



Instructor Guide: Social Punishment and Formal Sanctions

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 social punishment and formal sanctions games available through 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 classroom use	1
1.1 Access through the updated laboratory website	1
1.2 Purpose and game summary	3
1.3 Access through the laboratory	7
1.4 Preparing the session	7
2 Rules, screens, and accounts	9
2.1 The order of decisions	9
2.2 Closing and accumulated payoffs	11
2.3 Individual payoffs and the extended account	14
3 Incentives and theoretical predictions	16
3.1 A monetary benchmark	16
3.2 Why costly responses may occur	17
3.3 Substitution, complementarity, and mechanical effects	17
3.4 Repetition, history, and limits	18
4 Institutional extensions	19
4.1 A map of twelve treatments	19
4.2 Norms, fines, and social responses	19
4.3 The recipient as responder	21
4.4 Compensation and the choice of response	21
4.5 Visible history and continuity	24
4.6 A fourth participant and second-order sanctions	24
4.7 Connections to experimental research	25
5 The instructor report	27
5.1 Reveal results at a deliberate moment	27
5.2 Indicators, denominators, and zeros	27
5.3 Distributions, roles, and round patterns	29
5.4 Export and reconstruction	30
6 Classroom use and exercises	31
6.1 A sixty-minute lesson	31
6.2 A ninety-minute lesson	31
6.3 Exercises with worked answers	32
6.3.1 The threshold and monetary incentives	32
6.3.2 Fine, reduction, and the recipient's payment	32
6.3.3 A capped reduction and additional expenditure	32
6.3.4 Choosing between two responses	33

6.3.5	An empty denominator	33
6.3.6	Interaction and compliance costs	33
6.3.7	Second-order accounting	33
6.4	Assessing learning	34
6.5	Interruptions and responsible closure	34

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	Current basic-edition instructions, with initial resources and experimental units. Continuation 1/2.	5
1.8	Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.	6
2.1	Welcome and general rules.	9
2.2	Current comprehension check with an incorrect answer, its specific explanation, and retained responses.	10
2.3	Transfer decision by A.	11
2.4	Preview of the response.	12
2.5	Results and payments by role.	13
2.6	Closure and accumulated payment.	14
4.1	Social response after the fine.	20
4.2	B responds with a capped reduction.	22
4.3	One budget for both responses.	23
4.4	Fixed-group history.	24
4.5	Second-order decision.	25
5.1	Report with hidden aggregates.	27
5.2	Indicators and denominators.	28
5.3	Transfer distribution.	29
5.4	Payments by role.	30
5.5	Round evolution.	30

Preparation and classroom use

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.

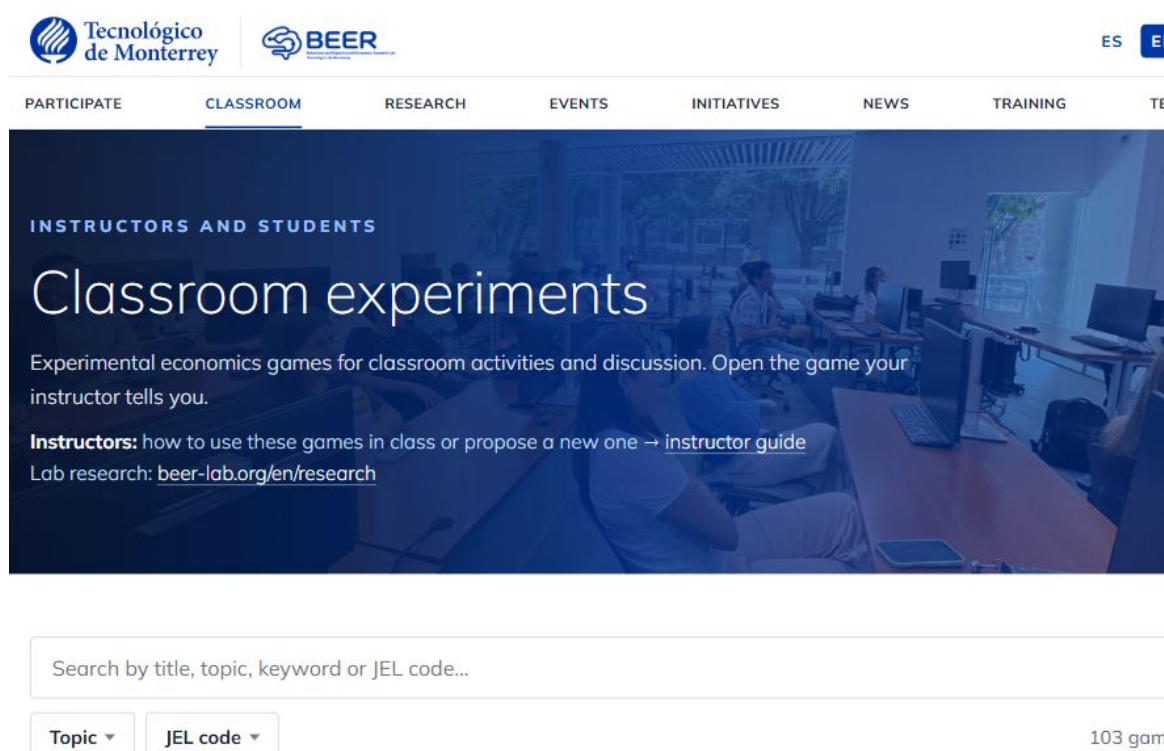


Figure 1.2: Current catalog: search by title, topic, or JEL code. Then locate this family's card and select language and edition. Captured on October 7, 2026.

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.

al y sanción formal — una interacción

al y sanción formal — 10 rondas, grupo fijo

al y sanción formal — 10 rondas, reasignación

ctendidas — Castigo de terceros: control

ctendidas — Regla sin multa

ctendidas — Regla y multa moderada

ctendidas — Regla y multa disuasiva

ctendidas — Multa sin respuesta social

ctendidas — Regla sin multa ni respuesta social

ctendidas — Respuesta del receptor

ctendidas — Compensación por un tercero

ctendidas — Elección entre reducción y compensación

ctendidas — Grupo fijo con historial

ctendidas — Cuarto participante: control

ctendidas — Sanciones de segundo orden

Castigo social y sanción formal / Social Punishment and Formal Sanctions — BEER Lab

Selecione una variante de este juego. Select a version of this game.

Figure 1.3: Demo contains only the configurations of this game family, in both languages and editions. Continuation 1/2.

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.

ctions — Fixed group with history

1.2 Purpose and game summary

ctions — Second-order sanctions

A rule can influence behavior through several channels. It can state what people are expected to do, attach a monetary consequence to a violation, or provide other people with a reason to intervene. These channels need not reinforce each other. A fine may increase a transfer while reducing an observer's willingness to punish. It may also reduce the balance available for subsequent punishment without changing anyone's preferences. This activity makes those possibilities visible

BE	ctendidas — Respuesta del receptor
	ctendidas — Compensación por un tercero
	ctendidas — Elección entre reducción y compensación
	ctendidas — Grupo fijo con historial
	ctendidas — Cuarto participante: control
	ctendidas — Sanciones de segundo orden
	hment and Formal Sanctions — one interaction
	hment and Formal Sanctions — 10 rounds, fixed group
	hment and Formal Sanctions — 10 rounds, rematching
	nctions — Third-party punishment: control
	nctions — Rule without a fine
	nctions — Rule and mild fine
	nctions — Rule and deterrent fine
	nctions — Fine without a social response
	nctions — Rule without a fine or social response
	nctions — Recipient response
	nctions — Third-party compensation
	nctions — Choice between reduction and compensation
	nctions — Fixed group with history
	nctions — Fourth participant: control
	nctions — Second-order sanctions

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

Rooms

Current rooms:

[ES] Castigo social y sanción formal
[EN] Social Punishment and Formal Sanctions
[ES] Castigo social y sanción formal — Extendido
[EN] Social Punishment and Formal Sanctions — Extended

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



Activity instructions

Resources and unit of account

Each round A receives 100 ECU and everyone else receives 20 ECU each.

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

This session does not convert ECU into real money.

Your role: Person C · Round 1 / 1

Each round A receives 100 ECU and everyone else receives 20 ECU each. A sends B an integer amount between 0 and 100. This amount is subtracted from A and added to B. Your role stays the same throughout the session.

This session has no required minimum transfer and no automatic deductions.

