



Instructor's Guide to Evidence, Deliberation and Collective Decisions

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 preparation, implementation, and classroom discussion of the family of evidence, deliberation and collective decisions prepared for the Behavioral and Experimental Economics Research Lab (BEER Lab) at the Department of Economics of Tecnológico de Monterrey. It is intended for instructors with different levels of familiarity with experimental and behavioral economics.

The document offers three reading paths. The **quick guide** is enough to launch a session. The operational chapters explain each screen used by instructors and students. The conceptual chapters present the theories, predictions, and empirical findings needed to lead the post-game discussion.

Access instructions and common screens are explained once. Each variant then focuses only on what changes relative to the basic game: configuration, group size, decisions, predictions, relevant outcomes, and academic literature.

Author

This guide was prepared by Stanislaw 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

This guide follows the English-language interface. Spanish configurations use the same logic. Never project individual codes or connect outcomes to names, student identification numbers, or seats in the room.

Contents

How to use this guide	i
1 Preparation and participant workflow	1
1.1 Access through the updated laboratory website	1
1.2 Teaching question and scope	5
1.3 Institutional access and stable routes	6
1.4 Preparing groups, devices, and timing	6
1.5 Welcome, instructions, and comprehension	7
1.6 Evidence, probability, and vote	12
1.7 Waiting, results, and closure	13
2 Information, errors, and strategic voting	17
2.1 Why a panel can improve on an individual signal	17
2.2 False positives, false negatives, and denominators	18
2.3 Bayesian updating through worked examples	18
2.4 Being pivotal and voting strategically	19
2.5 Error costs and choices at a given belief	20
2.6 Academic evidence and limits of the adaptation	20
3 Matching and institutional extensions	21
3.1 Designing an interpretable comparison	21
3.2 One case, fixed panels, and reassigned panels	21
3.3 Unanimity without messages	22
3.4 Messages under majority and unanimity	22
3.5 Asymmetric error costs under both rules	24
3.6 An additional public source and copies of one source	24
3.7 Public and private heterogeneous reliability	26
3.8 Review with and without cost	26
4 Report, discussion and work with data	31
4.1 Progress and controlled disclosure	31
4.2 Metrics, distributions and stage comparisons	32
4.3 Probabilities and calibration	33
4.4 A fully worked report	34
4.5 A seventy-five-minute class	35
4.6 Export, privacy and operational difficulties	35

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. Continuation 1/2.	3
1.5	Rooms contains the four rooms for this family: basic and extended editions in Spanish and English.	4
1.6	Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.	4
1.7	Distribute the room route and use each person's current link.	7
1.8	Keep administrative access outside the projected view.	8
1.9	Confirm the general rules before beginning the case.	9
1.10	Current basic-edition instructions, with initial resources and experimental units. Continuation 1/2.	10
1.12	Identify accuracy points and possible adjustments.	11
1.13	Current comprehension check with an incorrect answer, its specific explanation, and retained responses.	11
1.14	Inspect private evidence before reporting a probability.	12
1.15	Report a private probability about the true state.	12
1.16	Cast A's vote without observing other votes.	13
1.17	Position B follows the same rule with private evidence.	14
1.18	Wait for the panel to complete the required responses.	14
1.19	Reconstruct the result using the state, resolution, and ledger.	15
1.20	Closure confirms recorded decisions and the cumulative result.	16
3.1	Send a signal claim and recommendation through the enabled channel.	23
3.2	Distinguish claims from true signals before reassessing the fact.	24
3.3	Observe the extra public evidence while counting it as one source.	25
3.4	Do not count three copies as independent draws.	25
3.5	Public reliability identifies each position's precision.	26
3.6	Private reliability shows own precision and the common assignment procedure. . .	27
3.7	The second panel combines a true record with new private evidence.	28
3.8	Cast a review vote before the true state is revealed.	29
3.9	Separate the initial resolution, final resolution, and applied cost.	30
4.1	Check progress before revealing decisions to the class.	31
4.2	Read each statistic together with its denominator.	32
4.3	Inspect the vote distribution before reducing it to an average.	32
4.4	Compare the four initial and final outcome counts.	33
4.5	Keep each active case and its denominators visible.	33
4.6	Distinguish reviews, reversals, corrections and deteriorations.	34

Preparation and participant workflow

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.

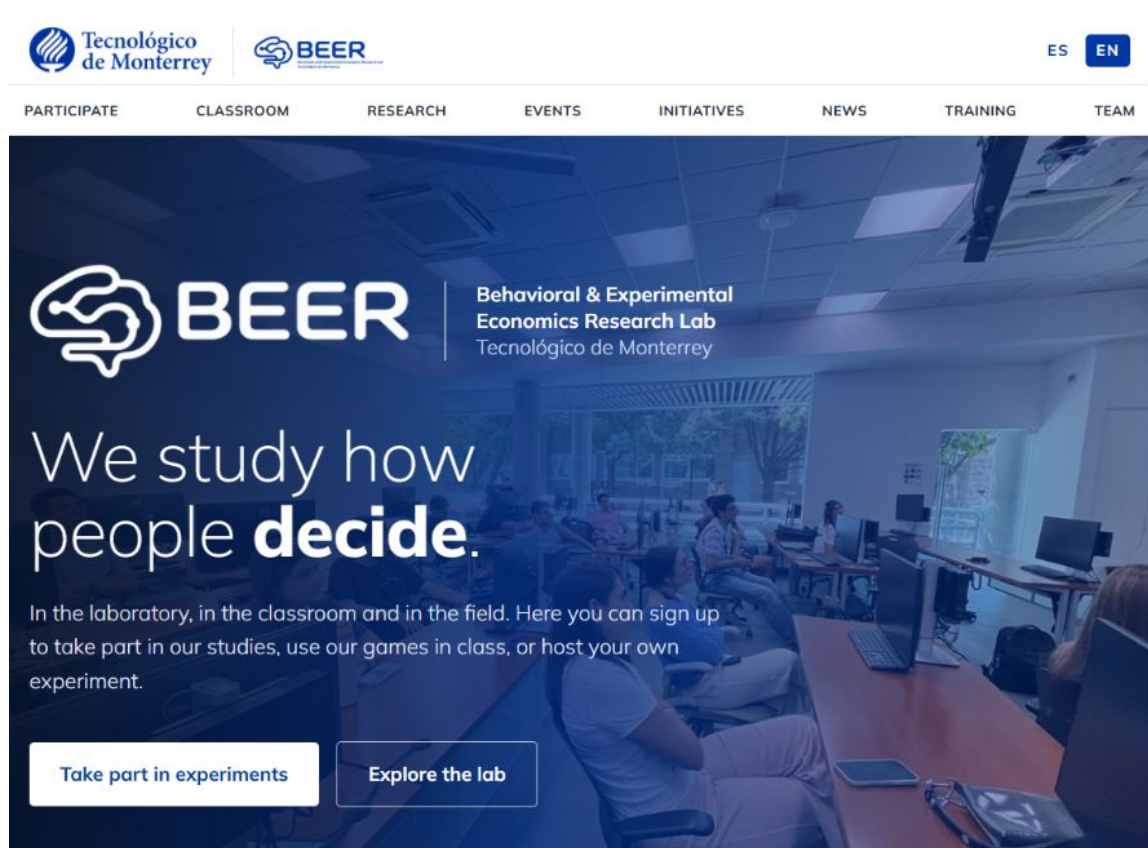


Figure 1.1: Current institutional website. Select CLASSROOM in the top navigation menu to open the teaching experiments. Captured on October 7, 2026.



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.

	Prueba, deliberación y decisión
Paneles fijos	Prueba, deliberación y decisión colectiva / Evidence, Deliberation and Collective Decisions — BEER Lab Seleccione una variante de este juego. Select a version of this game.
Paneles reasignados	
Unanimidad sin mensajes	
Mensajes y mayoría	
Mensajes y unanimidad	
Errores asimétricos y mayoría	
Errores asimétricos y unanimidad	
Fuente pública adicional	
Copias de una misma fuente	
Fiabilidad heterogénea pública	
Fiabilidad heterogénea privada	
Revisión sin costo	
Revisión con costo	
Fixed panels	
Reassigned panels	
Unanimity without messages	
Messages and majority	
Messages and unanimity	
Symmetric errors and majority	
Symmetric errors and unanimity	
Additional public source	
Copies of a single source	
Public heterogeneous reliability	
Private heterogeneous reliability	
Costless review	
Costly review	

Figure 1.4: Demo contains only the configurations of this game family, in both languages and editions. Continuation 2/2.

Rooms

Current rooms:

[ES] Prueba y deliberación: básica
[EN] Evidence and deliberation: basic
[ES] Prueba y deliberación: extendida
[EN] Evidence and deliberation: extended

Figure 1.5: Rooms contains the four rooms for this family: basic and extended editions in Spanish and English.

Room: [EN] Evidence and deliberation: basic

Create a new session

Session Config

Number of participants

Complete groups of 3 participants. Use a multiple of 3; minimum 3.

[Configure session](#) 

App sequence

jury_evidence_game (10 rounds)

Figure 1.6: 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.

1.2 Teaching question and scope

This family examines how a procedure transforms incomplete information into a collective resolution. Three members observe signals about an experimental fact and cast votes. The resolution depends on how many present votes the rule requires. Under the basic scoring table, everyone prefers the resolution to match the true state. The problem arises because nobody knows that state before deciding and because a person's vote can have different consequences depending on the others' choices. The activity connects evidence, uncertainty, and collective judgment. It does not represent a contest among candidates or a choice among personally preferred policies.

Begin by distinguishing four objects: state, signal, vote, and resolution. The state is the fact generated by the program. The signal is an imperfect observation about that fact. The vote is a participant's choice and may or may not match the signal. The resolution is the outcome produced by combining votes under the rule. Keeping these objects separate makes it possible to discuss errors without assuming that every vote is sincere or that three messages constitute three independent sources. A successful class ends with students able to explain what changes when a rule changes and what information is needed to evaluate its consequences.

The game belongs to the laboratory's Law collection. It retains an abstract case: a fact is present or absent. There is no human defendant, no record of a real crime, no punishment imposed on students, and no recommendation about the practices of a particular jurisdiction. Instructors may connect the results to safeguards, burdens of proof, and review after decisions are complete, while identifying dimensions of procedure that the exercise does not include. Statistical accuracy measures performance in this environment; it is not a legal definition of justice. The institutional questions become more useful when these limits are explicit.

