

Ecolabels promote sustainable behavior in a dynamic common-pool resource dilemma

Julian Kirschner¹, Natalie Lee¹, Jan Engelmann^{1,2 *} & Jan Hausfeld^{1,2 *}

Abstract:

Sustaining shared resources and coordinating collectively for sustainable behavior are essential to human economic activities and environmental conservation. Ecolabels are widely used to signal sustainable practices and to influence individual consumer behavior, yet their role as a coordination device in common-pool resource problems remains largely unexplored. We test the impact of ecolabels on behavior in a dynamic common-pool resource dilemma with real economic and environmental incentives. In our lab experiment, groups of participants trade off short-term individual profits from resource extraction against long-term collective gains and resource preservation, a tension that induces high strategic uncertainty. We find that ecolabels significantly increase sustainable behavior, extending resource survival by 20-23% while also raising participants' earnings. Results show that ecolabels set standards that facilitate early-on coordination on sustainable behavior, which is then maintained through conditional cooperation. Importantly, these effects arise even when ecolabels are purely symbolic, providing no direct monetary benefit, and they persist when participants interact with computerized others. Our findings uncover underexplored mechanisms of ecolabels that can affect individual and collective decision making, and suggest that, beyond consumer nudges, ecolabels can serve as policy instruments for promoting sustainable behavior and establishing shared sustainable standards.

¹ University of Amsterdam, Center for Research in Experimental Economics and Political Decision (CREED)

² The Tinbergen Research Institute

* Shared senior authors; address correspondence to j.hausfeld@uva.nl or j.b.engelmann@uva.nl

1. Main

Sustaining shared resources is one of the oldest and most consequential collective action problems societies face. From forests and fisheries to the global climate, common-pool resources are depleted when individuals pursue short-term gains by consuming beyond sustainable benchmarks, usually at the expense of long-term collective welfare¹⁻³. Current practices of production and over-consumption accelerate this depletion, with potentially irreversible consequences for ecosystems, human health, and global economies and societies⁴⁻⁷. These common-pool resource dilemmas are difficult not only because individuals may prefer immediate private gains, but also because sustainable behavior often depends on expectations about what others will do. When actors are uncertain whether others will preserve the resource, even those willing to be conditionally sustainable may instead choose to extract unsustainably.

As shown by decades of research on commons governance in the tradition of Ostrom's seminal work⁸⁻¹⁰, economic games are powerful tools to model conservation and socio-ecological conflicts¹¹. Using economic games, research has identified a broad range of institutions that, if implemented, can address commons dilemmas^{12,13}. Communication is one major institution^{14,15} which raises cooperation substantially¹⁶. Monitoring paired with sanctioning or punishment can sustain cooperation in both laboratory and field commons, especially when incentives to defect are strong¹⁷⁻²⁰. Signaling based social interventions that help reduce the tragedy of the commons include providing normative information^{21,22}, making individual contributions coupled with social approval visible²³, and allowing reputation formation²⁴. Importantly, these institutions can operate through different channels: structural (by changing what people consider in their decisions, payoffs or rules), psychological (by changing perceptions and expectations), and interactional ones (by working through observation of and response to others)^{13,12}. Yet many institutions require costly monitoring, enforcement, repeated communication or formal collective decision-making, despite not addressing multiple of these channels. This raises the question whether other lower-cost instruments, that address multiple of these channels and make sustainable behavior and its standards visible, can facilitate sustainable behavior in a dynamic resource dilemma.

Ecolabels (and carbon labels) are such an instrument. They have been extensively studied as cost-effective, scalable, and feasible interventions to promote sustainable consumption^{25,26}. Some studies show that labels increase consumers' willingness to pay for sustainable products

and influence purchase behavior^{26,27}. However, evidence on their effectiveness is mixed in recent meta-analytic discussions and field experiments, and effect sizes are often small^{26,28–30}. Moreover, targeting consumers alone can be insufficient: even when consumers fully internalize an externality, market outcomes do not necessarily reach socially optimal outcomes³¹. Yet, most existing research studies individual consumer decisions, often in static, applied settings^{25,29,32,33} that do not capture repeated strategic interdependence, i.e. individuals cannot observe others' choices or infer whether others will adopt sustainable behavior. This leaves an important question open: do ecolabels affect sustainability not only by informing isolated choices, but also by helping groups to coordinate in a commons dilemma?

Our study addresses two underexplored mechanisms of ecolabels. First, ecolabels can serve as a focal point that establishes visible standards^{34–37}, which can reshape prevailing norms, coordinate behavior and catalyze sustainable change^{25,37–42}. These standard-setting effects of labels have been observed in consumer markets but also in whole industry sectors (e.g., ISO-labels^{43,44} and electronics^{38,45,46}; in forestry, ecolabels improved ecological outcomes and market competitiveness⁴⁷). Second, ecolabels can serve as a communication device, enabling individuals to display their adherence to sustainable standards to others with the potential to counter misperceptions^{48,49}. For example, labels disclose how frequent a certain behavior is^{25,37}, thereby shaping individuals' perceptions^{50,51}. These two mechanisms would make ecolabels a coordination device in a group setting without directly restricting actions or imposing costs.

In this study, we focus on the strategic uncertainty and miscoordination embedded in common-pool resource dilemmas and test whether ecolabels can mitigate them by serving as a coordination device. In many social-dilemma games, a substantial share of participants is classified as conditional cooperators^{14,52–55}, i.e. cooperation that depends on others' willingness to cooperate⁵⁵, and conditionality is commonly emphasized in public climate debates^{56–58}. Yet groups may fail to reach cooperative outcomes even when they broadly agree that cooperation is desirable⁵⁹. This failure stems from people being uncertain about others likely actions^{48,49,60,61}, or because they fail to adopt institutions in the first place^{62,63}. Conditional cooperation, as well as conformity⁶⁴ (a related tendency to match others), can sustain sustainable behavior only once a sustainable pattern has emerged. Neither determines which pattern emerges initially. We propose that visible ecolabels can supply the focal sustainable benchmark on which conditional cooperators can coordinate. Accordingly, we hypothesize that

when the tragedy of the commons is driven primarily by miscoordination rather than by unwillingness to cooperate, ecolabels can help players coordinate on a cooperative equilibrium.

A crucial question that arises is whether this coordination mechanism requires ecolabels to provide any additional monetary incentives, or whether symbolic ecolabels are sufficiently effective when standards are set^{63,65}. Symbolic ecolabels could then provide a cheaper and scalable alternative when other interventions such as monitoring, or sanctions are too costly or difficult to enforce. Nevertheless, their potential for governing the commons remains untested⁶⁶. On the other hand, symbolic ecolabels could be less effective if the tragedy of the commons is driven by strong deviation incentives or a general lack of willingness to cooperate. Thus, we study how and through what mechanism ecolabels can help groups coordinate on sustainable extraction in a dynamic common-pool resource dilemma. More broadly, our study responds to recent calls for behavioral science to move beyond individual consumer choices and toward broader, system-wide effects such as coordination^{25,67–72}.

To empirically investigate the causal effects and mechanisms of ecolabels on sustainable behavior and coordination, we use an experimental adaptation of the common-pool resource dilemma^{3,73–75}. Our multi-round experiment, the “Social Forest Game”, is framed as a relatable simulation of a common resource, i.e. withdrawing trees from a shared forest^{76,77}. In our controlled lab experiment, each participant decides how many trees to withdraw from a forest shared with three others. The forest regenerates at a fixed, limited rate (Fig. 1). Withdrawing more trees yields higher short-term profits, but it increases the risk of irreversibly depleting the forest, which undermines one’s own and others’ future profits. The game is calibrated to allow for both sustainable and unsustainable equilibria, thereby capturing the strategic and temporal uncertainties that often hinder sustainable transitions^{78–80}.

To identify the effect of ecolabels, we compare the individual behavior in a baseline condition without ecolabels (“NoEco”-treatment) with two conditions in which participants can earn ecolabels for sustainable behavior, defined as withdrawing no more than 10 trees (the share of regrowth rate) in each prior round. We consider two forms of ecolabels: a purely “Symbolic” label without monetary benefits and a “Monetary” label that increases profits by 25%⁴⁰. This comparison assesses whether label effects require material incentives or can already arise from a visible sustainability standard. Additionally, we implement a computer-player version of the “Social Forest Game”, in which the other group members are computerized players whose

behavior is based on human decisions. Comparing the human-player and computer-player rounds reduces the scope for explanations based on interaction with, or concern for, human peers.

Because the relationship between environmental attitudes and behavior is not always direct or uniform^{81,82}, recent work in environmental psychology and behavioral economics has emphasized the importance of real and environmentally consequential incentives in experiments^{83–85}. Accordingly, our experiment combines monetary incentives (money) with environmental consequences linked to real-world tree planting (via onetreeplanted.org), allowing us to observe decisions with actual consequences for oneself, others, and the environment, rather than relying on hypothetical choices or self-reported intentions^{84,86}. Last, we elicit pro-environmental attitudes and record eye movements to explore psychological and cognitive channels: If symbolic labels work by making the sustainability standard salient, they should strengthen the link between pro-environmental attitudes and behavior, engaging intrinsic environmental motivation. Monetary labels, by contrast, do not necessarily need to strengthen this link, and should instead draw relatively more attention toward payment-relevant information⁸⁷. Our results are consistent with both predictions.

The framework used in our study distills the commons-failure aspect of the resource-management problem into a controlled environment, which allows us to make clean causal inferences that would otherwise be difficult to identify amid the full complexity of market decisions. Our design also helps separate the motives that may drive label effects. The fixed horizon makes end-of-game behavior diagnostic: trees left standing in the final round, when no future extraction is possible, reflect donative or environmental rather than strategic or monetary motives, a distinction that cannot be identified under an indefinite horizon. Contrasting the Symbolic with the Monetary label separates purely symbolic standard-setting from monetary incentives. Comparing rounds played with human and with computer-player group members separates effects that depend on interaction with, or concern for, human peers from those that operate through the label itself. First-round choices are informative because they occur before participants have received any feedback about others; differences at this stage are therefore consistent with ecolabel availability affecting expectations of others or the salience of the sustainable benchmark. The development of cooperation over subsequent rounds allows us to observe how information about others' behavior shapes sustainable behavior over time. To assess the effects of the labels, we complement preregistered analyses with a robustness

analysis. Overall, the experiment allows us to understand if and how ecolabels change behavior dynamically in a common-pool resource dilemma.

