



Instructor Guide to Expectations, Goals, and Reference Points

Laboratory implementation, theoretical foundations,
experimental evidence, and variants

Behavioral and Experimental Economics Research Lab
Department of Economics



Revised edition | October 2026

<https://beer-lab.org/>

How to use this guide

This guide supports the selection, implementation, and discussion of activities on expectations and goals as reference points.

The basic edition separates outcome evaluation, goal evaluation, and effort provision; the core adds a monetary choice.

The extended edition provides stand-alone modules on near goals, saving, income targets, goal origin and difficulty, expectation precision, and multiple reference points.

Author

This guide was prepared by Stanislao Maldonado, research professor in the Department of Economics at Tecnológico de Monterrey, Monterrey Campus. For questions or suggestions about its content, please contact the author at stanislao.maldonado@tec.mx.

For the instructor

All activities are individual. References are assigned in balanced fashion within each session. Do not reveal other treatments before the activity ends: every participant must respond only to the expectation or goal shown on screen.

Contents

How to use this guide	i
1 Quick reference	1
1.1 Access through the updated laboratory website	1
1.2 Learning objectives	5
2 Theoretical foundation	7
2.1 Reference-dependent utility	7
2.2 Goals as reference points	7
2.3 Expectations, effort, and income targets	7
2.4 Origin, difficulty, and proximity	7
2.5 Savings goals, subgoals, and multiple references	8
3 Access and preparation	9
4 Participant sequence	11
5 Instructor report	26
5.1 Basic edition	31
5.2 Expectations	31
5.3 Goals	31
5.4 Effort	31
5.5 Core	31
5.6 Extended edition	31
5.7 Near goal	32
5.8 Saving and subgoals	32
5.9 Income target	32
5.10 Origin and difficulty	32
5.11 Expectation information	32
5.12 Multiple references	32
5.13 Complete diagnostic	32
5.14 Configuration catalog	33
5.15 Data, limitations, and troubleshooting	33
6 Expanded teaching manual for nonspecialist instructors	34
6.1 Conceptual map before opening the session	34
6.1.1 Unit of analysis	34
6.1.2 Benchmark prediction	34
6.1.3 Alternative mechanisms	35
6.2 Worked example and explanation	35
6.3 Reading the instructor report	35
6.4 Step-by-step debrief script	36

6.5	Comparable class plans	36
6.5.1	A 50-minute class	36
6.5.2	An 80-minute class	37
6.5.3	A 110-minute class	37
6.6	Exercises with answer guidance	37
6.6.1	1. Reconstructing the prediction	37
6.6.2	2. Auditing a percentage	37
6.6.3	3. Reading heterogeneity	37
6.6.4	4. Separating explanations	37
6.6.5	5. Interpreting no difference	38
6.6.6	6. Transferring carefully	38
6.7	Frequently asked questions and model answers	38
6.8	Diagnostic cases for interpretation practice	39
6.8.1	Agreement with the benchmark	39
6.8.2	Two groups with opposite behavior	39
6.8.3	A large difference with few units	40
6.8.4	A null difference	40
6.9	Explanatory sequence for a complete class	40
6.9.1	Step 1: pose the problem without labels	40
6.9.2	Step 2: reconstruct rules and incentives	40
6.9.3	Step 3: derive the benchmark	40
6.9.4	Step 4: record predictions	40
6.9.5	Step 5: reveal in layers	41
6.9.6	Step 6: work through one sequence	41
6.9.7	Step 7: compare institutions	41
6.9.8	Step 8: transfer and close	41
6.9.9	A short script to read aloud	41
6.10	Assessment bank with solutions	41
6.11	Advanced seminar: from evidence to mechanism	42
6.12	Quality control before teaching	43
6.12.1	Full rehearsal	43
6.12.2	Fifteen control questions	43
6.12.3	A micro-rehearsal of the debrief	43
6.12.4	After class	44

List of Figures

1.1	Current institutional website. Select CLASSROOM in the top navigation menu to open the teaching experiments. Captured on October 7, 2026.	1
1.2	Current catalog: search by title, topic, or JEL code. Then locate this family's card and select language and edition. Captured on October 7, 2026.	2
1.3	Demo contains only the configurations of this game family, in both languages and editions.	3
1.4	Rooms contains the four rooms for this family: basic and extended editions in Spanish and English.	4
1.5	Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.	4
3.1	Public entry page for the extended English room.	9
3.2	Links for an extended session with six participants.	10
4.1	Institutional welcome and laboratory rules.	11
4.2	Current instructions for the basic edition: initial resources, decision rules, and experimental units.	12
4.3	Current comprehension check with an incorrect answer, its specific explanation, and retained responses. Continuation 1/2.	13
4.5	The same outcome under different expectations.	14
4.6	The same performance under different goals.	15
4.7	Effort choice under different earnings expectations.	16
4.8	Safe option and lottery under an assigned goal.	17
4.9	An outcome just below or above a goal.	18
4.10	A total savings goal or weekly subgoal.	19
4.11	Work decision relative to an income target.	20
4.12	A self-set or assigned goal.	21
4.13	A moderate or stretch goal.	22
4.14	A point expectation or an interval.	23
4.15	Coexisting expectation and goal with different salience.	24
4.16	Explicit closing screen and ECU result when applicable.	25
5.1	Hidden report while responses are pending.	26
5.2	Complete report before disclosure.	26
5.3	Aggregate report after disclosure. Continuation 1/3.	27
5.6	Means, medians, and histograms by treatment.	30
5.7	Descriptive contrasts and ECU-balance distribution.	30

Quick reference

1.1 Access through the updated laboratory website

Start at <https://beer-lab.org/en/>. Select **CLASSROOM** in the top navigation menu; it opens <https://teaching.beer-lab.org/en/>. In the **Classroom experiments** catalog, use the search box to locate the activity by name. You can also filter by **Topic** or **JEL code**. Clear any filters left from an earlier search before looking for another game.

Each family's card distinguishes **Español** and **English** access links. The **Show extended version** checkbox selects that edition before following a link. Catalog language and session language are separate choices: verify both. Students use the room specified by the instructor; the instructor prepares and monitors the session through this family's own administrative pages.



Figure 1.1: Current institutional website. Select CLASSROOM in the top navigation menu to open the teaching experiments. Captured on October 7, 2026.

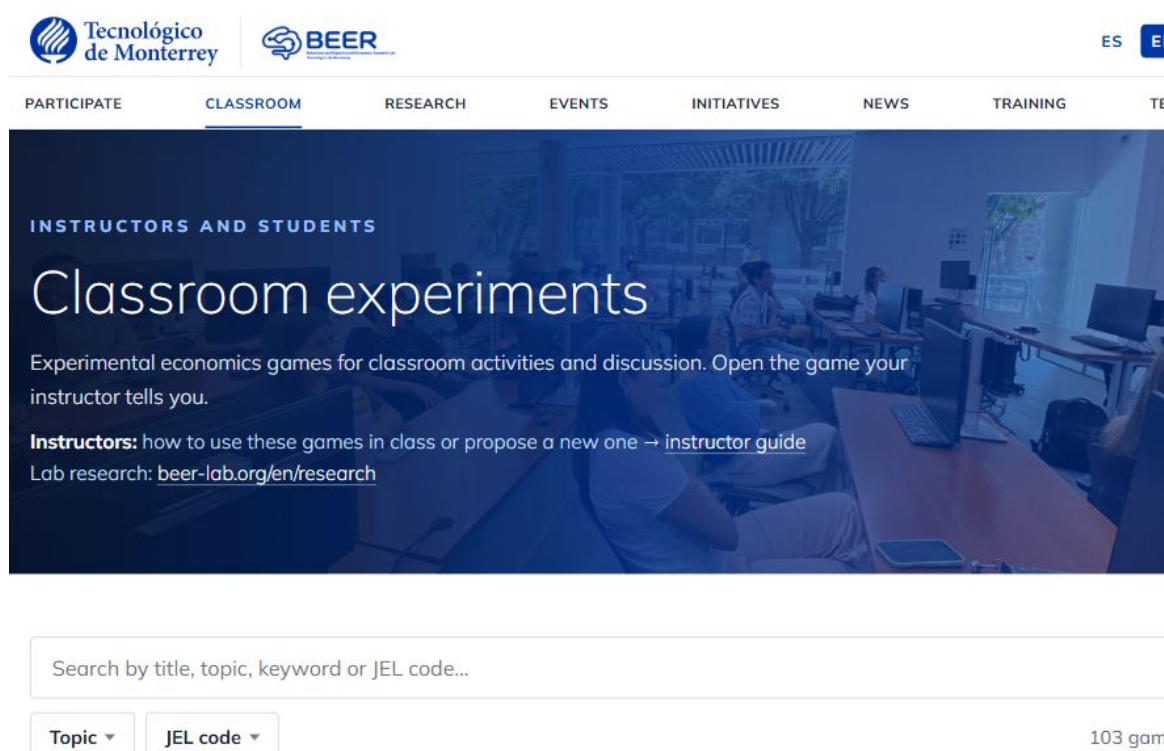


Figure 1.2: Current catalog: search by title, topic, or JEL code. Then locate this family's card and select language and edition. Captured on October 7, 2026.

Checking resources, comprehension, and participant requirements

Before opening the room, check the selected configuration and the number of students present. The session-creation form states the group size and minimum required by that configuration. An individual activity accepts any positive participant count; an interactive activity must satisfy its stated requirement: complete groups when size is fixed, or a minimum participant count when size is flexible. Requirements may change across treatments, for example when adding a third party or changing an institution's size. If you edit a grouping parameter, check the requirement for the edited configuration, rather than relying only on the demo's suggested count.