Purpose	Examine evidence aggregation, errors, and collective decision rules.
Participants	Multiples of three; multiples of six for review.
Suggested duration	Estimated activity of 25–40 minutes; 60–80 with discussion. Classroom pilot pending.
Experimental unit	An independent case decided by a panel of three; two connected panels for review.
Basic edition	One case, ten cases with fixed panels, or ten with reassignment.

Extended edition	Eleven variants of rules, messages, error costs, sources, reliability, and review.
Decision	Private probability, a message where applicable, and a simultaneous vote.
Payoff	100 ECU for a correct resolution; specific tables and adjustments in variants.
Topics	evidence; deliberation; collective decisions; private information; decision errors; review.
JEL	C92; D71; D82; K41.

For the instructor

Choose a question and a comparison before opening the session. Majority versus unanimity examines a rule; messages versus silence examines a channel; review with and without a cost uses the same procedure. There is no need to run all eleven variants in one class.

1.3 Institutional access and stable routes

The route begins at <https://beer-lab.org/> and continues to <https://teaching.beer-lab.org/>. The directory supports finding game families by topic or JEL code. This family's proposed domain is <https://evidencejury.beer-lab.org/>. After deployment, the instructor can open Demo to select a configuration and Rooms to create a session associated with a stable room. The verified local instance is <http://localhost:8081/demo>. Local access works on the same computer and asks for the local administrative login. A working local demonstration does not establish that the public subdomain has been deployed.

The Spanish basic route is https://evidencejury.beer-lab.org/room/prueba_es and the English route is https://evidencejury.beer-lab.org/room/evidence_en. Extended rooms use https://evidencejury.beer-lab.org/room/prueba_extendido and https://evidencejury.beer-lab.org/room/evidence_extended. Preserve these names when connecting the family to the catalog. A room is a common entry address; an individual link belongs to one person in a particular session. The report link is administrative and must remain private. Project the aggregate results, not the page containing individual links. Students should not have to use the administrator's address to take part.

The family instance contains twenty-eight configurations and exactly four rooms. Each basic room offers three options in its language, while each extended room offers eleven. The first basic-room option is a single basic case, never an extended treatment. If another family appears, stop preparing that session and check the instance. Configuration filtering prevents mixing games; a separate database prevents sessions and data from other families appearing in the administrative lists. Both conditions matter. A correct menu alone does not guarantee that the underlying session records are isolated.

1.4 Preparing groups, devices, and timing

The basic unit requires three participants. To observe differences across panels and use reassignment, select at least six. Review variants need six people per group because they connect an initial panel to a second panel. Do not complete a group with fabricated votes, fictitious passive participants, or an automatic absent decision. The program requires a valid multiple and keeps a case pending

when necessary responses are missing. A class with extra students can assign observation duties for the discussion or choose a compatible count before creating a session. Do not change panel size while a case is underway.

Anonymous positions are assigned randomly at the beginning and remain stable. Basic panels use A, B, and C. In review, A, B, and C form the first instance; D, E, and F form the second. Fixed panels preserve membership. Reassignment mixes participants across cases while retaining positions and avoids repeating an entire panel when alternatives exist. A single panel cannot be reassigned to different members. Explain this restriction when selecting a mode. A three-person session cannot be interpreted as a comparison involving repeated encounters with new participants.

Allow a few minutes for distributing links, reading the rules, and answering comprehension questions. Ten cases take time to inspect information, report a probability, vote, and read the result. Messages and review add stages. The stated durations are operational estimates awaiting classroom piloting and depend on connectivity and participants' experience. Do not accelerate decisions by announcing the outcome the class is expected to produce. Prepare the report projection and a place to write numerators and denominators. Discussion improves when students reconstruct the accounts before comparing percentages.

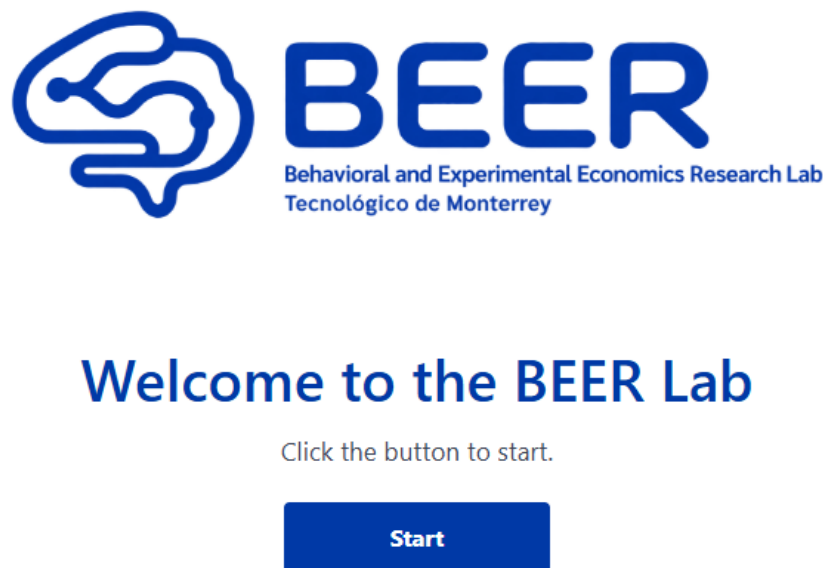


Figure 1.7: Distribute the room route and use each person's current link.

1.5 Welcome, instructions, and comprehension

The welcome introduces the laboratory and general rules. Each person confirms having read them before moving forward. External communication remains prohibited during the activity. Message variants enable only a structured form at a specific stage, and that exception is stated visibly. A free conversation would alter the treatment by letting participants transmit information or coordinate votes through an unrecorded channel. Record such an incident and avoid interpreting that session as an observation of the intended protocol. Instructor assistance should address operational problems without supplying substantive recommendations.

oTree
Demo
Sessions
Rooms
Data
Server Check
Logout

[EN] One case: session '55slmb26' (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:

Split-screen view
[Launch](#)

Grid view
[Launch](#)

Session-wide testing link

Open the below link in up to 6 browser tabs.

<http://127.0.0.1:8081/join/sigozami>

Single-use links

Open each link in its own browser tab.

P1	http://127.0.0.1:8081/InitializeParticipant/wwgfoehi
P2	http://127.0.0.1:8081/InitializeParticipant/954ca253
P3	http://127.0.0.1:8081/InitializeParticipant/16sx8kp8
P4	http://127.0.0.1:8081/InitializeParticipant/fz098h5y
P5	http://127.0.0.1:8081/InitializeParticipant/skdaqrnn
P6	http://127.0.0.1:8081/InitializeParticipant/44n06gz9

Figure 1.8: Keep administrative access outside the projected view.

Instructions describe the equiprobable state, number of cases, panel rule, signal sources, and points table. In the basic design, each signal has seventy percent accuracy and is independent of the others conditional on the state. Conditional independence does not mean that signals never agree. It means that conditioning on the true state removes statistical dependence among their errors. An exploratory demonstration need not begin with a full lesson on Bayes if the purpose is to observe initial reasoning. It must still explain the terms required for informed participation, especially the difference between a signal and the fact it imperfectly describes.



Welcome to the BEER Lab

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

- Remain silent and do not communicate with other participants.
- Make your decisions individually and privately.
- Follow the instructions on the screen.
- Do not refresh, close, or leave the page.
- If there is a problem, raise your hand or contact the staff privately.

☐ I have read the rules and am ready to start.

Start

Powered by [oTree](#)

Figure 1.9: Confirm the general rules before beginning the case.

Comprehension asks for the required number of present votes, accuracy points for a correct final resolution before adjustments, and whether the true state is known before voting. Basic answers are two, one hundred, and zero; unanimity changes the first answer to three. The program returns an error and permits correction without deducting points or replacing an answer. Passing checks establishes minimum reading rather than full understanding of strategic voting. Later, ask for a verbal example involving two present votes and one absent vote under each rule. This reveals whether students understand how the same ballots can produce different resolutions.

Checkpoint

Ask what happens with two present votes and one absent vote. Majority resolves present; unanimity resolves absent. In this activity, lack of unanimity does not produce a mistrial, an unresolved case, or an additional discussion round. Clarify this feature before the first vote.



Activity 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.

Position B · First panel · Case 1 / 1

Each case has an experimental fact: it can be present (1) or absent (0). The program selects both states with equal probability. Your result depends on your group's final resolution. There is no human defendant or real punishment.

You will participate in 1 case(s). Each case starts from zero; your ECU results are added at the end.

Positions and group membership remain the same across cases.

Each first-panel member receives a signal that matches the true state in 70 out of 100 cases. Signals are independent of each other conditional on the state. A signal can be wrong.

The first panel resolves fact present if at least two of its three votes indicate present; otherwise it resolves absent.

Individual result per case

True state	Resolution	Accuracy points
Fact absent (0)	Fact absent (0)	100 ECU
Fact absent (0)	Fact present (1)	0 ECU
Fact present (1)	Fact absent (0)	0 ECU
Fact present (1)	Fact present (1)	100 ECU

All members receive the same points table and, where applicable, the same adjustment.

The true state and all positions' signals are revealed only after the case closes. Before voting, you do not know the other members' votes. Probability responses are private, unpaid, and do not affect your result.

Continue

Con tecnología de oTree

Figure 1.11: Current basic-edition instructions, with initial resources and experimental units.
Continuation 2/2.

True state	Resolution	Accuracy points
Fact absent (0)	Fact absent (0)	100 ECU
Fact absent (0)	Fact present (1)	0 ECU
Fact present (1)	Fact absent (0)	0 ECU
Fact present (1)	Fact present (1)	100 ECU

Figure 1.12: Identify accuracy points and possible adjustments.



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

Check the rules

Position B · First panel · Case 1 / 1

How many present votes does the first present resolution require?

Your answer: 3. The correct answer is 2. Review the instructions: the threshold depends on the rule, a correct resolution gives 100, and the true state remains hidden before voting.

How many accuracy points do you receive for a correct final resolution, before adjustments?

Do you know the true state before voting? Enter 0 for no and 1 for yes.

