

AI Role Play Scenario Template

Design simulations that change behavior on real calls. Includes a scenario brief, persona card, weighted scoring rubric, difficulty ladder, calibration log, and launch checklist. Every field is fillable.

Companion to: AI Role Play Simulator: 2026 Guide for Training Leaders

How to use this template

- 1 Pick one moment that matters.**
Pull last quarter's QA scores, escalations, and low-CSAT contacts. Choose the call type where mistakes cost the most.
- 2 Write the scenario brief and persona card (page 2).**
Make the AI customer behave like your hardest real customers, not a quiz.
- 3 Build the rubric from your QA scorecard (page 3).**
Weight must-say, must-do, soft skills, and outcome. Weights total 100%.
- 4 Set four difficulty levels and calibrate (page 4).**
Have two QA reviewers score 10 practice attempts and compare with the AI.
- 5 Test with five tenured agents before launch.**
If the persona gives in faster than real customers do, make it tougher.

Worked example: duplicate charge dispute

Who they are	Small-business owner, customer for six years
Goal	Refund of a duplicate \$240 charge
Mood at start	Irritated, short on time
What they know	Has the bank statement; doesn't know the invoice number
Triggers	Being put on hold without explanation; hearing "policy"
Hidden objection	Plans to cancel if the refund takes more than three days
What earns trust	A clear timeline and a confirmation reference

Scenario brief and persona card

Fill in one card per scenario. Keep it short: the AI performs best with specific, concrete detail.

Scenario brief

Scenario name

e.g., Duplicate charge dispute

Learning objective

What the agent must be able to do after practice

Call type and channel

Voice, chat, or both

Systems and screens needed

Only the screens this moment uses

Who practices

Role, tenure, team

Successful outcome

What "resolved" looks like for the customer

Persona card

Who they are

Role, relationship with you, relevant history

Goal

What they want from this contact

Mood at start

How they sound in the first 30 seconds

What they know

Facts they have, and facts they don't

Triggers

What makes them angrier or less cooperative

Hidden objection

What they won't say unless the agent asks well

What earns trust

What the agent must say or do to turn the call

Opening line

The first thing the AI customer says

STEP 3

Scoring rubric

Start from your QA scorecard so practice and live evaluation measure the same things. Weights across the four parts must total 100%.

1. Must-say

Required statements, scored verbatim (disclosures, verification language)

Weight %

ITEM	WHAT GOOD LOOKS LIKE	POINTS

2. Must-do

Correct process steps and accurate entries in the system

Weight %

ITEM	WHAT GOOD LOOKS LIKE	POINTS

3. Soft skills

Empathy, confidence, clarity, pacing

Weight %

ITEM	WHAT GOOD LOOKS LIKE	POINTS

4. Outcome

Was the customer's issue resolved?

Weight %

ITEM	WHAT GOOD LOOKS LIKE	POINTS

Total weight % (must equal 100)

Pass score to certify / 100

Tip: score must-say items pass or fail. One missed required disclosure should fail the attempt in regulated scenarios.

STEP 4

Difficulty ladder, calibration, and launch

New hires start at level 1 and move up after reaching the pass score. Tenured agents can start at level 4.

Difficulty ladder

LEVEL	HOW THE PERSONA CHANGES AT THIS LEVEL	SCORE TO ADVANCE
1 Cooperative Patient, gives facts when asked		
2 Confused Unclear on the issue, mixes up details		
3 Frustrated Interrupts, pushes back on hold time		
4 Hostile Threatens to cancel or escalate		

Calibration log

Compare AI scores with two QA reviewers on 10 attempts before launch.

#	AI SCORE	REVIEWER 1	REVIEWER 2	LARGEST GAP	NOTES
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Launch checklist

- ☐ Persona tested against recordings of real difficult calls
- ☐ Rubric items match the QA scorecard
- ☐ Scenario owner named, with a review date after each product or policy change
- ☐ System replica uses fictional data only, with no customer PII
- ☐ AI and reviewer scores within the agreed range
- ☐ Learners told the customer is AI and that practice scores are for development

Build this scenario in minutes, with no code.

See the Mindtickle AI Role Play Simulator: mindtickle.com/ai-role-play-simulator