In the instructions, ensure that students can identify what each role receives, what they can decide, and how the outcome is calculated. An initial account, a distributable budget, and final earnings need not be the same amount. Refer to the resources panel and the example for the selected treatment. Do not carry the basic edition's endowment into a treatment that changes accounts or introduces additional positions.

ECU stands for *Experimental Currency Units*: the activity's unit of account. An outcome of 100 ECU does not automatically equal 100 pesos or 100 dollars. Conversion depends on the session configuration. Before announcing a payment, confirm with the laboratory whether money is paid, the conversion rate, and which decisions count for payment. In hypothetical modules, describe amounts as outcomes in the stated situation.

When a comprehension check is included, an incorrect response receives an explanation beside that question, identifying the submitted answer and the correct answer. Other answers remain in the form. Ask the student to return to the example, locate the relevant rule, and redo the calculation before correcting the response. Avoid revealing answers to the entire class for an individual error, and do not interpret a comprehension mistake as an economic choice in the game.

, metas y puntos de referencia — expectativas	Expectativas, metas y puntos de referencia / Expectations, Goals, and Reference Points — BEER Lab Seleccione una variante de este juego. Select a version of this game.
, metas y puntos de referencia — metas	
, metas y puntos de referencia — esfuerzo	
, metas y puntos de referencia — núcleo	
s, Goals, and Reference Points — expectations	
s, Goals, and Reference Points — goals	
s, Goals, and Reference Points — effort	
s, Goals, and Reference Points — core	
ferencia extendidos — meta cercana	
ferencia extendidos — meta de ahorro	
ferencia extendidos — ingreso objetivo	
ferencia extendidos — meta autoimpuesta	
ferencia extendidos — dificultad de la meta	
ferencia extendidos — información de expectativas	
ferencia extendidos — referencias múltiples	
ferencia extendidos — diagnóstico completo	
ference Points — near goal	
ference Points — savings goal	
ference Points — income target	
ference Points — self-set goal	
ference Points — goal difficulty	
ference Points — expectation information	
ference Points — multiple reference points	
ference Points — complete diagnostic	

Figure 1.3: Demo contains only the configurations of this game family, in both languages and editions.

Rooms

Current rooms:

[ES] Expectativas, metas y puntos de referencia
[EN] Expectations, Goals, and Reference Points
[ES] Expectativas, metas y puntos de referencia: extendido
[EN] Expectations, Goals, and Reference Points: extended

Figure 1.4: Rooms contains the four rooms for this family: basic and extended editions in Spanish and English.

Room: [EN] Expectations, Goals, and Reference Points

Create a new session

Session Config

[EN] Expectations, Goals, and Reference Points — expectations ▾

Number of participants

Individual decisions: any positive participant count is allowed.

Create

[Configure session](#) ⚙️

App sequence

reference_points_game (11 rounds)

Individual classroom activities about expectations and goals as reference points. The family includes outcome evaluation, effort, savings, labor supply, goal origin and difficulty, information precision, and competing references.

Figure 1.5: Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.

Downloading a session with several rounds

For the instructor

Open **Data** within the relevant session and use that session's CSV or CSV-for-Excel download link. The general export has one row per participant and columns for the different rounds. Ten rounds do not necessarily mean ten rows per person. When the family provides its own long-format export, observations may instead be organized by participant and round or another unit defined by the design.

Save the original file together with the configuration name, session code, and active-round count. Check that columns for the last active round are present and that the data belong to the selected session. Empty cells can represent inactive rounds, a pending decision, or a field that does not apply to a role; do not automatically recode them as zeros. To compare learning, preserve round and group identifiers alongside participant identifiers. Share only anonymous aggregates in class, and record session incidents before constructing derived tables.

Purpose	Show how expectations and goals alter evaluation, effort, and other decisions even when observable options are the same.
Participants	Any number; six or more provide balanced arms and useful denominators.
Basic edition	Expectations, goals, effort, or integrated core.
Extended edition	Near goal, saving, income target, goal origin, difficulty, information, multiple references, or complete diagnostic.
Interaction	Individual; balanced assignment to two treatments.
Payoff	Vignettes are hypothetical. Only the safe choice or lottery can generate an ECU balance.
Students	https://referencepoints.beer-lab.org/room/reference_points_en
Instructor	https://referencepoints.beer-lab.org/demo
Local test	http://localhost:8017/demo with <code>GAME_FAMILY=referencepoints</code> .
Experimental unit	Individual decisions; each person completes their own tasks.
Decision	Choose or value outcomes presented with the treatment's reference point.
Topics	reference points; expectations; goals; reference dependence; behavioral economics.
JEL	C91; D84; D91.

Attention

The `referencepoints.beer-lab.org` domain is a proposed route. The laboratory must create DNS, an isolated process, and a separate PostgreSQL database before classroom use. The local version was verified on port 8017.

1.2 Learning objectives

By the end of the session, students should be able to:

- define a reference point and distinguish it from an outcome's absolute level;
- explain how rational expectations can become endogenous reference points;
- show why the same outcome can be evaluated as a gain or loss under different expectations or goals;
- relate goals to satisfaction, effort, persistence, saving, and labor supply;
- distinguish a total goal from a subgoal, a self-set from an assigned goal, and a point forecast from a range;
- recognize that reference points can coexist and that their salience may change evaluation;
- read means, medians, proportions, denominators, and histograms without assigning individual parameters.

Theoretical foundation

2.1 Reference-dependent utility

Prospect theory evaluates outcomes as gains or losses relative to a reference and gives greater weight to losses than comparable gains ([Kahneman and Tversky, 1979](#)). In a general representation, utility combines consumption and change relative to the reference, $U(c | r) = u(c) + \mu(c - r)$. Knowing c is therefore insufficient; a prediction also requires r .

[Kőszegi and Rabin \(2006\)](#) propose that the reference can be determined by recent rational expectations. Expectations formed before an outcome then become part of the later experience. The basic modules hold the final outcome constant while changing only an expectation or goal, focusing discussion on $c - r$.

2.2 Goals as reference points

[Heath et al. \(1999\)](#) argue that goals inherit properties of the value function: performance below a goal is experienced as a loss, sensitivity diminishes with distance, and loss aversion can affect effort and persistence. The monetary module holds constant a safe 80 ECU and a 120/70 ECU lottery but assigns a goal of 70 or 100 ECU. The lower goal makes 80 a gain; the higher goal leaves it below the reference.

2.3 Expectations, effort, and income targets

In a real-effort experiment, [Abeler et al. \(2011\)](#) manipulate expected payment and find more effort under high expectations. The effort module reproduces the comparative logic with identical accumulated earnings and marginal compensation; only expected earnings vary.

Goals have also been applied to labor supply. [Crawford and Meng \(2011\)](#) model income and hours targets as rational expectations to explain taxi-driver decisions. Our vignette contrasts an extra hour below and above an income target. Accumulated income changes between arms, so the activity demonstrates the mechanism but is not a *ceteris-paribus* manipulation.

2.4 Origin, difficulty, and proximity

Self-set goals can operate as commitment devices. [Kaur et al. \(2015\)](#) show that workers demand commitment contracts and that those contracts raise output. Extremely difficult goals can instead

discourage effort; the difficulty module supports discussion of a potentially non-monotonic relation between aspiration and effort.

Proximity also matters. [Pope and Simonsohn \(2011\)](#) document higher effort just below round-number goals. The near-goal module compares 79 and 81 around a goal of 80. Because absolute outcomes differ by two points, interpret any discontinuity with that difference in mind.

2.5 Savings goals, subgoals, and multiple references

Dividing a broad goal into subgoals can make immediate action more salient. [Colby and Chapman \(2013\)](#) find that weekly subgoals can increase judged likelihood of saving. The savings module holds constant the total goal and shortfall while adding or omitting a weekly subgoal.

A person can simultaneously have an expectation and a goal. The multiple-reference module keeps both values and the outcome fixed while changing which reference appears first. It supports discussion of salience and competing references; it does not identify a single psychologically true reference for each participant.

Theoretical foundation

A classroom difference is descriptive evidence. A causal conclusion also requires correct assignment, sufficient sample size, no communication, and a prespecified analysis. A null result remains useful for discussing power, comprehension, heterogeneity, and whether an assigned reference was adopted.

Access and preparation

Start at <https://beer-lab.org/>, open <https://teaching.beer-lab.org/>, and select Expectations, Goals, and Reference Points once the family is deployed. Proposed routes are:

Basic Spanish: https://referencepoints.beer-lab.org/room/puntos_referencia_es

Basic English: https://referencepoints.beer-lab.org/room/reference_points_en

Extended Spanish: https://referencepoints.beer-lab.org/room/puntos_referencia_extendido

Extended English: https://referencepoints.beer-lab.org/room/reference_points_extended

Instructors: <https://referencepoints.beer-lab.org/demo>

The isolated instance lists 24 configurations and four rooms, with no other game families.

