



Kirkpatrick Evaluation Plan Template

Updated August 8, 2026

Start with Level 4, then work backward

Most evaluation plans fail because they start at Level 1 and never get past it. This template runs the other direction. Fill it in top-down, starting with the business result.

1. Name the business result the program supports. Get it from the stakeholder who owns the number, not from the training request.
2. Define the critical behaviors that drive that result. These become your Level 3 targets.
3. Decide what people must know, do, and believe to perform those behaviors. These become your Level 2 targets.
4. Design the experience so people can engage with it. Level 1 checks whether the conditions for learning were in place.

Why backward

If you start at Level 1, you measure what is easy instead of what matters. Starting at Level 4 forces every lower level to justify its place in the plan.

Program name: _____

Business result this program supports: _____

Stakeholder who owns that result: _____

Evaluation owner: _____

Baseline data available (Y/N) and where it lives: _____

Planning tables

Fill in one row per success criterion. Work down the page: Level 4 first, Level 1 last.

Level 4: Results

What business outcome will move, and what leading indicators will you track along the way? Common methods: business dashboards, operational reports, quality scores, customer metrics.

Success criteria	Planned method	Timing (month/quarter)	Owner (name)

Level 3: Behavior

Which critical on-the-job behaviors will you measure, and how? Common methods: manager observation, self-report pulse surveys, work product review, system data.

Success criteria	Planned method	Timing (30/60/90 day)	Owner (name)

Level 2: Learning

What must people know, be able to demonstrate, and feel confident and committed to do? Common methods: knowledge checks, skill demonstrations, role plays, simulations, confidence and commitment items.

Success criteria	Planned method	Timing (pre/during/post)	Owner (name)

Level 1: Reaction

Was the experience relevant, engaging, and applicable in the eyes of the learner? Common methods: post-session survey, pulse questions mid-program, facilitator observation.

Success criteria	Planned method	Timing (during/post)	Owner (name)

Question banks

Use these as written or swap in your own program language where you see brackets.

Level 1: Reaction survey items

Items 1 through 7 use a 1 to 5 agreement scale. Items 8 through 10 are open-ended.

1. The training was relevant to my job.
2. The content was organized in a way that was easy to follow.
3. The experience kept me engaged.
4. The pace was right for me.
5. I had enough opportunities to practice or participate.
6. I know how to apply what I learned to my work.
7. I would recommend this training to a colleague.
8. What was the most useful part of this training?
9. What would you change about this training?
10. What support do you need to apply this on the job?

Level 2: Learning prompts

Knowledge prompts

- You are in [realistic job situation]. What do you do first, and why?
- A colleague handles [situation] by doing [common mistake]. What would you do differently?
- Which of the following steps comes next in [process]? (Multiple choice with plausible wrong answers pulled from real errors.)
- Explain [concept] in your own words as if teaching a new hire.

Demonstration prompts

- Demonstrate [skill] in a role play or simulation while an observer scores against the rubric.
- Complete [task] in the practice environment without a job aid.
- Review this sample work product and identify what is wrong and how you would fix it.

Confidence and commitment prompts

- How confident are you that you can [critical behavior] on the job? (1 to 5)
- How committed are you to applying [critical behavior] in the next 30 days? (1 to 5)
- What could get in the way of applying this, and what support would remove that barrier?

Level 3: Behavior measures

Self-report items (send at 30, 60, and 90 days)

1. In the past 30 days, how often have you [critical behavior]? (Never / Sometimes / Often / Always)
2. Rate how successfully you have applied [critical behavior] on the job. (1 to 5)
3. What has helped you apply what you learned?
4. What has made it hard to apply what you learned?
5. What one change would help you apply it more often?

Level 4: ROE stakeholder questions

Ask these in the kickoff meeting and record the answers word for word.

1. What result does this program need to influence for you to call it a success?
2. How is that result measured today, and who owns the number?
3. What would good look like six to twelve months from now?
4. Which on-the-job behaviors most directly drive that result?
5. What else could move this number at the same time as the training?
6. What evidence would you find credible when we report progress?
7. How often do you want updates, and in what format?

Level 3 observation checklist

Use one copy of this page per person observed. Print this page again if you are observing more than one person.

Person observed: _____

Observer: _____

Date: _____

Critical behavior	Observed (Y/N)	Notes

Biggest barrier this person faces: _____

Follow-up action and owner: _____

ROE stakeholder kickoff checklist

Run this before design starts. Return on expectations is set in this meeting, not reported after the fact.

- ☐ Identify the stakeholder who owns the business result
- ☐ Schedule the kickoff before any design work begins
- ☐ Ask the seven ROE stakeholder questions and record the answers
- ☐ Agree on the leading indicators you will track between now and the result
- ☐ Confirm baseline data exists, or assign someone to collect it
- ☐ Define the critical behaviors with the stakeholder, not for them
- ☐ Agree on required drivers and name who provides each one
- ☐ Set the reporting cadence, format, and first report date
- ☐ Send a written summary of expectations back for sign-off

Required drivers

Training alone rarely changes behavior. Required drivers are the on-the-job systems that reinforce, encourage, reward, and monitor the critical behaviors: manager coaching, job aids, recognition, and dashboards that make the behavior visible. Assign an owner to each driver at kickoff. A program without drivers is a program betting on willpower.

30/60/90-day Level 3 follow-up checklist

Day 30

- ☐ Send the self-report pulse survey to all participants
- ☐ Ask managers to complete one check-in conversation per participant
- ☐ Review early leading indicator data against the baseline
- ☐ Pick the most common reported barrier and remove it
- ☐ Log results in the Level 3 planning table

Day 60

- ☐ Complete an observation round using the observation checklist
- ☐ Compare observation results to the 30-day self-report data
- ☐ Reinforce weak behaviors with a job aid, refresher, or practice activity
- ☐ Send the stakeholder a short progress update in the agreed format

Day 90

- ☐ Send the second self-report pulse survey
- ☐ Pull the leading indicator and business result data
- ☐ Draft the Level 4 progress summary for the stakeholder
- ☐ Decide with the stakeholder: sustain the program as is, adjust it, or expand it
- ☐ Schedule the next measurement point

Have any questions? Email me at devlin@devlinpeck.com