



Instructor's Guide to Bomb Risk Elicitation Task

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 implementation and discussion of the Bomb Risk Elicitation Task, or BRET, even when the instructor has not previously taught experimental risk elicitation.

The opening chapters explain the economic logic, platform flow, and instructor report.

Later chapters develop predictions, variants, evidence, and inference limits. An introductory class can use one mode; an advanced class can compare treatments.

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

BRET is an individual activity. There are no partners, social roles, or wait pages. Each person chooses exposure to risk, and the computer resolves an independent loss position. The instructor report displays aggregates and stays hidden until all active decisions are complete.

Contents

How to use this guide	i
1 Quick implementation guide	1
1.1 Access through the updated laboratory website	1
1.2 Learning objectives	5
1.3 Minimum preparation	5
2 Economic structure and predictions	6
2.1 The decision	6
2.2 Expected utility and risk aversion	6
2.3 What the task does not identify alone	7
2.4 Access and family architecture	7
3 Participant flow	8
4 Instructor report and debrief	14
4.1 Basic modes	14
4.2 Static choice	16
4.3 Dynamic collection	16
4.4 Static–dynamic comparison	16
4.5 Original evidence and adaptations	16
4.6 Data, privacy, and troubleshooting	16
5 Extended variants and their teaching use	18
5.1 High stakes	18
5.2 5 by 5 grid	18
5.3 20 by 20 grid	19
5.4 Random order	19
5.5 Fast collection	19
5.6 Five decisions	19
5.7 Neutral wording	20
6 Expanded teaching manual for nonspecialist instructors	22
6.1 Conceptual map before opening the session	22
6.1.1 Unit of analysis	22
6.1.2 Benchmark prediction	22
6.1.3 Alternative mechanisms	23
6.2 Worked example and explanation	23
6.3 Reading the instructor report	23
6.4 Step-by-step debrief script	24
6.5 Comparable class plans	24
6.5.1 A 50-minute class	24

6.5.2	An 80-minute class	25
6.5.3	A 110-minute class	25
6.6	Exercises with answer guidance	25
6.6.1	1. Reconstructing the prediction	25
6.6.2	2. Auditing a percentage	25
6.6.3	3. Reading heterogeneity	25
6.6.4	4. Separating explanations	25
6.6.5	5. Interpreting no difference	26
6.6.6	6. Transferring carefully	26
6.7	Data, inference, and learning assessment	26
6.7.1	A minimum results table	26
6.7.2	Comparing without overstatement	26
6.7.3	Assessing understanding	26
6.7.4	Closing protocol	27
6.8	Frequently asked questions and model answers	27
6.9	Diagnostic cases for interpretation practice	28
6.9.1	Agreement with the benchmark	28
6.9.2	Two groups with opposite behavior	28
6.9.3	A large difference with few units	28
6.9.4	A null difference	28
6.10	Explanatory sequence for a complete class	29
6.10.1	Step 1: pose the problem without labels	29
6.10.2	Step 2: reconstruct rules and incentives	29
6.10.3	Step 3: derive the benchmark	29
6.10.4	Step 4: record predictions	29
6.10.5	Step 5: reveal in layers	29
6.10.6	Step 6: work through one sequence	29
6.10.7	Step 7: compare institutions	30
6.10.8	Step 8: transfer and close	30
6.10.9	A short script to read aloud	30
6.11	Assessment bank with solutions	30
6.12	Advanced seminar: from evidence to mechanism	31
6.13	Quality control before teaching	31
6.13.1	Full rehearsal	31
6.13.2	Fifteen control questions	32
6.13.3	A micro-rehearsal of the debrief	32
6.13.4	After class	32

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.	3
1.5	Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.	4
3.1	Institutional cover and general laboratory rules.	9
3.2	Current instructions for the basic edition: initial resources, decision rules, and experimental units.	10
3.3	Current comprehension check with an incorrect answer, its specific explanation, and retained responses.	11
3.4	Static choice with the information needed for calculation.	12
3.5	Explicit completion screen and individual outcome.	13
4.1	Report status before aggregate revelation.	14
4.2	Revealed report with distribution and denominators.	15
5.1	Dynamic collection on the 20 by 20 grid.	21

Quick implementation guide

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.

go con bomba — decisión estática	Tarea de riesgo con bomba / Bomb Risk Elicitation Task — BEER Lab Seleccione una variante de este juego. Select a version of this game.
go con bomba — recolección dinámica	
go con bomba — comparación estática y dinámica	
licitation Task — static choice	
licitation Task — dynamic collection	
licitation Task — static and dynamic comparison	
go con bomba extendida — incentivos altos	
go con bomba extendida — rejilla 5 × 5	
go con bomba extendida — rejilla 20 × 20	
go con bomba extendida — orden aleatorio	
go con bomba extendida — recolección rápida	
go con bomba extendida — cinco decisiones	
go con bomba extendida — encuadre neutral	
omb Risk Elicitation Task — high stakes	
omb Risk Elicitation Task — 5 × 5 grid	
omb Risk Elicitation Task — 20 × 20 grid	
omb Risk Elicitation Task — random collection order	
omb Risk Elicitation Task — fast collection	
omb Risk Elicitation Task — five decisions	
omb Risk Elicitation Task — neutral frame	

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

Rooms

Current rooms:

[ES] Tarea de riesgo con bomba
[EN] Bomb Risk Elicitation Task
[ES] Tarea de riesgo con bomba: extendida
[EN] Bomb Risk Elicitation Task: extended

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

Room: [EN] Bomb Risk Elicitation Task

Create a new session

Session Config

[EN] Bomb Risk Elicitation Task — static choice

Number of participants

Individual decisions: any positive participant count is allowed.