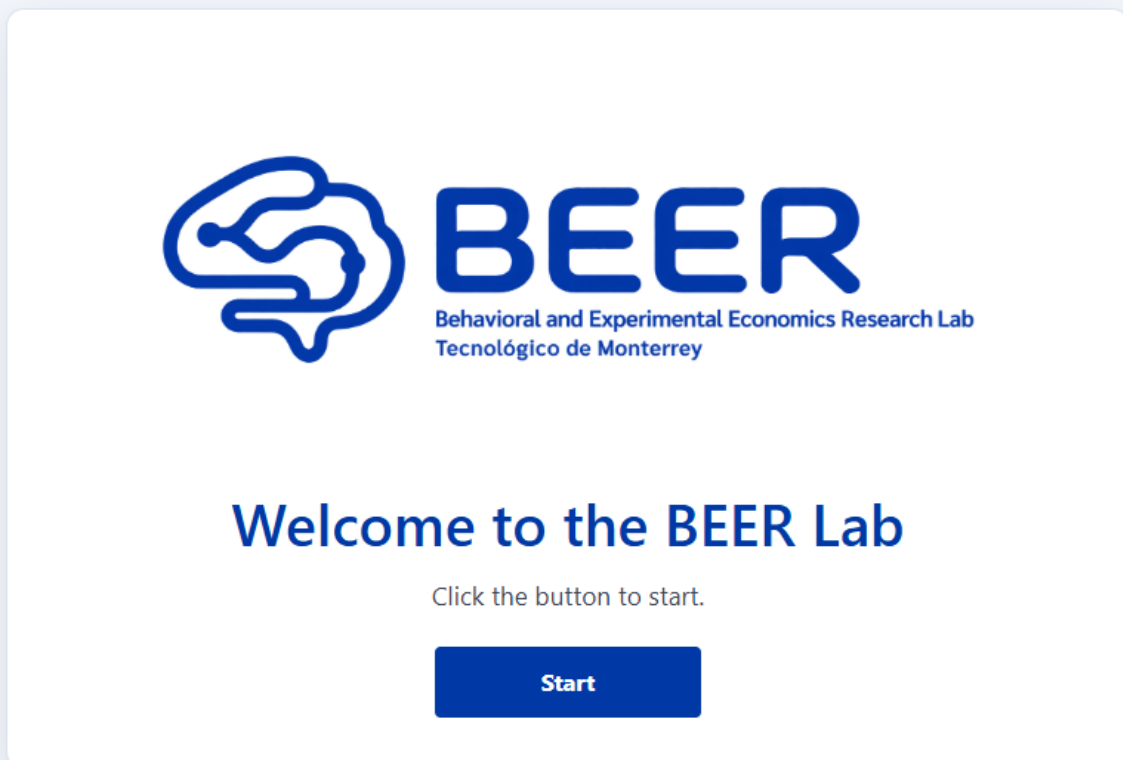


Figure 3.1: Public entry page for the extended English room.

Choose the configuration before distributing links. Open **Monitor** and **Report** on the instructor computer. Use even participant counts; six participants yield three observations per arm.

oTree Demo Sessions Rooms Data Server Check

[EN] Extended Reference Points — complete diagnostic: session '2qpz5aor' (demo)

New Links Monitor Data Payments Report Description

Below are temporary links for testing and demonstration. To launch a real study, either create persistent links by setting up a [room](#), or create a session through the [sessions page](#).

Choose one of the following options to play the session:

Combined views

Split-screen view [Launch](#) Grid view [Launch](#)

Session-wide testing link

Open the below link in up to 6 browser tabs.

<http://127.0.0.1:8017/join/tarebudo>

Single-use links

Open each link in its own browser tab.

P1	http://127.0.0.1:8017/initializeParticipant/nknab9j3
P2	http://127.0.0.1:8017/initializeParticipant/y2x12aw9
P3	http://127.0.0.1:8017/initializeParticipant/3j7cpqw7
P4	http://127.0.0.1:8017/initializeParticipant/nf81pscy
P5	http://127.0.0.1:8017/initializeParticipant/wbjahus7
P6	http://127.0.0.1:8017/initializeParticipant/r7w5zuux

Figure 3.2: Links for an extended session with six participants.

Participant sequence



Welcome

The **Behavioral and Experimental Economics Research Lab** of the Department of Economics at Tecnológico de Monterrey is a space designed to learn about economic decision making through controlled activities.

In this activity, you will examine how expectations and goals influence outcome evaluation and decisions. Read each situation carefully and respond independently.

General laboratory rules

- Remain quiet throughout the session.
- Do not talk or communicate with other participants.
- Make your decisions individually and privately.
- Follow the instructions shown on the screen.
- Do not refresh, close, or leave this page.
- If you have a question or technical problem, raise your hand and wait for assistance.

☐ I have read these rules and agree to follow them.

Start

Figure 4.1: Institutional welcome and laboratory rules.



Instructions

1 decision(s)

Resources and unit of account

ECU stands for Experimental Currency Units. It is the unit of account for this activity.

This session does not convert ECU into real money.

Basic version · Expectations as reference points

Each situation presents an outcome and a relevant reference, such as an expectation, a goal, or an income target. References may vary across participants.

1

Situations and treatments

The system may assign different versions of the same situation. Use only the information shown on your screen.

2

Private decisions

Do not compare answers or try to coordinate. The activity is individual and does not require waiting for anyone else.

3

Decisions and payments

Most situations are hypothetical. Only a decision marked in green can determine an ECU balance.

Continue

Con tecnología de [oTree](#)

Figure 4.2: Current instructions for the basic edition: initial resources, decision rules, and experimental units.



Review the marked questions. Each shows your answer and an explanation; your other answers are retained.



Review the marked questions. Each shows your answer and an explanation; your other answers are retained.

Comprehension questions

Resources and unit of account

ECU stands for Experimental Currency Units. It is the unit of account for this activity.
This session does not convert ECU into real money.

Expectations as reference points

Answer correctly before proceeding to the decisions.

How should you respond?

- ☐ Individually and privately
- ☒ By comparing answers with other participants

Your answer: By comparing answers with other participants. The correct answer is Individually and privately.
Review the instructions: decisions are private, only green decisions can pay, and the reference may vary.

Which decisions can determine an ECU balance?

- ☒ Only those marked in green
- ☐ All situations

Can the expectation or goal vary across participants?

- ☒ Yes, the system may assign different versions
- ☐ No, everyone always receives the same reference

Check and continue

Con tecnología de [oTree](#)

Figure 4.4: Current comprehension check with an incorrect answer, its specific explanation, and retained responses. Continuation 2/2.

The interface visually separates the reference, in blue, from the observed outcome, in amber.

Outcome evaluation

Decision 1 of 11

Expectations

Expectations

Before learning the outcome, you expected to obtain 60 points. The final outcome was 80 points.

REFERENCE POINT

Prior expectation

60 points

OBSERVED SITUATION

Final outcome

80 points

How satisfied are you with the outcome?

This situation is hypothetical. Respond as you would if the decision were real.

Not at all satisfied

Very satisfied



50 points

Confirm decision

Con tecnología de [oTree](#)

Figure 4.5: The same outcome under different expectations.

Evaluation relative to a goal

Decision 2 of 11

Goals

Goals

A goal of 70 points was set before the task. Final performance was 80 points.

REFERENCE POINT

Goal

70 points

OBSERVED SITUATION

Final performance

80 points

How satisfied are you with your performance?

This situation is hypothetical. Respond as you would if the decision were real.

Not at all satisfied

Very satisfied

50 points**Confirm decision**Con tecnología de [oTree](#)

Figure 4.6: The same performance under different goals.

Expectations and additional effort

Decision 3 of 11

Effort provision

Effort provision

You have accumulated 70 ECU. Before starting, you expected to finish with 70 ECU. You may classify up to 10 additional records; each pays 3 ECU.

REFERENCE POINT

Expected earnings

70 ECU**OBSERVED SITUATION**

Accumulated earnings

70 ECU

How many additional records do you want to classify?

This situation is hypothetical. Respond as you would if the decision were real.

None

10 records

**0** records**Confirm decision**Con tecnología de [oTree](#)

Figure 4.7: Effort choice under different earnings expectations.

Goal and ECU decision

Decision 4 of 11

Monetary decision

Monetary decision

Your goal for this decision is to finish with at least 70 ECU. The options are identical for everyone.

REFERENCE POINT

Goal

70 ECU

OBSERVED SITUATION

Balance before choosing

Not determined

Which option do you choose?

This decision is marked in green and can determine your ECU balance.

Option A



Receive 80 ECU for certain

- Final balance: 80 ECU.

Option B



Choose a lottery

- 50% chance of 120 ECU.
- 50% chance of 70 ECU.

Confirm decisionCon tecnología de [oTree](#)

Figure 4.8: Safe option and lottery under an assigned goal.

Outcome near a goal

Decision 5 of 11

Near goals

Near goals

Your goal was 80 points and you obtained 81. You may complete an additional unpaid practice task.

REFERENCE POINT

Goal

80 points**OBSERVED SITUATION**

Outcome

81 points

How willing are you to complete the additional practice?

This situation is hypothetical. Respond as you would if the decision were real.

Not at all willing

Very willing

**50** points**Confirm decision**Con tecnología de [oTree](#)

Figure 4.9: An outcome just below or above a goal.

Savings goal

Decision 6 of 11

Goals and subgoals

Goals and subgoals

You have saved 240 ECU toward a semester goal of 300 ECU. In addition, your subgoal for this week is to save the remaining 60 ECU. Today you receive 100 ECU that you may spend or save.

REFERENCE POINT

Weekly goal

60 ECU

OBSERVED SITUATION

Gap to semester goal

60 ECU

How many of today's 100 ECU do you want to save?