Continue

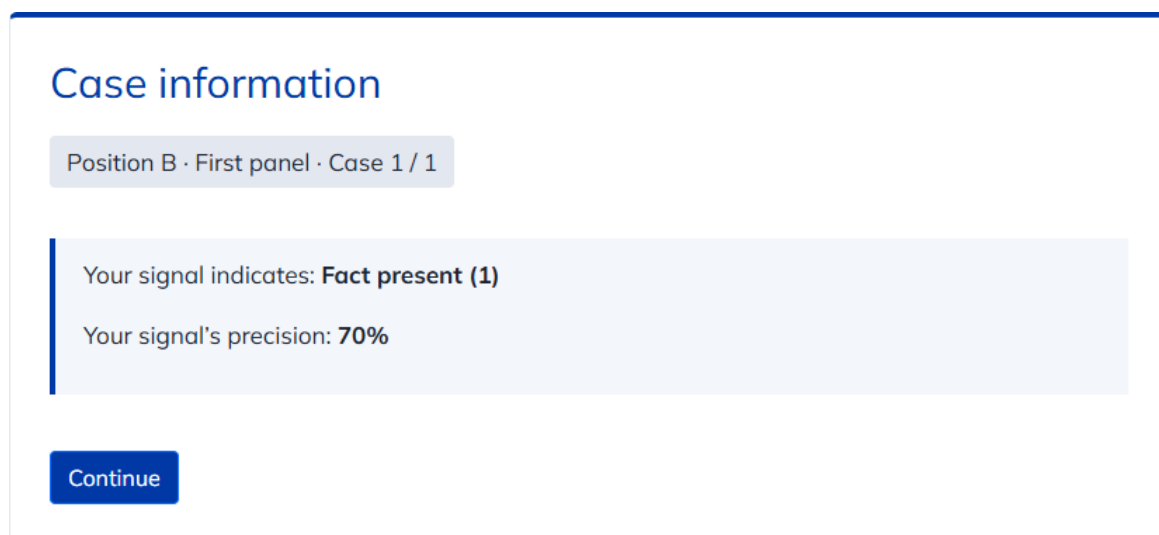
Con tecnología de [oTree](#)

Figure 1.13: Current comprehension check with an incorrect answer, its specific explanation, and retained responses.

1.6 Evidence, probability, and vote

Every case starts with information authorized for the position. The person's signal and its precision appear before the probability response. Additional public evidence, common copies, and public precision assignments are displayed only in the variants that enable them. Private reliability does not show a table of other positions' precision. Students still know the assignment procedure and their own accuracy, allowing some inference about other members. Hiding a table does not remove every piece of knowledge about the group; the guide's comparisons recognize this distinction.

The probability response is an integer from zero to one hundred about the fact being present. It is private, unpaid, and recorded before voting. Explain that fifty expresses intermediate uncertainty rather than an obligation to split votes. Nor does it predict the share of classmates who will vote present: its object is the true state of the current case. The program rejects values outside the interval. Resolve interpretive questions without suggesting a particular probability. Responses can be compared with the later outcome, but one wrong outcome does not prove that a probabilistic assessment was poorly formed.



Case information

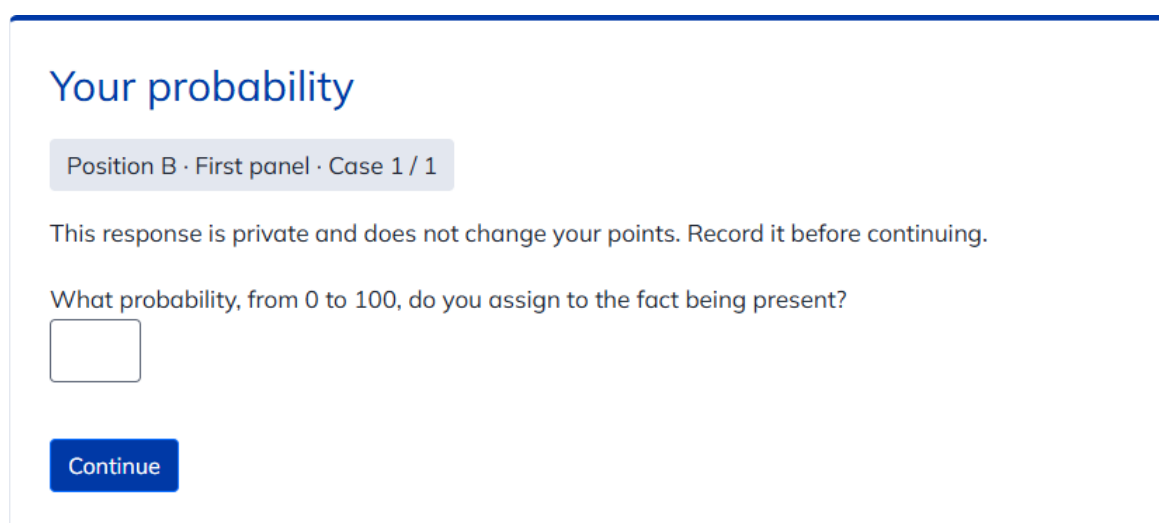
Position B · First panel · Case 1 / 1

Your signal indicates: **Fact present (1)**

Your signal's precision: **70%**

Continue

Figure 1.14: Inspect private evidence before reporting a probability.



Your probability

Position B · First panel · Case 1 / 1

This response is private and does not change your points. Record it before continuing.

What probability, from 0 to 100, do you assign to the fact being present?

Continue

Figure 1.15: Report a private probability about the true state.

The vote selects present or absent. Decisions are simultaneous in the informational sense: nobody observes another member's vote before confirming their own. Arriving at the page first does not

turn a vote into a public announcement or create a timing advantage. The case waits for the necessary responses to be completed. A, B, and C have the same voting authority. Reliability differences change information, not how many votes a person controls. A highly accurate signal does not automatically become a weighted vote; its influence must operate through participants' choices or the information they communicate.

Record your vote

Position A · First panel · Case 1 / 1

Your signal indicates: **Fact present (1)**

Your signal's precision: **70%**

Votes are simultaneous and private until this stage closes. Select one of the two options.

Which resolution do you choose for this case?

☐ Fact present (1) ☐ Fact absent (0)

Continue

Figure 1.16: Cast A's vote without observing other votes.

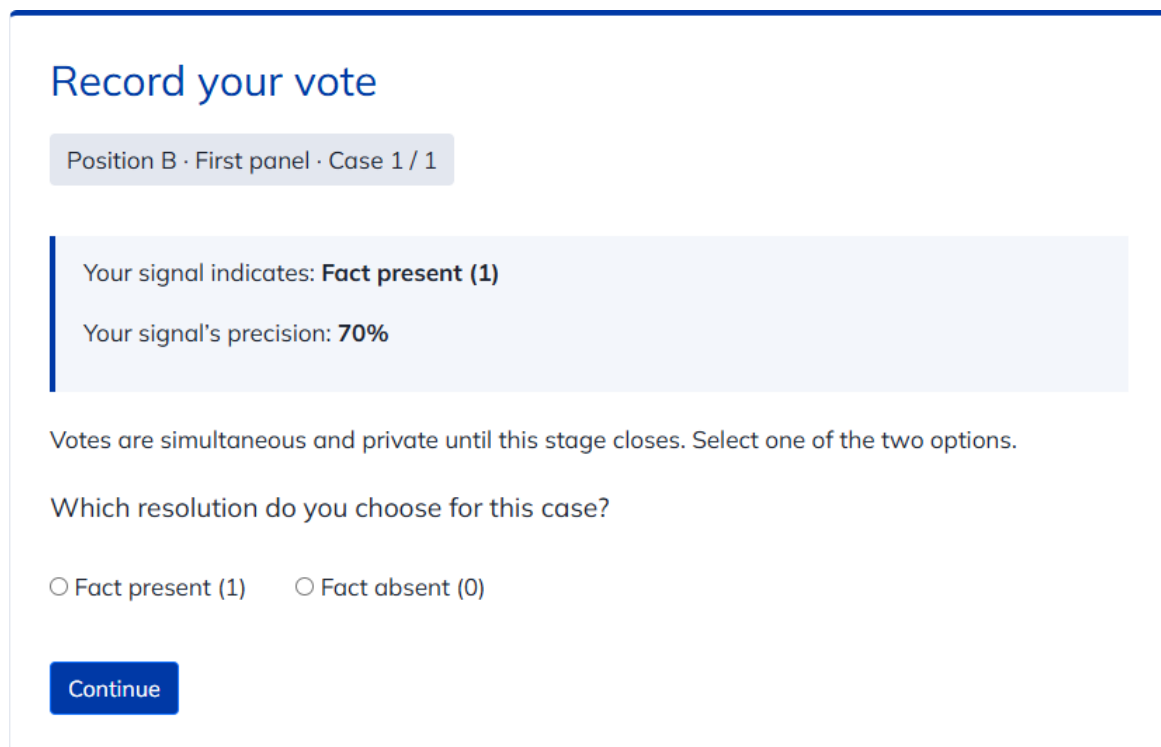
What participants will see

Before voting, each person knows their signal, precision, and public rules. They do not know the true state or other positions' votes. Message treatments display claims that may be inaccurate. Review provides the true first-panel record only to the second panel after an initial present resolution.

1.7 Waiting, results, and closure

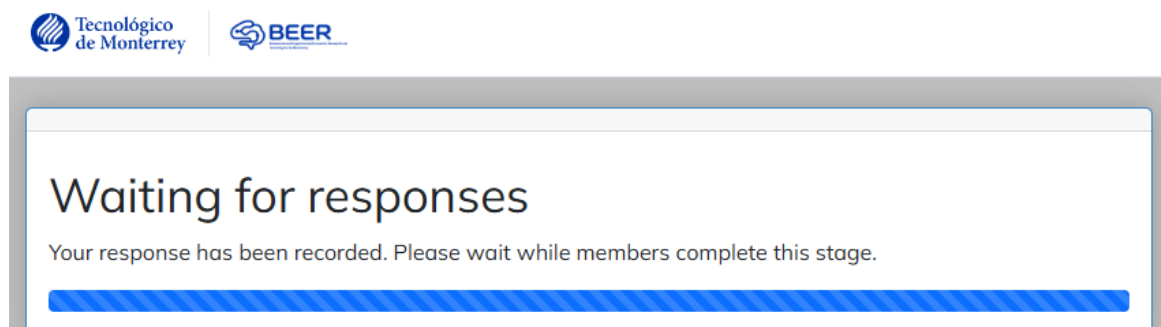
Waits synchronize stages. They do not convert silence into a vote or create an answer for a disconnected person. If someone leaves, the case remains pending and the operational problem must be resolved with that participant. The true state is not revealed to complete a panel artificially. Pending cases also affect the report: an unclosed case does not contribute an accuracy observation even if some votes have been recorded. Identify this distinction before comparing the expected number of cases with the observed number. Leaving out incomplete outcomes is a reporting rule, not evidence that the missing outcomes would resemble the completed ones.