Create

[Configure session](#) 

App sequence

bomb_risk_game (5 rounds)

Bilingual static and dynamic Bomb Risk Elicitation Task for BEER Lab teaching.

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	Make visible the tradeoff between an increasing potential reward and an increasing probability of losing it.
Experimental unit	Person and decision. Participants do not interact.
Basic edition	Static choice, dynamic collection, and a comparison of both modes.
Extended edition	High stakes, 25- and 400-box grids, random order, faster collection, five decisions, and neutral wording.
Payoff	One of N positions contains the bomb. Collecting k boxes pays $k\gamma$ ECU if the bomb is excluded and 0 ECU if it is included.
Participants	Any number. Ten or more support a useful histogram, but a smaller class can still support descriptive discussion.
Students	/room/bomb_risk_en
Instructor	/demo

Local test	http://localhost:8025/demo with <code>GAME_FAMILY=bombrisk</code> .
Decision	Choose how many boxes to collect or when to stop collecting.
Topics	risk attitudes; risk elicitation; decisions under uncertainty; behavioral economics.
JEL	C91; D81; D91.

Attention

The `bombrisk.beer-lab.org` domain is the proposed production route. It requires dedicated DNS, an oTree process, and a PostgreSQL database. The local address works only on the computer running oTree.

1.2 Learning objectives

Students should be able to distinguish potential reward, loss probability, and expected payoff; calculate the risk-neutral benchmark; connect utility concavity with lower exposure; interpret a distribution without diagnosing individuals; compare static and dynamic elicitation; and propose a treatment that separates rival explanations. The exercise does not seek a morally correct number. It studies how one choice summarizes a probability–prize tradeoff.

1.3 Minimum preparation

Before class, choose language, edition, and treatment; decide whether ECU are hypothetical or converted; create the session; open *Monitor* and *Report*; and test one link. Write two predictions in advance, such as the median share and the share below 50 percent. Precommitting these quantities keeps the theoretical benchmark separate from the observed result.

Economic structure and predictions

2.1 The decision

There are N boxes and exactly one loss position. Every position has probability $1/N$. The participant chooses $k \in \{0, 1, \dots, N\}$ boxes. Each box is worth γ ECU. If the bomb is among the collected boxes, payoff is zero; otherwise payoff is $k\gamma$. Hence, the probability of keeping the reward is $(N - k)/N$ and the probability of losing it is k/N .

Expected monetary payoff equals

$$E[\pi(k)] = \frac{N - k}{N} k \gamma.$$

This expression is a parabola. It starts at zero when $k = 0$, reaches its maximum around the midpoint, and returns to zero at $k = N$. A risk-neutral participant has benchmark $k = N/2$ when N is even. With 100 boxes worth one ECU, collecting 50 has expected payoff 25 ECU. Collecting 40 and 60 each has expected payoff 24 ECU, although the two choices combine reward and probability differently.

Worked example

With 100 boxes and 1 ECU per box, choosing $k = 35$ produces a 35 percent bomb probability, a conditional reward of 35 ECU, and expected payoff $35 \times .65 = 22.75$ ECU. At $k = 50$, conditional reward rises to 50 and expected payoff to 25. At $k = 70$, potential reward is higher but the chance of keeping it is only 30 percent, so expected payoff is 21 ECU. The neutral benchmark evaluates the full distribution rather than the largest possible prize.

2.2 Expected utility and risk aversion

If payoff utility is $u(x)$, the participant chooses k to maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$. Normalize $u(0) = 0$ and suppose $u(x) = x^r$. The continuous solution is $k^* = Nr/(1 + r)$. When $r = 1$, the benchmark is $N/2$. When $0 < r < 1$, utility is concave and $k^* < N/2$. When $r > 1$, utility is convex and $k^* > N/2$. Discrete boxes require rounding, so one choice identifies at most an interval of compatible values.

The derivation is informative but not exhaustive. Choices can also reflect focal numbers, collection speed, interface features, comprehension, attention, framing, or decision error. A choice below 50 percent is descriptively consistent with risk aversion in this task. It does not establish a stable trait across scales and domains. Comparative evidence finds that different elicitation methods can yield different measures even in homogeneous samples (Crosetto and Filippin, 2016).