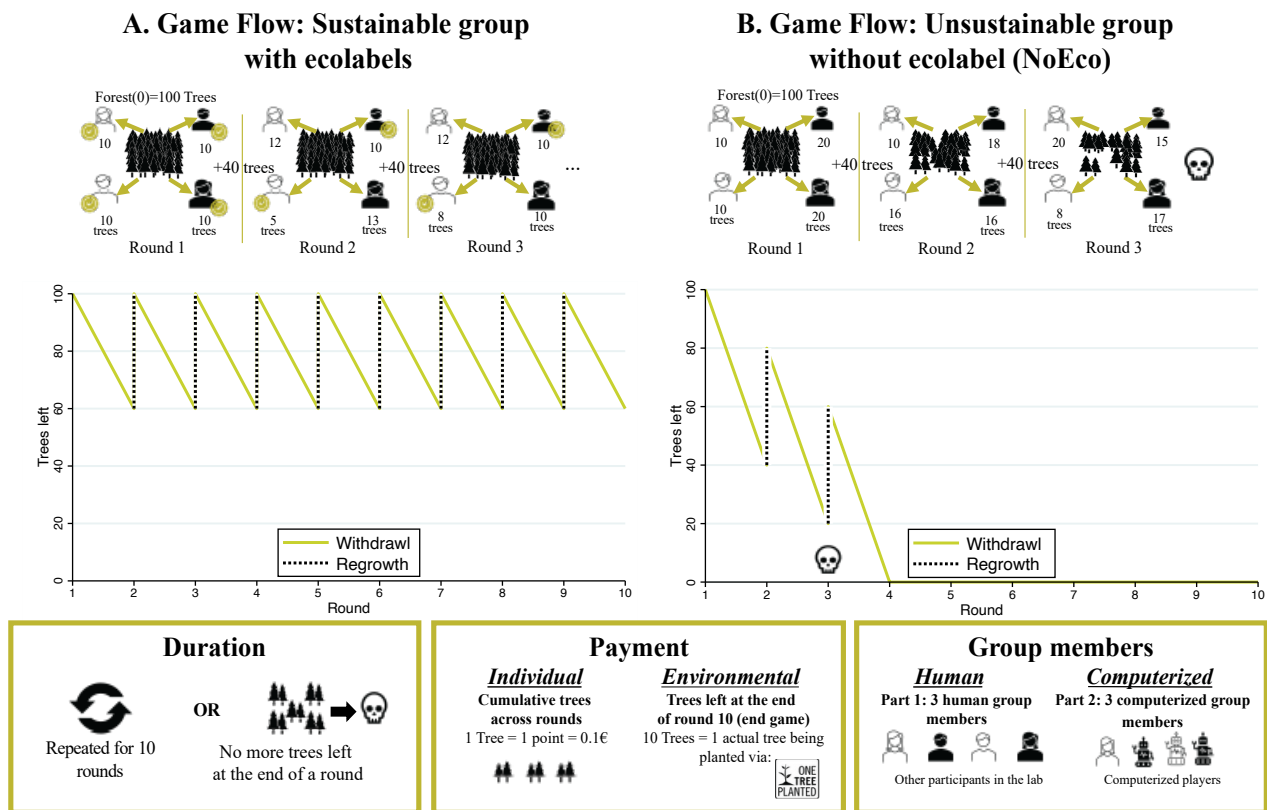


Fig. 1: Schematic representation of “The Social Forest Game”. Illustrative game flows for (A) a sustainable group under either ecolabel treatment (Symbolic or Monetary) and (B) an unsustainable group in the control (NoEco) treatment. The top row shows examples of individual decisions made within a group; the middle row illustrates how group withdrawal and regrowth impact resource levels across rounds. The game starts with a forest of 100 trees. Each of 4 participants can withdraw 0-20 trees worth 10 cents each. Withdrawing trees reduces the resource quantity (Trees left in graph). If at least 1 tree remains, the next round starts with 40 new trees in addition to the remaining ones: otherwise, the game ends early. This cycle lasts up to 10 rounds, and remaining trees at the end of round 10 are planted in real life (10 in-game trees = 1 real tree). If the forest is depleted at any point, participants earn no points in that round, the game ends and no trees will be planted. All points accumulated until the game ends are paid out. In the ecolabel treatments, participants receive a label for sustainable behavior (light green label in the top-left schematic) that is irrecoverably lost as soon as participants behave unsustainably. Middle-Left: A sustainable group would survive all 10 rounds and most keep their ecolabel. Light green lines represent the extraction per round for the whole group, dashed black lines show the 40 trees that are added at the end of each round if the forest is not depleted. Middle-Right: A group with several unsustainable players (no ecolabel can be earned in the NoEco condition) and the game would end after round 3. Bottom-Left: The game lasts 10 rounds or until the forest is depleted. Bottom-Middle: Each participant earns 10 cents per tree withdrawn. Remaining trees at the end of round 10 were planted through onetreeplanted.org. Bottom-Right: In Part 1, participants interact with 3 other human participants in real-time. In Part 2, each participant interacts with 3 computer players which mimic behavior of participants in a pilot study. All subjects first play Part 1 (human rounds) followed by Part 2 (computerized rounds). Subjects are randomly allocated to one of two ecolabel treatments (Symbolic, and Monetary, see Introduction and Methods) where participants can earn ecolabel for sustainable production (withdrawing up to 10 trees) or the control treatment without ecolabels.

2. Results

Ecolabels lead to more sustainable behavior

We first analyze how ecolabels affect sustainable behavior and long-term resource outcomes in the Social Forest Game. Fig. 2 compares the control condition (“NoEco”) with the “Symbolic” and “Monetary” ecolabel treatments on preregistered outcomes (AsPredicted#120226): group-level measures related to forest longevity (Fig. 2A-B) and individual tree-withdrawal in early and later rounds (Fig. 2C-D). Table 1 summarizes outcomes by conditions in the human and computer version, while Fig. 3 then visualizes the temporal development of treatment differences across rounds and different outcome indicators.

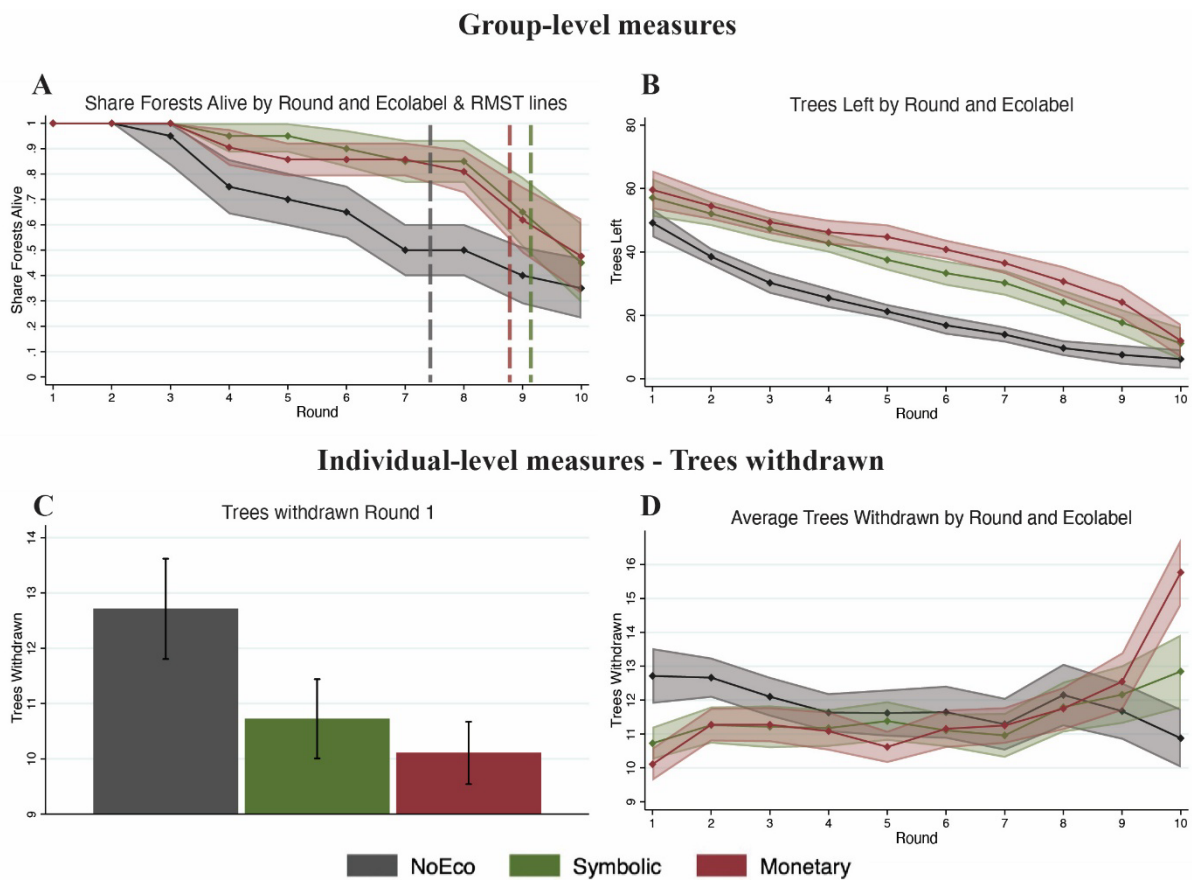


Fig. 2: Group- and individual-level behavior is more sustainable in ecolabel conditions. A,B: Group-level measures. (A) Share of forests alive at the end of a round averaged across groups per round and ecolabel condition (horizontal lines) and Restricted Mean survival Time (RMST, vertical lines). RMST values reflect average survival durations (Rounds Alive) up to round 10; the ratio between the gray (No Eco) and red (Monetary) and green (Symbolic) vertical lines reflects the protection factor (PF) yielded by ecolabels, which ranges from 1.20 to 1.23, indicating an increase in forest survival time by 20-23% (Supplementary Results 1.1). (B) Number of trees left averaged across groups per round and ecolabel condition, groups that depleted their forest have 0 trees left. C, D: Individual-level measures. (C) Average number of trees withdrawn in round 1. (D) Average number of trees withdrawn per round and ecolabel condition. A and B use Cousineau-Morey within-groups confidence intervals, D uses Cousineau-Morey within-subjects confidence intervals, all figures use 90% CIs.

Both ecolabel conditions increase resource longevity at the group and individual levels compared to the NoEco condition (Fig. 2A, RMST Analysis below, Table 1). Restricted Mean Survival Time (RMST) analyses at the group level show that both ecolabel conditions significantly delay forest depletion relative to NoEco (Supplementary Results 1.1): Forests survive on average 1.70 rounds longer in the Symbolic condition (RMST-ratio=1.23, 95%-CI[1.04,1.46], $p=0.018$) and 1.45 rounds longer in the Monetary condition (RMST-ratio=1.20, 95%-CI[1.001,1.434], $p=0.0498$). Thus, both ecolabels increase expected forest survival, reflecting a protection level between 20-23% (ratio between vertical lines Fig. 2A). Similarly, groups in both ecolabel treatments preserve more trees, compared to the NoEco treatment, throughout most of the game (Fig. 2B; rounds 1-8, Mann-Whitney-U test (MWU) $p_{\text{uncorrected}} < 0.05$ for all but one comparison; Fig. 3; Supplementary Table S1; for BH-corrected results see Supplementary Results 1).

Differences in sustainable outcomes emerge from the beginning. In round 1 (Fig. 2C), before participants received any information about others, participants withdraw fewer trees under both ecolabels than under NoEco (Means: $m_{\text{NoEco}}=12.71$, $m_{\text{Symbolic}}=10.73$ & $m_{\text{Monetary}}=10.11$; $\text{MWU}_{\text{NoEco-Symbolic}} p < 0.01$, $\text{MWU}_{\text{NoEco-Monetary}} p < 0.01$, $\text{MWU}_{\text{Symbolic-Monetary}} p = 0.32$). This early round difference is important because the initial behavior strongly predicts subsequent decisions and outcomes as shown later. Additionally, given our dynamic setup, early-round behavior is more comparable across groups, whereas later-round behavior is history- and state-dependent (i.e. on near-depleted forests).

As the game progresses, group dynamics diverge: the most unsustainable group depletes its forest as early as round 3, whereas fully sustainable groups maintain it at the initial level. Because unsustainable groups drop out early, later rounds over-represent sustainable groups. To avoid this survivor bias, we first restrict the comparisons to rounds 1–3, in which all groups are still active. Here, participants withdraw fewer trees in both ecolabel conditions than in NoEco (Fig. 2D; means across first 3 rounds: $m_{\text{NoEco}}=12.49$, $m_{\text{Symbolic}}=11.07$ & $m_{\text{Monetary}}=10.88$; Kruskal-Wallis Test: $\chi^2_{\text{ties}}(2)=41.91$, $p < 0.01$; $\text{MWU}_{\text{NoEco-Symbolic}} p < 0.01$, $\text{MWU}_{\text{NoEco-Monetary}} p < 0.01$, $\text{MWU}_{\text{Symbolic-Monetary}} p = 0.75$). Consequently, by the end of round 3, forests under both ecolabel conditions had significantly more trees left (Fig. 2B, $m_{\text{NoEco}}=30.25$, $m_{\text{Symbolic}}=47.20$ & $m_{\text{Monetary}}=49.38$; OLS-regression F-test=4.68, $p=0.01$; $p_{\text{NoEco-Symbolic}}=0.02$, $t_{\text{NoEco-Symbolic}}=2.46$; $p_{\text{NoEco-Monetary}} < 0.01$, $t_{\text{NoEco-Monetary}}=2.81$; $p_{\text{Symbolic-Monetary}}=0.75$, $t_{\text{Symbolic-Monetary}}=0.32$). Figure 3 summarizes these effects across rounds and measures.