After observing the transfer and automatic deduction, Person C may spend between 0 and 10 ECU to reduce A's balance. Each ECU spent reduces A by 3 ECU, up to A's remaining balance. The full chosen cost is paid even when the reduction is capped. The reduced amount is not given to anyone.

Group decisions and balances are shown at the end of the round, identified only by role. Names and information from other groups are not displayed.

The session has 1 round(s). The group stays fixed. All your payoffs from those rounds are added together.

Calculation example

These numbers illustrate the accounts; they do not recommend a choice. Expenditures are deducted from the person choosing them.

Transfer from A to B: **20 ECU** · Automatic deduction: **0 ECU**

Spending to reduce A: 5 ECU · Spending to increase B: 0 ECU

Effective reduction of A: 15 ECU · Externally funded increase for B: 0 ECU

Balances by role: Person A = 65 ECU Person B = 40 ECU Person C = 15 ECU

Continue

Con tecnología de [oTree](#)

Figure 1.7: Current basic-edition instructions, with initial resources and experimental units.
Continuation 2/2.

Room: [EN] Social Punishment and Formal Sanctions

Create a new session

Session Config

[EN] Social Punishment and Formal Sanctions — one interaction

Number of participants

Use a multiple of 3 participants.

Create

[Configure session](#) 

App sequence

social_sanctions_game (10 rounds)

Social punishment and formal sanctions

Figure 1.8: Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.

through separate decisions and accounts.

The game belongs to the economic analysis of law. Its central question concerns the relationship between an explicit norm, an automatic sanction, and a costly social response. It does not represent the law of a particular jurisdiction. A sanction here is a monetary consequence inside the activity; it is not a criminal sentence, a judicial finding, or a measure of culpability. The legal discussion should examine compliance, distribution, compensation, proportionality, and the authority to impose consequences as distinct questions.

Purpose	Distinguish an explicit norm, a formal sanction, and a social response; compare punishment and compensation.
Experimental unit	One case is one group in one active round.
Basic edition	One interaction; ten rounds with a fixed group; ten rounds with reassignment.
Extended edition	Twelve selectable treatments, including controls for norms, fines, social response, and a fourth participant.
Decision	A transfers to B; B or C chooses response expenditures where applicable; D may respond in the second-order variant.
Payoff	Each round starts with 100 ECU for A and 20 for every other person. All active rounds are added.
Topics	third-party punishment; formal sanctions; legal norms; compensation; compliance; second-order sanctions.
JEL	C92; D91; K42.
Participants	Three per group; four for second-order sanctions and its control.
Suggested duration	30–45 minutes of activity; 60–90 including discussion.

The control uses a unilateral transfer followed by an opportunity to reduce the sender's balance. That mechanism also appears in the Dictator family. Here it provides a comparison for institutions that enforce rules. The distinctive material separates the announcement of a norm, the size of a fine, and the availability of social intervention. Instructors do not need to repeat the entire set of distribution games before using this family.

By the end, students should be able to reconstruct an individual account, distinguish a transfer from the destruction of resources, and explain why less punishment does not necessarily mean a weaker norm. They should also recognize that someone can comply with the announced rule while producing an unequal distribution. A sound explanation identifies the institutional change, the variable that responds, and the claim that remains outside the exercise's reach.

1.3 Access through the laboratory

Start at <https://beer-lab.org/> and continue to the teaching catalog at <https://teaching.beer-lab.org/>. Find Social Punishment and Formal Sanctions. The planned instructor entry is <https://sanctions.beer-lab.org/demo>, with the room directory at <https://sanctions.beer-lab.org/rooms>. The English basic room is:

https://sanctions.beer-lab.org/room/sanctions_en

The extended room is:

https://sanctions.beer-lab.org/room/sanctions_extended.

Spanish rooms end in `sanciones_es` and `sanciones_extendido`. These are the product's intended stable routes; public availability depends on laboratory deployment.

Demo should display thirty configurations from this family. Rooms should contain four rooms, corresponding to the basic and extended editions in both languages. Each basic room offers three modes; each extended room offers twelve treatments. The room determines which options are available. Confirm language, edition, and treatment before creating a session. If unrelated families appear, consult the laboratory team before distributing links: the instance's catalog is not correctly isolated.

A room link assigns a place in the active session. An individual participation link belongs to one place and should not be shared among students. The instructor panel requires authentication; its report is not a participation room. A local preview uses the port assigned to its instance. A localhost address works on the computer running the server and is not automatically accessible from students' phones or other computers.

1.4 Preparing the session

Choose a participant count divisible by three for ordinary treatments. Use a multiple of four for the fourth-person control and second-order sanctions. The room page updates its group-size instruction when a treatment is selected. Do not create a session for a multiple of three and then attempt to convert it to four-person groups after participants have entered. If attendance changes, prepare a separate session with a valid size before beginning decisions.

In each round A begins with 100 experimental currency units, abbreviated ECU. B and C each begin with 20; D also receives 20 when present. Roles remain fixed across rounds. This makes the role-specific payment patterns interpretable but means the activity does not rotate everyone through each position. Explain that role assignment is part of the design and that a high starting endowment is not a reward for effort, ability, or prior conduct.

The basic edition offers one interaction, ten rounds with a fixed group, and ten rounds with reassignment. With at least two groups, reassignment avoids reproducing the complete previous group. It does not guarantee that every individual meets entirely new people: some pairwise encounters may recur. With a single group, different partners are unavailable, so the group necessarily remains unchanged. Do not describe that situation as a test of stranger matching.

Extended treatments all use ten rounds and fixed groups. This holds the horizon and continuity constant when comparing their rules. The platform retains ten technical rounds even in a one-interaction session, but only the configured active rounds are shown, paid, and reported. A participant should reach a deliberate closing screen after the last active round. The exercise does not ask students to click through empty rounds.

For the instructor

Before the class, select a comparison and write down its outcome of interest. For a norm question, compare the control with the explicit norm without a fine. For formal enforcement, retain the same norm while changing the fine or social-response opportunity. For second-order sanctions, compare two four-person conditions. A menu of treatments is useful only when its contrasts answer identifiable questions.

Use one device or browser profile for each participant. Keep individual links private and give everyone the same explanation of the conversion between ECU and any classroom reward. The game records ECU payoffs; it does not promise an external cash payment. Announce the payment procedure before decisions and avoid changing it after observing outcomes. If participation is instructional without payment, state that arrangement clearly.

The instructor should open the private report before starting and verify that its aggregates are hidden. Invite written predictions before revealing any class data. Early display can alter subsequent choices, especially when groups are still working. If the lesson deliberately provides public feedback, document its timing as part of the procedure rather than treating it as an invisible administrative action.

Checkpoint

Confirm the language, treatment, active horizon, matching mode, group size, participant count, and payment instructions. Test a participation link and the private report from separate browser contexts. Check that students see descriptive titles and their assigned roles. Keep production labels out of slides, spoken instructions, and handouts. Record any departure from the planned classroom procedure.

Rules, screens, and accounts

2.1 The order of decisions

Participants first see the institutional welcome screen and acknowledge the instructions. They then read the rules and answer a short comprehension check. Incorrect answers produce a message and another opportunity to respond. The check asks about the starting endowment, the fine specified by the treatment, and whether B receives money removed by a social reduction. It does not ask for a preferred transfer or imply that a particular social decision is correct.



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. Decisions use ECU, the accounting units of this activity.

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

☐ I have read the general rules.

Start

Powered by [oTree](#)

Figure 2.1: Welcome and general rules.

At the beginning of each active round, the role and horizon are visible. A chooses an integer transfer x from zero to one hundred. B receives that amount in addition to the initial twenty. In the control, no minimum transfer is imposed or announced. In norm treatments, the announced threshold is thirty. A transfer below thirty triggers the treatment's fine; a transfer equal to or



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

Check the rules

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.

Your role: Person C · Round 1 / 1

How many ECU does A receive at the start of each round?

Your answer: 101. The correct answer is 100. Review the instructions and try again. This attempt does not deduct ECU.

How much is automatically deducted from A if A sends zero?

When someone reduces A's balance, how much does B receive from that reduction?

Continue

Con tecnología de [oTree](#)

Figure 2.2: Current comprehension check with an incorrect answer, its specific explanation, and retained responses.

above thirty does not. The fine is automatic and certain within the activity.

Let f denote the fine specified in the treatment. Its realized amount is $F = f \mathbf{1}\{x < 30\}$ when the norm applies and zero otherwise. In the explicit-norm condition without a monetary fine, $f = 0$. In the moderate and formal-only conditions, $f = 20$; in the strong condition, $f = 40$. The fine leaves the participants' accounts. It is not transferred to B and does not count as compensation.