Results reveal the true state, initial and final resolutions, real signals, precisions, and votes. A ledger separates accuracy points, additional balance, and review cost. Ask students to find the state and reconstruct their payoff: an incorrect resolution does not always produce zero in asymmetric-error variants. Both panels receive the final outcome in review. When a second stage did not open, reviewers have no vote and the page says so explicitly. A blank or inapplicable review response should never be interpreted as a vote for absence.



The screenshot shows a web interface for recording a vote. At the top, the title "Record your vote" is displayed in a large, dark blue font. Below the title, a grey box contains the text "Position B · First panel · Case 1 / 1". A light blue box highlights the signal information: "Your signal indicates: **Fact present (1)**" and "Your signal's precision: **70%**". Below this, a message states: "Votes are simultaneous and private until this stage closes. Select one of the two options." The question "Which resolution do you choose for this case?" is followed by two radio button options: "Fact present (1)" and "Fact absent (0)". At the bottom left, there is a blue button labeled "Continue".

Figure 1.17: Position B follows the same rule with private evidence.



The screenshot shows a web interface for waiting for responses. At the top, the title "Waiting for responses" is displayed in a large, dark blue font. Below the title, a message states: "Your response has been recorded. Please wait while members complete this stage." Below the message is a blue progress bar with a diagonal hatching pattern.

Figure 1.18: Wait for the panel to complete the required responses.

Case result

Position B · First panel · Case 1 / 1

True state: **Fact absent (0)**

Initial resolution: **Fact present (1)** · First-panel present votes: 3

Final resolution: **Fact present (1)**

Accuracy points	0 ECU
Additional balance	0 ECU
Review cost	0 ECU
Your result in ECU	0 ECU

Signals and votes at closure

Position	True signal	Your signal's precision	Vote
A	Fact absent (0)	70%	Fact present (1)
B	Fact present (1)	70%	Fact present (1)
C	Fact present (1)	70%	Fact present (1)

[Continue](#)

Figure 1.19: Reconstruct the result using the state, resolution, and ledger.

After the last case, an explicit final screen confirms recorded decisions and shows the sum of results from all active cases. It does not display a button to leave the activity or end on a generic page without directions. Students remain there until discussion. Restarting the server must preserve sessions. If connectivity is interrupted, recover through the current individual link; do not create fictitious identities to rescue an old code. Reinitializing the database breaks old link associations and requires a new session. Such a reset is not part of an ordinary server update, and should not be used to refresh the preview after changes to presentation or instructions.

Activity completed

Your decisions have been recorded. Remain on this screen and wait for the instructor's directions.

Your cumulative result in ECU: **0**

Figure 1.20: Closure confirms recorded decisions and the cumulative result.

Information, errors, and strategic voting

2.1 Why a panel can improve on an individual signal

A seventy-percent accurate signal is correct seven times out of ten in a long sequence of cases produced by the same procedure. That frequency neither makes a particular signal certainly true nor requires exactly seven correct signals in ten classroom cases. The program draws states and signals, so a small sample can differ substantially from the reference frequency. To understand the benefit of combining information, first assume that everyone votes exactly according to their signal and that errors are independent conditional on the state. This assumption defines a mechanical reference rather than a necessary description of human choices.

With three equally accurate signals, majority is correct when all three signals are correct or when exactly two are correct. The first event has probability 0.7 cubed. The second has three possible combinations, each with probability 0.7 squared times 0.3. Adding them gives 0.784. The improvement over an individual signal comes from allowing two correct observations to offset one isolated error. It does not come from the number three itself. If all screens reproduce the same source's error, this offsetting mechanism disappears, however many people receive the copies.

$$P(\text{correct majority}) = q^3 + 3q^2(1 - q) = 0.784, \quad q = 0.7.$$

The reference requires accuracy above one half, conditional independence, and votes that follow signals. If participants copy opinions, receive dependent evidence, or change votes strategically, the argument must be reconsidered. This is a useful occasion to teach what an assumption does. Saying that a majority knows more is insufficient: identify the additional information and how it enters the resolution. Nor should a panel be compared with a single person who observes the entire record. That person has different information from the individual private signal used in the benchmark calculation.

Theoretical foundation

The 78.4% reference applies to three independent signals of 70% accuracy and signal-following votes under majority. It is a mechanical benchmark of the design. A different observed rate can reflect small-sample variation, strategic voting, information processing, or a protocol deviation. The rate alone does not identify the explanation.

2.2 False positives, false negatives, and denominators

A false positive occurs when the fact is absent but the resolution declares it present. A false negative occurs when the fact is present but the resolution declares it absent. Both reduce accuracy, but an institution need not value them equally. Total accuracy divides correct cases by all completed cases. The false-positive rate divides false positives by cases with true state zero. The false-negative rate divides false negatives by cases with true state one. These denominators answer different questions and should be written down before interpreting a percentage.

Under sincere majority voting with seventy-percent accurate signals, the false-positive rate is 0.216 and the false-negative rate is also 0.216. With equiprobable states, each error contributes 0.108 to the set of all cases. Saying that false positives account for 10.8% of all cases is different from saying that 10.8% of absent-state cases are incorrectly resolved present. The latter conditional rate is 21.6%. This distinction arises even without selection problems or missing records, making it a good starting point for discussing denominators. It also clarifies why the confusion matrix should accompany average accuracy.

Sincere unanimity resolves present when all three signals indicate one. Given state one, that occurs with probability 0.343; the false-negative rate is therefore 0.657. Given state zero, three erroneous signals jointly indicate one with probability 0.027, the false-positive rate. Overall accuracy is 0.658 when the states are equiprobable. A stricter rule reduces one error and increases the other under this reference. Institutions cannot be ranked only by the number of required votes, and a lower false-positive rate does not demonstrate that every relevant outcome has improved.

$$FP_U = (1 - q)^3 = 0.027, \quad FN_U = 1 - q^3 = 0.657, \quad A_U = \frac{1}{2}(1 - FP_U) + \frac{1}{2}(1 - FN_U) = 0.658.$$

For practice, suppose twenty cases are complete: ten with state zero and ten with state one. Two false positives and four false negatives imply fourteen correct cases out of twenty, or seventy-percent accuracy. Conditional error rates are two out of ten, twenty percent, and four out of ten, forty percent. The same two false positives constitute ten percent of all cases. The confusion matrix keeps all four counts visible. If no state-zero case occurred, the corresponding error rate is not zero: it could not be computed, and the report displays a dash. A missing denominator cannot support a claim that a procedure avoided that error.

2.3 Bayesian updating through worked examples

Updating a probability means combining what was known beforehand with a new observation. The prior probability that the fact is present is one half. A positive signal from a seventy-percent accurate source is more likely if the fact is present than if it is absent. Bayes compares those possibilities. After a positive signal, the posterior probability is seventy percent; after a negative signal, it is thirty percent. These values use the equiprobable prior and a symmetric source. Changing prior frequency or introducing asymmetric source errors would change the meaning of the same observed signal.

$$P(\theta = 1 | s) = \frac{P(s | \theta = 1)P(\theta = 1)}{P(s | \theta = 1)P(\theta = 1) + P(s | \theta = 0)P(\theta = 0)}.$$

Here θ denotes the true state and s denotes a signal. The numerator is the joint probability of presence and the observed signal. The denominator adds the two ways the signal can appear, with the fact present or absent. Combining independent sources multiplies their conditional probabilities.

For equally accurate signals one, one, and zero, their joint likelihood is 0.7 times 0.7 times 0.3 if the state is one and 0.3 times 0.3 times 0.7 if it is zero. Normalizing these likelihoods gives a posterior of 0.7. The contrary evidence offsets one of the favorable signals.

With three independent positive signals, the posterior is 0.343 divided by 0.343 plus 0.027, approximately 92.7%. With three copies of one positive signal, the posterior remains seventy percent. The likelihood of one observation cannot be multiplied three times as though three new draws occurred. This example explains why the number of documents or messages is not automatically the amount of information. The source procedure is known in the exercise. Identifying dependence in an actual case requires additional investigation that this classroom environment does not simulate, and the mathematical example should not conceal that practical difficulty.

Heterogeneous reliability supplies another comparison. Suppose one positive signal has ninety-percent precision and two negative signals have sixty-percent precision. Joint likelihoods are 0.9 times 0.4 times 0.4 and 0.1 times 0.6 times 0.6. The posterior probability of presence is 0.144 divided by 0.18, or eighty percent. Two negative signals do not automatically outweigh one positive signal. Majority of votes is a different operation from selecting the most likely state after pooling observations of unequal quality. This distinction motivates examining what members know about one another's signal precision and how such knowledge affects their choices.

2.4 Being pivotal and voting strategically

A vote is pivotal when changing it, while holding the others constant, changes the resolution. Under three-person majority, this happens if the other two votes are split. Under unanimity for presence, it happens if both others vote present. A person seeking the correct resolution should assess what that event implies about the state, as well as use their own signal. Conditioning on being pivotal does not assume the person observes other votes: someone can reason about the event before knowing whether it actually occurs.

Suppose the other two members follow their signals sincerely. Under majority, the pivotal event includes one positive and one negative signal. With symmetric precision, they offset each other, and the private signal still tips the choice. Under unanimity, being pivotal implies two positive signals from the others. A member with a negative signal then combines one, one, and zero; the posterior probability of presence is seventy percent. With symmetric scoring, that member would prefer voting present in the pivotal event despite their negative signal. Signal-following behavior is therefore not an equilibrium of this unanimity rule under those assumptions.