This situation is hypothetical. Respond as you would if the decision were real.

0 ECU

100 ECU

50 ECU**Confirm decision**Con tecnología de [oTree](#)

Figure 4.10: A total savings goal or weekly subgoal.

Income target and work decision

Decision 7 of 11

Income target

Income target

A driver set a daily income target of 100 ECU. The driver has earned 90 ECU and must decide about one additional hour.

REFERENCE POINT

Income target

100 ECU**OBSERVED SITUATION**

Accumulated earnings

90 ECU

What should the driver do?

This situation is hypothetical. Respond as you would if the decision were real.

Option A**End the shift**

- No additional effort.
- Final earnings: 90 ECU.

Option B**Work one more hour**

- One hour of additional effort.
- Expected additional earnings: 20 ECU.

Confirm decisionCon tecnología de [oTree](#)

Figure 4.11: Work decision relative to an income target.

Origin of the goal

Decision 8 of 11

Self-set goals

Self-set goals

Before starting, the platform assigned you a goal of completing 8 tasks. You have completed 4 and may complete up to 10 more.

REFERENCE POINT

Assigned goal

8 tasks

OBSERVED SITUATION

Completed tasks

4 tasks

How many additional tasks do you want to complete?

This situation is hypothetical. Respond as you would if the decision were real.

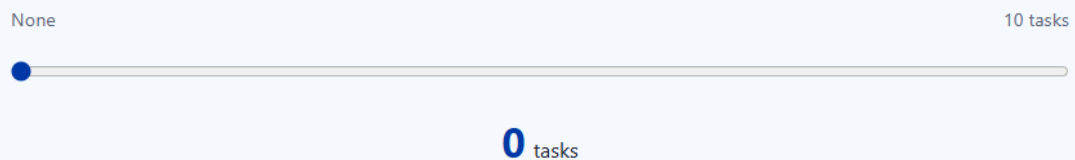
**Confirm decision**Con tecnología de [oTree](#)

Figure 4.12: A self-set or assigned goal.

Goal difficulty

Decision 9 of 11

Difficulty and effort

Difficulty and effort

You have completed 4 tasks. The stated goal is to complete 12 in total. You may complete up to 10 additional tasks; each requires the same effort.

REFERENCE POINT

Goal

12 tasks

OBSERVED SITUATION

Completed tasks

4 tasks

How many additional tasks do you want to complete?

This situation is hypothetical. Respond as you would if the decision were real.

None

10 tasks

**0** tasks**Confirm decision**Con tecnología de [oTree](#)

Figure 4.13: A moderate or stretch goal.

Expectation precision

Decision 10 of 11

Expectation information

Expectation information

Before receiving the outcome, the available forecast indicated a range from 70 to 110 points, with midpoint 90. The final outcome was 80.

REFERENCE POINT

Forecast

70–110 points

OBSERVED SITUATION

Outcome

80 points

How satisfied are you with the outcome?

This situation is hypothetical. Respond as you would if the decision were real.

Not at all satisfied

Very satisfied

50 points**Confirm decision**Con tecnología de [oTree](#)

Figure 4.14: A point expectation or an interval.

Conflicting expectation and goal

Decision 11 of 11

Multiple reference points

Multiple reference points

The goal was 90 points and you expected 70. The final outcome was 80. When evaluating it, consider first the goal.

REFERENCE POINT

Highlighted goal

90 points

OBSERVED SITUATION

Outcome (expectation: 70)

80 points

How satisfied are you with the outcome?

This situation is hypothetical. Respond as you would if the decision were real.

Not at all satisfied

Very satisfied

50 points**Confirm decision**Con tecnología de [oTree](#)

Figure 4.15: Coexisting expectation and goal with different salience.



Activity completed

Your decisions were recorded successfully.

ECU decision outcome

Selected decision

Goal and ECU decision

Chosen option

Choose a lottery

Balance

70 ECU

Remain quiet and follow the instructor's directions.

Con tecnología de [oTree](#)

Figure 4.16: Explicit closing screen and ECU result when applicable.

Instructor report

Aggregates remain hidden while responses are pending. Once everyone finishes, the instructor controls disclosure.

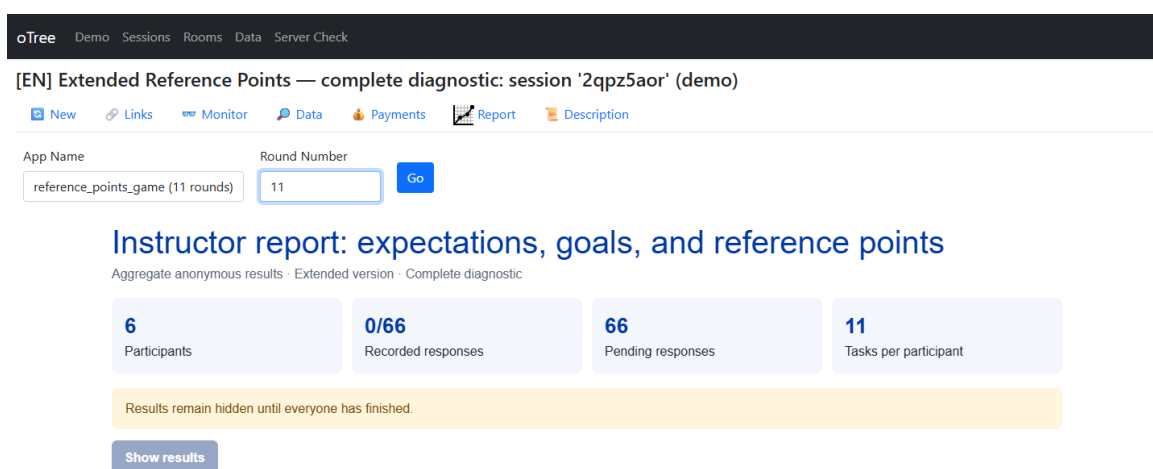


Figure 5.1: Hidden report while responses are pending.

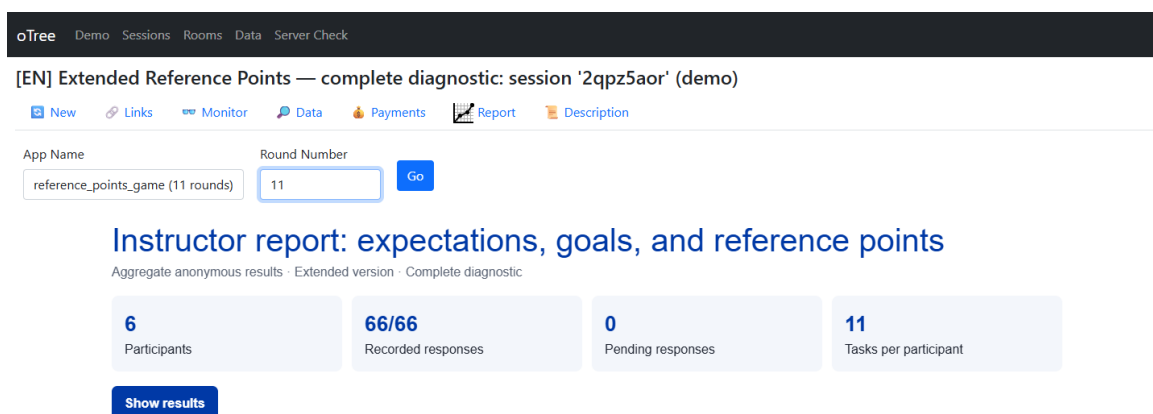


Figure 5.2: Complete report before disclosure.

The report shows option-B shares, 0–100 response distributions, effort from 0–10, saving from 0–100 ECU, arm contrasts, and balances of 70, 80, or 120 ECU. Every row reports its denominator, and no identifiers appear.

[EN] Extended Reference Points — complete diagnostic: session '2qpz5aor' (demo)

New Links Monitor Data Payments Report Description

App Name

Round Number

reference_points_game (11 rounds)

11

Go

Instructor report: expectations, goals, and reference points

Aggregate anonymous results · Extended version · Complete diagnostic

6

Participants

66/66

Recorded responses

0

Pending responses

11

Tasks per participant

Hide results

Binary decisions

Percentage choosing option B in each treatment.

Selected decision	Treatment	n	%
Goal and ECU decision	Goal: 70	2/3	66,7%
Goal and ECU decision	Goal: 100	1/3	33,3%
Income target and work decision	Below income target	1/3	33,3%
Income target and work decision	Above income target	2/3	66,7%

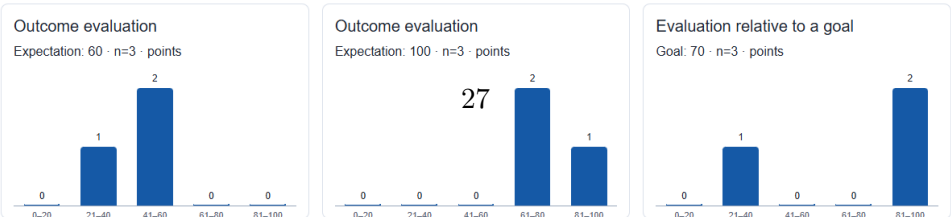
Quantitative responses

Means and medians by treatment; every row reports its own denominator.