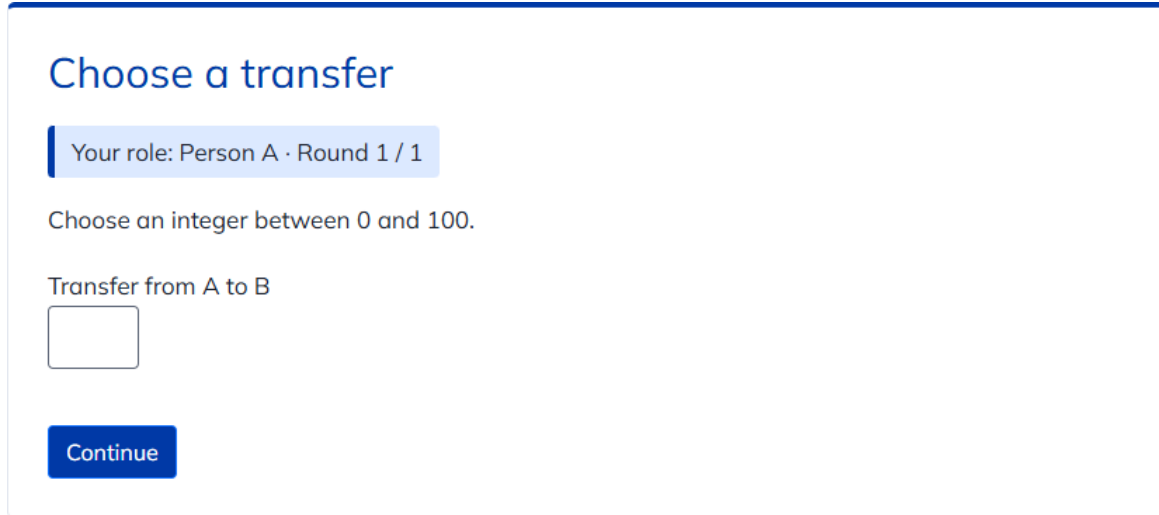


Figure 2.3: Transfer decision by A.

After the transfer and any formal deduction, the eligible responder sees the transfer, the fine, and the balance available for a further reduction. C is the responder in the control. C may spend an integer z between zero and ten. Each unit costs C one unit and nominally reduces A's balance by three. The effective reduction is $L = \min\{3z, 100 - x - F\}$. The account never becomes negative. The entire chosen expenditure is still paid if the cap prevents some of the nominal reduction.

In the recipient-response variant, B makes that decision instead. In compensation, C spends c to add $3c$ to B, financed from outside the group. In the choice variant, C may do both with $z + c \leq 10$. In formal-only and norm-without-sanctions treatments, no social response is available. The interface displays only decisions supported by the selected treatment. A missing response stage is an institutional condition, not a recorded refusal to intervene.

In second-order sanctions, a fourth person D acts after C. D observes A's transfer and C's expenditure, then chooses expenditure w from zero to ten to reduce C's remaining balance. The reduction is $M = \min\{3w, 20 - z\}$. D pays the entire w . This stage is absent in the otherwise comparable fourth-person control, where D remains passive. The institution does not tell D that C has violated a duty by spending zero.

Wait pages synchronize members of the group. They ensure that later decisions use completed earlier decisions. A passive participant may therefore wait while someone else legitimately acts. When all required decisions are submitted, the game settles the accounts once and displays the results. Refreshing a page does not create an additional fine or multiply the same expenditure. Participants then proceed to the next active round or to the final summary.

2.2 Closing and accumulated payoffs

The results screen lists the group's decisions, formal deduction, nominal and effective reductions, compensation where applicable, intervention expenditure, destruction, and totals. It also shows role-specific payoffs and the participant's own round payment. These entries let students reconstruct the outcome without having to infer other people's decisions from a single final balance. The same

Choose your spending

Your role: Person C · Round 1 / 1

Transfer from A to B: **20 ECU** · Automatic deduction: **0 ECU**

A's balance before the response: **80 ECU**

Spending to reduce A

Preview of your decision

Your cost: **5 ECU**

Effective reduction of A: **15 ECU**

Target's balance after the reduction: **65 ECU**

The full cost is paid even when the balance limits the reduction.

Continue

Figure 2.4: Preview of the response.

Round result

Your role: Person A · Round 1 / 1

Your balance: **65 ECU**

Transfer from A to B	20 ECU
Automatic deduction	0 ECU
Spending to reduce A	5 ECU
Requested reduction of A	15 ECU
Effective reduction of A	15 ECU
Spending to increase B	0 ECU
Externally funded increase for B	0 ECU
Total response expenditure	5 ECU
Resources destroyed	20 ECU
Sum of participant balances	120 ECU
Total including the external account, net of funding	120 ECU

Balances by role

Person A	65 ECU
Person B	40 ECU
Person C	15 ECU

Continue

Figure 2.5: Results and payments by role.

account is available to the instructor through the export.

All active rounds contribute to the accumulated payment. A one-interaction session pays its single active round; a ten-round session sums all ten. The final screen confirms that the activity is complete and shows the accumulated ECU amount. There is no additional decision button and no generic empty-page ending. Before leaving, participants should verify that they reached that screen rather than closing a browser while their group is still waiting.

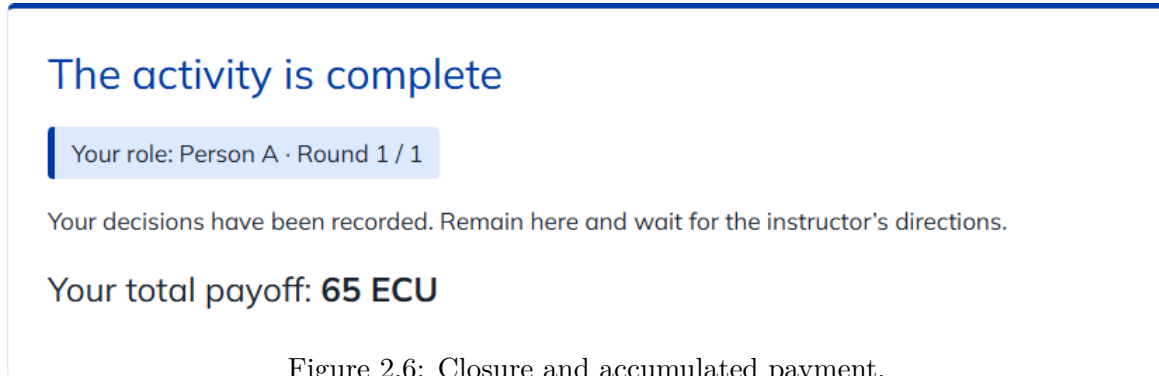


Figure 2.6: Closure and accumulated payment.

2.3 Individual payoffs and the extended account

For the usual three-person conditions with C responding, the payoffs are

$$\pi_A = 100 - x - F - L, \quad \pi_B = 20 + x + 3c, \quad \pi_C = 20 - z - c.$$

When B responds, its payoff also subtracts z and C retains twenty. In second-order sanctions, subtract M from C and give D $20 - w$. Compensation is absent in that variant. The initial participant total is 140 with three people and 160 with four. Endowments are renewed at the beginning of every active round rather than carried forward as a spending constraint.

Theoretical foundation

Keep three operations separate. A transfer moves resources between participants. A fine moves resources outside their accounts. A reduction destroys resources without crediting another participant. If T is the sum of participant payoffs, the extended account is $T + F - 3c$. It equals the initial endowment total minus response expenditures and effective reductions: $E - z - c - w - L - M$. This identity is an accounting check, not a welfare theorem.

For a transfer of twenty and a punishment expenditure of five without a fine, A finishes with sixty-five, B with forty, and C with fifteen. Their total is 120. Five units were spent and fifteen were removed from A. If the same decisions occur with a fine of twenty, A finishes with forty-five, while B and C remain at forty and fifteen. Participant payoffs total 100; adding the twenty collected outside the group gives an extended total of 120.

For a transfer of ninety-five and C's expenditure of ten, only five units remain available for reduction. A receives zero, B receives 115, and C receives ten. The total is 125. It is incorrect to subtract thirty from A and report a negative balance. It is also incorrect to reimburse C for the unused nominal reduction. The live preview presents both the personal cost and effective consequence so that this feature is known before submission.