This reasoning does not classify every contrary-signal vote as an error or deception. It may reflect strategic choice, communicated information, misunderstanding, or preferences that the scoring table does not capture. The report shows signal following and votes conditional on signal and precision. For an advanced equilibrium exercise, suppose a positive signal always produces a present vote, while a negative signal produces a present vote with probability x . Under unanimity, the negative-signal member must be indifferent in the pivotal event. Likelihood ratios express that condition without requiring the instructor to introduce an entire model of jury procedure.

$$\frac{0.3}{0.7} \left(\frac{0.7 + 0.3x}{0.3 + 0.7x} \right)^2 = 1.$$

Inside the parentheses, the numerator is the probability that another person votes present under state one; the denominator is that probability under state zero. The square reflects the two other members. The initial factor incorporates the person's own negative signal. An interior solution describes a symmetric strategic reference for this simplified design, assuming independence and common preferences. It is not an estimate of how every student will vote. The teaching value lies

in recognizing that a voting rule changes the information relevant to a consequential choice even when the accuracy-points table stays the same.

2.5 Error costs and choices at a given belief

The asymmetric table awards one hundred for a correct resolution, zero for a false positive, and sixty for a false negative. Relative to being correct, losses are one hundred for the first error and forty for the second. At probability p of presence, resolving present yields one hundred times p . Resolving absent yields one hundred times one minus p plus sixty times p . The first alternative is better when p exceeds five sevenths. A probability above one half is insufficient because a false positive is more costly in points than a false negative.

$$EU(1) = 100p, \quad EU(0) = 100(1 - p) + 60p, \quad EU(1) > EU(0) \Leftrightarrow p > \frac{5}{7}.$$

This comparison concerns choosing a resolution rather than an automatic instruction for every vote. In a collective choice, p should be the probability relevant when the person's vote changes the outcome. Even with the same ordinary belief, different rules can change the belief conditional on being pivotal. Five sevenths must not be equated with reasonable doubt or another legal standard. It follows from an experimental scoring table selected to teach relative error costs. Feedback should distinguish an ECU loss from legal reasons for valuing a safeguard even when it does not maximize this particular score.

Attention

Do not interpret fewer present resolutions as greater accuracy. Compare all four counts and both conditional error rates, and preserve the loss table. Changing the rule and changing error costs are separate treatments. Asymmetric variants include both voting rules so that these changes can be examined separately.

2.6 Academic evidence and limits of the adaptation

Guarnaschelli, McKelvey, and Palfrey (2000) compare group size, rule, and deliberation. They find strategic votes under unanimity but fewer outcomes analogous to convicting innocents than under majority, despite the motivating theoretical comparison. This supports measuring individual behavior and collective errors separately. Our scores and protocols are classroom adaptations, not a replication or a transfer of their effect sizes.

Feddersen and Pesendorfer (1998) show that strategic voting can alter advantages attributed to unanimity: requiring more votes need not reduce errors. Their argument has defined assumptions and does not establish that every unanimous tribunal is inferior. Coughlan (2000) adds communication and mistrials, changing the comparison. Our channel is limited, and lack of unanimity resolves absent; it does not implement his mistrial mechanism.

Source, precision, and review extensions are original comparisons whose predictions follow from explicit parameters. We do not attribute empirical findings to articles that did not test them. A classroom pattern can fit a mechanism; a causal claim requires comparable variation, sufficient independent observations, and control of order. Documenting these conditions connects institutions, behavior, and inference without turning a demonstration into a universal conclusion.

Matching and institutional extensions

3.1 Designing an interpretable comparison

An extension is useful when it changes an identifiable mechanism and permits a question that the basic game cannot answer. The eleven extended configurations are not difficulty levels or a required sequence. Some supply necessary controls: communication is crossed with two rules, asymmetric losses also use both rules, precision assignment is public or private, and review has versions with and without cost. Choosing only one may demonstrate a procedure, but does not separate the effects of dimensions that change together. Let the teaching question determine the comparison before following the menu's order.

Keep language, case count, the relevant scoring table, and preparation constant when comparing groups. If the same people run majority before unanimity, they carry experience from the first block. Varying order across groups can help examine that issue, but a short class does not eliminate every experience effect. Do not present the difference between two sessions as a precise causal effect without comparable assignment. Record matching and order alongside external communication or connectivity incidents. Transparent records are part of experimental learning, particularly when the apparent result is surprising or differs from a reference prediction.

3.2 One case, fixed panels, and reassigned panels

A single case shows the complete sequence without prior within-session experience. It is suitable for introducing signals, votes, and resolutions, checking devices, and practicing ledger interpretation. One outcome cannot estimate a rule's accuracy or compare error rates reliably. If two panels solve the same type of problem, their states and signals are drawn independently; they are not necessarily examining the same record. An observed difference combines procedure, the particular information received, and sampling variation. Students should understand this before comparing two panels' decisions as if only their reasoning differed.

Ten fixed-panel cases preserve members and positions. Facts and signals are drawn anew, and past outcomes do not change the distribution of current cases. Members may learn the rule, interpret one another's behavior, or develop conventions without seeing current votes before choosing. The report retains evolution by case to describe such patterns. Do not attribute a trend to learning without considering changing states and signals. Repetition does not turn the previous signal into evidence about the current state, and repeated observations from one panel do not create ten independent groups of people.

Reassignment keeps positions while mixing members between cases. It prevents an entire panel immediately recurring when alternative groups exist; it does not promise that every pair of people differs in every case. Comparing it with fixed panels opens a question about shared experience

and coordination of votes. It leaves signal precision, the rule, and scoring unchanged. With only three participants, composition cannot change, so the substantive comparison disappears. Explain matching accurately instead of labeling an activity anonymous when students can recognize partners from the physical organization of the room.

3.3 Unanimity without messages

Unanimity requires three present votes instead of two. Other conditions match the basic majority game: independent seventy-percent accurate signals, equiprobable states, and symmetric scores. Every combination with fewer than three present votes resolves absent. This operational detail prevents confusing an asymmetric veto with a procedure that requires unanimity for either resolution. The variant examines consequences of an asymmetric vote requirement; it does not fully reproduce a criminal safeguard or a process with repeated deliberation until agreement.

The sincere reference lowers false positives from 21.6% to 2.7% while raising false negatives from 21.6% to 65.7%. Participants may nevertheless change their ballots when the rule changes. Inspect present votes conditional on a negative signal before examining collective errors. An increase in those votes is compatible with the pivotal reasoning developed earlier. Ask whether the rule changed relevant information even though the signal and scoring table stayed the same. The experimental jury literature motivates this question; it does not determine the error rates a particular class will produce.

3.4 Messages under majority and unanimity

Both communication variants add the same stage. Each member states which signal they claim to have received and which resolution they recommend. Messages are submitted without seeing others and appear together once the panel finishes sending them. Each person then reports another probability and votes. The program permits a claim to differ from the actual signal and does not pay for truthfulness. There is no free conversation, successive exchange of arguments, moderation, or bargaining over agreement. Deliberation in the catalog refers to this structured protocol, and that definition should remain explicit in discussion.

The channel may share information, coordinate votes, or transmit claims that are not credible. Enabling a form guarantees none of these functions. If everyone truthfully communicates independent signals, the panel can reason from a larger record. If claims are strategic, interpretation depends on procedure and expectations about others. Compare silence and messages while retaining the same voting rule. Comparing majority and unanimity with messages asks whether the channel changes responses to the rule; it need not imply that communication has the same effect under both institutions. The two comparisons address different mechanisms.

The report adds the proportion of claims matching actual signals and the mean change in reported probability. Claim accuracy uses messages from completed cases as its denominator; it does not measure the quality of a conversation that never occurred. A zero average change can hide increases for some people and decreases for others. Use exports to study dispersion or absolute changes when that question matters. In feedback, ask what new information arrived, which part could be verified before voting, and how a false claim might affect another member. Do not identify participants or turn an aggregate indicator into a moral judgment of the group.

Send your message

Position C · First panel · Case 1 / 10

Your signal indicates: **Fact present (1)**

Your signal's precision: **70%**

Others will see both responses next to your position. They will not see your probability. The claim is not verified before voting and is unpaid.

Which signal do you claim to have received?

☐ Fact present (1) ☐ Fact absent (0)

Which resolution do you recommend?

☐ Fact present (1) ☐ Fact absent (0)

Continue

Figure 3.1: Send a signal claim and recommendation through the enabled channel.

Panel messages

Position C · First panel · Case 1 / 10

These are participants' claims; they may differ from their true signals. Record your probability again and then cast your vote.

Position	Claimed signal	Recommendation
A	Fact present (1)	Fact absent (0)
B	Fact present (1)	Fact absent (0)
C	Fact present (1)	Fact absent (0)

What probability, from 0 to 100, do you assign to the fact being present?

Continue

Figure 3.2: Distinguish claims from true signals before reassessing the fact.

3.5 Asymmetric error costs under both rules

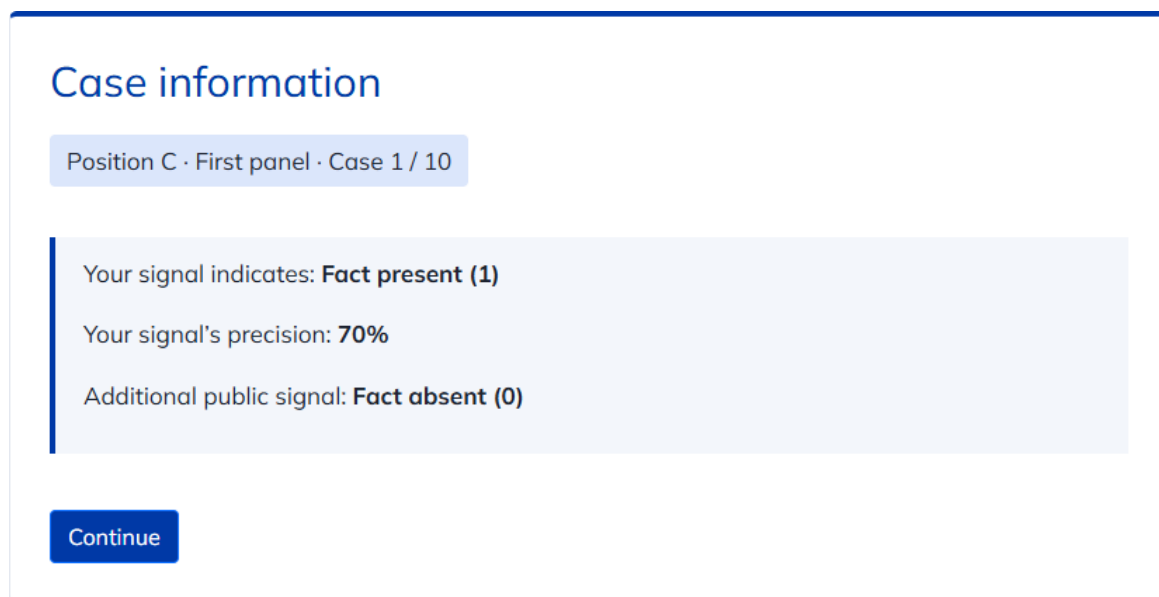
These variants change the loss table while preserving sources. Either correct resolution awards one hundred ECU. A false positive awards zero and a false negative sixty. Majority requires two present votes and unanimity requires three. The controls allow comparing one table across rules or one rule across tables. The change does not assert that participants have identical personal preferences about errors; it supplies a common, verifiable incentive in the experimental case. Preferences concerning real legal outcomes are broader than the points assigned here.