2.3 What the task does not identify alone

BRET does not separately identify utility curvature, probability weighting, emotion generated by the bomb metaphor, and choice noise. It is not an ambiguity task because the position distribution is known. It does not identify loss aversion unless a credible reference point and losses are introduced. The report bands —below 45, between 45 and 55, and above 55 percent— organize discussion; they are not structural estimates or personal diagnoses.

2.4 Access and family architecture

Begin at <https://beer-lab.org/>, open the teaching catalog at <https://teaching.beer-lab.org/>, and select the Bomb Risk Elicitation Task card. The family site should expose only its 20 configurations and four rooms.

Basic Spanish	/room/bomb_risk_es
Basic English	/room/bomb_risk_en
Extended Spanish	/room/bomb_risk_extendido
Extended English	/room/bomb_risk_extended
Instructor	/demo

The room is a stable route; each session creates temporary individual links. If the database is rebuilt, an old link may say that the participant no longer exists. Create a new session and distribute its current links. Do not fabricate participants or reuse obsolete links.

Participant flow

The cover identifies the laboratory, explains the purpose, and requires acceptance of silence, privacy, and proper workstation use. Instructions describe the grid, equiprobable position, payoff formula, collection order, and active mode. The context panel names the treatment so that conditions are not mixed. The check requires three facts: 30 of 100 boxes means a 30 percent bomb probability; collecting the bomb pays zero; and 50 boxes maximize expected monetary payoff under canonical parameters. In static mode, a slider selects k . The grid, bomb probability, conditional reward, and expected payoff update together. Participants can compare choices before confirming. The final screen reports boxes, drawn position, outcome, and payoff. It asks the participant to keep the page open and prevents the generic *No more pages left* ending.



Welcome to the BEER Lab

The Behavioral and Experimental Economics Research Lab in the Department of Economics at Tecnológico de Monterrey is a space for learning about economic decisions through controlled activities.

This activity studies a choice between a potential gain and the risk of losing it.

General laboratory rules

- Remain silent throughout the activity.
- Do not talk or look at other participants' decisions.
- Use only the assigned page and do not refresh the browser.
- Raise your hand if you have a question.

I understand and will follow these rules.

- ☐ Sí
- ☒ No

Start

Con tecnología de [oTree](#)

Figure 3.1: Institutional cover and general laboratory rules.



Instructions

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.

100 × 1,0 ECU

[EN] Bomb Risk Elicitation Task — static choice

The screen contains a grid of boxes. One box contains a bomb and every position is equally likely.

1

If the bomb is outside the collected boxes, you receive the stated value for every box collected.

2

If you collect the box containing the bomb, your payoff from that decision is 0 ECU.

3

Under ordered collection, boxes are collected from the upper-left corner. Under random collection, the computer determines the order without changing the probability of collecting the bomb.

4

In dynamic mode, press Start and then Stop when your preferred number has been collected.

With n boxes, k collected boxes, and v ECU per box, expected monetary payoff is $k \times v \times (n-k)/n$.

Each decision is resolved with an independent bomb position. Aggregate results remain hidden until the instructor reveals them.

Continue

Con tecnología de [oTree](#)

Figure 3.2: 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.

Comprehension check

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.

If you collect 30 of 100 boxes, what is the percentage probability of collecting the bomb?

%

Your answer: 31. The correct answer is 30. Review the rules: risk is k/n , the bomb leaves 0 ECU, and expected monetary payoff is maximized at half the grid.

If you collect the bomb, how much do you earn in that decision?

☒ 0 ECU☐ The number of collected boxes

With 100 boxes worth 1 ECU each, how many boxes maximize expected monetary payoff?

Continue

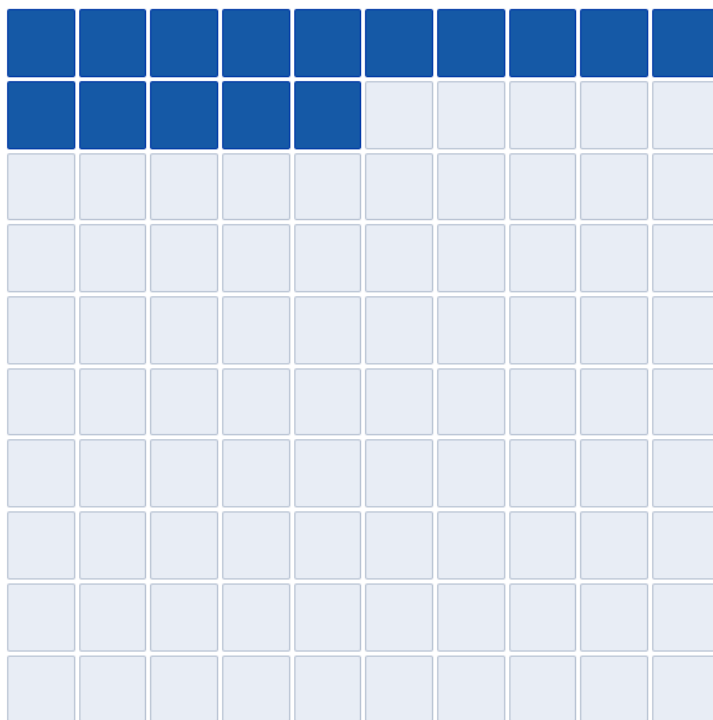
Con tecnología de [oTree](#)

Figure 3.3: Current comprehension check with an incorrect answer, its specific explanation, and retained responses.