Selected decision	Treatment	n	Mean	Median
Outcome evaluation	Expectation: 60	3	43 points	43 points
Outcome evaluation	Expectation: 100	3	76 points	76 points
Evaluation relative to a goal	Goal: 70	3	72 points	83 points
Evaluation relative to a goal	Goal: 90	3	61 points	61 points
Expectations and additional effort	Expected earnings: 70	3	5,7 tasks	6 tasks
Expectations and additional effort	Expected earnings: 100	3	5,3 tasks	5 tasks
Outcome near a goal	One point below goal	3	82 points	82 points
Outcome near a goal	One point above goal	3	46,3 points	45 points
Savings goal	Total goal only	3	63,3 ECU	59 ECU
Savings goal	Weekly subgoal	3	41,7 ECU	33 ECU
Origin of the goal	Self-set goal	3	6,3 tasks	9 tasks
Origin of the goal	Assigned goal	3	3,7 tasks	2 tasks
Goal difficulty	Moderate goal	3	1,3 tasks	1 tasks
Goal difficulty	Stretch goal	3	7 tasks	9 tasks
Expectation precision	Point forecast	3	54,3 points	47 points
Expectation precision	Range forecast	3	50,7 points	58 points
Conflicting expectation and goal	Expectation highlighted	3	65 points	65 points
Conflicting expectation and goal	Goal highlighted	3	54 points	43 points

Distributions

Histograms by task and treatment.



BE

nts

Selected decision	Treatment	n	Mean	Median
Outcome evaluation	Expectation: 60	3	43 points	43 points
Outcome evaluation	Expectation: 100	3	76 points	76 points
Evaluation relative to a goal	Goal: 70	3	72 points	83 points
Evaluation relative to a goal	Goal: 90	3	61 points	61 points
Expectations and additional effort	Expected earnings: 70	3	5,7 tasks	6 tasks
Expectations and additional effort	Expected earnings: 100	3	5,3 tasks	5 tasks
Outcome near a goal	One point below goal	3	82 points	82 points
Outcome near a goal	One point above goal	3	46,3 points	45 points
Savings goal	Total goal only	3	63,3 ECU	59 ECU
Savings goal	Weekly subgoal	3	41,7 ECU	33 ECU
Origin of the goal	Self-set goal	3	6,3 tasks	9 tasks
Origin of the goal	Assigned goal	3	3,7 tasks	2 tasks
Goal difficulty	Moderate goal	3	1,3 tasks	1 tasks
Goal difficulty	Stretch goal	3	7 tasks	9 tasks
Expectation precision	Point forecast	3	54,3 points	47 points
Expectation precision	Range forecast	3	50,7 points	58 points
Conflicting expectation and goal	Expectation highlighted	3	65 points	65 points
Conflicting expectation and goal	Goal highlighted	3	54 points	43 points

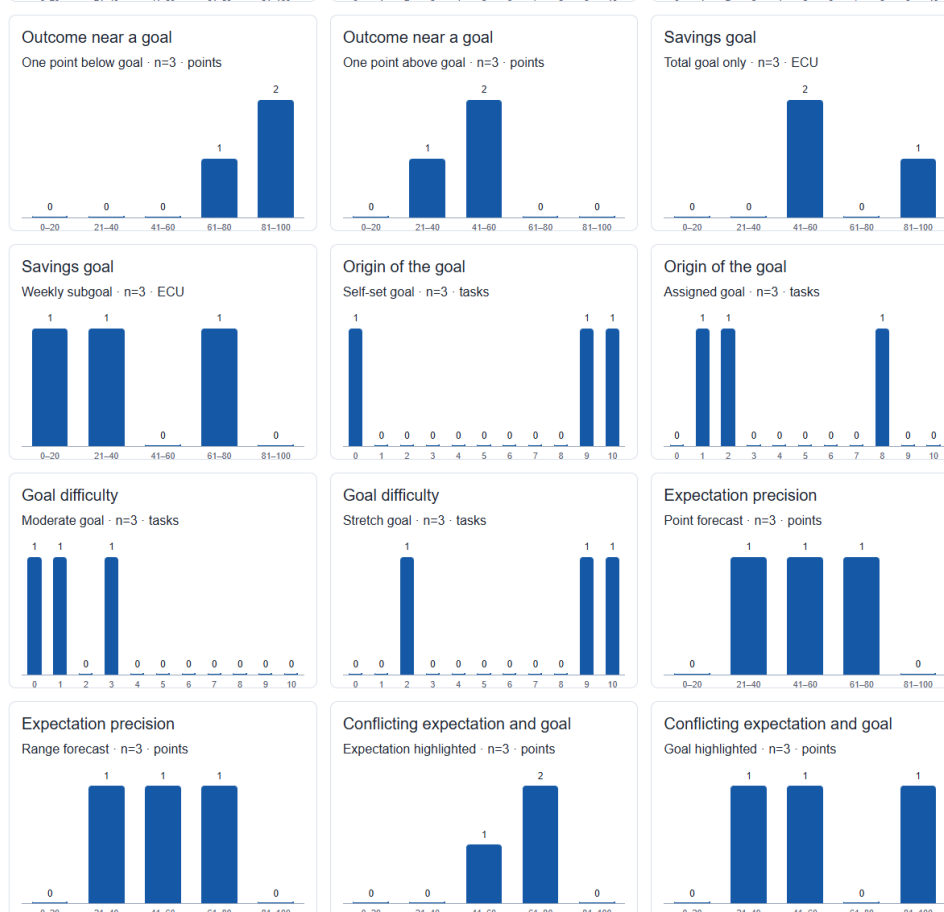
Distributions

Histograms by task and treatment.



Descriptive contrasts

Differences between treatments. They are descriptive and do not replace statistical inference.



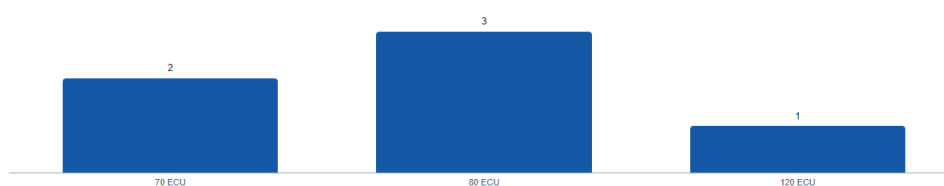
Descriptive contrasts

Differences between treatments. They are descriptive and do not replace statistical inference.

Selected decision	Measure	Contrast	Difference
Outcome evaluation	Mean	Expectation: 100 – Expectation: 60	33 points
Evaluation relative to a goal	Mean	Goal: 90 – Goal: 70	-11 points
Expectations and additional effort	Mean	Expected earnings: 100 – Expected earnings: 70	-0,4 tasks
Goal and ECU decision	% option B	Goal: 100 – Goal: 70	-33,4 p.p.
Outcome near a goal	Mean	One point above goal – One point below goal	-35,7 points
Savings goal	Mean	Weekly subgoal – Total goal only	-21,6 ECU
Income target and work decision	% option B	Above income target – Below income target	33,4 p.p.
Origin of the goal	Mean	Assigned goal – Self-set goal	-2,6 tasks
Goal difficulty	Mean	Stretch goal – Moderate goal	5,7 tasks
Expectation precision	Mean	Range forecast – Point forecast	-3,6 points
Conflicting expectation and goal	Mean	Goal highlighted – Expectation highlighted	-11 points

Distribution of ECU balances

Recorded monetary decisions: 6.



Suggested interpretation: compare the full distributions and sample sizes first; then discuss whether the reference point changed how the same situation was evaluated.

Figure 5.5: Aggregate report after disclosure. Continuation 3/3.

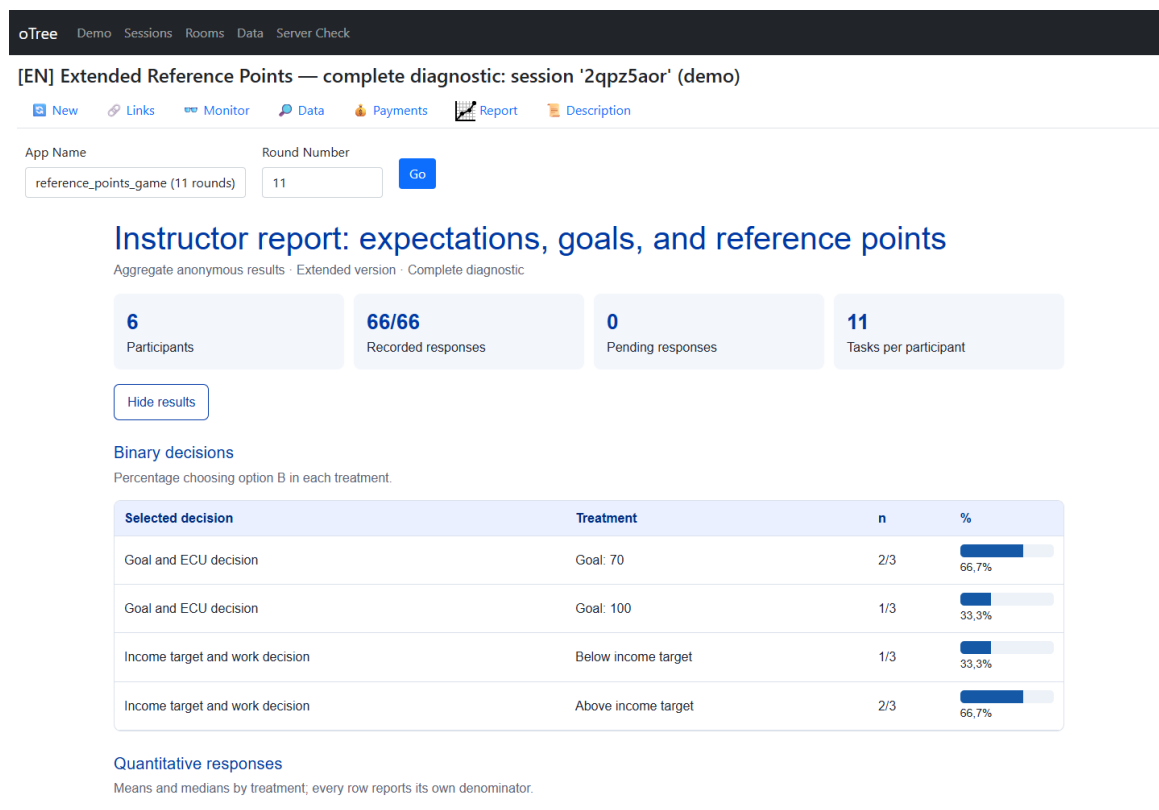


Figure 5.6: Means, medians, and histograms by treatment.