These early differences set groups on diverging trajectories, and because unsustainable groups deplete their forest and exit the game, the composition of surviving groups shifts toward sustainable groups as rounds progress. This attrition produces a distinctive and predictable pattern across outcome measures. First, groups in the label conditions kept extracting sustainably from living forests for more rounds, resulting in higher cumulative trees withdrawn and earnings per participant under both labels (Table 1, Fig. 3). This reflects longer resource availability, not higher per-round extraction. Second, per-round extraction is mostly different in the early rounds and converge as unsustainable NoEco groups deplete their forest; averaging across all rounds, the difference is accordingly marginal (Supplementary Results 2: OLS clustered by group and individual: $F=1.87$; $\beta_{\text{NoEco-Symbolic}}=-0.54$, $p=0.053$; $\beta_{\text{NoEco-Monetary}}=-0.43$, $p=0.15$), because this average pools rounds with very different surviving populations. Third, by the final rounds the surviving groups in every condition are predominantly sustainable, and end-game defection compresses differences further: seven of the eight forests depleted during round 10 were in label conditions, and when excluding final-round depletions, depletion rates are 0.60 under NoEco versus 0.35 and 0.38 under Symbolic and Monetary. End-of-game trees remaining (Avg. Trees Left) and forest survival (Share Forests Alive) therefore no longer differ from NoEco (Table 1A), even though both labels significantly extend resource longevity and preserve more trees through most of the game (Figs. 2 and 3).

Variable	A: Part 1			B: Part 2		
	Human Rounds (SE)			Computer Rounds (SE)		
	NoEco (N)	Symbolic (S)	Monetary (M)	NoEco (N)	Symbolic (S)	Monetary (M)
Round 1	12.71	10.73	10.11	14**	11.26†	10.55†
Withdrawal	(0.55)	(0.43)	(0.34)	(0.58)	(0.43)	(0.33)
	<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.32</i>			<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.08</i>		
Rounds	3.71	6.80	6.87	3.90	6.49	6.57
Sustainable	(0.34)	(0.33)	(0.33)	(0.38)	(0.40)	(0.38)
	<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.85</i>			<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.83</i>		
Total Trees	80.05	97.10	94.67	59.94***	77.11***	80.30***
withdrawn	(4.38)	(3.09)	(3.28)	(2.51)	(2.33)	(2.51)
per person	<i>p_{N-S}: 0.00 p_{N-M}: 0.01 p_{S-M}: 0.71</i>			<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.16</i>		
Avg. Num.	7.45	9.15	8.90	6.33**	7.9***	8.11**
Rounds	(0.59)	(0.36)	(0.44)	(0.31)	(0.29)	(0.28)
Alive	<i>p_{N-S}: 0.04 p_{N-M}: 0.08 p_{S-M}: 0.78</i>			<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.61</i>		
Avg. Trees	6.20	11.15	11.95	5.25	8.78	6.90*
Left	(2.86)	(4.48)	(4.45)	(2.02)	(1.42)	(1.26)
	<i>p_{N-S}: 0.47 p_{N-M}: 0.31 p_{S-M}: 0.89</i>			<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.38</i>		
Share	0.35	0.45	0.48	0.19*	0.45	0.40
Forests Alive	(0.11)	(0.11)	(0.11)	(0.04)	(0.06)	(0.05)
	<i>p_{N-S}: 0.75 p_{N-M}: 0.62 p_{S-M}: 1.00</i>			<i>p_{N-S}: 0.00 p_{N-M}: 0.00 p_{S-M}: 0.67</i>		

Table 1: Overview of Outcome variables by Condition. (A) Outcome variables by ecolabel condition in Part 1, in which participants interact with other humans. (B) Outcome variables by ecolabel condition in Part 2, in which participants interact with three computer players; hence group outcomes are driven only by individual choices of the player. For each variable, the first row shows the mean and the standard errors in brackets. The second row (blue/white background) shows the exact p-values from Mann-Whitney U tests comparing the ecolabel conditions, with blue indicating significant differences between ecolabels and the NoEco control, with significant results highlighted in bold for $p \leq 0.05$. In (B), stars indicate statistically significant differences between the human and computer rounds within an ecolabel condition (Wilcoxon signed-rank test). Note that Avg. Num. Rounds Alive refers to rounds in which participants could decide. Regressions with interaction effects between condition and game version can be found in Supplementary Results 3. For rows 4, 5, 6 (last 3 variables) analyses for human rounds at group level ($N=61$). P-values abbreviations in second rows: p_{N-S} : NoEco vs. Symbolic; p_{N-M} : NoEco vs. Monetary; p_{S-M} : Symbolic vs. Monetary. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Ecolabel effects are robust across rounds and outcome measures

Fig. 3 summarizes the differences between NoEco and each ecolabel condition by round, across measures including economic and resource outcomes (uncorrected p-values; BH-corrected results in Supplementary Results 1.2; see legend). In both game versions, ecolabel conditions

show lower initial extractions, and more sustainable rounds and greater resource preservation (trees left and forests alive) during the middle of the game. In the human rounds, differences of the resource outcomes narrow toward round 10. In the computer rounds, where peer behavior is standardized, both labels result in more trees left at the end and a higher share of forests alive than NoEco (Table 1B; Fig. 3). For example, on average participants leave significantly different amounts of trees alive ($m_{\text{NoEco}}=5.25$, $m_{\text{Symbolic}}=8.78$ & $m_{\text{Monetary}}=6.9$, both $p<0.01$). The ecolabel effects on early extraction, sustainable rounds, resource longevity, and total trees withdrawn are consistent across the human and computer rounds (Fig. 3; Table 1 & Supplementary Results 3).

The main ecolabel effects are consistent throughout and the reversals shown in Fig. 3 do not contradict the ecolabel effects. Under the Monetary label, extraction exceeds NoEco in the final human round, this reversal reflects the end-game effect described in the next section. In the computer rounds, per-round extraction in the labels conditions exceeds that of NoEco in several middle rounds. This reflects a selection effect, i.e. the selfish have already depleted the forest, among surviving NoEco participants. Many participants that are still active have reduced extraction after experiencing that over-extraction accelerates depletion as the conditional-cooperator computer player reciprocates unsustainable behavior. They extract more conservatively, even below the 10-tree benchmark, than participants in the label conditions, who withdrew closer to the sustainable benchmark from round 1 onward (Supplementary Results 4, Fig. S3D).

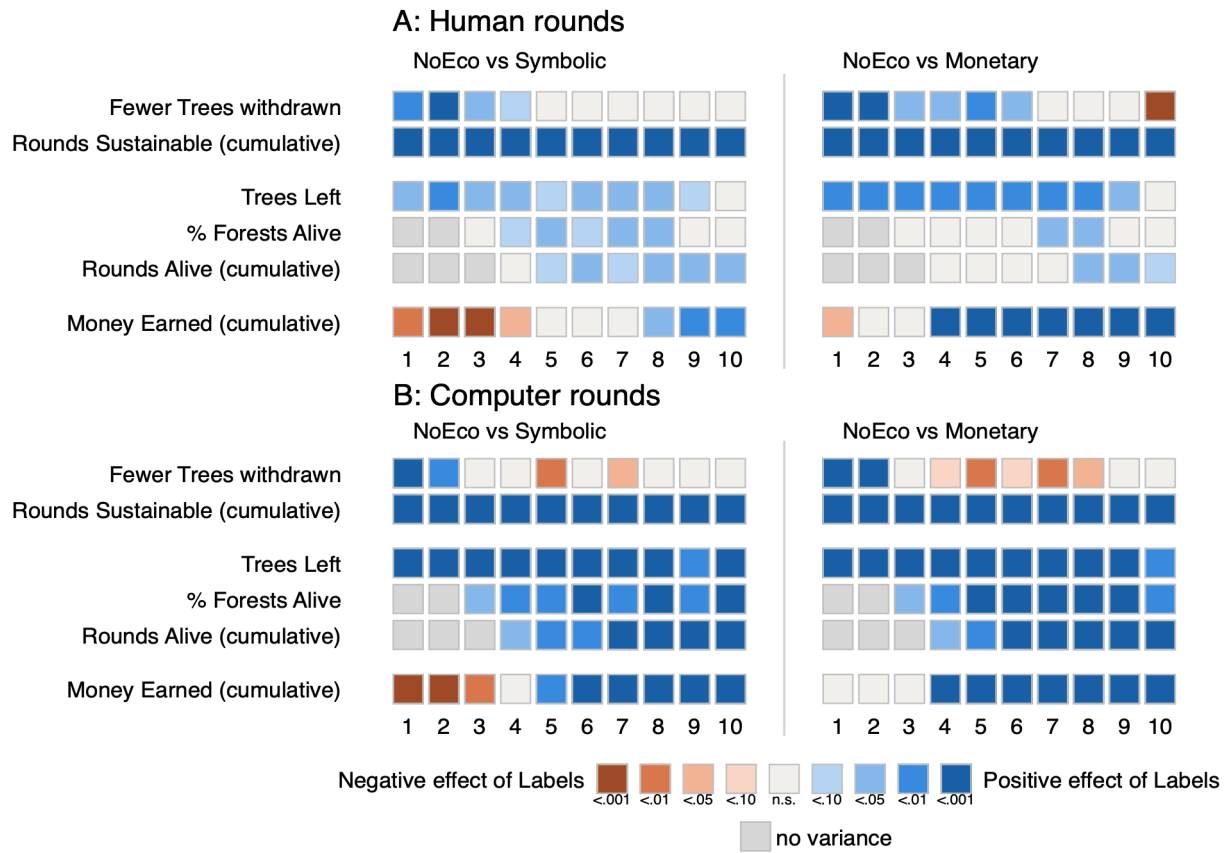


Fig. 3 Ecolabel effects across outcomes, rounds and game versions. Each cell reports a two-sided Mann-Whitney U test comparing the NoEco condition with one ecolabel condition in a single round. Rows are the six outcome measures; columns are rounds 1-10. Panel A shows Part 1, in which participants play with three other humans; Panel B shows Part 2, in which they play with computerized players. The left block compares NoEco with Symbolic, the right block NoEco with Monetary; the comparison of Symbolic and Monetary is presented in Supplementary Results 1.2. The first 2 rows describe individual behavior ($n = 80$ NoEco, 80 Symbolic, 84 Monetary). Variables: Fewer trees withdrawn is the trees withdrawn in that round, restricted to forests still alive at the start of the round. Rounds Sustainable (cumulative) is each participant's running counter of rounds in which they withdrew ten trees or fewer. The three middle rows describe the state of the forest and are measured at the group level in Part 1, one observation per group ($n = 20, 20, 21$); in Part 2 each participant plays with computerized players and constitutes their own group. Trees left is the standing counter, Forests Alive is the share of forests not yet depleted, and Rounds alive (cumulative) is the number of rounds the forest has survived up to and including that round. The final row is the economic outcome: Money earned (cumulative) is each participant's cumulative payoff, which includes the bonus paid in the Monetary condition and stops accruing once the forest is exhausted. Shading intensity gives the p-value (from $p < 0.10$ to $p < 0.001$). Blue indicates that the difference runs in the positive direction for that measure (labels have a positive effect), orange that it runs in the opposite direction. Gray cells could not be tested because the measure had no variance at that point: no forest had yet been exhausted in rounds 1-2 (Forests Alive) or rounds 1-3 (Rounds Alive). P-values are uncorrected, see Supplementary Results 1.2 for BH-corrected figures including the ecolabel comparisons.

Symbolic and Monetary ecolabels show similar main outcomes

Both types of ecolabels improve environmental and economic outcomes relative to the NoEco condition. Comparing the Monetary and the Symbolic condition, we find no significant differences for economic and environmental outcomes, i.e. forest longevity and first round behavior (Fig. 2, Table 1, Supplementary Results 1.2). However, we do observe higher total payouts in the Monetary ecolabel condition ($m_{\text{NoEco}}=80.05$, $m_{\text{Symbolic}}=97.10$ & $m_{\text{Monetary}}=109.93$; Supplementary Results 1.2) which is mechanically driven by the per-tree multiplier, rather than a difference in the number of trees withdrawn (Table 1: Total Trees withdrawn per person). Overall, the purely symbolic ecolabel performs similar on the main outcomes at no financial cost.