Your decision

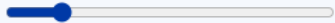
Decision 1 of 1

Choose how many boxes you want to collect.



Confirm decision

15
Boxes collected



Boxes not collected **85**

Bomb probability **15.0%**

Payoff if the bomb is avoided **15.0 ECU**

Expected monetary payoff **12.8 ECU**

Con tecnología de [oTree](#)

Figure 3.4: Static choice with the information needed for calculation.



Activity complete

These are your decisions and outcomes. Keep this page open until the instructor closes the session.

#	Mode	Boxes collected	Bomb position	Outcome	Payoff
1	Static	15/100	71	Bomb avoided	15,0 ECU
Payoff: 15,0 ECU					

Con tecnología de [oTree](#)

Figure 3.5: Explicit completion screen and individual outcome.

Instructor report and debrief

Before revelation, the report displays participants, expected decisions, completed decisions, and pending decisions. The reveal action becomes available only when coverage is complete. After revelation, it shows the mean, median, mean expected payoff, realized bomb rate, normalized histogram, variant comparison, evolution by decision, and descriptive categories.

The histogram unit is a completed decision. In a one-decision session this equals the participant count. In the five-decision treatment, each participant contributes five observations, so the instructor must also report the number of people and the by-decision chart. Realized bomb frequency is noisy in a small class and should not be used to grade rationality. The primary measure is the share of boxes selected, which makes the 25-, 100-, and 400-box grids comparable.

Feedback script

Reveal in four layers. First confirm coverage. Second ask students to describe the histogram without assigning motives. Third locate the 50 percent benchmark and compare the mean, median, and evolution by decision. Fourth introduce mechanisms: utility curvature, focal points, speed, stakes, and interface. End by asking what treatment would distinguish two rival explanations.

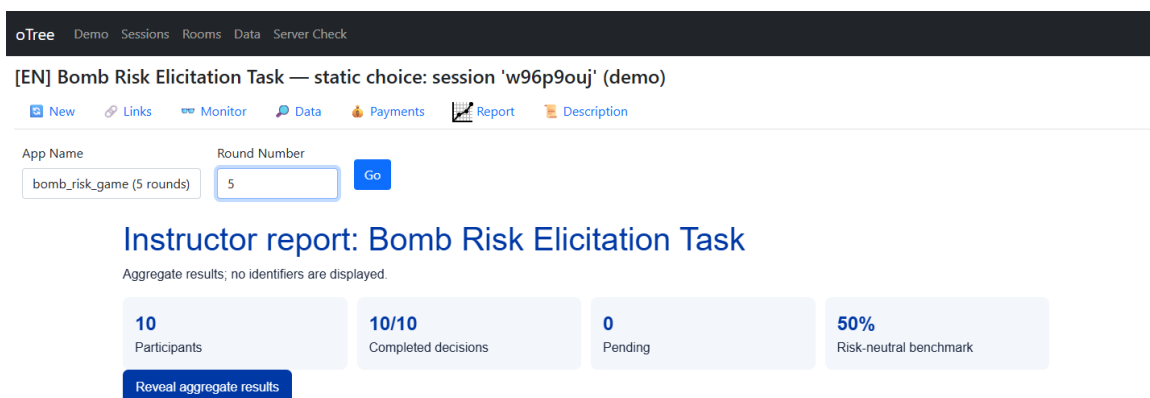


Figure 4.1: Report status before aggregate revelation.