Descriptive contrasts

Differences between treatments. They are descriptive and do not replace statistical inference.

Selected decision	Measure	Contrast	Difference
Outcome evaluation	Mean	Expectation: 100 – Expectation: 60	33 points
Evaluation relative to a goal	Mean	Goal: 90 – Goal: 70	-11 points
Expectations and additional effort	Mean	Expected earnings: 100 – Expected earnings: 70	-0,4 tasks
Goal and ECU decision	% option B	Goal: 100 – Goal: 70	-33,4 p.p.
Outcome near a goal	Mean	One point above goal – One point below goal	-35,7 points
Savings goal	Mean	Weekly subgoal – Total goal only	-21,6 ECU
Income target and work decision	% option B	Above income target – Below income target	33,4 p.p.
Origin of the goal	Mean	Assigned goal – Self-set goal	-2,6 tasks
Goal difficulty	Mean	Stretch goal – Moderate goal	5,7 tasks
Expectation precision	Mean	Range forecast – Point forecast	-3,6 points
Conflicting expectation and goal	Mean	Goal highlighted – Expectation highlighted	-11 points

Distribution of ECU balances

Recorded monetary decisions: 6.



Suggested interpretation: compare the full distributions and sample sizes first; then discuss whether the reference point changed how the same situation was evaluated.

Figure 5.7: Descriptive contrasts and ECU-balance distribution.

Feedback script

1. Before disclosure, write the standard prediction: if only the absolute outcome matters, the reference should not alter evaluation or choice.
2. Elicit separate predictions for satisfaction, effort, and lottery choice.
3. Reveal distributions first; ask for direction, magnitude, and denominator before offering theory.
4. Ask whether the same outcome was a gain or a loss relative to each reference.
5. Compare modules that hold outcomes constant with income target and near goal, where the state also changes.
6. Discuss reference adoption, heterogeneity, and power before interpreting a null effect.
7. Close with applications to work, saving, and goal design, including risks from unattainable or manipulative goals.

5.1 Basic edition

5.2 Expectations

The outcome stays at 80 while expectations are 60 or 100. Lower satisfaction under the high expectation is consistent with expectation-dependent evaluation.

5.3 Goals

Performance stays at 80 while goals are 70 or 90. Compare the mean, median, and distribution rather than only the sign of the contrast.

5.4 Effort

Accumulated earnings remain 70 ECU and each additional record pays 3 ECU; the expectation is 70 or 100. More records in the second arm are consistent with effort to reduce a loss relative to expectations.

5.5 Core

The core combines all three modules and the monetary choice. In the final task, options are identical and only the goal varies. Allow 15–20 minutes for responses and 20–30 minutes for discussion.

5.6 Extended edition

5.7 Near goal

The module compares 79 and 81 around a goal of 80. Willingness to practice may rise just below the goal, but interpretation must acknowledge that the outcome also differs.

5.8 Saving and subgoals

The total goal is 300 ECU, prior saving is 240, and 100 is available. One arm adds a weekly subgoal of 60. Compare amounts and clustering at 60.

5.9 Income target

Earnings are 90 or 110 relative to a target of 100. Continuing has the same cost and expected return, but the starting point changes.

5.10 Origin and difficulty

Origin keeps an eight-task goal and changes whether it was chosen or assigned. Difficulty keeps four completed tasks and compares total goals of six and twelve.

5.11 Expectation information

Outcome and midpoint remain fixed, but the expectation is stated as exactly 90 or an interval of 70–110. The contrast asks whether precision changes the reaction to a shortfall.

5.12 Multiple references

Expectation 70, goal 90, and outcome 80 remain fixed. Only the highlighted reference changes. Use it to discuss coexistence and salience, without assigning a unique cognitive process.

5.13 Complete diagnostic

The battery contains eleven decisions. Allow 25–35 minutes. For an ordinary class, prefer a focal module to reduce fatigue and multiple comparisons.

5.14 Configuration catalog

Module	Spanish / English configuration
Expectations	puntos_referencia_es_expectativas / reference_points_en_expectations
Goals	puntos_referencia_es_metas / reference_points_en_goals
Effort	puntos_referencia_es_esfuerzo / reference_points_en_effort
Core	puntos_referencia_es_nucleo / reference_points_en_core
Near goal	puntos_referencia_extendido_meta_cercana / reference_points_extended_near_goal
Saving	puntos_referencia_extendido_meta_ahorro / reference_points_extended_savings_goal
Income target	puntos_referencia_extendido_ingreso_objetivo / reference_points_extended_income_target
Self-set goal	puntos_referencia_extendido_meta_autoimpuesta / reference_points_extended_self_set_goal
Difficulty	puntos_referencia_extendido_dificultad_meta / reference_points_extended_goal_difficulty
Information	puntos_referencia_extendido_informacion_expectativas / reference_points_extended_expectation_information
Multiple references	puntos_referencia_extendido_referencias_multiples / reference_points_extended_multiple_reference_points
Complete diagnostic	puntos_referencia_extendido_diagnostico_completo / reference_points_extended_full_diagnostic

5.15 Data, limitations, and troubleshooting

The export contains one row per participant and decision: task, module, treatment, response, monetary eligibility, lottery outcome, and balance. Do not average distinct tasks as if they measured a single parameter. The scales capture responses within a situation and do not estimate individual loss-aversion coefficients.

Effort tasks ask how many activities a participant wishes to complete but do not implement a subsequent real-effort task; they are stated choices. The income-target and near-goal modules also change the observed outcome. State these features explicitly during feedback.

If *Please wait* appears, inspect Monitor: this family never waits for another participant. If a link fails after a database rebuild, create a new session. If Demo lists other games, stop the instance and check `GAME_FAMILY=referencepoints`. If the report remains locked, confirm that each person used a distinct link and reached the closing page.

Expanded teaching manual for nonspecialist instructors

This part brings together what an instructor needs to turn Reference points into a learning experience. It complements the operational material above. It can be read before the first session and reused as a script during the discussion of results. The central question is how expectations, goals, and initial states change the evaluation of the same outcome. The guide always separates four levels: rules, benchmark prediction, observed behavior, and interpretation. Moving directly from a bar in a chart to a psychological explanation confuses those levels.

6.1 Conceptual map before opening the session

6.1.1 Unit of analysis

The relevant unit is each participant-item observation together with the assigned reference point. One data row is not always one independent person: the same person may decide repeatedly, and interacting participants share information, payoffs, or history. The instructor should therefore state what enters the numerator, what enters the denominator, and whether observations are people, decisions, pairs, groups, or rounds. This prevents thirty repeated decisions from being presented as thirty independent participants.

The task is to evaluate satisfaction, effort, saving, or continuation under different goals and expectations. Before presenting theory, ask students to restate the rule and identify what information is available when they decide. A correct calculation does not guarantee understanding of the strategic or behavioral structure. Check available actions, information, payoffs, order, and feedback timing separately.

6.1.2 Benchmark prediction

The starting point is: an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes. The benchmark is a disciplined comparison, not a moral statement about what a person ought to do. First write the assumptions that support it, then derive the action or pattern, and only then compare it with the class. If information, repetition, incentives, matching, or the reference point changes, derive the prediction again instead of carrying it over automatically.

Keep equilibrium, efficiency, and fairness separate. An action can be a best response and generate a poor joint outcome; another action can raise surplus and remain vulnerable to unilateral deviation. Likewise, an equal outcome need not be efficient, and an efficient distribution does not reveal

the process that produced it. Asking the class to label each statement as positive, normative, or strategic reduces confusion.

6.1.3 Alternative mechanisms

The literature and treatments permit discussion of gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references. These mechanisms are not interchangeable labels. Two can predict the same baseline action and separate only when a treatment changes. The informative comparison therefore asks not only how much behavior occurred but also what experimental change should increase, reduce, or leave the pattern unchanged if one mechanism matters.

Organize discussion in three columns. Write the observed fact in the first, such as a share, median, or trajectory. List compatible mechanisms in the second. Record the additional evidence needed to distinguish them in the third. The third column is essential because a classroom session rarely identifies a stable preference or causal relationship by itself.

6.2 Worked example and explanation

Two people receive 80 points. One expected 60 and the other 100. The outcome is identical, but the first exceeds the reference by 20 and the second falls 20 short. A satisfaction difference is consistent with expectation dependence. It does not by itself identify loss aversion because surprise, attribution, and expectation credibility may also differ.