Compensation requires a separate entry for outside financing. With a transfer of twenty and compensation expenditure of five, the final payoffs are eighty, fifty-five, and fifteen. Participants receive 150 in total. The external contribution is fifteen, so the extended total is 135. Participant

income rises by ten, but that increase is not net resource creation when the financing account is included. The five-unit difference from the initial extended total is C's expenditure.

Instructors can draw accounts on the board using one row per person and another for outside transfers. This makes it harder to mistake fines for compensation or expenditures for reductions. Always state whether a reported total includes the outside account. Two students can calculate different totals correctly if they use different boundaries; the task is to make those boundaries explicit before comparing them.

Incentives and theoretical predictions

3.1 A monetary benchmark

Begin with a narrow reference model: each person values only their own monetary payoff, understands the rules, and correctly anticipates the later choices. In a final response stage, expenditure reduces the responder's own payment and does not generate a private monetary return. The responder therefore chooses zero. This applies to punishment, compensation, and D's second-order response. It also applies to B's response after the transfer, because reducing A cannot recover B's own expenditure.

Anticipating zero social expenditure, A compares the amount retained after the transfer and formal fine. Without a fine, sending zero maximizes A's payoff. With a fine of twenty below thirty, sending zero yields eighty and sending thirty yields seventy. The monetary benchmark remains zero. With a fine of forty, sending zero yields sixty and sending thirty yields seventy, so A prefers the threshold. Sending more than thirty lowers A's payoff without avoiding any further formal deduction.

Transfers below the threshold do not gradually reduce the fine. Sending twenty-nine still incurs the entire deduction; sending thirty avoids it. This discontinuity makes the threshold an especially useful classroom example. Ask students to compare the payoff just below and at the rule rather than treating the fine as a continuous tax rate. Their prediction should follow the implemented schedule rather than an intuitive interpretation of the word penalty.

The announced norm and an equal division are different reference points. A starts with one hundred and B with twenty, so their balances become equal at a transfer of forty before sanctions. The explicit rule asks for thirty. A transfer of thirty complies but leaves A with seventy and B with fifty. An observer might evaluate compliance, equality, intentions, or some combination. The activity does not identify those considerations merely by observing that C spends a positive amount.

Feedback script

Ask students to complete two sentences before revealing results: "Under the monetary benchmark, I expect ... because ..." and "A different pattern could arise if ...". The first sentence should state assumptions and incentives. The second should identify an alternative mechanism without claiming it was measured. Return to both sentences after inspecting the distribution, not only its mean.

3.2 Why costly responses may occur

A person may value the recipient's payoff, dislike a distribution, want to enforce a perceived norm, or react to the sender's conduct. Such motives can make a costly response meaningful to that person. They are not interchangeable explanations. Punishment lowers A without increasing B. Compensation raises B without lowering A. A participant who cares primarily about helping B may therefore choose differently from someone who wants to impose a cost on A, even when both object to the original transfer.

A simple illustration gives C utility equal to C's own payoff plus a weight α on B's payoff. One unit of compensation costs one and raises B's payment by three. Its utility change is $-1 + 3\alpha$, so compensation is attractive if $\alpha > 1/3$, indifferent at equality, and unattractive below it in this stylized specification. This is an instructor's derivation for the implemented technology. The game does not estimate that parameter or claim that every choice follows this utility function.

Punishment does not raise B's payment, so the same simple concern for B alone cannot explain positive punishment after the transfer. An explanation would need an additional preference or an expected future consequence. That observation helps students see why a label such as prosocial behavior is insufficient to describe all costly interventions. Different actions affect different accounts and may respond to different features of the initial decision.

The presence of compensation also changes the accounting boundary. An increase in participant income is partly financed outside the group. A welfare argument would need to specify the value of redistribution, the opportunity cost of outside funds, and any consequences omitted from the activity. The extended total records the financing transparently. It does not prove that compensation is undesirable simply because intervention has a resource cost.

3.3 Substitution, complementarity, and mechanical effects

Suppose punishment expenditure is lower when a formal fine is present. One possibility is substitution: C regards the formal response as sufficient and reduces personal intervention. Another is that A sends more, leaving fewer cases that provoke a response. A third is mechanical: the fine leaves less balance available for a reduction, altering what expenditure can achieve. These explanations can coexist. The report separates expenditure and effective loss, but it does not ask the responder which explanation applies.

If a fine raises transfers more when social punishment is available, the two institutions may appear complementary with respect to that outcome. If its incremental effect is smaller, they may appear substitutable. These are outcome-specific comparisons. A pair of institutions could increase compliance while lowering private intervention expenditure. It is therefore misleading to describe the relationship with one word without specifying the outcome and contrast.

Conditioning responses on transfer intervals helps describe whether punishment differs at similar observed transfers. It is not a complete causal solution, since transfers themselves may change under the treatment. Cases that fall in the same interval under different institutions can come from different types of participants or different histories. Explain this selection issue in ordinary language: comparing people who made the same observed choice does not ensure that they would have made the same choice under the other rule.

3.4 Repetition, history, and limits

Ten rounds with the same people permit shared experience and expectations about future interactions. Under common knowledge of purely monetary preferences, the final costly response is still zero. The backward-induction argument can then proceed through the finite horizon. Uncertainty about preferences, learning, or conditional strategies can support different observed paths. These are possible explanations, not variables automatically identified by the game's repeated structure.

The history treatment displays the transfer, C's expenditure, and A's effective loss in previous rounds of the same group. Its appropriate comparison is the fixed-group ten-round control, which preserves partners without supplying that exact table. Comparing history with reassignment would change both continuity and visible records. Nor is this a public reputation market: outsiders and future groups do not see the record.

Repeated observations should not be mistaken for independent samples. One group contributes ten related cases in a ten-round session. A changing average may reflect learning, strategic response, fatigue, or a few unusual trajectories. The class can inspect those possibilities and propose additional evidence. It cannot infer a stable behavioral law from a handful of repeated groups simply because the number of rows is large.

Attention

Do not use the exercise to conclude that a real fine is proportionate, that a person deserves punishment, or that a student who does not intervene lacks moral commitment. The choices come from a simplified monetary environment. The lesson examines mechanisms and arguments; it does not assess culpability or replace a legal justification for an institution.

Institutional extensions

4.1 A map of twelve treatments

Treatments are selected when a session is created. They are not stages that become active during a game or levels that participants must complete in sequence. Each extended session retains its parameters and fixed groups for ten rounds. Instructors can select a pair for a short lesson or allocate several conditions across a larger class. Comparisons should preserve language, initial information, duration, and payment procedures. Prior exposure to other treatments is also part of the conditions of a later activity.

The first treatment is the third-party punishment control. A chooses a transfer and C can spend ten units to reduce A's balance. The second announces a threshold of thirty without a fine. The third adds a fine of twenty, and the fourth raises it to forty. The fifth applies the twenty-unit fine without a social response. The sixth retains the norm but removes both the fine and social-response opportunity. The remaining treatments change the responder, allow compensation, offer a choice between punishment and compensation, display history, add a passive fourth person, or permit a second-order response.

Controls count as treatments because they make differences interpretable. Removing a control to add a more dramatic mechanism can weaken the lesson. For instance, comparing a second-order condition with a three-person group changes both D's authority and the composition of the group. The four-person control preserves composition while leaving D without a decision. Specify that contrast before seeing the outcomes, rather than searching for a comparison that supports an appealing interpretation afterward.

4.2 Norms, fines, and social responses

The control does not announce an obligation to transfer thirty. Initial endowments and the possibility of punishment may generate expectations, but the program does not present those expectations as an explicit rule. Norm without Fine does announce the threshold. Compare transfers and C's responses between those conditions. The contrast supports a discussion of expressing a common expectation, subject to treatment assignment and the small scale of a class. It does not by itself measure legitimacy or internal acceptance of the norm.

The moderate-fine variant retains the norm and deducts twenty when the transfer is below thirty. The deduction occurs before C decides. C therefore observes the transfer, formal consequence, and available balance. The strong-fine variant retains this procedure but deducts forty. Comparing those two conditions changes a single monetary parameter. In the absence of a social response, the moderate fine leaves zero attractive to a purely monetary A, whereas the stronger fine makes the threshold preferable.

Do not expect social punishment to rise uniformly with a fine. C may think the institution has already responded, focus on the distribution received by B, or face a smaller balance available for reduction. The interface displays the effective reduction before expenditure is confirmed. Expenditure and loss are reported separately, but the program does not identify which reason motivated C. Encourage students to propose compatible explanations and then identify evidence that would distinguish them.