4.1 Basic modes

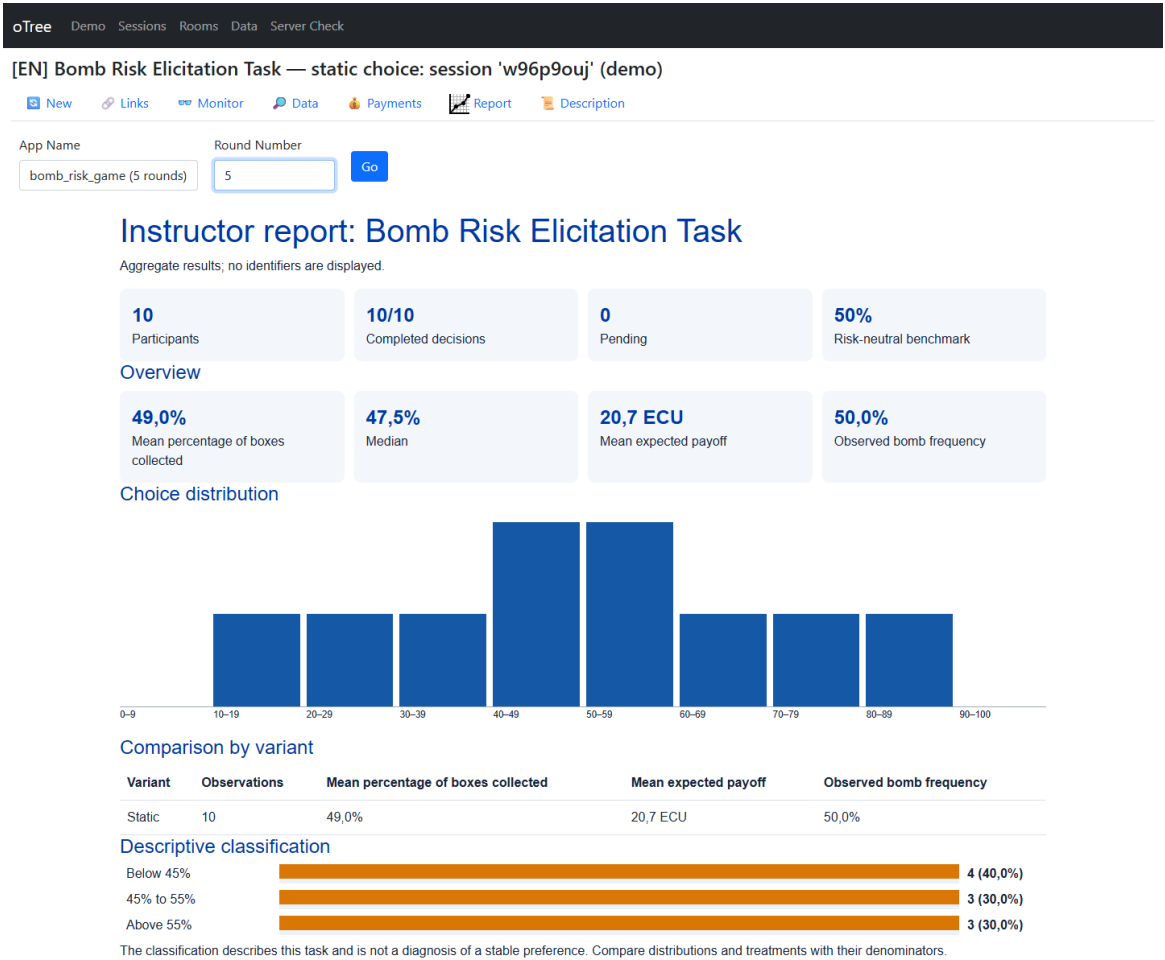


Figure 4.2: Revealed report with distribution and denominators.

4.2 Static choice

The participant moves a slider and observes all consequences immediately. This mode separates choice from timing pressure. It is well suited to deriving expected payoff and showing why choices symmetric around 50 have the same expected monetary payoff. It also makes clear that the expected-payoff maximum is not the largest possible prize.

4.3 Dynamic collection

The participant presses Start and watches boxes accumulate until pressing Stop. The activity becomes a stopping decision. Sequence can create salience, inertia, anticipation, or small timing errors. For classroom use, collection is compressed while the choice set and economic formula remain unchanged.

4.4 Static–dynamic comparison

The comparison configuration presents a static choice and then a dynamic choice with an independent bomb position. It is a useful within-person procedure comparison, but order is fixed. A difference can therefore combine mode, experience, and order. Stronger causal comparison would counterbalance order or assign modes between participants.

4.5 Original evidence and adaptations

[Crosetto and Filippin \(2013\)](#) introduce BRET as an intuitive, fine-grained risk elicitation method. The canonical design uses 100 boxes, linear earnings per box, and one equiprobable bomb. The paper compares static and dynamic versions and studies robustness to stakes, grid size, order, speed, and repetition. The baseline dynamic mean was below 50 boxes with substantial heterogeneity. That estimate is prior evidence, not a target every class should reproduce.

The authors report sensitivity to higher stakes and the smaller grid. The 20 by 20 grid was close to baseline, whereas the 5 by 5 grid concentrated choices and raised the mean. Fast collection was similar in average, although speed can create small stopping errors. Random order tests whether collection geometry drives the response. Repetition studies stability and experience.

[Holzmeister and Pfurtscheller \(2016\)](#) present a configurable, multilingual oTree implementation. The BEER Lab family follows the shared-logic principle and adds an institutional cover, comprehension, four isolated rooms, explicit closure, a projectable report, and instructor documentation. Grid timing is compressed for class: 25, 100, and 400 boxes each take roughly 40 seconds to traverse. Neutral wording is a teaching adaptation. It replaces the word bomb with a random loss position and is not described as an exact replication of a published treatment.

4.6 Data, privacy, and troubleshooting

The export includes treatment, round, boxes chosen, grid size, value per box, drawn position, bomb indicator, safe probability, expected payoff, realized payoff, decision time, and mode. These fields reproduce the report. Participant codes are operational identifiers and should not be projected or published.