Notably, the two label conditions differ in the final round among groups where forests are still alive until the end. In round 10, participants in Monetary withdraw significantly more trees than those in Symbolic (Fig. 2D, $m_{\text{NoEco}}=10.875$ (N=32), $m_{\text{Symbolic}}=12.846$ (N=52) & $m_{\text{Monetary}}=15.769$ (N=52); Kruskal-Wallis Test: $\chi^2_{\text{ties}}(2)=14.33$, $p<0.01$; $\text{MWU}_{\text{Symbolic-Monetary}} p<0.01$; $\text{MWU}_{\text{NoEco-Symbolic}} p=0.09$; $\text{MWU}_{\text{NoEco-Monetary}} p<0.01$; this also holds when looking at the last active round per group prior to forest depletion: $m_{\text{NoEco}}=12.74$ (N=80), $m_{\text{Symbolic}}=12.85$ (N=80) & $m_{\text{Monetary}}=15.39$ (N=84); Kruskal-Wallis Test: $\chi^2_{\text{ties}}(2)=11.54$, $p<0.01$; $\text{MWU}_{\text{Symbolic-Monetary}} p<0.01$; $\text{MWU}_{\text{NoEco-Symbolic}} p=0.66$; $\text{MWU}_{\text{NoEco-Monetary}} p<0.01$). This end-round effect suggests that earlier cooperative behavior in the Monetary condition is at least partially motivated by monetary incentives, and stronger monetary incentives may be associated with reduced cooperation in later rounds. This comparison of the two labels is consistent with prior findings showing that i) sustainability is possible despite greed⁷⁴ and ii) extrinsic incentives can crowd out intrinsic motivation, particularly in ethical decision making and coordination contexts^{88,89}.

Treatment effects persist with computerized players

To test whether the ecolabel effects depend on interacting with human peers, we use the computer rounds in the second part of our experiment. Here, each participant interacts with three computerized players implementing three withdrawal strategies commonly observed in human play (see Methods). This version complements the human rounds in two ways. Peer behavior is standardized: every participant faces the three preprogrammed strategies, so differences in outcomes reflect participants' own choices and the preprogrammed responses to

them, rather than the composition of their group. Moreover, each participant's forest is independent, since choices neither affect nor are observed by other human players. Replacing the human group members with computer players removes the scope for explanations based on concern for human peers, reputation, social approval and other-regarding preferences, that could otherwise drive the observed effects⁹⁰.

The main treatment effects persist in the computer rounds. In addition, and unlike the human rounds, sustainable outcomes differ throughout until the end of the game (forest alive, trees left, see Table 1B and Fig. 3): 19% of NoEco forests survive to the end, compared with 45% under Symbolic and 40% under Monetary (both $p < 0.01$, Table 1B, Fig. 3). The overall level differences between game versions are also by construction: group withdrawals are higher in the computer rounds because the preprogrammed strategies withdraw more in combination (Round 1: 40 plus own extraction) than three human players do on average (33.5). Accordingly, no fully sustainable groups remain in the computer rounds. Participants' own withdrawals are only modestly higher (Table 1, row 1). However, the ecolabel effects do not differ substantially between game versions, indicated by an absence of treatment by game-version interaction (Supplementary Results 3). The persistence of the ecolabel effects when human players are removed suggests that the mechanism through which ecolabels operate goes beyond social preferences. We therefore examine two components of the proposed coordination mechanism: (1) that ecolabels improve initial coordination on sustainable behavior, and (2) that this coordination is then sustained through conditional cooperation.

Ecolabels facilitate initial coordination

The first-round results are consistent with the presence of ecolabels shaping initial coordination. In round 1, both ecolabels lead to reduced withdrawals in the first round compared to NoEco (Fig. 2C), despite the sustainable benchmark being clearly communicated in each condition ("10 trees is considered sustainable"). Further, we find that in the ecolabel conditions, a larger share adheres to the sustainable behavior benchmark of up to 10 trees (Fig. 4A) or exactly 10 trees (Fig. 4B; round 1: NoEco=42.5%, Symbolic=76.25% & Monetary=82.1%; Kruskal-Wallis Test: $\chi^2_{ties}(2)=33.49$, $p < 0.01$; $MWU_{NoEco-Symbolic} p < 0.01$, $MWU_{NoEco-Monetary} p < 0.01$, $MWU_{Symbolic-Monetary} p > 0.3$). Within-group variance in round 1 withdrawals are also significantly smaller in both ecolabels than in NoEco (variance ratio test: NoEco_vs_Symbolic: $F(79,79)=1.60$, $p=0.04$; NoEco_vs_Monetary: $F(79,83)=2.46$, $p < 0.01$). These first-round

patterns also appear in the computer rounds (Supplementary Results 6.1) again emphasizing a mechanism beyond social preferences, signaling or status-seeking.

Conditional cooperation sustains initial coordination

Because ecolabel conditions significantly enhance sustainable withdrawals in round 1, we next examine how participants' subsequent behavior responds to their own and their group members' prior sustainable behavior. Across conditions, many participants show a pattern consistent with conditional cooperation: for every additional group member who withdrew sustainably in the previous round, a participant's odds of withdrawing sustainably increase (Odds ratio (OR)=1.28, 95%-CI[1.08,1.51], $p<0.01$, Supplementary Results 8.2). This association is even stronger when considering the number of group members who have been sustainable in all previous rounds (OR =1.96, 95%-CI[1.54,2.47], $p<0.01$).

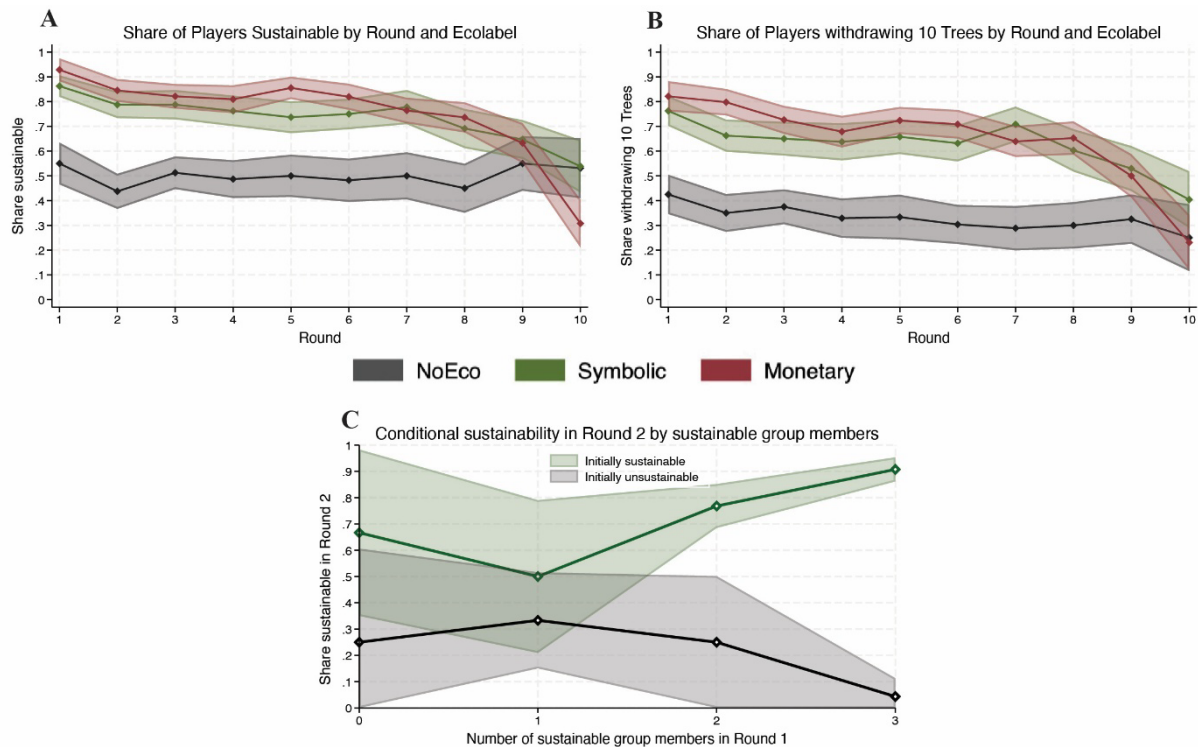


Fig. 4: Evidence for standards and conditional cooperation. (A) shows the proportion of sustainable players (i.e. withdrawing up to 10 trees) per round and ecolabel condition. (B) shows the proportion of players withdrawing exactly 10 trees per round and ecolabel. In both (A) and (B), there are significantly fewer sustainable players in the NoEco condition (see Supplementary Results 5 for computer rounds). (C) shows the proportion of players being sustainable in round 2 as a function of the number of other group members who were sustainable in round 1, reflecting players' early-game behavioral reactions to other group members' round-1 behavior. Players are separated by whether they were initially sustainable (green; $N=191$) or initially unsustainable (black; $N=53$). A and B use Cousineau-Morey within-subjects confidence intervals. All shaded areas represent 90% confidence intervals. Supplementary Results 5 replicates this figure in computer rounds.

Since the initially sustainable participants constitute a majority of the sample (191 initially sustainable vs. 53 initially unsustainable more details in Supplementary Results 8), their responses largely determine the aggregate dynamics. Initially sustainable participants showed significantly different withdrawal patterns from participants that were initially unsustainable after feedback about the number of sustainable group members at the beginning of round 2 (Fig. 4C; Table S10; interaction OR = 4.71, $p < 0.001$). Specifically, among the initially sustainable participants, the probability of sustainable withdrawal in round 2 increases with the number of sustainable group members in round 1 (Fig. 4C; initially sustainable participants only; OR=2.18, 95%-CI[1.34,3.54], $p<0.01$, Supplementary Results 8.2). By contrast, participants who did not behave sustainably in round 1 reduce their cooperation as most others cooperate (Fig. 4C; initially unsustainable only; OR = 0.46, $p = 0.022$). Thus, initially sustainable participants showed a pattern consistent with conditional cooperation, while initially unsustainable participants showed a pattern more consistent with freeriding⁵⁴.

The same pattern of reactivity to group behaviors unfolds when conditioning on the number of trees remaining at the start of round 2 instead of the number of sustainable others (Supplementary Fig. S5C). Because of this conditionality, early withdrawals are consequential for later outcomes. Groups in which all members act sustainably in round 1 survive 9.78 rounds on average. For each additional tree withdrawn above the sustainable threshold (>10) in round 1, forests survive approximately 0.2 rounds less (OLS, $t=-3.54$, $p<0.01$), and the odds of eventual depletion rise by 11% (logistic regression, OR=1.11, 95%-CI[1.01,1.23], $p=0.04$; Table S5). Across groups, a greater number of initially sustainable members is associated with more rounds alive (OLS, $t=3.82$, $p<0.01$) and directionally with lower estimated odds of eventual forest depletion (OR=0.52, 95%-CI[0.24,1.09], $p=0.08$). Full analyses and replications in the computer rounds are reported in Supplementary Results 6.2, Tables S4-S6.

As a descriptive complement to these analyses, we classify participants according to strategy types as done in previous studies on social dilemma (e.g. defectors, conditional cooperators, unconditional cooperators). Most participants are classified as conditional cooperators 69.26%, followed by 24.18% that are unconditional cooperators (UC) and 6.56% that are unconditional defectors (UD), Supplementary Fig. S6). The resulting distribution of types predicts long-term outcomes in the expected direction (Supplementary Table S9): groups with at least one unconditional defector (UD; always withdrawing >10 trees) survive an average of 2.9 fewer

rounds (mean rounds alive= 6.29, $t=4.16$, $p<0.01$) compared to groups without this type of player, emphasizing that group composition affects how conditional cooperators behave⁹¹.