Choose your spending

Your role: Person C · Round 1 / 10

Transfer from A to B: **20 ECU** · Automatic deduction: **20 ECU**

A's balance before the response: **60 ECU**

Spending to reduce A

5

Preview of your decision

Your cost: **5 ECU**

Effective reduction of A: **15 ECU**

Target's balance after the reduction: **45 ECU**

The full cost is paid even when the balance limits the reduction.

Continue

Figure 4.1: Social response after the fine.

Four treatments form a factorial comparison with the same thirty-unit norm. All four cells retain the same endowments and horizon. The strong-fine condition addresses a different question and should not replace one of these cells without changing the interpretation of the contrast.

	No social response	Social response available
No fine	Norm without sanction mechanisms	Norm without Fine
Fine of twenty	Formal Sanction Only	Moderate Formal Sanction

To study transfers, first calculate the change associated with adding a fine when no social response is available. Then calculate the change when C can respond. The difference between those changes is a descriptive interaction. Suppose the four cell means, read across the first row and then the second, were ten, twenty, twenty-five, and twenty-eight. The fine would raise the transfer by fifteen without social response and by eight with it. The interaction would be minus seven. These are arithmetic examples, not session results or published empirical estimates.

To study C's expenditure, use only the cells in which C can respond. Absence of an opportunity to punish is not a voluntary decision to spend zero. Compliance can rise while social expenditure falls. That pattern is compatible with substitution, but a causal claim requires defensible assignment and attention to the transfers being observed. Conditioning responses on the same transfer interval helps describe the pattern; it does not remove every identification problem caused by selecting cases on a treatment-affected decision.

The factorial contrast can also be calculated for the extended total, the effective reduction, or participant payoffs. Each outcome answers a different question. A higher transfer might coexist with lower total resources if punishment becomes more frequent. A lower private expenditure might coexist with more money leaving the group through fines. Present those consequences together rather than selecting only the outcome that makes one institution appear successful.

4.3 The recipient as responder

In Recipient Punishment, B takes the decision assigned to C in the control. B starts with twenty, receives the transfer, and can spend up to ten. C keeps twenty and remains passive. The reduction technology is unchanged, and there is no norm or fine. B's expenditure reduces B's own payment. With a transfer of twenty and expenditure of five, A receives sixty-five, B thirty-five, and C twenty. Their total is 120.

The common expenditure limit prevents a difference in formal punishment capacity. It does not equalize wealth at the decision stage: B has received the transfer and C has not. It also does not equalize their relationship to A's choice. B is directly affected, while C observes the distribution. A treatment difference therefore combines position, receipt of the transfer, and available income. Discuss it as a contrast between second-party and third-party responses, not an isolated estimate of impartiality.

Consider a transfer of ninety-five followed by B's expenditure of ten. A has five left and loses only five; B pays the full ten. Final payments are zero for A, 105 for B, and twenty for C. The decision may seem surprising but is admissible. The program does not presume that punishment always targets a low transfer, and it does not delete choices because an instructor finds them morally inconsistent. This freedom should be explained before students treat observed responses as evidence of a particular norm.

4.4 Compensation and the choice of response

In Compensation, C can spend up to ten to raise B's balance. Each unit creates three units for B through external financing. A does not lose money from this intervention. With a transfer of twenty and compensation expenditure of five, final payments are eighty, fifty-five, and fifteen. Their sum is 150. Subtracting the outside contribution of fifteen gives an extended total of 135. The intervention redistributes with a five-unit resource cost when the financing account is included.

This condition separates helping the recipient from reducing the income of the initial decision maker. It does not declare either response legally superior or reproduce a complete judicial remedy. The increase can exceed a shortfall relative to a reference point chosen by students. There is no assessment of proven harm, party request, evidentiary hearing, or adjudication. Interpret the outcomes as allocation decisions under a known response technology.

In Choice between Punishment and Compensation, two expenditure fields appear. C allocates a single budget of ten between them and may use either, both, or neither. The sum cannot exceed ten. The interface does not double the budget, and the server rejects ten plus one. With a transfer

Choose your spending

Your role: Person B · Round 1 / 10

Transfer from A to B: **95 ECU** · Automatic deduction: **0 ECU**

A's balance before the response: **5 ECU**

Spending to reduce A

Preview of your decision

Your cost: **10 ECU**

Effective reduction of A: **5 ECU**

Target's balance after the reduction: **0 ECU**

The full cost is paid even when the balance limits the reduction.

Continue

Figure 4.2: B responds with a capped reduction.

of twenty, punishment expenditure of three, and compensation expenditure of two, A receives seventy-one, B forty-six, and C fifteen. Participant payoffs sum to 132. Subtracting six in outside financing yields an extended total of 126.

Choose your spending

Your role: Person C · Round 1 / 10

Transfer from A to B: **20 ECU** · Automatic deduction: **0 ECU**

A's balance before the response: **80 ECU**

Spending to reduce A

Spending to increase B

Preview of your decision

Your cost: **5 ECU**

Effective reduction of A: **9 ECU**

Target's balance after the reduction: **71 ECU**

Externally funded increase for B: **6 ECU**

The full cost is paid even when the balance limits the reduction.

Continue

Figure 4.3: One budget for both responses.

A useful comparison examines total expenditure in this condition against expenditure when only one response is available. Another examines the share allocated to compensation among people who spend a positive amount. For that question, exclude nonspenders from the share's denominator. If the question instead concerns how many eligible responders choose any compensation, include all eligible responders. Both statistics are legitimate, but they describe different margins of behavior.

Changes in the response menu may also change A's expectations. A might anticipate that an observer will compensate B rather than reduce A's own balance. The resulting transfer could then change even if the responder's total expenditure remains similar. The sequence of moves matters for interpretation: the menu is announced before A acts. Do not treat C's later decision as an intervention that was unknown to A unless a different information structure has actually been implemented.

4.5 Visible history and continuity

Visible History retains the control's ten rounds and fixed group. From the second round onward, it adds a table with the transfer, C's expenditure, and A's effective loss in every earlier completed round. All group members see the same record. It does not display other groups' choices or personal names. It does not rank students, award prestige points, or disclose responses to future groups.

Compare this treatment with the fixed-group control. Both permit participants to remember earlier interaction; the treatment supplies an exact record. A difference may reflect assisted memory, attention to a trend, or greater ease in conditioning a subsequent choice. It is not the pure effect of moving from anonymity to publicity. If reputation is discussed, define it as a history within a stable relationship of known duration.

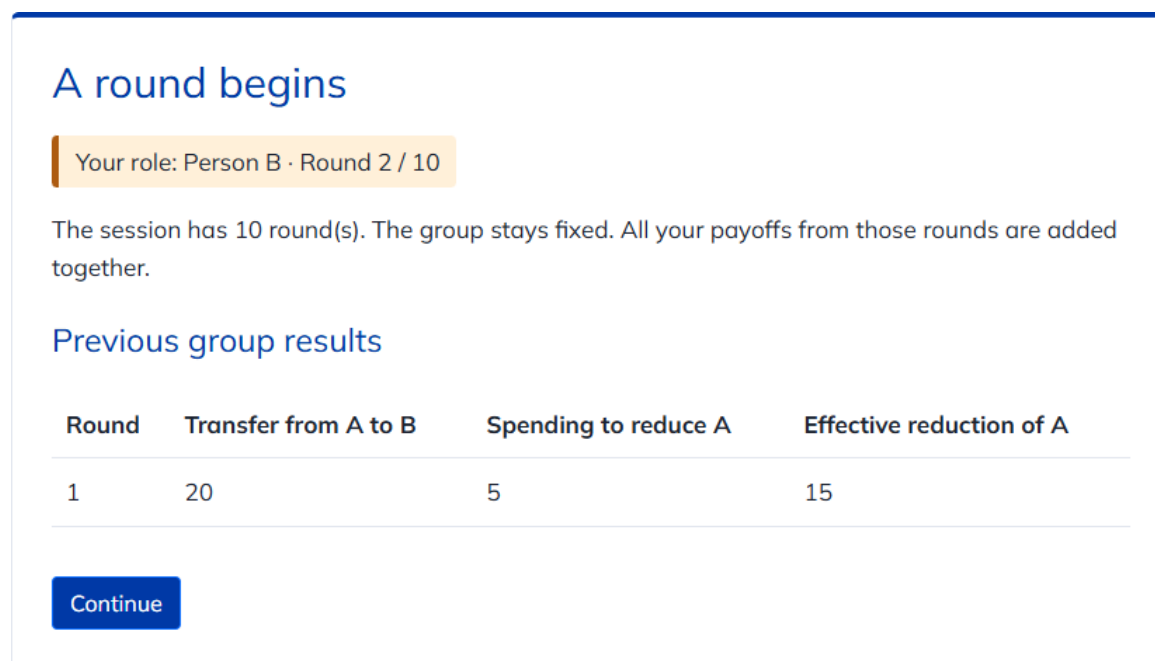


Figure 4.4: Fixed-group history.