If *Please wait* appears, confirm that the session belongs to this family: no page depends on another participant. If the grid does not update, check that JavaScript is enabled and refresh only with staff support. If Demo lists another game, stop and verify `GAME_FAMILY=bombrisk`. If the report stays locked, use Monitor to locate pending decisions. If a student opened one link in two tabs, close one and continue only in the other.

Checkpoint

- Demo lists 20 BRET configurations and no other family.
- Rooms lists four rooms; the basic English room has three basic options.
- Every student receives a distinct current link.
- Students understand risk, zero payoff, and the 50 percent benchmark.
- The instructor waits for completion before revealing aggregates.
- The debrief states numerators, denominators, and inference limits.
- The export is archived without sharing identifiers.

Extended variants and their teaching use

The variants change one dimension at a time whenever possible. State a prediction before opening the report and hold unmanipulated parameters constant. Comparisons across separate classes remain descriptive without common assignment and protocol.

5.1 High stakes

Mechanism. The grid retains 100 positions and dynamic collection, but each box is worth 5 ECU instead of 1. The change increases the payoff scale without changing the 50 percent risk-neutral benchmark.

Prediction. Linear expected utility predicts the same share. Stronger effective curvature, attention, or a scale effect can reduce it.

Evidence and interpretation. The original study reports lower collection under higher stakes. In class, compare the distributions rather than only the mean; tails may carry the important pattern.

Report measures. Use the mean share, median, share below 45 percent, and expected payoff. Keep stake effects separate from cohort differences.

Teaching adaptation. ECU avoid claiming that five points automatically constitute high monetary stakes. With hypothetical payoffs, call this a points-scale comparison.

5.2 5 by 5 grid

Mechanism. There are 25 boxes worth 4 ECU each, and one box changes risk by four percentage points. The choice set is coarser and does not include exactly 50 percent.

Prediction. Power utility predicts approximately the same share, apart from rounding. Greater salience per box and focal points can alter the distribution.

Evidence and interpretation. Crosetto and Filippin report a higher mean and tighter concentration for the small grid. Instrument mechanics can affect measurement.

Report measures. Compare shares rather than raw counts. Explain that 12 boxes equal 48 percent and 13 equal 52 percent.

Teaching adaptation. The interface compresses the total duration to about 40 seconds and therefore does not literally reproduce four seconds per box.

5.3 20 by 20 grid

Mechanism. There are 400 boxes worth .25 ECU. One box changes risk by .25 percentage points, creating a fine support.

Prediction. If preference and comprehension are format invariant, the share should resemble baseline. A dense display can also create noise or effort.

Evidence and interpretation. The original evidence found results close to the 10 by 10 grid. This is a useful robustness comparison when the support becomes finer.

Report measures. Normalize $k/400$ and compare it with $k/100$. Inspect dispersion even when means are similar.

Teaching adaptation. The sequence uses 100 milliseconds per box to preserve approximate duration; this is a documented classroom adaptation.

5.4 Random order

Mechanism. The computer builds a random visible sequence. Collected boxes appear dispersed, but the probability of including the bomb remains k/N .

Prediction. A model using only probabilities and rewards predicts the same share as geometric order. Visual contiguity, perceived control, or spatial perception may generate differences.

Evidence and interpretation. Random order is one of the original robustness treatments. It shows that equivalent representations can feel different.

Report measures. Compare mean and shape. Ask whether dispersion changes even though the probability schedule does not.

Teaching adaptation. Bomb position is interpreted as rank in the collection sequence. The display neither reveals the position nor changes probability.

5.5 Fast collection

Mechanism. The sequence advances four times faster than basic dynamic collection. The feasible k values and payoffs are unchanged.

Prediction. Without attention costs or timing error, choice should remain stable. Limited time can create overshooting and dispersion.

Evidence and interpretation. The original fast treatment had a similar mean, with possible trembles near the intended stopping point. A small class can move in either direction.

Report measures. Use decision time and the full distribution. Do not interpret a one- or two-box difference as a preference change.

Teaching adaptation. Speed is relative to a classroom baseline that is already compressed. Record milliseconds when comparing sessions.

5.6 Five decisions

Mechanism. Each participant faces five independent resolutions with the same parameters and no aggregate feedback between decisions.

Prediction. Stable preference suggests similar choices. Interface learning, realized wealth, focal points, and noise can produce a path.

Evidence and interpretation. The original repetition was unannounced to study stability. The teaching interface states the total number for transparency.

Report measures. Plot choice by round in addition to the pooled histogram. The independent unit remains the person rather than the row.

Teaching adaptation. Because repetition is announced and outcomes are summarized, this is not an exact replication of the unannounced treatment.

5.7 Neutral wording

Mechanism. The word bomb becomes random loss position while probabilities and payoffs remain identical.

Prediction. Standard expected utility predicts invariance. Emotion, availability, or narrative interpretation can produce a difference.