Ecolabels shift strategy types and facilitate conditional cooperation

The distribution of these types differs across treatments. We find shifted distributions of strategic types in both ecolabel conditions compared to NoEco ($\chi^2(4)=22.38$, Fisher's exact $p<0.01$), reflected by an increased number of unconditional cooperators ($m_{\text{NoEco}}=11.25\%$ vs $m_{\text{Symbolic}}=37.5\%$ & $m_{\text{Monetary}}=23.8\%$; $\chi^2(2) = 15.04$, $p<0.01$) and fewer unconditional defectors ($m_{\text{NoEco}}=13.75\%$ vs $m_{\text{Symbolic}}=2.5\%$ & $m_{\text{Monetary}}=3.57\%$; $\chi^2(2)=10.13$, $p<0.01$) while the number of conditional cooperators differs only marginally ($\chi^2(2) = 4.91$, $p=0.09$). Finally, treatments differ in how participants react to other participants' behavior. While the conditionality effects above are pooled across ecolabel conditions, adding the label conditions, shows an interaction effect and different patterns of conditionality for the label conditions compared to the NoEco condition. This difference becomes clear when restricting the analysis to the early rounds (where more labels are present) and/or to participants for whom conditional cooperation rather than free riding is the relevant response (i.e. that were sustainable in round 1; Fig. 4C). The association between the number of sustainable group members and subsequent sustainable behavior is only positive in both ecolabel conditions (Rounds 2-4: Symbolic: OR=2.14, 95%-CI[1.14,4.01], $p=0.02$; Monetary: OR=2.57, 95%-CI[1.35,4.87], $p<0.01$; but not in NoEco: OR=0.88, 95%-CI[0.56,1.38], $p=0.58$; Supplementary Results 8.2). These results suggest that ecolabels affect both the visibility and prevalence of sustainable behavior by communicating others' behavior more clearly. Rather than demonstrating that ecolabels create cooperators, the pattern indicates that they seem to reduce the likelihood of free-riding by making sustainable behavior a more salient standard to match.

Environmental attitude, Symbolic labels, attention and behavior

Individual and group-level environmental attitudes are consistently positively associated with sustainable behavior across the experiment, reflecting the relevance of donative motives for individual decision making^{81,86}. Participants and groups with stronger pro-environmental attitude (1-item; 5-point Likert-scale; "I find the environment..."; see Methods) withdraw fewer trees in round 1 ($\beta=-0.765$, $p=0.04$), remain alive for more rounds ($\beta=0.499$, $p<0.01$), and leave more trees at the end ($\beta=2.30$, $p<0.01$; Supplementary Results 9).

When differentiating between ecolabel conditions, we only observe a significant correlation and an interaction effect between the Symbolic ecolabel condition and environmental attitudes. In the Symbolic ecolabel condition, stronger environmental attitudes are associated with the number of rounds alive (interaction: $\beta_{\text{Symbolic}}=1.21$, $p_{\text{Symbolic}}<0.01$) and the total number of trees withdrawn (interaction: $\beta_{\text{Symbolic}}=9.76$, $p_{\text{Symbolic}}=0.01$; Supplementary Table S11). We observe no significant interactions and no significant independent association (correlations split by condition) under the NoEco ($p>0.5$) and Monetary condition ($p>0.5$). Follow-up analyses for both game versions separately and a replication with group-averaged environmental attitudes are presented in Supplementary Results 9.

Analyses of eye-tracking data (recorded for a subset of 100 participants; Supplementary Results 10) suggest that ecolabels can also increase the salience of incentives. Specifically, participants in the Monetary condition allocate a larger share of their attention to payoffs compared to the Symbolic condition. This effect is consistent across two different analyses that ensure that all groups are included in the analyses: (1) before any of the groups deplete the forest (first 3 rounds; attention to points earned: $\beta_{\text{Round1-3}}=0.110$, $p_{\text{Round1-3}}=0.005$) and (2) in the last 3 rounds before forest depletion for each group (attention to points earned: $\beta_{\text{Last3Rounds}}=0.114$, $p_{\text{Last3Rounds}}<0.01$). This effect is enhanced for attention to the participants own earnings in the early rounds (attention to own points: $\beta_{\text{Round1-3}}=0.071$, $p_{\text{Round1-3}}=0.01$; $\beta_{\text{Last3Rounds}}=0.07$, $p_{\text{Last3Rounds}}=0.11$; see all comparisons also to the NoEco in Supplementary Results 10). Taken together, the attitude and attention findings provide secondary evidence for how ecolabels might work: the correlations between environmental attitudes and sustainable behavior and the attentional evidence suggest that symbolic ecolabels reinforce intrinsic environmentally-oriented motivation, while the monetary ecolabel can crowd these out and instead shift attention toward individual financial incentives.

3. Discussion

Our findings from “The Social Forest Game” highlight the underexplored potential of ecolabels as coordination devices. We show that ecolabels foster longer forest survival (Fig. 2), lower initial resource extraction (Fig. 2 & 3) and more coordinated early behavior (Fig. 4). These patterns are consistent with ecolabels enhancing the salience of a sustainable benchmark and, consequently, helping groups coordinate on the sustainable equilibria, rather than selfish ones. Once labels are in place and can be observed, they can provide additional information about others’ previous behavior, thereby likely reducing strategic uncertainty, especially in early rounds. The collective sustainable equilibria are then more likely to be maintained through conditional cooperation (Fig. 4C). Together, these findings support our hypothesis that ecolabels earned for sustainable behavior improve outcomes across both environmental and economic measures. This pattern is broadly consistent across measures and game versions (Fig.3).

Remarkably, a purely symbolic label (“Symbolic”) which carried no monetary reward, produces comparable outcomes to those of the monetary label (“Monetary”). If financial motives dominated, the Monetary ecolabel would consistently outperform the Symbolic one. Instead, both ecolabels yield similar effects across behavioral and environmental outcomes, which is consistent with ecolabels operating by signaling salient behavioral standards. Importantly, we find that the effects of ecolabels also persist in the computer rounds, where participants’ choices neither affect the outcomes of, nor are observed by, human peers. Ecolabel effects are already present in the first round, before participants had received feedback about others’ behavior or label status. This comparison reduces the likelihood that ecolabels work mostly based on social or status concerns.

While we find only minor differences between the labels in outcomes, one part of the underlying mechanisms seems to differ. The differences between the Symbolic and the Monetary label found in our study aligns with prior theory and evidence on intrinsic and extrinsic motivation⁸⁹. Symbolic labels strengthen the attitude-behavior link: pro-environmental attitudes predict more sustainable choices and outcomes only in the Symbolic condition⁸². These findings are compatible with the Focus Theory of Normative Conduct⁹² and the Norm Activation Model^{93,94}, which suggest that symbolic ecolabels make environmental norms both psychologically accessible and situationally salient. In contrast, participants in the

Monetary condition allocate more attention towards payoff information and withdraw more trees in the final rounds than participants in the Symbolic condition.

Our results connect economic and psychological theories, underscoring the interdisciplinary relevance. Ecolabels can provide a focal point⁹⁵ and facilitate conditionality⁵⁵, thereby helping groups solve the coordination problem inherent in dynamic common-pool settings. At the same time, by aligning norms and attention, ecolabels help bridge the value-action gap via the value-belief-norm pathway that underlies pro-environmental behavior⁹⁶. On the other hand, we show that adding monetary incentives to ecolabels, can bias attention and motives towards payouts and in course lead to crowding out effects undercutting the norm activation^{88,89}. More broadly, the findings are compatible with labels operating through the multiple influences identified by Van Lange et al. (2013)¹³: (1) psychological influence, i.e. by making a sustainable standard salient affecting initial behavior; (2) dynamic-interaction influence, i.e. by providing visible and continuous information about the sustainable behavior of peers; and (3) structural influence i.e. by altering material incentives (Monetary label).

The strength of this coordination depends on who is coordinating. Participants who behaved sustainably in round 1 become more likely to remain sustainable as more of their group did so. On the flip side, those who did not behave sustainably in round 1 become less likely to behave sustainably as several others behave sustainably, free-riding on others' restraint rather than reciprocating (Fig. 4C). The ecolabels make sustainable behavior more visible for either group and their effectiveness can depend on the distributions of types within a group. Here, ecolabels can have two positive effects. First, in our sample, initially sustainable participants predominate (78.3% are sustainable in round 1). The distribution of types in both ecolabel conditions contain a higher share of unconditional cooperators and a reduced share of unconditional defectors, hence increasing the chance to observe more sustainable behavior. Second, among initially sustainable participants, the association between others' sustainable behavior and subsequent own sustainable behavior is positive, whereas this association is not present or reversed in the NoEco condition. This indicates that without a visible standard and clear indication of others' adherence, even willing cooperators might behave more selfishly. Without labels, over-harvesting is visible in the state of the forest, while other's restraint is difficult to infer. Ecolabels supply this missing signal of (at least partial) restraint. Our results also suggest when labels should work best: they are most promising where many actors are willing to cooperate conditionally but fail to coordinate.

While the effects of ecolabels are consistent, in the hypothesized direction, and clear across all measures in the computer rounds, two end-state outcomes in the human rounds, trees left (and hence trees planted) and the share of forests depleted, show no significant differences across conditions. These are the outcomes most affected by well-documented end-game effects in finitely repeated games^{97–99}, in which participants extract more in the final rounds. In our data, this falls disproportionately on the ecolabel conditions, whose forests are the ones still alive to be depleted: seven of the eight forests depleted in round 10 were in the label conditions, and Monetary participants withdrew significantly more than Symbolic or NoEco participants in the final round. Together with attrition (unsustainable groups deplete the forest early, leaving predominantly sustainable groups by the end), this accounts for the end-state convergence in the human rounds. The realized end-state consequence is moreover only one of two sustainability-relevant outcomes that our design captures. The other is resource persistence, reflecting whether the shared resource survives continued extraction. This is arguably the more important outcome, as it defines a commons dilemma and is the criterion by which institutions are judged in the commons literature^{8,100}. It also matters beyond the model: depletion is irreversible, and delaying it preserves options for recovery, adaptation, or institutional and technological change¹⁰¹. On this outcome ecolabels have clear effects: forests survive significantly longer under both labels. We therefore interpret the results as evidence that ecolabels improved resource persistence and early sustainable behavior within the experimental commons dilemma, rather than as a demonstration of a direct real-world environmental effect.

Several limitations apply. We relied on a student sample which might differ in strategic decision making from firms or other resource users but has the advantage of having more experimental control. To get broader insights into the underlying processes, future work could elicit beliefs or expectations about other's strategies directly. We chose to not elicit them directly since this might interfere with choices in our dynamic and multi-round setting^{102,103}. Additionally, we find consistent results across subject pools at two different universities and countries. Further, as discussed in the methods, the known fixed horizon differs from several real resource management settings, but allows for more control and removes complexity. In addition, our paradigm also yields an equilibrium-selection problem mimicking studies with probabilistic ending periods. Finally, given the mixed field evidence of labels, one should be careful with taking the effect sizes too literal, as the effect might differ by salience of the

settings (such as observability or relevance of the labels) and other context factors (such as salience of environmental consequences, subject pool, strength of the label signal).

Our findings carry policy implications: in markets where individual actors exert substantial environmental impact but face strategic uncertainty, ecolabels can help to resolve that uncertainty and steer standards towards sustainability. Future work could further explore whether ecolabels can work effectively as a scalable and politically viable instrument for advancing the goals of the EU Climate Law and the Paris Agreement. Importantly, the Symbolic ecolabel produces comparable effects to those of the Monetary, despite not providing a 25%-increase in payments, while also not showing the crowding out effects. The purely symbolic label requires no subsidy, no transfer, and no payment infrastructure; it requires only a rule for what counts as sustainable, and a way of displaying who meets it. In our experiment, labels were awarded truthfully and at no cost. Real-world effectiveness, however, will hinge on certification credibility, monitoring frequency, and appropriately chosen thresholds. These issues merit further investigation²⁵, especially as field evidence on consumer effects of ecolabels is mixed, ecolabels can be misused for greenwashing¹⁰⁴, overload the consumer¹⁰⁵, or fulfil very low standards³⁰. Here, the emerging European large-scale policies to increase transparency and labelling for sustainable practices, such as the Digital Product Passport and the CSDDD, can serve as case studies and springboards for widespread labelling policies.