Solve the example in six steps. First list actions and possible states. Second calculate payoffs or consequences without interpreting motives. Third derive the theoretical benchmark under its assumptions. Fourth identify the minimum observation needed to evaluate that benchmark. Fifth compare treatments while stating what is held constant and what changes. Sixth formulate a limited conclusion: “the pattern is consistent with” is more accurate than “the experiment proves” when rival explanations remain.

Then change one parameter and ask for a new prediction before calculating. This pause turns the example into a reasoning tool. It also reveals whether students memorized an answer or understood the mechanism. If answers diverge, ask each group to identify the assumption producing the difference.

6.3 Reading the instructor report

The central measures are difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference. Write the denominator next to every indicator. A rate can divide by eligible participants, completed decisions, formed pairs, groups, or rounds; these alternatives are not equivalent. If records are incomplete, show how many were excluded and why.

Use this reading order:

1. Verify participants, groups, and completed decisions. Do not interpret percentages before checking coverage.
2. Inspect the full distribution. A mean can hide two groups with opposing responses or a concentration at focal values.
3. Compare the median, range, and theoretically relevant categories. Explain whether the graph

counts people or decisions.

4. Inspect time paths only when rounds are comparable. Changes in matching or treatment must be marked.
5. Compare treatments with absolute differences and denominators, not merely with bars of different heights.
6. Select one anonymous case or sequence connecting rules and the aggregate. Never use it to identify or ridicule a student.

A visible difference in a small class is descriptive evidence from that session. Do not call it a causal effect without appropriate assignment, adequate size, and a comparison defined before seeing the data. Across classes, composition, time, and prior experience may change together. The pedagogical value lies in stating a prediction, confronting it with evidence, and asking what design would teach us more.

6.4 Step-by-step debrief script

Begin without revealing results. Ask for the benchmark prediction and which assumption students find most demanding. Record two quantitative class predictions. Next reveal the observation count and distribution, initially withholding the mean if it would distract from heterogeneity. Ask students to describe the chart without explaining causes.

In a second stage, show aggregate measures and ask: which observation contradicts the simplest prediction, which remains compatible with it, and what heterogeneity disappears in the average? Accept multiple explanations, but require each to yield a testable prediction. Introduce the variants: expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references. For each variant, identify the single changed element, predicted direction, and report measure to inspect.

Close by connecting the laboratory to performance goals, saving, sales, work, and satisfaction with outcomes. Point out which elements of the real setting are absent: institutions, selection, repetition, long-run consequences, or rich information. The analogy is useful when it isolates a mechanism; it becomes misleading if the class assumes that the laboratory reproduces the full context.

A verbal sequence that works well is: “first describe,” “now derive,” “then interpret,” and “finally ask what evidence is missing.” This discipline allows a nonspecialist instructor to lead a rigorous discussion without treating every departure as a bias or every match as definitive confirmation.

6.5 Comparable class plans

6.5.1 A 50-minute class

Use one mode. Spend five minutes on rules and comprehension, five on private predictions, ten on the activity, five collecting predictions about the aggregate, fifteen revealing the report in layers, and ten connecting the result to theory. The goal is not to cover every variant but to ensure that every student can explain how expectations, goals, and initial states change the evaluation of the same outcome. End with one written sentence separating outcome from mechanism.

6.5.2 An 80-minute class

Run the baseline and one variant from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references. Allow ten minutes for instructions, fifteen for each treatment, twenty for comparing distributions, and fifteen for theory and applications. Before the second treatment, request a directional prediction. If the result points the other way, do not dismiss it: inspect comprehension, composition, denominators, and descriptive power.

6.5.3 A 110-minute class

Combine experience, derivation, and design. After two treatments, give small groups a rival explanation and ask them to propose a third condition that distinguishes it. Each group must define the manipulated variable, outcome, unit of analysis, and prediction. Close by comparing proposals with existing variants and with performance goals, saving, sales, work, and satisfaction with outcomes. This stage turns the student from participant into experimental designer.

6.6 Exercises with answer guidance

6.6.1 1. Reconstructing the prediction

Question. State the benchmark and three indispensable assumptions. What observation would be a departure?

Guidance. Begin from an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes. Mention incentives, information, and horizon, and define the observation at the correct level. A departure is inconsistent with the full assumption set; it does not automatically identify its cause.

6.6.2 2. Auditing a percentage

Question. The report displays 60 percent. Propose two possible denominators and explain how interpretation changes.

Guidance. It could be 60 percent of eligible people or of completed decisions. If some people did not decide, the rates differ. A complete answer names numerator, denominator, exclusions, and aggregation level.

6.6.3 3. Reading heterogeneity

Question. Two classes have the same mean, but one is concentrated around it and the other has two modes. What does the instructor learn?

Guidance. The mean does not exhaust the distribution. The second class may combine different strategies or interpretations. Compare the median, dispersion, and categories linked to difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference before explaining mechanisms.

6.6.4 4. Separating explanations

Question. Select two mechanisms from this list: gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references. Design a contrast that produces different

predictions.

Guidance. Change one relevant element, hold others constant, and state each directional prediction. A variant from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references may help, but the justification must connect manipulation and report measure.

6.6.5 5. Interpreting no difference

Question. Two treatments produce similar means. Does the manipulation not matter?

Guidance. Not necessarily. There may be weak comprehension, a small sample, offsetting responses, manipulation failure, or effects in tails and dispersion. Inspect descriptive uncertainty and the full distribution before concluding.

6.6.6 6. Transferring carefully

Question. Relate the activity to performance goals, saving, sales, work, and satisfaction with outcomes and identify two contextual differences.

Guidance. A convincing application states the common mechanism and the limits. Real consequences, selection, experience, institutions, and repetition can change behavior. The conclusion should be an informed hypothesis rather than a universal prediction.

6.7 Frequently asked questions and model answers

Should theory be explained before play? In a first session, collect predictions without revealing the full solution. Explain rules, payoffs, and information. Then derive an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes. In a second session, reverse the order and assess whether theory improves prediction or only changes vocabulary.

Is there a correct answer? Separate a normative or equilibrium solution from a behavioral prediction. Theory supplies a conditional benchmark. The laboratory reveals what occurs and why a reasonable person may choose differently.

Does a departure prove irrationality? No. It may reflect omitted preferences, beliefs, error, weak comprehension, or a different institution. Check controls and incentives. Then treat gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references as hypotheses requiring further comparisons.

Why show the distribution? The same mean can arise from moderate consensus or two opposing groups. A distribution reveals focal values, boundaries, and heterogeneity. In Reference points, connect those shapes to difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference.

How many observations are enough? There is no universal threshold for debriefing. A small class is useful when described cautiously. Treatment comparisons require several independent units, clear assignment, and common protocols. Repeated decisions do not create new participants.

Can sessions be pooled? Yes, if parameters and definitions remain fixed. Retain session identifiers and inspect each cohort first. If instructions, composition, or incentives change, document the heterogeneity.

What if results contradict the literature? Check rules, treatment, comprehension, exclusions, and calculation. If no error exists, preserve the result. A class need not reproduce a published

estimate. Ask what design would distinguish chance, composition, and mechanism.

How can embarrassment be avoided? Project aggregates and anonymous sequences. Avoid personal labels. Describe actions and consequences. The laboratory studies decisions under specific rules; it does not diagnose traits.

Should the decision be graded? Do not grade agreement with a prediction. Assess understanding: derive the benchmark, calculate a measure, explain a limitation, and design a comparison. Experimental incentives and academic assessment serve different purposes.

How should a variant be chosen? Begin with the mechanism to separate. Select from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references, state a direction and measure, and hold other elements constant. Do not run more conditions than can be analyzed.

What conclusion is appropriate? Use three sentences: what was observed, which mechanism it is consistent with, and what evidence is missing. The connection to performance goals, saving, sales, work, and satisfaction with outcomes should preserve those limits.

What should students remember? They should reconstruct the central tension, explain the benchmark, read the main denominator, and propose an institution that could change the outcome.

6.8 Diagnostic cases for interpretation practice

6.8.1 Agreement with the benchmark

The distribution is concentrated near the prediction derived from an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes. Do not conclude that everyone shared the same motive. Some calculated, others followed a signal, and others chose intuitively. The evidence supports the aggregate prediction under session conditions but does not identify individual reasoning. Ask what belief, process, or response-time measure would distinguish those routes.

Then introduce a variant from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references. If the pattern remains when the seemingly obvious mechanism weakens, the original explanation needs revision. A correct prediction does not automatically validate all its assumptions.

6.8.2 Two groups with opposite behavior

Half the class is at one extreme and half at the other; the mean lies where almost nobody chose. Showing only the average would create a fictional description. Calculate the median, modes, range, and share in theoretical categories. Ask whether heterogeneity is related to role, round, treatment, or comprehension.

Consider gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references. Write one additional observation for each. If it predicts role differences and none appear, it receives less relative support. If it predicts learning, inspect the trajectory. Do not fill missing measurement with a compelling story.

6.8.3 A large difference with few units

One treatment apparently doubles the main measure but contains two groups and the comparison contains three. Display each group before pooling. One extreme unit can create the difference. Twenty decisions nested in two groups are not twenty independent units.

Ask students to plan the next session: groups, assignment, primary measure, and exclusions. The answer should mention difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference and incomplete data. The experience thereby introduces design and transparency.