The history display also distinguishes expenditure from consequence. Two identical expenditures can produce different reductions when balances differ. A participant observing the record can therefore respond to the cost accepted by C or the loss imposed on A. The activity records both but does not determine which feature participants attended to. This makes it useful to ask students what a history table communicates and what remains unobserved.

4.6 A fourth participant and second-order sanctions

The four-person control adds D with twenty units and no decision. A and C keep the control's rules. In Second-order Sanctions, D can spend ten units to reduce C's balance after observing the transfer and C's expenditure. Each unit costs D one and reduces C by up to three, without creating a negative balance. D can respond after zero or positive expenditure by C. The activity does not announce a duty for C to punish or classify abstention as a violation.

With a transfer of twenty and C spending five, A receives sixty-five and B forty. C retains fifteen before D acts. If D spends ten, the nominal reduction of thirty is capped at fifteen. C finishes with zero and D with ten. The final total is 115, compared with an initial 160. C and D together

spent fifteen; the two effective reductions destroyed thirty. Their sum of forty-five equals the total resource loss.

The treatment permits observation of responses to abstention, intervention, or both. It does not assume that D sustains a shared norm. To interpret a sanction against C, inspect the initial transfer and earlier intervention. Punishing nonintervention after a low transfer and punishing intervention after a high transfer can have different meanings. The report separates zero and positive expenditure by C, but that division alone does not establish motivation.

Decide about C's balance

Your role: Person D · Round 1 / 10

Transfer from A to B: 20 ECU · Spending to reduce A: 5 ECU

C's balance before your decision: **15 ECU**

Spending to reduce C

Preview of your decision

Your cost: **10 ECU**

Effective reduction of C: **15 ECU**

The full cost is paid even when the balance limits the reduction.

Continue

Figure 4.5: Second-order decision.

When comparing second order with its control, the extra participant and initial total are held constant. Comparing it with a three-person condition can still be informative descriptively, but it combines institutional authority with group composition. The passive D control is therefore part of the substantive design. Explain why a person with no decision may still be essential to identifying the comparison of interest.

4.7 Connections to experimental research

Fehr and Fischbacher experimentally examine third-party punishment and social norms. Their design provides a foundation for discussing responses by people who did not directly receive an allocation. This activity uses its own teaching endowments, menus, and conditions and is not an exact replication. Its control introduces the mechanism before additional institutions are considered.

Kube and Traxler study the interaction between legal and social norm enforcement. Their evidence of partial displacement of social punishment motivates a question, not a result the classroom must reproduce. Here the question is approached through controls separating norms, fines, and intervention opportunities. The numerical parameters and factorial illustration are specific to this teaching adaptation.

Leliveld, van Dijk, and van Beest compare punishment and compensation and examine empathic concern. Related response options appear here, but this activity does not measure empathy. Kiyonari and Barclay investigate second-order incentives and emphasize differences between reward and punishment. Their research cautions against assuming systematic punishment of nonpunishers. The four-person allocation variant is an adaptation rather than a replication of their social dilemmas.

The instructor report

5.1 Reveal results at a deliberate moment

The report is part of the session's private instructor panel. Its aggregates are initially hidden. Instructors can reveal them after predictions have been recorded and hide them again to continue discussion. This presentation control does not alter decisions or payoffs. Avoid projecting individual links or participant lists while discussing aggregates. The aim is to inspect patterns without identifying the student who made a particular transfer.

The progress counter shows completed cases out of possible cases. Six participants in groups of three over ten rounds generate twenty possible cases. If one group completes its first round while the other is still waiting, there is one completed case. The mean uses that case and does not treat unfinished decisions as zeros. Eight participants in groups of four also produce twenty cases over ten rounds. Participants, cases, and response opportunities are different quantities.

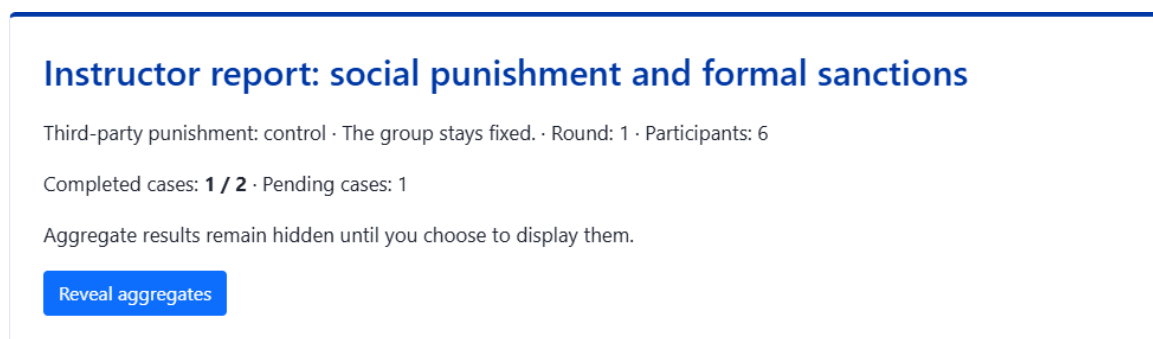


Figure 5.1: Report with hidden aggregates.

Technical rounds after a one-interaction activity are excluded. The report traverses active rounds and can be consulted from different administrative pages within the session. Participants do not need to advance through inactive rounds to make their payments appear. If all participants have reached the closing screen but the report shows fewer cases than expected, first check that the correct session is open and that the expected count reflects its active horizon and group size.

5.2 Indicators, denominators, and zeros

The violation rate is the share of cases with a transfer below thirty when an explicit norm exists. Without that norm the indicator is inapplicable and appears as a dash. An analyst can later use thirty as a common reference point across conditions, but doing so does not retroactively make the control an announced-obligation treatment. Distinguish the institution participants experienced

from a threshold used after the fact to summarize outcomes.

The punishment rate divides positive punishment decisions by opportunities to punish. The capped-reduction rate divides punishments whose effective loss is less than three times expenditure by positive punishments. If nobody punishes, the second denominator is zero and the report displays a dash. A dash does not mean lost data or an estimated zero. In Formal Sanction Only, the absence of a social-response opportunity likewise produces inapplicable indicators.

For second order, rates show the frequency of positive expenditure by D conditional on C having spent zero or a positive amount. Each cell displays its own case count. If C never abstains, the first rate is undefined. Do not fill it with zero to make a visually symmetrical comparison. An empty category is information that students should learn to recognize rather than conceal.