Both of these policy instruments target producers rather than consumers. Here, the standard-setting and coordination effect could matter significantly as producer decisions carry greater environmental impact and signaling power than individual choices^{25,37,71,106}. The mechanisms discussed may therefore translate to firms as well as households and prosumers¹⁰⁷. Like participants in our game, actors operating in the same market face strategic uncertainty about competitors' willingness to adopt sustainable practices and what sustainable practices are relevant for consumers and policymakers. A credible, visible standard can reduce this uncertainty in the same way it does in our experiment: it defines what counts as sustainable and reveals if others are adhering to it. We do not claim to capture the full complexity of organizational decision-making; certification choices by firms are also driven by strategic motives such as legitimacy, costs, differentiation, and market access^{37,108,109}. Yet, firms often adopt standards for symbolic and conformity reasons rather than purely material ones, so the coordination and standard-setting channel that we isolate may be one ingredient of why labelling regimes gain traction in markets¹¹⁰. Households and prosumers, in turn, may be the

most direct analogue of our setting: their environmentally consequential choices, such as home-energy investments, i) combine private returns with collective effects, ii) are made by individuals rather than organizations, and iii) are known to respond to labels and disclosure^{111–113}. Where such choices become visible to peers, as with rooftop solar, the same interplay of standards and observability that drives coordination in our experiment, could operate at the local scale.

By demonstrating that ecolabels can promote sustainable behavior even without financial incentives, our results broaden the potential role of labeling beyond influencing isolated consumer choices. Well-designed ecolabels can serve as standard-setting mechanisms and catalysts of sustainability transitions which can complement other cooperation interventions including communication, monitoring and sanctioning. Where actors face strategic uncertainty about others' willingness to adopt sustainable behavior, credible and observable labels can help groups coordinate quickly on sustainable behavior. The real-world effectiveness will depend on several factors (verification, realistic thresholds, or contexts), but labels can provide a comparatively accessible and scalable complement to other approaches required for climate mitigation and resource management.

4. Methods

Laboratory participants (N=244; 100 with eye-tracking in Konstanz; 144 in Amsterdam) made decisions with real monetary and environmental consequences in the “Social Forest Game” (Fig. 1). We pre-registered a target of about 48 participants per ecolabel condition (AsPredicted #120226) and did not specify an a-priori power analysis. Data collection at two sites increased the final sample size to 244 participants (Symbolic n=80, Monetary n=84, NoEco n=80). Since participants made decisions in groups of four, the sample size differs across levels of analysis: individual vs group and human vs computer rounds. At the individual level, the sample offers a minimum detectable effect of $d=0.4$ ($\alpha = .05$, power = .80), while at the group-level (groups of 20, 21 and 20 per condition) the analysis is less sensitive for the main Mann-Whitney U tests, with a minimum detectable effect of $d=0.82$. To check that pooling across sites did not introduce systematic differences, we tested condition by site interactions and found no relevant differences (Supplementary Results 11). In both locations, each session involved 8 to 20 participants and lasted around 50 minutes on average. Even though choices were made anonymously, we chose to have at least 8 people per session to ensure that participants did not know with certainty who was in their group. Participants received around 13.50 Euros on average. The experiment was coded in zTree5.1¹¹⁴. Data were analyzed using Stata¹¹⁵ and R¹¹⁶. We received ethics approval prior to conducting the study from the Economics and Business Ethics Committee (EBEC) at the University of Amsterdam.

The Social Forest Game

In the “Social Forest Game”, participants acted in groups of four, with roles framed as producer decisions. Each group was initially endowed with a common forest of 100 trees. In each round, each participant decided how many trees to withdraw, choosing between 0 and 20. Each withdrawn tree yielded €0.10 in profit. If at least one tree remained after all group members’ withdrawals in a round, the forest regrew: 40 trees were added to the remaining trees for the next round. We chose a constant regrowth rate to set a clear sustainable benchmark, and to lower the computational difficulties for the participants, as opposed to prior studies that used stochastic regrowth rates to model environmental uncertainties⁷⁵. Note that this might make sustainable choices also in the condition without ecolabel more salient, which might have reduced our effects. If no trees remained, the forest would be depleted permanently, and no profit was generated for that round and the game ended.

The decision repeated for up to ten rounds, or until the forest was depleted. We chose a finitely repeated design deliberately, because the resulting end-game behavior is diagnostic of players' motives, which is central to our research question. In our game, leaving trees in the final rounds reveals an intention to donate real trees, since no strategic reason for future extraction remains. In an indefinitely repeated game, trees left in any given round could instead reflect preservation for later profit, so donative and strategic motives could not be separated. A known final round also keeps the sustainable benchmark clearly computable for participants, whereas probabilistic continuation would confound continuation beliefs with treatment effects and add complexity to an already dynamic task. A finitely repeated common-pool resource game normally has a unique non-cooperative equilibrium, while cooperation can be sustained under indefinite repetition. We parametrized our game to admit multiple equilibria, both cooperative and non-cooperative, so that its predictions are qualitatively similar to those of an indefinitely repeated game while preserving the inferential leverage that a known final round provides.

At the end of the experiment, profits accumulated across rounds were paid to the subjects. All trees remaining after 10 rounds were aggregated across groups and we planted a total of 209 trees (through onetreepanted.org) for the main experiment and the pilot. Every 10 in-game trees corresponded to one real tree, which costs approximately €1.

Ecolabel treatments

The game was implemented under three between-subjects ecolabel conditions: NoEco, Symbolic, and Monetary. In NoEco, participants could not obtain ecolabels but were clearly instructed that a withdrawal of 10 trees or less is considered sustainable (see instructions and all other materials on [OSF](#)). In the Symbolic and Monetary conditions, participants received an ecolabel if they withdrew 10 or fewer trees per round. Withdrawing more than 10 trees permanently removed the ecolabel. The Symbolic ecolabel did not affect profits per tree. In contrast, while maintaining the Monetary ecolabel, profits increased by 25%. This 25%-increase was based on existing ecolabels thresholds⁴⁰ and pilot participants' support for this multiplier as reasonable. The participant's own ecolabel status and the number of producers holding ecolabels in the group were shown on the decision screen.

Experimental Procedure

Each experimental session consisted of three parts. In Part 1, each participant played up to ten rounds of the "Social Forest Game" with three other human participants. In Part 2, participants

played up to ten rounds of the same game, this time with three computerized players. In Part 3, participants answered a questionnaire including Likert-scale measures of environmental attitudes and demographics ([OSF](#)). More specifically, for our analyses we used the following statement which was rated on a 5-point Likert scale (from “not important” to “very important”): “I find the environment...”. All other questionnaires can be found on the [OSF](#). Note that we additionally recorded eye movements for 100 subjects (Tobii EyeX, see Supplementary Results 10).

Computerized Players

The computerized players in Part 2 replicated strategies commonly used by human subjects in our pilot study (N=36). Specifically, each group contained one sustainable computer player (withdrawing 5 to 10 trees per round), one selfish computer player (withdrawing 11 to 20 trees per round), and one conditional sustainable computer player who adapted withdrawals based on the human participant’s withdrawal choice in the previous round (withdrawing 5 to 20 trees per round). In the pilot, 39% were classified as selfish types (starting by withdrawing more than 10 trees and continuing to do so for additional rounds), 30.5% behaved sustainably (never withdrawing more than 10 trees), and the remaining 30.5% were conditional cooperators (withdrawing between 5-20 trees depending on others’ behavior during a given round).

Participants were told that the computerized players’ strategies were modeled after real human behavior observed in an earlier experiment (including the month of that earlier experiment). This design preserved participants’ expectations about others’ behavior in their group while removing the direct influence of human peers⁵⁵.

Data availability

All data and code used for the analyses are available on the project’s [OSF page](#).

Acknowledgments

We are grateful for funding provided by the *A Sustainable Future* initiative of the faculty of Economics and Business, University of Amsterdam.

Author Contributions

JBK: Implementation, Analysis, Writing original draft

NL: Methodology, Analysis support, Writing-Review & Editing

JBE: Conceptualization, Analysis support, Writing-Review & Editing, Supervision

JH: Conceptualization, Implementation, Analysis support, Writing-Review & Editing, Supervision, Funding

Use of LLMs

During revision, the authors used Claude (Anthropic; Claude Fable 5.1, September 2026) to assist with language and structural editing of the Introduction and Results, to check the internal consistency of reported statistics across text, tables, and figures, and to review the Stata code. All analyses were run by the authors, and the model did not generate data, figures, or scientific conclusions. All suggested text was reviewed and edited by the authors, who take full responsibility for the content.

5. References

1. Hardin, G. The Tragedy of the Commons. *Science* **162**, 1243–1248 (1968).
2. Ostrom, E. The value-added of laboratory experiments for the study of institutions and common-pool resources. *J. Econ. Behav. Organ.* **61**, 149–163 (2006).
3. Ostrom, E. A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science* **325**, 419–422 (2009).
4. Adger, W. N., Barnett, J., Heath, S. & Jarillo, S. Climate change affects multiple dimensions of well-being through impacts, information and policy responses. *Nat. Hum. Behav.* **6**, 1465–1473 (2022).
5. Cai, W. *et al.* Views on climate change and health. *Nat. Clim. Change* **14**, 419–423 (2024).
6. Sellers, S., Ebi, K. L. & Hess, J. Climate Change, Human Health, and Social Stability: Addressing Interlinkages. *Environ. Health Perspect.* **127**, 45002 (2019).
7. Stern, N. The Economics of Climate Change. *Am. Econ. Rev.* **98**, 1–37 (2008).
8. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*. (Cambridge University Press, Dallas, TX, 1990).
9. De Moor, T. *The Dilemma of the Commoners: Understanding the Use of Common-Pool Resources in Long-Term Perspective*. (Cambridge University Press, Cambridge, 2015). doi:10.1017/CBO9781139135450.
10. Gardner, R., Ostrom, E. & Walker, J. M. The Nature of Common-Pool Resource Problems. *Ration. Soc.* **2**, 335–358 (1990).
11. Redpath, S. M. *et al.* Games as Tools to Address Conservation Conflicts. *Trends Ecol. Evol.* **33**, 415–426 (2018).
12. Ostrom, E. *The Drama of the Commons*. (National Academy Press, Washington, DC, 2002).

13. Van Lange, P. A. M., Joireman, J., Parks, C. D. & Van Dijk, E. The psychology of social dilemmas: A review. *Organ. Behav. Hum. Decis. Process.* **120**, 125–141 (2013).
14. Chaudhuri, A. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Exp. Econ.* **14**, 47–83 (2011).
15. Mantilla, C. Communication networks in common-pool resource games: Field experimental evidence. *J. Econ. Behav. Organ.* **118**, 215–226 (2015).
16. Balliet, D. Communication and Cooperation in Social Dilemmas: A Meta-Analytic Review. *J. Confl. Resolut.* **54**, 39–57 (2010).
17. Balliet, D., Mulder, L. B. & Van Lange, P. A. M. Reward, punishment, and cooperation: a meta-analysis. *Psychol. Bull.* **137**, 594–615 (2011).
18. Nikiforakis, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *J. Public Econ.* **92**, 91–112 (2008).
19. Ostrom, E., Walker, J. & Gardner, R. Covenants with and without a Sword: Self-Governance Is Possible. *Am. Polit. Sci. Rev.* **86**, 404–417 (1992).
20. Rustagi, D., Engel, S. & Kosfeld, M. Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science* **330**, 961–965 (2010).
21. Bergquist, M., Nilsson, A. & Schultz, W. P. A meta-analysis of field-experiments using social norms to promote pro-environmental behaviors. *Glob. Environ. Change* **59**, 101941- (2019).
22. Biel, A. & Thøgersen, J. Activation of social norms in social dilemmas: A review of the evidence and reflections on the implications for environmental behaviour. *J. Econ. Psychol.* **28**, 93–112 (2007).
23. Rege, M. & Telle, K. The impact of social approval and framing on cooperation in public good situations. *J. Public Econ.* **88**, 1625–1644 (2004).

