **Designing Dashboards with AI Playbook:** dashboards hands-on in Claude, and the DESIGN loop for any tool.

Not legal or tax advice. Check contract, registration and tax rules with a qualified adviser where you work.



THE HUBFORGE PLAYBOOK SERIES

Consulting in Monitoring & Evaluation with AI Playbook



Win the right assignments. Contract them well. Deliver evidence that holds up, and see it used.



Consulting in Monitoring & Evaluation with AI Playbook

Made by HubForge Global Impact Network Foundation, which builds evidence systems and ecosystems for the MERL sector. Share it freely with credit.

write to hubforgeglobal@gmail.com

visit www.hubforgeglobal.com

EDITION 3 • 2026