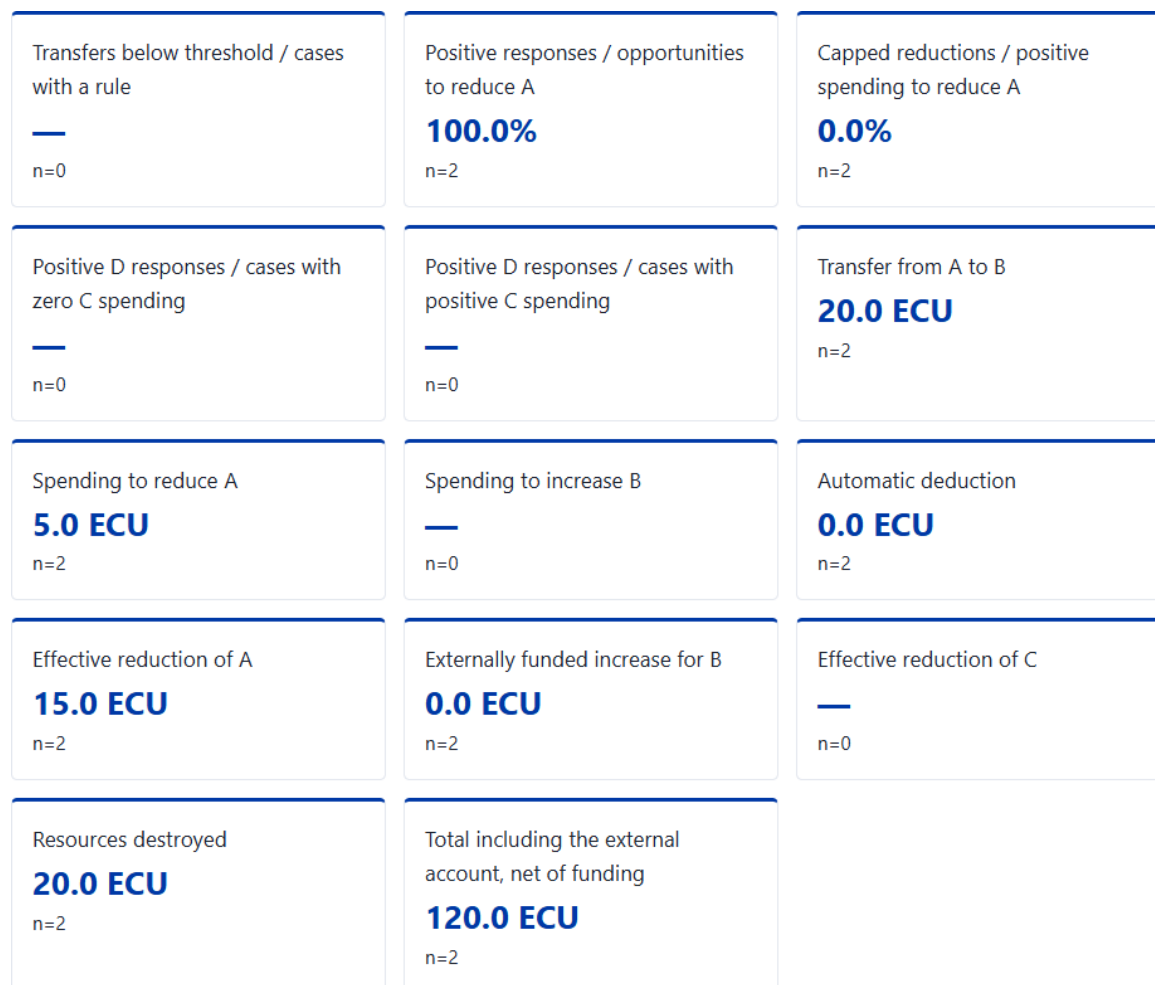


Figure 5.2: Indicators and denominators.

Monetary means include transfers, available response expenditures, fines, effective reductions, compensation, and the extended total. The resources-destroyed indicator sums intervention expenditure and effective reductions from A and C; it does not include the fine transferred outside the group. Expenditure is also displayed separately to distinguish personal cost from the effect on another account. Do not add it again to the resources-destroyed indicator. The extended total adds fines to participant payments and subtracts outside compensation financing. This creates a consistent account without assigning equal social value to every account holder or consequence.

An indicator's unit matters as much as its denominator. A percentage describes frequency, a mean expenditure describes resources committed, and an effective reduction describes resources removed from another account. A high punishment rate can coexist with low expenditure if many people spend one unit. A lower rate can coexist with larger total losses if a few people spend heavily. Ask

students to use the distribution before treating any one summary as a complete description.

5.3 Distributions, roles, and round patterns

The transfer histogram uses integer intervals from zero through nine, ten through nineteen, and so on; its final interval runs from ninety through one hundred. Both admissible endpoints are included. Frequencies sum to completed cases. Bar width represents the interval's share of cases rather than a cumulative frequency. A mean of thirty can arise when everyone sends thirty or from a mixture of extreme choices. The histogram helps separate those situations.

The punishment-expenditure histogram has categories from zero through ten. Its denominator is the set of eligible cases where punishment is available. The conditional-response table reports mean punishment and compensation expenditure within transfer intervals. Empty intervals show dashes for their means. When comparing treatments, keep cell counts visible: a mean based on one response can change entirely when one additional case is observed.

Transfer interval	Cases	Percentage
0–9	0	0.0%
10–19	0	0.0%
20–29	2	100.0%
30–39	0	0.0%
40–49	0	0.0%
50–59	0	0.0%
60–69	0	0.0%
70–79	0	0.0%
80–89	0	0.0%
90–100	0	0.0%

Figure 5.3: Transfer distribution.

Role-specific payment summaries show mean, median, range, and count. A and B are assigned positions, not demographic groups. Roles remain fixed across rounds in this family. Their payment differences can follow mechanically from endowments and rules. Do not interpret a higher income for A as evidence of greater skill, and do not compare C's income across treatments without considering whether C had an expenditure opportunity.

The round table records mean transfers, punishment expenditures, fines, and completed-group counts. It supports description of adaptation over time. The same fixed groups produce several observations, so ten rounds are not ten independent samples of students. An apparent trend can be dominated by one group or by uneven progress through the session. Wait for completion before describing learning, and inspect trajectories as well as overall averages.

To compare treatments, use their separate completed-session reports or an export with a treatment column. The on-screen report describes the session that is open; it does not silently pool other sessions. Cross-treatment analysis should keep the treatment labels, matching conditions, language, and active horizon explicit. A common chart is useful only after the analyst has decided which

cases belong in each comparison.

Your role	Mean	Median	Minimum–maximum	n
Person A	65.0 ECU	65.0	65–65	2
Person B	40.0 ECU	40.0	40–40	2
Person C	15.0 ECU	15.0	15–15	2

Figure 5.4: Payments by role.

Round	n	Transfer from A to B	Mean spending to reduce A	Automatic deduction
1	2	 20.0	5.0	0.0

Figure 5.5: Round evolution.

5.4 Export and reconstruction

The custom export contains one row per participant per active round. It identifies session, round, group, position, role, treatment, matching mode, and active horizon. It then records the individual transfer field, group transfer, expenditures, eligibility, fine, payment, and completion status. It also preserves a full group ledger and parameter set in structured fields. A person who is not A may retain a technical no-decision value in the individual transfer field; analyze transfers using the group field and one unique row per case.

The complete ledger is repeated for each member of a group. For case-level analysis, remove repetitions using session, round, and group together. Do not sum a repeated ledger three or four times. For individual payment analysis, retain each person's row and role. Eligibility identifies whether an expenditure could be chosen; completion identifies whether final settlement has occurred. A default zero in an unfinished row is not a completed decision to abstain.

Before making additional charts, verify three identities. Group transfers and fines must agree across members of the same case. Individual payments must match their positions in the group ledger. The extended total must equal initial resources minus expenditures and effective reductions. If a spreadsheet violates one of these identities, check row selection, numeric conversion, and duplicated ledgers before attributing the discrepancy to participant behavior or the payment logic.

A classroom archive can consist of the export plus a short note with date, treatment, language, group count, and interruptions. Keep individual records in instructor storage. Use anonymous examples when projecting data and avoid linking session identifiers with attendance lists. The activity does not collect personal characteristics to explain punishment. Adding such information would require a separate purpose and design, beyond the documented teaching exercise.

Classroom use and exercises

6.1 A sixty-minute lesson

Use the first ten minutes to explain roles, units, and horizon. Work through a neutral numerical example before asking for predictions. Have each student write down an expected transfer and response expenditure. Do not present a single behavioral prediction as the answer to the comprehension check. That check verifies mechanics; the prediction records an expectation that can later be compared with data.

Use the next fifteen minutes for a basic session and, where time permits, a condition with a norm and fine. A class with enough groups can allocate treatments simultaneously to reduce order effects. A small class may run a second session with the same students, but then experience and knowledge of the lesson also change. Make that limitation explicit. Do not restart a session because its first round failed to produce the expected pattern.

Reserve ten minutes for account reconstruction before comparing means. Select a completed case and ask for three independent calculations: A's balance, the responder's payment, and the extended total. Reveal the indicators and distribution afterward. Ask students first to describe what happened, then to propose a mechanism. This sequence prevents a plausible interpretation from being presented as if it were directly observed.

During the following fifteen minutes, compare institutions and examine an alternative explanation. If C's expenditure falls, check whether transfers rose, the fine reduced the balance available for punishment, or response expenditure is lower within comparable transfer intervals. If compliance rises, ask about its associated costs. Finish with ten minutes of short writing: one conclusion supported by the class data and one claim requiring additional evidence.

What participants will see

Before viewing the report, predict A's transfer and the eligible responder's expenditure. State an observation that would make you change your explanation. After viewing the report, distinguish a classroom regularity, a prediction derived from a model, and a judgment about the rule. Each requires a different kind of argument; none automatically supplies the others.