Evidence and interpretation. This is a pedagogical adaptation based on the logic of framing research. It is not the original Explosion or Loss treatment.

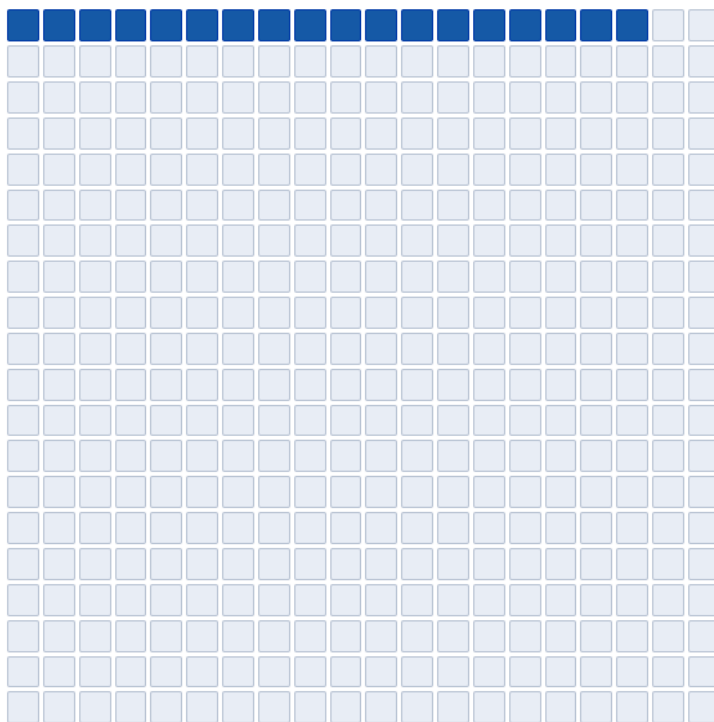
Report measures. Compare groups assigned to bomb and neutral wording. Within-class comparison requires clear assignment.

Teaching adaptation. Do not describe a difference as proof of fear or a null result as equivalence. Record comprehension and use limited claims.

Your decision

Decision 1 of 1

Start the sequence and stop it at your preferred number of boxes.

**18**

Boxes collected

Start collecting

Stop

Reset

Boxes not collected **382**Bomb probability **4.5%**Payoff if the bomb is
avoided **4.5
ECU**Expected monetary
payoff **4.3
ECU****Confirm decision**Con tecnología de [oTree](#)

Figure 5.1: Dynamic collection on the 20 by 20 grid.

Expanded teaching manual for nonspecialist instructors

This part brings together what an instructor needs to turn the Bomb Risk Elicitation Task 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 to connect box choice with loss probability, potential reward, expected payoff, and utility curvature. 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 and each choice of a number of boxes. 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 choose how many boxes to collect before resolving an equiprobable loss position. 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: maximize $[(N-k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k . 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 utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives. 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

With 100 boxes worth one ECU, choosing 50 creates a potential reward of 50, a one-half probability of keeping it, and expected payoff 25. Choosing 40 yields expected payoff 24; choosing 60 also yields 24, although the second choice combines a larger prize with more risk. One choice describes this environment and need not generalize to other scales or domains.

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 box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark. 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: static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording. For each variant, identify the single changed element, predicted direction, and report measure to inspect.

Close by connecting the laboratory to insurance, investment, entrepreneurship, inventory management, and safety decisions. 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 to connect box choice with loss probability, potential reward, expected payoff, and utility curvature. End with one written sentence separating outcome from mechanism.

6.5.2 An 80-minute class

Run the baseline and one variant from static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording. 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 insurance, investment, entrepreneurship, inventory management, and safety decisions. 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 maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k . 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 box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark before explaining mechanisms.

6.6.4 4. Separating explanations

Question. Select two mechanisms from this list: utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives. Design a contrast that produces

different predictions.

Guidance. Change one relevant element, hold others constant, and state each directional prediction. A variant from static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording 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 insurance, investment, entrepreneurship, inventory management, and safety decisions 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 Data, inference, and learning assessment

6.7.1 A minimum results table

Prepare a table with treatment, eligible participants, completed decisions, independent unit, primary measure, median, and dispersion. For BRET, the primary measure should come from box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark. Add a note describing exclusions. With rounds, report people and groups as well as decisions; with roles, present each role separately.

6.7.2 Comparing without overstatement

An internal comparison requires assigned treatments under the same protocol. A comparison across classes can generate questions but may mix cohort, time, instructor, and experience. Report absolute differences and, when accessible, an uncertainty interval. Do not interpret nonsignificance as equality or statistical significance as automatic economic importance.

Do not select the metric after seeing which one produces the most attractive story. Define a primary measure in advance and use the others diagnostically. With many variants, some differences arise by chance. In a teaching session, transparency about this issue is more valuable than simulating research precision.