An ordinary seventy-percent belief would justify selecting present in an individual decision with symmetric losses, but falls below the five-sevenths threshold for asymmetric losses. A collective ballot must also consider beliefs in the pivotal event. Ask what changes before asking which procedure is better. Inspect both error rates and each person's score: accuracy and points can rank the same sequence differently. Fewer false positives accompanied by more false negatives may increase points when the table penalizes the errors unequally. That accounting result is informative about the specified objective, not a complete normative assessment.

3.6 An additional public source and copies of one source

Additional public evidence preserves three independent private signals and adds a common seventy-percent accurate signal, independent of private sources conditional on the state. Everyone sees the same public value. Each person can combine two available sources: private and public. If both indicate present, the ordinary posterior is approximately 84.5%; if they disagree, it is fifty percent. Choices may become more similar because everyone received common information. Such similarity does not prove that external communication occurred or that every person ignored their

own signal.



Case information

Position C · First panel · Case 1 / 10

Your signal indicates: **Fact present (1)**

Your signal's precision: **70%**

Additional public signal: **Fact absent (0)**

Continue

Figure 3.3: Observe the extra public evidence while counting it as one source.

Copies of one source removes independent private signals: all three positions receive the same seventy-percent accurate draw. The interface states that these are copies. Voting according to them under majority produces seventy-percent accuracy rather than 78.4%. This is a teaching comparison about complete dependence, not evidence that shared information is always inferior. Additional public evidence adds a source, whereas the copy variant replaces the source structure. They therefore do not form a comparison that changes publicity alone. Their difference should be described as changing available information and dependence.



Case information

Position A · First panel · Case 1 / 10

Your signal indicates: **Fact absent (0)**

Your signal's precision: **70%**

This is a copy of the single source shared by all three members.

Continue

Figure 3.4: Do not count three copies as independent draws.

Ask students to reconstruct the posterior after three independent positive signals and compare it with three positive copies. The contrast is approximately 92.7% versus seventy percent. Ask which assumption is violated by multiplying a copy's likelihood three times. The report shows behavior and accuracy but does not directly observe how a person counted sources. Probability responses provide another description, and discussion may explore interpretations. No single indicator proves that every participant used the same internal reasoning, and the guide should not

treat a convenient explanation as an observed fact.

3.7 Public and private heterogeneous reliability

In every case, one signal has ninety-percent precision and the other two sixty-percent precision. Assignment is drawn again while positions remain stable. In the public version everyone knows the precision of A, B, and C. In the private version each person sees their own precision, while knowing that one high and two low signals exist. A member assigned ninety can infer that the others receive sixty. A member assigned sixty knows the others include one of each type without knowing their positions. This changes informational detail rather than removing all knowledge of the panel.

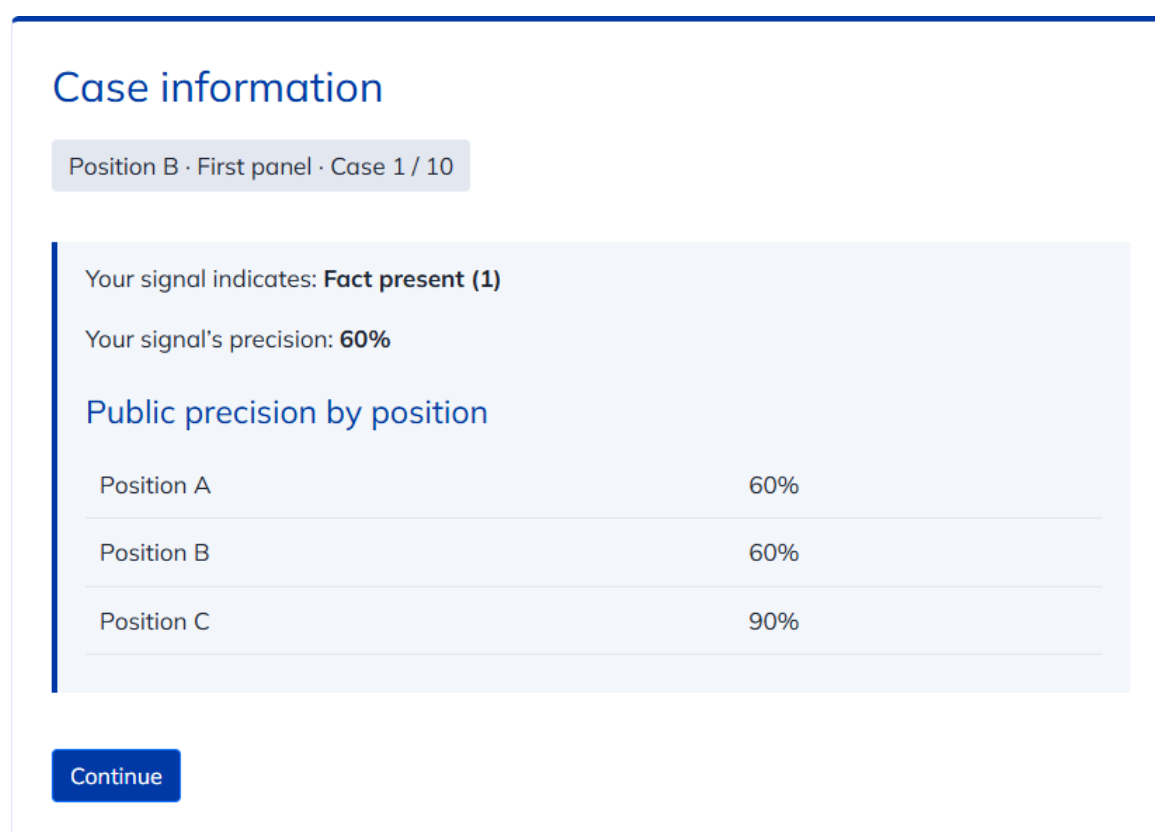


Figure 3.5: Public reliability identifies each position's precision.

Sincere majority with these precisions has a mechanical accuracy reference of 79.2%. That does not describe the best possible use of the entire record. A ninety-percent positive signal can outweigh two sixty-percent negative signals in the joint posterior, as the earlier worked example shows. Without communication, members reason about their influence and others' information. Inspect votes conditional on precision and signal, not only position. The most reliable signal moves between positions across cases; A's average does not describe the behavior of a permanent expert. Publicly identifying the precision assignment can matter without giving extra formal voting authority to that member.

3.8 Review with and without cost

Review connects two three-person panels, for six participants in each case. The first instance uses majority and independent signals. An initial present resolution automatically opens review; an

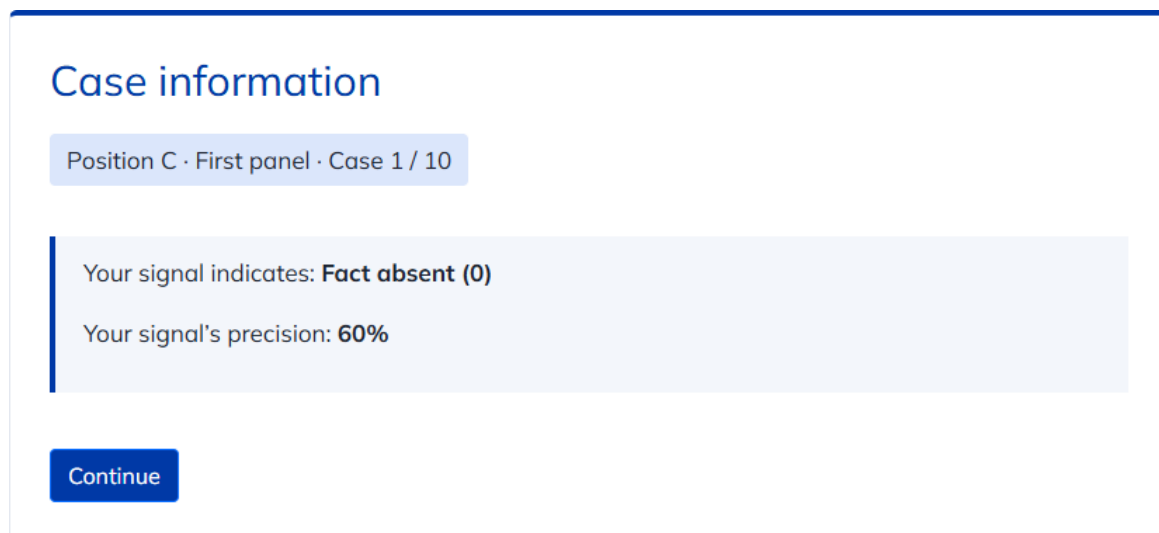


Figure 3.6: Private reliability shows own precision and the common assignment procedure.

absent resolution becomes final. The second panel observes the true signals, precisions, and votes from the first instance, and each reviewer receives a new seventy-percent accurate signal. The state remains hidden. The second majority can affirm or reverse. Both panels receive points based on the final resolution, giving the review stage actual consequences for all six people.