6.2 A ninety-minute lesson

With more time, preserve the initial sequence and devote twenty additional minutes to a preselected contrast. For formal institutions, use the four cells with a common norm. For distributive responses, compare punishment, compensation, and choice. For second order, compare the four-person control

with the condition where D decides. Do not attempt to run every treatment simply to cover the menu. The lesson needs an identifiable comparison and enough time to inspect its accounts.

Assign one student group to derive the monetary benchmark, another to describe distributions, and a third to develop alternative explanations. Give them the same evidence and require denominators for quantitative claims. Rotate responsibilities so that analysis does not become defense of an initial story. Assessment should reward coherence between rule, data, and conclusion even when the observed pattern contradicts the instructor's expectation.

Use the final fifteen minutes to design an additional study. To separate wealth from position in B's punishment decision, students must explain how they would equalize available resources without inadvertently changing the allocation being evaluated. To study legitimacy, they must specify the information or procedure that would change it and how it would be measured. The present program does not automatically implement these proposals. Their purpose is to identify what a new design would need.

6.3 Exercises with worked answers

6.3.1 The threshold and monetary incentives

Assume no one spends on social responses and A values only personal income. Compare transfers of zero and thirty when the fine is twenty. A receives eighty in the first case and seventy in the second. With a fine of forty, the corresponding payments are sixty and seventy. The monetary choice therefore switches to the threshold under the stronger fine. A transfer of twenty-nine still incurs the full deduction and is not a way to avoid part of the fine.

This answer does not claim everyone will make the benchmark choice. It derives a reference from specified preferences and expectations. A larger transfer may be consistent with concern for B, expected punishment, or a different interpretation of the rule. Distinguishing those explanations requires additional variation or measurement. An observed departure should be described before it is treated as a rejection of all purposeful choice.

6.3.2 Fine, reduction, and the recipient's payment

With a transfer of twenty, a fine of twenty, and C's expenditure of five, calculate final payments. A retains sixty after the transfer and fine, loses fifteen more, and finishes with forty-five. B receives forty and C keeps fifteen. Participant payments total one hundred. Adding twenty collected outside the group produces an extended total of 120, twenty below the initial 140. The difference is five in expenditure plus fifteen in reduction. B does not receive either the fine or the money removed from A.

A common error adds the fine to B's payoff because it is interpreted as a remedy. The implemented rule does not make that transfer. Another subtracts five from A instead of fifteen, confusing the responder's cost with the effect on another account. Draw separate labeled arrows to distinguish those quantities. A legal institution that transfers the fine to B could be studied, but its rule would have to be announced and implemented explicitly.

6.3.3 A capped reduction and additional expenditure

With a transfer of ninety-five, no fine, and C's expenditure of ten, A loses five and ends at zero; B receives 115 and C receives ten. The total is 125. If C had spent two, the reduction would also have exhausted A's remaining five, while C would have retained eighteen. Spending ten costs eight

more without increasing the reduction. The form permits that choice and previews its consequence. A maximum budget is not a recommendation for efficient expenditure.

The cap also limits interpretations of punishment effectiveness. A small observed reduction need not mean weak willingness to punish; the responder may have spent heavily against a small balance. Conversely, expenditure can have a large nominal effect that is never realized. Use expenditure to describe resources committed and effective reduction to describe the monetary consequence. Neither alone identifies the responder's intended message.

6.3.4 Choosing between two responses

For a transfer of twenty, punishment expenditure of three, and compensation expenditure of two, reconstruct payments of seventy-one, forty-six, and fifteen. Verify total response expenditure of five, effective reduction of nine, and outside financing of six. The extended total is 126. Compare this with compensation expenditure of five: payments would be eighty, fifty-five, and fifteen, and the extended total would be 135. Both choices cost C five but have different distributional and destructive effects.

The accounting comparison does not determine which distribution should be preferred. Students can evaluate alternatives using different criteria, but they must state those criteria. A claim about equality requires examining final balances. A claim about resource preservation requires defining the financing boundary. A claim about responsibility requires an argument beyond the balance sheet. Keeping these questions separate makes the discussion more precise.

6.3.5 An empty denominator

Suppose every completed second-order case has positive expenditure by C, and D responds in half of those cases. The rate conditional on positive expenditure is fifty percent. The rate conditional on zero expenditure is undefined because there are no such cases. It is not zero percent. If a new case has zero expenditure by C and no response by D, the rate becomes zero percent with one observation. Its evidential weight remains very limited and its count should be displayed.

6.3.6 Interaction and compliance costs

Return to cell means of ten, twenty, twenty-five, and twenty-eight. Calculate the interaction of minus seven and explain that the increase associated with the fine is smaller when social response is available. Now suppose the fine reduces social expenditure from eight to two. This might reduce private intervention costs, but it does not settle the institutional evaluation. Effective reductions, fines, payment distributions, and any outside administrative costs still matter. The game does not charge a separate budget for administering automatic fines.

6.3.7 Second-order accounting

Take a transfer of twenty, C's expenditure of five, and D's expenditure of ten. Before D acts, C holds fifteen. Its effective loss is therefore fifteen rather than thirty. Final payments are sixty-five, forty, zero, and ten. Check the total of 115 against an initial 160. Fifteen was spent and thirty was removed through effective reductions. The forty-five-unit difference is fully accounted for. The example also shows why a high expenditure by D is not evidence of a proportionally high consequence for C.

6.4 Assessing learning

A satisfactory answer identifies the treatment, states the sequence of decisions, and reconstructs a correct account. A stronger answer distinguishes expenditure, reduction, and outside transfers and uses appropriate denominators. The strongest explanation offers an alternative interpretation and a comparison that could distinguish it. Do not award higher grades for transferring more or punishing less. Those decisions provide material for the activity rather than answers to a moral examination.

Use four assessment dimensions: accounting accuracy, design interpretation, use of data, and inferential restraint. Each can be classified as insufficient, adequate, or strong. Elaborate legal vocabulary does not repair an incorrect account. A correct calculation does not justify a causal claim without a design. Require every quantitative statement to identify its unit and sample, and every evaluative statement to identify its normative criterion.

6.5 Interruptions and responsible closure

If a participant closes the browser, they should return through their own participation link and recover the same place. Opening someone else's link creates confusion rather than solving it. Wait pages synchronize the group and do not imply a need for continuous refreshing. First check which person must decide and whether a visible form error is preventing submission. A passive role can legitimately wait while another participant makes a choice.

Do not enter a student's choice on their behalf without a procedure announced beforehand. If a session remains incomplete, preserve its data and document the interruption. Completed cases are still observable, but they need not be a random selection of all planned cases. A new session creates new decisions and identifiers; it does not retrospectively complete the old one. Avoid joining the two as a continuous ten-round trajectory.

Do not alter treatment, language, group size, or endowments during an activity. Before a real class, verify a valid participant count, accessible participation links, and the private instructor report. A local environment supports review; classroom distribution requires an address reachable from student devices and the corresponding deployment. Printed institutional routes describe intended access and do not by themselves prove that the new family is already publicly available.

At closure, explain how active rounds were added and confirm that participants reached the final screen. Any conversion from ECU to another reward should have been announced before the activity and should match the session configuration. Save the export, treatment information, and incident notes before closing the panel. Descriptive game titles, role names, and the documented rules are sufficient for instruction; internal production identifiers have no place in participant materials or classroom interpretation.

Bibliography

- Fehr, E., and Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25, 63–87. [https://doi.org/10.1016/S1090-5138\(04\)00005-4](https://doi.org/10.1016/S1090-5138(04)00005-4).
- Kube, S., and Traxler, C. (2011). The Interaction of Legal and Social Norm Enforcement. *Journal of Public Economic Theory*, 13(5), 639–660. <https://doi.org/10.1111/j.1467-9779.2011.01515.x>.
- Leliveld, M. C., van Dijk, E., and van Beest, I. (2012). Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *European Journal of Social Psychology*, 42(2), 135–140. <https://doi.org/10.1002/ejsp.872>.
- Kiyonari, T., and Barclay, P. (2008). Cooperation in social dilemmas: Free riding may be thwarted by second-order reward rather than by punishment. *Journal of Personality and Social Psychology*, 95(4), 826–842. <https://doi.org/10.1037/a0011381>.
- American Economic Association. JEL Classification System. Classification consulted for the teaching catalog. <https://www.aeaweb.org/econlit/jelCodes.php?view=jel>.