6.7.3 Assessing understanding

Use three products: a prior prediction, a posterior explanation, and a design proposal. The prediction assesses whether the student derives the benchmark; the explanation assesses whether evidence and mechanism are separated; the design shows understanding of identification. A short

rubric can award one point each for correct rules, prediction, denominator, limited interpretation, and a relevant treatment.

6.7.4 Closing protocol

Before ending, confirm that nobody remains on a wait page, close the session, preserve the export, and record treatment, date, and incidents. Do not project individual codes. Review open text before displaying it. Documenting parameter changes allows the next instructor to compare sessions responsibly.

6.8 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 maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k . 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 utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives 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 BRET, connect those shapes to box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark.

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 static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording, 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 insurance, investment, entrepreneurship, inventory management, and safety decisions 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.

6.9 Diagnostic cases for interpretation practice

6.9.1 Agreement with the benchmark

The distribution is concentrated near the prediction derived from maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k . 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 static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording. 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.

6.9.2 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 utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives. 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.

6.9.3 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 box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark and incomplete data. The experience thereby introduces design and transparency.

6.9.4 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.

6.10 Explanatory sequence for a complete class

6.10.1 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 to connect box choice with loss probability, potential reward, expected payoff, and utility curvature. Collect private responses so the first speaker does not become a focal point.

6.10.2 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 BRET, the decision is to choose how many boxes to collect before resolving an equiprobable loss position. Do not move to interpretation until the class can calculate a simple case.

6.10.3 Step 3: derive the benchmark

Write maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k . 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.

6.10.4 Step 4: record predictions

Request one prediction about the mean, another about the distribution, and a third about a variant. Predictions should use box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark. Preserve the record; later comparison reveals learning and reduces hindsight.

6.10.5 Step 5: reveal in layers

Show coverage first, then distribution, measures, and comparisons. At every layer request description before explanation. If someone proposes utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives, ask what observation would distinguish it from another mechanism.

6.10.6 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.

6.10.7 Step 7: compare institutions

Choose one condition from static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording. 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.

6.10.8 Step 8: transfer and close

Connect the mechanism to insurance, investment, entrepreneurship, inventory management, and safety decisions. 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.

6.10.9 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.

6.11 Assessment bank with solutions

Exercise 1. List actions, information, and payoffs. **Solution:** reconstruct choose how many boxes to collect before resolving an equiprobable loss position without adding information unavailable to the participant.

Exercise 2. Derive the material benchmark. **Solution:** begin from maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k 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 utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives. **Solution:** design a condition changing one prediction but not the other, with an outcome from box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark.

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 static choice; dynamic collection; stakes; 5 by 5 and 20 by 20

grids; random order; speed; repetition; and neutral wording. **Solution:** identify what changes, what remains constant, predicted direction, and an interpretive threat.

Exercise 10. Transfer to insurance, investment, entrepreneurship, inventory management, and safety decisions. **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.

6.12 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 maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k ; the second uses utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives; the third selects a variant from static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording; 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 insurance, investment, entrepreneurship, inventory management, and safety decisions, 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.

6.13 Quality control before teaching

6.13.1 Full rehearsal

Open a test session and complete one full participant path. Check cover, language, comprehension, decision, waiting, result, and close. In BRET, verify especially that the participant understands how to choose how many boxes to collect before resolving an equiprobable loss position. Open the report before class, keep it hidden, and confirm that box share, mean, median, expected payoff, realized bomb frequency, and categories around the 50 percent benchmark appear with their denominators.

6.13.2 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 insurance, investment, entrepreneurship, inventory management, and safety decisions state explicit limits?
15. Will the export and incidents be documented?

6.13.3 A micro-rehearsal of the debrief

Practice a three-minute explanation. Begin with maximize $[(N - k)/N]u(k\gamma) + [k/N]u(0)$ over the box count k . Then show a hypothetical distribution and state two compatible interpretations from utility curvature, focal points, probability weighting, stopping errors, comprehension, scale, and incentives. Identify which variant among static choice; dynamic collection; stakes; 5 by 5 and 20 by 20 grids; random order; speed; repetition; and neutral wording would distinguish them. End with one limitation. If the sequence cannot fit in three minutes, simplify the session objective.

6.13.4 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

- Crosetto, P., and Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, 47(1), 31–65. [doi:10.1007/s11166-013-9170-z](https://doi.org/10.1007/s11166-013-9170-z).
- Crosetto, P., and Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, 19(3), 613–641. [doi:10.1007/s10683-015-9457-9](https://doi.org/10.1007/s10683-015-9457-9).
- Holzmeister, F., and Pfurtscheller, A. (2016). oTree: The “bomb” risk elicitation task. *Journal of Behavioral and Experimental Finance*, 10, 105–108. [doi:10.1016/j.jbef.2016.03.004](https://doi.org/10.1016/j.jbef.2016.03.004).
- Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1/2), 122–136. [doi:10.2307/1913738](https://doi.org/10.2307/1913738).
- von Neumann, J., and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.