24. Milinski, M., Semmann, D. & Krambeck, H.-J. Reputation helps solve the ‘tragedy of the commons’. *Nature* **415**, 424–426 (2002).
25. Taufique, K. M. R. *et al.* Revisiting the promise of carbon labelling. *Nat. Clim. Change* **12**, 132–140 (2022).
26. Potter, C. *et al.* The Effects of Environmental Sustainability Labels on Selection, Purchase, and Consumption of Food and Drink Products: A Systematic Review. *Environ. Behav.* **53**, 891–925 (2021).
27. Yokessa, M. & Marette, S. A Review of Eco-labels and their Economic Impact. *Int. Rev. Environ. Resour. Econ.* **13**, 119–163 (2019).
28. Grunert, K. G., Hieke, S. & Wills, J. Sustainability labels on food products: Consumer motivation, understanding and use. *Food Policy* **44**, 177–189 (2014).
29. Torma, G. & Thøgersen, J. A systematic literature review on meta sustainability labeling – What do we (not) know? *J. Clean. Prod.* **293**, 126194 (2021).
30. Harbaugh, R., Maxwell, J. W. & Roussillon, B. Label Confusion: The Groucho Effect of Uncertain Standards. *Manag. Sci.* **57**, 1512–1527 (2011).
31. Kaufmann, M., Andre, P. & Köszegi, B. Understanding Markets with Socially Responsible Consumers. *Q. J. Econ.* **139**, 1989–2035 (2024).
32. Darnall, N. & Aragón-Correa, J. A. Can Ecolabels Influence Firms’ Sustainability Strategy and Stakeholder Behavior? *Organ. Environ.* **27**, 319–327 (2014).
33. Liu, T., Wang, Q. & Su, B. A review of carbon labeling: Standards, implementation, and impact. *Renew. Sustain. Energy Rev.* **53**, 68–79 (2016).
34. Iraldo, F., Griesshammer, R. & Kahlenborn, W. The future of ecolabels. *Int. J. Life Cycle Assess.* **25**, 833–839 (2020).

35. Prieto-Sandoval, V., Alfaro, J. A., Mejía-Villa, A. & Ormazabal, M. ECO-labels as a multidimensional research topic: Trends and opportunities. *J. Clean. Prod.* **135**, 806–818 (2016).
36. Angel, D. P., Hamilton, T. & Huber, M. T. Global Environmental Standards for Industry. *Annu. Rev. Environ. Resour.* **32**, 295–316 (2007).
37. Kesselring, A. Energy labels in the European Union: Consumer inattention and producer responses. *Energy Econ.* **144**, 108275 (2025).
38. Büthe, T. The power of norms; the norms of power: who governs international electrical and electronic technology? in *Who Governs the Globe?* (eds Avant, D. D., Finnemore, M. & Sell, S. K.) 292–332 (Cambridge University Press, Cambridge, 2010). doi:10.1017/CBO9780511845369.012.
39. Constantino, S. M. *et al.* Scaling Up Change: A Critical Review and Practical Guide to Harnessing Social Norms for Climate Action. *Psychol. Sci. Public Interest* **23**, 50–97 (2022).
40. Giacomarra, M., Crescimanno, M., Vrontis, D., Miret Pastor, L. & Galati, A. The ability of fish ecolabels to promote a change in the sustainability awareness. *Mar. Policy* **123**, 104292 (2021).
41. Vandenbergh, M. P., Dietz, T. & Stern, P. C. Time to try carbon labelling. *Nat. Clim. Change* **1**, 4–6 (2011).
42. Nyborg, K. *et al.* Social norms as solutions. *Science* <https://doi.org/10.1126/science.aaf8317> (2016) doi:10.1126/science.aaf8317.
43. Ciravegna Martins Da Fonseca, L. M., Domingues, J. P., Machado, P. B. & Harder, D. ISO 9001:2015 adoption: A multi-country empirical research. *J. Ind. Eng. Manag.* **12**, 27 (2019).

44. Popa, D. I. & Dabija, D.-C. ISO 26000: A Brief Literature Review. in *ISO 26000 - A Standardized View on Corporate Social Responsibility: Practices, Cases and Controversies* (eds Idowu, S. O., Sitnikov, C. & Moratis, L.) 81–92 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-319-92651-3_6.
45. European Court of Auditors. *EU Action on Ecodesign and Energy Labelling : Important Contribution to Greater Energy Efficiency Reduced by Significant Delays and Non-Compliance*. (Publications Office, LU, 2020).
46. Schuitema, G., Aravena, C. & Denny, E. The psychology of energy efficiency labels: Trust, involvement, and attitudes towards energy performance certificates in Ireland. *Energy Res. Soc. Sci.* **59**, 101301 (2020).
47. So, H. W. & Laforteza, R. Reviewing the impacts of eco-labelling of forest products on different dimensions of sustainability in Europe. *For. Policy Econ.* **145**, (2022).
48. Andre, P., Boneva, T., Chopra, F. & Falk, A. Globally representative evidence on the actual and perceived support for climate action. *Nat. Clim. Change* **14**, 253–259 (2024).
49. Sparkman, G., Geiger, N. & Weber, E. U. Americans experience a false social reality by underestimating popular climate policy support by nearly half. *Nat. Commun.* **13**, 4779 (2022).
50. Tankard, M. E. & Paluck, E. L. Norm Perception as a Vehicle for Social Change. *Soc. Issues Policy Rev.* **10**, 181–211 (2016).
51. Dannenberg, A. *et al.* The Effects of Norms on Environmental Behavior. *Rev. Environ. Econ. Policy* **18**, 165–187 (2024).
52. Burlando, R. M. & Guala, F. Heterogeneous Agents in Public Goods Experiments. *Exp. Econ.* **8**, 35–54 (2005).
53. Casari, M. & Tavoni, A. Climate clubs in the laboratory. *J. Behav. Exp. Econ.* **110**, 102211 (2024).

54. Fischbacher, U. & Gächter, S. Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *Am. Econ. Rev.* **100**, 541–556 (2010).
55. Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404 (2001).
56. Battu, B. Co-evolution of conditional cooperation and social norm. *Sci. Rep.* **13**, 16625 (2023).
57. Carattini, S., Levin, S. & Tavoni, A. Cooperation in the Climate Commons. *Rev. Environ. Econ. Policy* **13**, 227–247 (2019).
58. Tingley, D. & Tomz, M. Conditional Cooperation and Climate Change. *Comp. Polit. Stud.* **47**, 344–368 (2014).
59. Bartling, B., Weber, R. & Yao, L. Do Markets Erode Social Responsibility? *Q. J. Econ.* **130**, 219–266 (2015).
60. Camilleri, A. R., Larrick, R. P., Hossain, S. & Patino-Echeverri, D. Consumers underestimate the emissions associated with food but are aided by labels. *Nat. Clim. Change* **9**, 53–58 (2019).
61. Claessens, S., Kelly, D., Sibley, C. G., Chaudhuri, A. & Atkinson, Q. D. Cooperative phenotype predicts climate change belief and pro-environmental behaviour. *Sci. Rep.* **12**, 12730 (2022).
62. Dannenberg, A. & Gallier, C. The choice of institutions to solve cooperation problems: a survey of experimental research. *Exp. Econ.* **23**, 716–749 (2020).
63. Velez, M. A., Stranlund, J. K. & Murphy, J. J. Preferences for government enforcement of a common pool harvest quota: Theory and experimental evidence from fishing communities in Colombia. *Ecol. Econ.* **77**, 185–192 (2012).

64. Velez, M. A., Stranlund, J. K. & Murphy, J. J. What motivates common pool resource users? Experimental evidence from the field. *J. Econ. Behav. Organ.* **70**, 485–497 (2009).
65. Gelcich, S., Guzman, R., Rodríguez-Sickert, C., Castilla, J. C. & Cárdenas, J. C. Exploring External Validity of Common Pool Resource Experiments: Insights from Artisanal Benthic Fisheries in Chile. *Ecol. Soc.* **18**, (2013).
66. Boldrini, M., Bosetti, V. & Nunnari, S. Negative Emission Technologies and Climate Cooperation. SSRN Scholarly Paper at <https://doi.org/10.2139/ssrn.4711262> (2024).
67. Stoddard, I. *et al.* Three Decades of Climate Mitigation: Why Haven't We Bent the Global Emissions Curve? *Annu. Rev. Environ. Resour.* **46**, 653–689 (2021).
68. Barahona, N., Otero, C. & Otero, S. Equilibrium Effects of Food Labeling Policies. *Econometrica* **91**, 839–868 (2023).
69. Chater, N. & Loewenstein, G. The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. *Behav. Brain Sci.* **46**, e147 (2023).
70. Whitmarsh, L., Poortinga, W. & Capstick, S. Behaviour change to address climate change. *Curr. Opin. Psychol.* **42**, 76–81 (2021).
71. Nielsen, K. S. *et al.* Realizing the full potential of behavioural science for climate change mitigation. *Nat. Clim. Change* **14**, 322–330 (2024).
72. Kukowski, C. A. *et al.* Leveraging agency for climate change mitigation. *Nat. Clim. Change* **16**, 627–629 (2026).
73. Brozyna, C., Guilfoos, T. & Atlas, S. Slow and deliberate cooperation in the commons. *Nat. Sustain.* **1**, 184–189 (2018).

74. Osten, F. B. von der, Kirley, M. & Miller, T. Sustainability is possible despite greed - Exploring the nexus between profitability and sustainability in common pool resource systems. *Sci. Rep.* **7**, 2307 (2017).
75. Malézieux, A. & Spiegelman, E. An anatomical review of the common pool resource game. *Exp. Econ.* **28**, 468–491 (2025).
76. Balboni, C., Berman, A., Burgess, R. & Olken, B. A. The Economics of Tropical Deforestation. *Annu. Rev. Econ.* **15**, 723–754 (2023).
77. Andersson, K. P. *et al.* Experimental evidence on payments for forest commons conservation. *Nat. Sustain.* **1**, 128–135 (2018).
78. Andrews, T. M., Delton, A. W. & Kline, R. High-risk high-reward investments to mitigate climate change. *Nat. Clim. Change* **8**, 890–894 (2018).
79. Marotzke, J., Semmann, D. & Milinski, M. The economic interaction between climate change mitigation, climate migration and poverty. *Nat. Clim. Change* **10**, 518–525 (2020).
80. Garcia, C. A. *et al.* Strategy games to improve environmental policymaking. *Nat. Sustain.* **5**, 464–471 (2022).
81. Bosshard, A. *et al.* Limited overlap between behavioral tasks, pro-environmental propensity, and carbon footprint. *J. Environ. Psychol.* 102297 (2024) doi:10.1016/j.jenvp.2024.102297.
82. ElHaffar, G., Durif, F. & Dubé, L. Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *J. Clean. Prod.* **275**, 122556 (2020).
83. Berger, S. & Wyss, A. M. Measuring pro-environmental behavior using the carbon emission task. *J. Environ. Psychol.* **75**, 101613 (2021).

84. Dablander, F., Lange, F., Brick, C. & Aron, A. R. Expressing intentions is not climate action. *Proc. Natl. Acad. Sci.* **122**, e2512457122 (2025).
85. Lange, F. Behavioral paradigms for studying pro-environmental behavior: A systematic review. *Behav. Res. Methods* **55**, 600–622 (2023).
86. Lange, F. *et al.* Beyond self-reports: A call for more behavior in environmental psychology. *J. Environ. Psychol.* **86**, 1–3 (2023).
87. Van Loo, E. J., Grebitus, C., Nayga, R. M., Verbeke, W. & Roosen, J. On the Measurement of Consumer Preferences and Food Choice Behavior: The Relation Between Visual Attention and Choices. *Appl. Econ. Perspect. Policy* **40**, 538–562 (2018).
88. Bruers, S. Review of behavioral economics models of the altruistic crowding-out effect from monetary incentives. *J. Econ. Surv.* **38**, 1656–1685 (2024).
89. Deci, E. L., Koestner, R. & Ryan, R. M. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol. Bull.* **125**, 627–668 (1999).
90. Engelmann, J. B., Meyer, F., Ruff, C. C. & Fehr, E. The neural circuitry of affect-induced distortions of trust. *Sci. Adv.* **5**, eaau3413 (2019).
91. Oliveira, A., Croson, R. & Eckel, C. One bad apple? Heterogeneity and information in public good provision. *Exp. Econ.* **18**, 116–135 (2015).
92. Cialdini, R. B., Reno, R. R. & Kallgren, C. A. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* **58**, 1015–1026 (1990).
93. Schwartz, S. H. Normative Influences on Altruism¹. in *Advances in Experimental Social Psychology* (ed. Berkowitz, L.) vol. 10 221–279 (Academic Press, 1977).