Both variants supply ten additional ECU per person in every case. Opening review deducts zero in the costless variant and five from each person in the costly variant. Without review, neither deducts anything. The common balance prevents negative scores and permits a cost comparison with the same structure. Cost does not transfer to another participant. In the second stage it has already been incurred by opening review, so affirming or reversing does not change it. In the first stage, a positive resolution does change whether the cost applies. Initial incentives may respond to that possibility rather than treating the deduction as irrelevant.

Review is selective and cannot correct an initial absent resolution. It can correct false positives but also introduce false negatives by reversing a correct present resolution. The report separates corrected errors and lost correct decisions among reviewed cases. Reversals alone do not establish an overall improvement. Comparing review with the basic game changes participant count, information, and sequence: it is an institutional package. Comparing costly with costless review preserves those elements and changes cost. These are original teaching adaptations for discussing procedure. They do not reproduce a real appeals system or a human defendant's choice to appeal.

Record and review evidence

Position E · Review panel · Case 1 / 10

The record contains the true signals recorded by the program. Your new signal is an additional source. The true state remains hidden. The second panel resolves by majority, and its resolution replaces the initial one.

Initial resolution: **Fact present (1)** · First-panel present votes: 3

First-panel record

Position	True signal	Your signal's precision	Vote
A	Fact absent (0)	70%	Fact present (1)
B	Fact absent (0)	70%	Fact present (1)
C	Fact absent (0)	70%	Fact present (1)

Your signal indicates: **Fact absent (0)**

Your signal's precision: **70%**

Continue

Figure 3.7: The second panel combines a true record with new private evidence.

Record your vote

Position E · Review panel · Case 1 / 10

Your signal indicates: **Fact absent (0)**

Your signal's precision: **70%**

Votes are simultaneous and private until this stage closes. Select one of the two options.

Which resolution do you choose for this case?

☐ Fact present (1) ☐ Fact absent (0)

Continue

Figure 3.8: Cast a review vote before the true state is revealed.

Case result

Position E · Review panel · Case 1 / 10

True state: **Fact absent (0)**

Initial resolution: **Fact present (1)** · First-panel present votes: 3

Review opened: Yes

Review present votes: 0

Final resolution: **Fact absent (0)**

Accuracy points	100 ECU
Additional balance	10 ECU
Review cost	5 ECU
Your result in ECU	105 ECU

Signals and votes at closure

Position	True signal	Your signal's precision	Vote
A	Fact absent (0)	70%	Fact present (1)
B	Fact absent (0)	70%	Fact present (1)
C	Fact absent (0)	70%	Fact present (1)
D	Fact absent (0)	70%	Fact absent (0)
E	Fact absent (0)	70%	Fact absent (0)
F	Fact present (1)	70%	Fact absent (0)

Continue

Figure 3.9: Separate the initial resolution, final resolution, and applied cost.

Report, discussion and work with data

4.1 Progress and controlled disclosure

The instructor report is available from the session's administrative link. On opening it, the instructor sees language, edition, treatment, participant count, completed cases and pending cases. Aggregates remain hidden until Reveal aggregates is pressed. Hide aggregates allows the class to return to a prediction without leaving outcomes projected. Disclosure belongs to that browser view; reopening the report starts with the results hidden. The report does not automatically send these statistics to participants, and its projectable tables contain no individual participant codes.

The outcome unit is a completed collective case. A basic session of six people contains two panels; ten active cases therefore yield twenty collective outcomes. It does not yield sixty independent outcomes, although sixty participant records exist. In the review protocols, six people form one connected decision unit, so ten cases yield ten final resolutions. The progress counter distinguishes these structures from the outset. A case is complete only after its applicable final stage. An initial resolution still awaiting review is not counted as a completed final outcome.

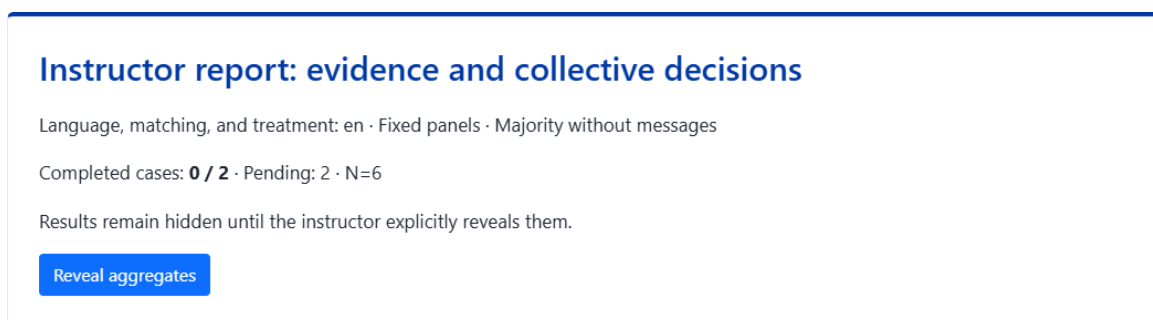


Figure 4.1: Check progress before revealing decisions to the class.

Open the report before the activity, during partial completion and after closure. A ten-case session can have its first case completed while nine remain pending for each panel. Percentages use closed cases rather than the planned horizon. Announce the observed n when revealing an interrupted session and explain why cases are missing. Do not supply zero outcomes for missing decisions or confuse an unanswered vote with a resolution of absent. The technical pages after a single-case configuration do not represent nine lost cases: only the configured active horizon belongs in the report.

4.2 Metrics, distributions and stage comparisons

Initial and final accuracy use completed cases as their denominator. Without review they coincide; with review they can differ. False-positive and false-negative rates use subsets defined by the true state. Initial disagreement counts panels with one or two present votes and divides by all completed initial panels. This measures ballot dispersion, rather than interpersonal conflict or belief disagreement. Three identical votes can arise from common sources, coordination, coincident independent signals or strategic reasoning. A unanimous observed ballot does not identify which mechanism generated it.

Signal-following divides initial votes matching their own signal by initial votes in completed cases. It excludes reviewers' ballots, since reviewers observe an additional record and face a different information set. Mean probability and mean squared error likewise describe the initial beliefs of the first panel. Conditional voting tables distinguish first panel from review, signal value and precision. They use actual ballots. A reviewer whose case never opens has no review ballot and is absent from that voting denominator; the export preserves the corresponding field as not applicable.

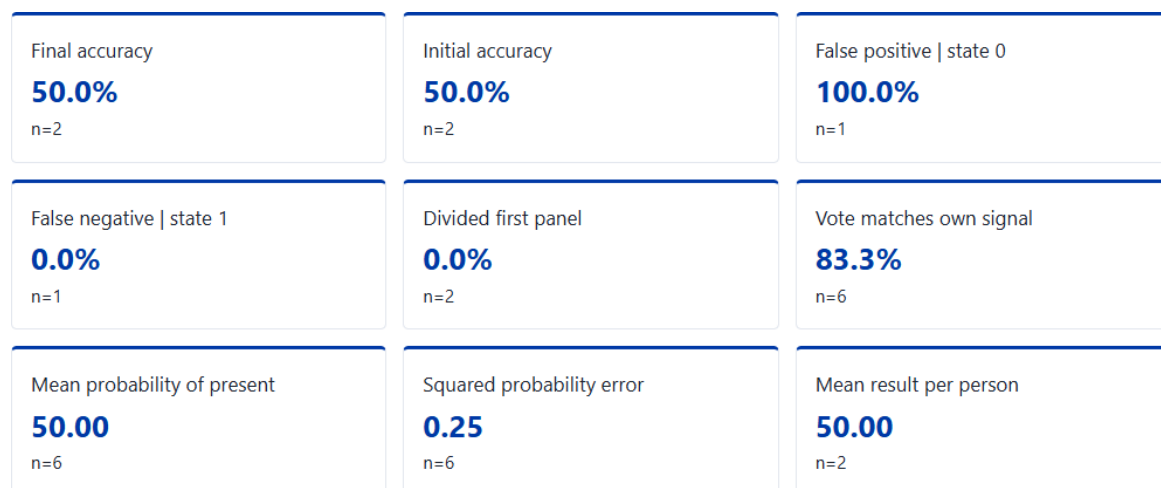


Figure 4.2: Read each statistic together with its denominator.

The vote distribution counts how many initial panels cast zero, one, two or three present votes. Each case contributes one count, independently of the number of records in its panel. The probability distribution groups initial responses into 0–24, 25–49, 50–74 and 75–100. These four bars describe concentration, extremes and diversity; they do not estimate a causal relationship. A mean near fifty can conceal half of the responses near zero and half near one hundred. Inspecting a distribution helps the class notice features that a single average suppresses.

Distribution of first-panel present votes

n=2

Observations	n	%
0	0	0.0%
1	0	0.0%
2	0	0.0%
3	2	100.0%

Figure 4.3: Inspect the vote distribution before reducing it to an average.

Resolution	Initial resolution	Final resolution
Present and present resolution	1	1
Absent and absent resolution	0	0
Absent and present resolution	1	1
Present and absent resolution	0	0

Figure 4.4: Compare the four initial and final outcome counts.

The evolution table retains one row per active case, displaying n, accuracy, conditional errors and payoff per person. Rows without observations display dashes. The series identifies descriptive changes, but the same participants return across cases and are not fresh independent samples. With review, the difference between initial and final accuracy measures the observed change in the complete sequence. Correction and deterioration rates divide by reviewed cases, a subset selected by an initial present resolution. They do not estimate what the second panel would have done with every possible case.

Case	n	Final accuracy	False positive state 0	False negative state 1	Mean result per person
1	2	50.0%	100.0%	0.0%	50.00

Figure 4.5: Keep each active case and its denominators visible.

4.3 Probabilities and calibration

A probability's squared error is its distance from the observed state, squared, after converting the percentage to a fraction between zero and one. A seventy-percent response in a present case produces 0.09; in an absent case it produces 0.49. Averaging these errors describes probabilistic performance. This indicator is unpaid and is not presented to participants as a prize. Leaving beliefs unincentivized reduces operational complexity, but limits what an instructor can infer about sincere probability reporting. Neither its classroom availability nor its numerical precision removes that limitation.

$$B_i = (p_i - \theta)^2, \quad p_i = \frac{\text{probability response}}{100}.$$

A perfectly calibrated seventy-percent forecast need not predict every case correctly. Across comparable cases reported at that value, the fact should be present approximately seven times out of ten. The error in one isolated case does not identify poor calibration. A low mean score also does not prove Bayesian updating or strategic rationality. Compare responses under equivalent information and acknowledge the available sample. In message protocols, the change between initial belief and belief after the transcript describes stated updating. Message truthfulness and factual accuracy remain different objects even when their aggregate values happen to coincide.