6.8.4 A null difference

Two conditions have the same mean. Before declaring that the manipulation does not matter, inspect compliance, comprehension, tails, and dispersion. Opposing subgroup effects can cancel. Define what range of differences would be pedagogically or economically meaningful and what size could distinguish it.

The cases teach four disciplines: do not infer motives from agreement, hide heterogeneity, confuse rows with units, or treat absence of evidence as evidence of absence.

6.9 Explanatory sequence for a complete class

6.9.1 Step 1: pose the problem without labels

Present the decision as a concrete situation and ask what information is missing. Do not announce the name of the bias, equilibrium, or preference at the start. The question is how expectations, goals, and initial states change the evaluation of the same outcome. Collect private responses so the first speaker does not become a focal point.

6.9.2 Step 2: reconstruct rules and incentives

Ask one student to describe actions and another to describe consequences. A third explains available information. Draw the sequence. In Reference points, the decision is to evaluate satisfaction, effort, saving, or continuation under different goals and expectations. Do not move to interpretation until the class can calculate a simple case.

6.9.3 Step 3: derive the benchmark

Write an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes. Underline each assumption. Ask what happens if one fails. Distinguish a logical derivation from an empirical expectation: an equilibrium strategy need not be played immediately by everyone.

6.9.4 Step 4: record predictions

Request one prediction about the mean, another about the distribution, and a third about a variant. Predictions should use difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference. Preserve the record; later comparison reveals learning and reduces hindsight.

6.9.5 Step 5: reveal in layers

Show coverage first, then distribution, measures, and comparisons. At every layer request description before explanation. If someone proposes gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references, ask what observation would distinguish it from another mechanism.

6.9.6 Step 6: work through one sequence

Use the worked example in this guide. Recalculate slowly and change one parameter. Ask each team to find one excessive conclusion and rewrite it with limited language. Editing claims is as important as calculation.

6.9.7 Step 7: compare institutions

Choose one condition from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references. Write a table with constant, change, prediction, and measure. If the variant changes several elements, acknowledge it. Ask whether a third treatment would separate explanations.

6.9.8 Step 8: transfer and close

Connect the mechanism to performance goals, saving, sales, work, and satisfaction with outcomes. Identify two differences between the laboratory and the real context. End with an exit response: prediction, evidence, compatible mechanism, and limitation. This structure assesses understanding without rewarding a particular decision.

6.9.9 A short script to read aloud

“Describe what happened before explaining why. Identify the prediction and its assumptions. Compare the distribution with the mean. Propose two compatible mechanisms. Finally, state which treatment or additional evidence could distinguish them.” The script keeps discussion in a common sequence across families.

6.10 Assessment bank with solutions

Exercise 1. List actions, information, and payoffs. **Solution:** reconstruct evaluate satisfaction, effort, saving, or continuation under different goals and expectations without adding information unavailable to the participant.

Exercise 2. Derive the material benchmark. **Solution:** begin from an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes and name at least two assumptions. For an interactive game, state best responses; for an individual task, state the decision rule.

Exercise 3. The report shows 40 percent. **Solution:** it cannot be interpreted without numerator, denominator, exclusions, and unit. Eligible people and completed decisions are alternative denominators.

Exercise 4. The mean matches the prediction. **Solution:** inspect the distribution. A correct mean can combine extremes, offsetting errors, or different strategies. Agreement does not identify reasoning.

Exercise 5. Select two mechanisms from gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references. **Solution:** design a condition changing one prediction but not the other, with an outcome from difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference.

Exercise 6. Compare treatments of different sizes. **Solution:** report counts and shares, preserve the independent unit, and do not compensate for a small group by multiplying rounds.

Exercise 7. Mean differences are absent. **Solution:** inspect dispersion, tails, compliance, and comprehension; define a relevant difference. A broad null does not establish equivalence.

Exercise 8. Propose a figure. **Solution:** use histograms or dots for distributions, lines for rounds, separate bars for roles, and bivariate plots when a response depends on a prior action.

Exercise 9. Explain a variant from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references. **Solution:** identify what changes, what remains constant, predicted direction, and an interpretive threat.

Exercise 10. Transfer to performance goals, saving, sales, work, and satisfaction with outcomes. **Solution:** describe the common mechanism and two institutional differences. Do not claim that one session determines policy.

Exercise 11. Write a three-sentence conclusion. **Solution:** observation with denominator; compatible mechanisms; evidence needed to distinguish them.

Exercise 12. Design a replication. **Solution:** preserve parameters, record assignment, announce the primary measure, set exclusions, and archive the export.

A twelve-point rubric awards one point for each correct component. For a short assignment use Exercises 2, 3, 5, and 11. For a project use all twelve and request a reproducible figure.

6.11 Advanced seminar: from evidence to mechanism

Divide the class into four teams: benchmark model, alternative preferences or beliefs, institutions, and identification critique. Each interprets the same report from its assigned role. The first derives an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes; the second uses gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references; the third selects a variant from expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references; the fourth identifies missing evidence.

In a replication round, every team writes a quantitative prediction and a condition that could refute it. It exchanges the proposal with another team, which audits unit, denominator, and manipulation. A useful explanation does not merely accommodate past outcomes; it risks a prediction about new evidence.

End with a one-page memo containing question, benchmark, descriptive finding, two mechanisms, discriminating design, application to performance goals, saving, sales, work, and satisfaction with outcomes, and limitation. Grade causal clarity, technical correctness, use of evidence, and inferential restraint.

For a quantitative extension, export data and reproduce three figures: individual distribution, aggregate by independent unit, and treatment comparison. All three require titles, denominators,

and exclusion notes. Compare how the narrative changes when the correct unit is used.

An ethics extension asks which decisions should remain private, when an incentive is proportionate, and how to communicate results without turning group patterns into personal judgments. The answer should separate participant protection, instructional integrity, and analytical transparency.

6.12 Quality control before teaching

6.12.1 Full rehearsal

Open a test session and complete one full participant path. Check cover, language, comprehension, decision, waiting, result, and close. In Reference points, verify especially that the participant understands how to evaluate satisfaction, effort, saving, or continuation under different goals and expectations. Open the report before class, keep it hidden, and confirm that difference in satisfaction; additional effort; saving; discontinuity near a goal; contrast by assigned reference appear with their denominators.

6.12.2 Fifteen control questions

1. Does the configuration belong only to this family and the correct language?
2. Can expected attendance form every required unit?
3. Do the rules describe actions, information, order, and payoffs?
4. Does the comprehension question detect the most likely mistake?
5. Is the benchmark derived rather than merely asserted?
6. Will students record a prediction before seeing the result?
7. Does the report show counts as well as percentages?
8. Does every rate state numerator and denominator?
9. Is the distribution readable from the classroom?
10. Are roles, treatments, and rounds labeled and colored consistently?
11. Was the main comparison selected before observing data?
12. Is there time to describe results before interpreting them?
13. Do questions avoid judgments about individual participants?
14. Does the link to performance goals, saving, sales, work, and satisfaction with outcomes state explicit limits?
15. Will the export and incidents be documented?

6.12.3 A micro-rehearsal of the debrief

Practice a three-minute explanation. Begin with an evaluation based only on the final outcome does not change when the outcome is fixed and only an irrelevant reference changes. Then show a hypothetical distribution and state two compatible interpretations from gain-loss utility, loss aversion, goals, rational expectations, salience, proximity, and multiple references. Identify which

variant among expectations; goals; effort; near goal; saving and subgoals; income target; origin and difficulty; multiple references would distinguish them. End with one limitation. If the sequence cannot fit in three minutes, simplify the session objective.

6.12.4 After class

Record attendance, completed decisions, parameters, and incidents. Note which explanation caused confusion and which graph worked. Preserve one question for the next cohort. Cumulative improvement depends on documenting experience rather than changing parameters without a record.

The closing condition is concrete: a person who did not attend should be able to read the report and know what was decided, what the benchmark was, how many units participated, what pattern appeared, and what cannot be concluded. If any element is missing, complete the instructor note before archiving the session.

Bibliography

- Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, 101(2), 470–492. doi:[10.1257/aer.101.2.470](https://doi.org/10.1257/aer.101.2.470).
- Colby, H., and Chapman, G. (2013). Savings, subgoals, and reference points. *Judgment and Decision Making*, 8(1), 16–24. doi:[10.1017/S1930297500004459](https://doi.org/10.1017/S1930297500004459).
- Crawford, V. P., and Meng, J. (2011). New York City cab drivers’ labor supply revisited. *American Economic Review*, 101(5), 1912–1932. doi:[10.1257/aer.101.5.1912](https://doi.org/10.1257/aer.101.5.1912).
- Heath, C., Larrick, R. P., and Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, 38, 79–109. doi:[10.1006/cogp.1998.0708](https://doi.org/10.1006/cogp.1998.0708).
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291. doi:[10.2307/1914185](https://doi.org/10.2307/1914185).
- Kaur, S., Kremer, M., and Mullainathan, S. (2015). Self-control at work. *Journal of Political Economy*, 123(6), 1227–1277. doi:[10.1086/683822](https://doi.org/10.1086/683822).
- Kőszegi, B., and Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4), 1133–1165. doi:[10.1093/qje/121.4.1133](https://doi.org/10.1093/qje/121.4.1133).
- Pope, D. G., and Simonsohn, U. (2011). Round numbers as goals. *Psychological Science*, 22(1), 71–79. doi:[10.1177/0956797610391098](https://doi.org/10.1177/0956797610391098).