For five responses of fifty percent, suppose the states are one, zero, one, zero and one. Every error is 0.25, and so is their average. For five responses of one hundred percent against these same states, three errors are zero and two are one, producing a mean of 0.4. The second series chose the more frequent binary category, present, but expressed stronger certainty than these realizations supported. This exercise separates binary decision accuracy from probabilistic performance without asking the instructor to recover a true distribution from five observations.

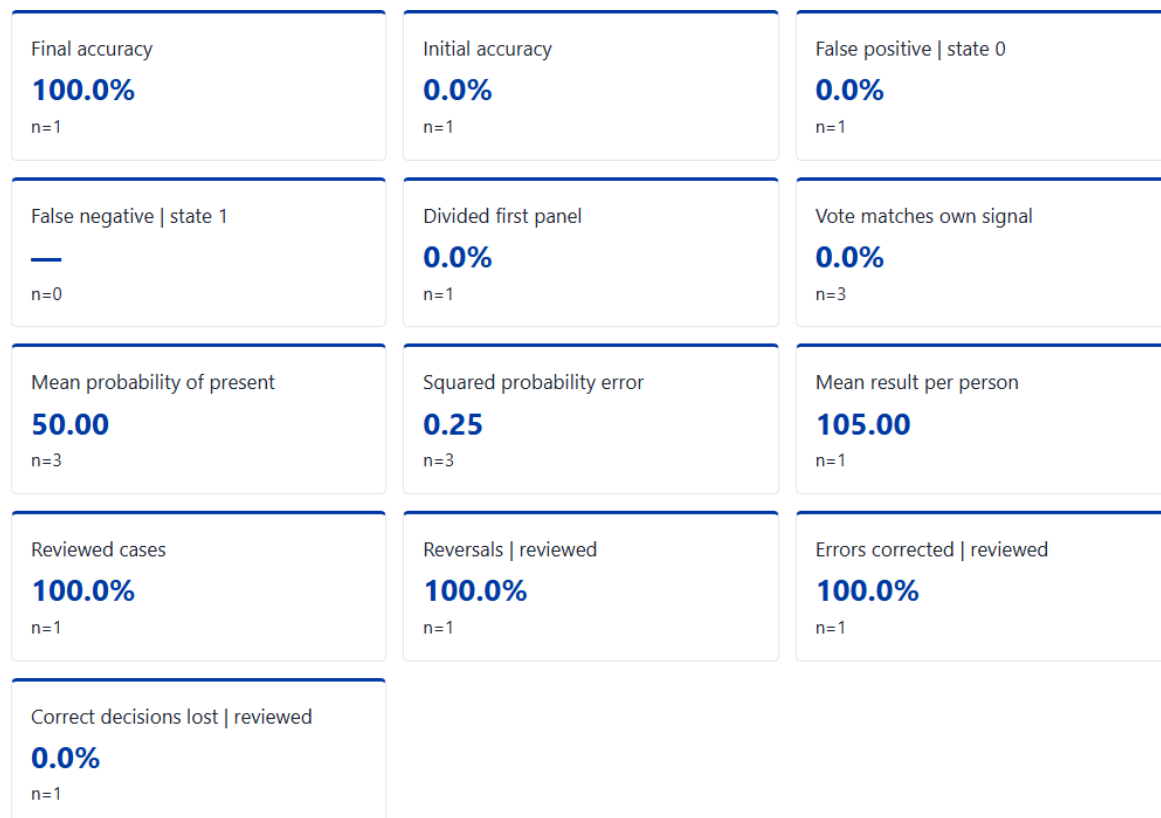


Figure 4.6: Distinguish reviews, reversals, corrections and deteriorations.

4.4 A fully worked report

Consider a hypothetical twelve-case example, with six state-zero cases and six state-one cases. The first panel has five true positives, four true negatives, two false positives and one false negative. These four counts sum to twelve; nine correct outcomes yield 75% accuracy. The false-positive rate is two out of six, or 33.3%, and the false-negative rate is one out of six, or 16.7%. The two false positives constitute 16.7% of all cases. This is an arithmetic exercise, not an observed implementation result or a finding reported by a research paper.

Selective review opens seven records, since five true positives and two false positives have an initial present resolution. Suppose review corrects both false positives but mistakenly reverses one true positive. The final matrix contains four true positives, six true negatives, zero false positives and two false negatives. Correct outcomes rise from nine to ten, producing 83.3% accuracy. The conditional false-positive rate falls to zero while the false-negative rate rises to 33.3%. Three of seven reviewed cases were reversed, two were corrected and one deteriorated. A reversal is therefore not synonymous with an improvement.

The accuracy increase is one net correction divided by twelve, approximately 8.3 percentage points. The correction rate of 28.6% divides two by seven; it is not the change in accuracy across the whole session. No initially absent record was reviewed, so the original false negative remained. Selection explains which improvement this institution can deliver and which error this protocol cannot remedy. Do not describe reviewers' performance on these seven selected records as their accuracy for all possible cases in the population.

Under costless review, mean final payoff per person is one hundred times ten divided by twelve plus the ten-ECU case balance, approximately 93.33 ECU. In the costly variant, seven of twelve cases deduct five ECU from each member. The average falls by thirty-five divided by twelve, approximately 2.92, to 90.42 ECU. This calculation holds states and ballots identical across variants:

it is a counterfactual accounting comparison used to explain the charge. Ballots themselves could change between institutions. Review can raise accuracy while reducing net payoff relative to a control, depending on the cost and the improvement's magnitude.

Feedback script

Start from the four outcome counts and their denominators. Ask what changed in information, decision rule or losses, and which alternative explanation could generate the same aggregate. In review, distinguish reversal from correction. Close by asking about procedural protections or costs that an ECU score alone cannot evaluate.

4.5 A seventy-five-minute class

An illustrative schedule reserves ten minutes for access, groups and comprehension; twenty for a basic activity or a short variant block; twenty-five for reconstructing the report and comparing benchmarks; fifteen for a worked exercise; and five for closure. Seventy-five minutes is a planning example rather than a verified duration for every treatment. Message and review protocols can require more operational time. For an introductory course, use one case and more discussion; if the objective concerns evolution across cases, retain ten cases and reduce the other activities in that class meeting.

Before beginning, ask whether a more demanding rule will change errors without announcing a preferred answer. Request a prediction of which error will decrease or increase and the assumption supporting it. During decisions, resolve access difficulties privately. After closure, project progress first and announce n ; then reveal the matrix and distribution. Ask a student to calculate each rate from the counts. Finally compare the sincere benchmark with voting conditional on signals, and discuss reasons for voting differently. This sequence connects arithmetic to procedure rather than leaving students with a memorized percentage.

For graduate students, derive the indifference condition under unanimity and examine what changes when losses change. For undergraduates, use the heterogeneous-signal example: one precise observation can shift a posterior more than two less precise observations. For a law course, close by distinguishing factual truth, error distribution, participation and legitimacy. The latter dimensions support discussion but are not measured by the game's accuracy score. A numerical payoff result cannot establish that a procedural guarantee lacks value. Students should state both what their evidence supports and what remains outside the modeled institution.

Three exercises check transfer. First, hold the total at twelve cases but change the number of state-zero cases: which rates can change even if the false-positive count stays fixed? The answer depends on the denominator and should be given as a fraction. Second, ask for the posterior associated with three positive copies of one source: under this design's prior, it is 70%, because there is only one observation. Third, propose a review that reverses two correct decisions and corrects one error: three reversals occur, but net accuracy worsens. The solution requires classifying changes, not merely counting different decisions.

4.6 Export, privacy and operational difficulties

The export retains session, individual code, case, group, position, panel, language, treatment, mode, state, sources, precision, probabilities, messages, ballots and balance. It includes completion status. Initial zero values for payoffs or decisions in an open case are not observed outcomes. Reviewers' inapplicable responses remain absent. For case-level analysis, retain one row per group

and case. Do not count a final resolution six times because six participant records carry it. For ballot analysis, retain the stage and the information available to that voter rather than pooling first-panel and reviewer choices indiscriminately.

Codes link responses and assist operational recovery, but they should neither be projected nor connected with names or student numbers during feedback. No demographic questions or free text are collected. Before using classroom data for research, establish the relevant protocol and consent. This guide does not declare nonexistent approvals or registrations. Classroom aggregates can support teaching without automatically forming a sample designed to produce a publishable estimate. Interpretation should reflect participation, interruptions, treatment assignment and the small number of collective cases actually observed.

If someone cannot advance, check whether the page is waiting for other panel members. For a stale route, use the current session link without deleting the database. If a selector shows another family, check the instance before creating sessions there. If a review stage appears missing, distinguish a record that never opened from an unanswered ballot: an initial absent resolution legitimately skips the second vote. Preserve incidents and sessions when updating the server. Operational recovery should retain the procedure rather than editing ballots to make a counter reach its planned total.

Checkpoint

Preparation is complete when language, treatment and participant count are correct; the instructor retains the private report; decisions and waits work; and a full case can close, including review where applicable. During class, record interruptions and reveal outcomes with explicit denominators.

Bibliography

- Guarnaschelli, Serena, Richard D. McKelvey and Thomas R. Palfrey. 2000. An Experimental Study of Jury Decision Rules. *American Political Science Review* 94(2): 407–423. <https://doi.org/10.2307/2586020>.
- Feddersen, Timothy and Wolfgang Pesendorfer. 1998. Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting. *American Political Science Review* 92(1): 23–35. <https://doi.org/10.2307/2585926>.
- Coughlan, Peter J. 2000. In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting. *American Political Science Review* 94(2): 375–393. <https://doi.org/10.2307/2586018>.
- American Economic Association. JEL classification: C92 laboratory group experiments; D71 committees and social choice; D82 private and asymmetric information; K41 litigation process. Catalog consultation: October 2026. <https://www.aeaweb.org/econlit/jelCodes.php?view=jel>.