94. Stern, P., Dietz, T., Abel, T., Guagnano, G. & Kalof, L. A Value-Belief-Norm Theory of Support for Social Movements: The Case of Environmentalism. *Hum. Ecol. Rev.* **6**, 81–97 (1999).
95. Schelling, T. C. *The Strategy of Conflict*. vii, 303 (Harvard Univer. Press, Oxford, England, 1960).
96. Stern, P. C. New Environmental Theories: Toward a Coherent Theory of Environmentally Significant Behavior. *J. Soc. Issues* **56**, 407–424 (2000).
97. Herr, A., Gardner, R. & Walker, J. M. An Experimental Study of Time-Independent and Time-Dependent Externalities in the Commons. *Games Econ. Behav.* **19**, 77–96 (1997).
98. Lugovskyy, V., Puzzello, D., Sorensen, A., Walker, J. & Williams, A. An experimental study of finitely and infinitely repeated linear public goods games. *Games Econ. Behav.* **102**, 286–302 (2017).
99. Normann, H.-T. & Wallace, B. The impact of the termination rule on cooperation in a prisoner's dilemma experiment. *Int. J. Game Theory* **41**, 707–718 (2012).
100. Cox, M., Arnold, G. & Villamayor Tomás, S. A Review of Design Principles for Community-based Natural Resource Management. *Ecol. Soc.* **15**, (2010).
101. Arrow, K. J. & Fisher, A. C. Environmental Preservation, Uncertainty, and Irreversibility. *Q. J. Econ.* **88**, 312–319 (1974).
102. Croson, R. Thinking like a game theorist: factors affecting the frequency of equilibrium play. *J. Econ. Behav. Organ.* **41**, 299–314 (2000).
103. Gächter, S. & Renner, E. The effects of (incentivized) belief elicitation in public goods experiments. *Exp. Econ.* **13**, 364–377 (2010).
104. Delmas, M. A. & Burbano, V. C. The Drivers of Greenwashing. *Calif. Manage. Rev.* **54**, 64–87 (2011).

105. Brécard, D. Consumer confusion over the profusion of eco-labels: Lessons from a double differentiation model. *Resour. Energy Econ.* **37**, 64–84 (2014).
106. Constantino, S. M., Pianta, S., Rinscheid, A., Frey, R. & Weber, E. U. The source is the message: the impact of institutional signals on climate change–related norm perceptions and behaviors. *Clim. Change* **166**, 35 (2021).
107. Bollinger, B. & Gillingham, K. Peer Effects in the Diffusion of Solar Photovoltaic Panels. *Mark. Sci.* **31**, 900–912 (2012).
108. Blyde, J. The ISO 14001 Environmental Standard and Exports. *Int. Trade J.* **39**, 115–128 (2025).
109. Iatridis, K. & Kesidou, E. What Drives Substantive Versus Symbolic Implementation of ISO 14001 in a Time of Economic Crisis? Insights from Greek Manufacturing Companies. *J. Bus. Ethics* **148**, 859–877 (2018).
110. Boiral, O. Corporate Greening Through ISO 14001: A Rational Myth? *Organ. Sci. Provid. RI* **18**, 127–146 (2007).
111. Brounen, D. & Kok, N. On the economics of energy labels in the housing market. *J. Environ. Econ. Manag.* **62**, 166–179 (2011).
112. Newell, R. G. & Siikamäki, J. Nudging Energy Efficiency Behavior: The Role of Information Labels. *J. Assoc. Environ. Resour. Econ.* **1**, 555–598 (2014).
113. Stadelmann, M. & Schubert, R. How Do Different Designs of Energy Labels Influence Purchases of Household Appliances? A Field Study in Switzerland. *Ecol. Econ.* **144**, 112–123 (2018).
114. Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* **10**, 171–178 (2007).
115. StataCorp. Stata Statistical Software: Release 18. StataCorp LLC (2023).

116. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2024).

Supplementary Results

1. *Forest longevity and consistency across rounds*

1.1 Restricted Mean Survival Time (RMST) Analysis

To quantify the protective effect of ecolabels on forest survival, we introduce a Protection Factor (PF_{50}) - an analogue of the LD_{50} concept in toxicology - defined as the ratio of median survival times, $PF_{50} = T_{50(\text{with ecolabel})} / T_{50(\text{without ecolabel})}$. Median survival times (T_{50}) were estimated at the group level for the human rounds from Kaplan–Meier curves that account for right-censoring. Median survival was 7 rounds for NoEco, whereas in both ecolabel conditions it was approximately 10 rounds, as fewer than half of the forests were lost before the end of the 10-round game ($T_{50} \geq 10$). These results imply a $PF_{50} \geq 1.43$ for each ecolabel, reflecting a 43% increase in forest survival length.

Given that medians in the ecolabel conditions were not reliably reached, the restricted mean survival time (RMST) ratio provides a more informative and stable measure. It serves as a continuous analogue to PF_{50} , capturing proportional changes in overall survival rather than only at the 50% threshold. Specifically, RMST is a non-parametric time-to-event measure that estimates the average survival time within a fixed observation window (here, $\tau = 10$ rounds) by integrating the Kaplan–Meier survival curve^{1,2}. RMST thus provides a robust, censoring-adjusted estimate of how interventions shift the entire survival distribution.

RMST analyses showed that forests in the NoEco condition survived on average 7.5 rounds (95%-CI[6.3, 8.6]) out of 10. The Monetary ecolabel extended average survival to 8.9 rounds (difference=1.45, RMST-ratio=1.20, $p=0.05$), and the Symbolic label to 9.2 rounds (difference=1.70, RMST-ratio=1.23, $p=0.018$). Pooling both ecolabel conditions yielded a mean survival of 9.0 rounds, 1.6 rounds longer than NoEco (RMST-ratio=1.21, $p=0.014$). Jointly, these results reflect an overall protection factor of $PF_{50} \geq 1.2$.

Table S1. MWU tests for the effect of ecolabel on Trees left by round

Round	NoEco vs Symbolic			NoEco vs Monetary		
	$z_{uncorrected}$	$p_{uncorrected}$	$p_{corrected}$	$z_{uncorrected}$	$p_{uncorrected}$	$p_{corrected}$
1	-2.065	0.039	0.065	-3.069	<0.01	0.011
2	-2.682	<0.01	0.057	-3.165	<0.01	0.011
3	-2.301	0.021	0.057	-2.665	<0.01	0.011
4	-2.210	0.027	0.057	-2.591	<0.01	0.012
5	-1.791	0.073	0.090	-2.676	<0.01	0.011
6	-2.001	0.045	0.065	-2.787	<0.01	0.011
7	-2.189	0.029	0.057	-2.917	<0.01	0.011
8	-2.368	0.018	0.057	-2.690	<0.01	0.011
9	-1.746	0.081	0.090	-1.984	0.047	0.053

Table S1. Pairwise Mann–Whitney U tests comparing the no ecolabel condition (NoEco) with Symbolic and Monetary ecolabel conditions on Trees Left at the end of each round (comparison at group level: NoEco: N=20, Symbolic: N=20, Monetary: N=21). The Benjamini-Hochberg correction was applied separately for the columns with corrected p-values (NoEco vs Symbolic; NoEco vs Monetary). Round 10 was excluded due to end game effects and non-significance.

1.2 Consistency across rounds

Figures S1 and S2 show the consistency of the treatment effects across the whole game (Fig. 3 in main text), now including the comparisons between the two ecolabels. For each outcome and the three pairwise condition contrasts, we ran a Mann-Whitney U test in every round and display the resulting grid of tests, separately for the human and computer rounds. Figure S1 reports uncorrected p-values and Figure S2 the Benjamini-Hochberg-corrected equivalents.

The pattern seems rather consistent throughout than driven by particular rounds: both ecolabels differ from NoEco on most outcomes across most of the game. The third block comparing the two ecolabels with each other, reveals mostly insignificant effect throughout both panels, indicating that the two ecolabel conditions rarely differ from each other in any round on any outcome. One difference is the money earned which stems from the monetary ecolabels increasing earnings per tree, i.e. 25%-increase while ecolabel is maintained.

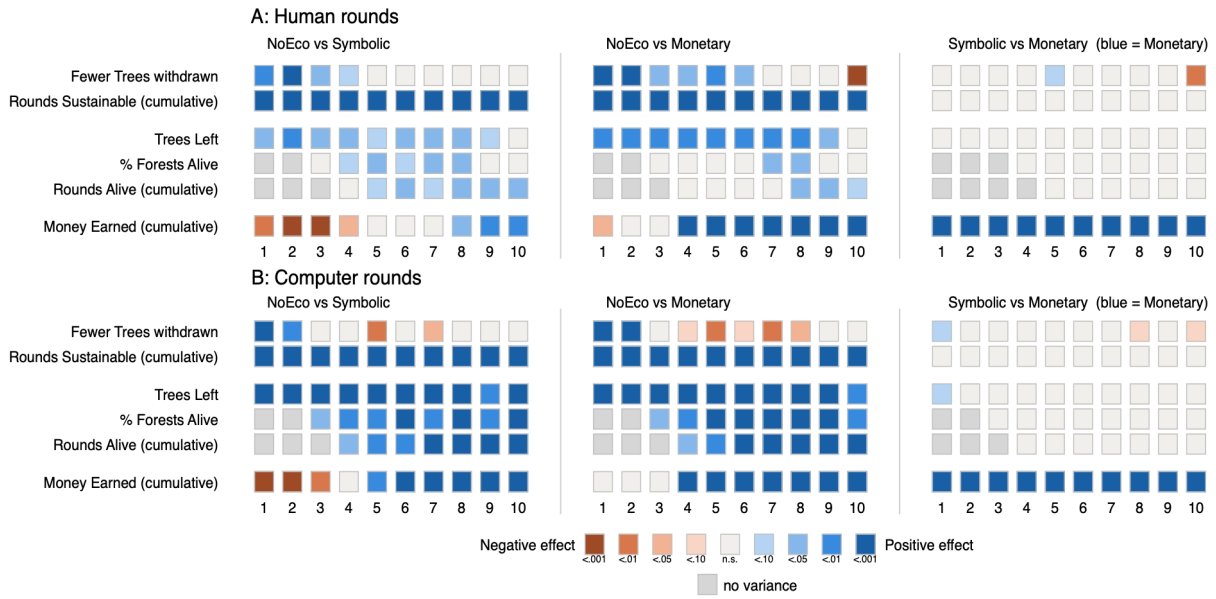


Fig. S2 Consistency of ecolabel effects across rounds (uncorrected), full Figure 3. Each cell reports a pairwise Mann–Whitney U test comparing two conditions on one outcome in one round. Rows are the six outcomes (trees withdrawn, reverse-coded so that blue always indicates the more sustainable outcome; cumulative rounds sustainable; trees left; % forests alive; cumulative rounds alive; cumulative money earned); columns are rounds 1–10. Panel A shows the human rounds, Panel B the computer rounds. The three blocks give the three pairwise contrasts: NoEco vs Symbolic, NoEco vs Monetary, and Symbolic vs Monetary. Color indicates the direction and strength of the effect, with blue favoring the second-named condition (for the third block, blue = Monetary) and orange the first; grey cells indicate no variance in the outcome, so no test could be run. Tests on group-level outcomes use one observation per group (NoEco $N = 20$, Symbolic $N = 20$, Monetary $N = 21$). p -values are uncorrected.

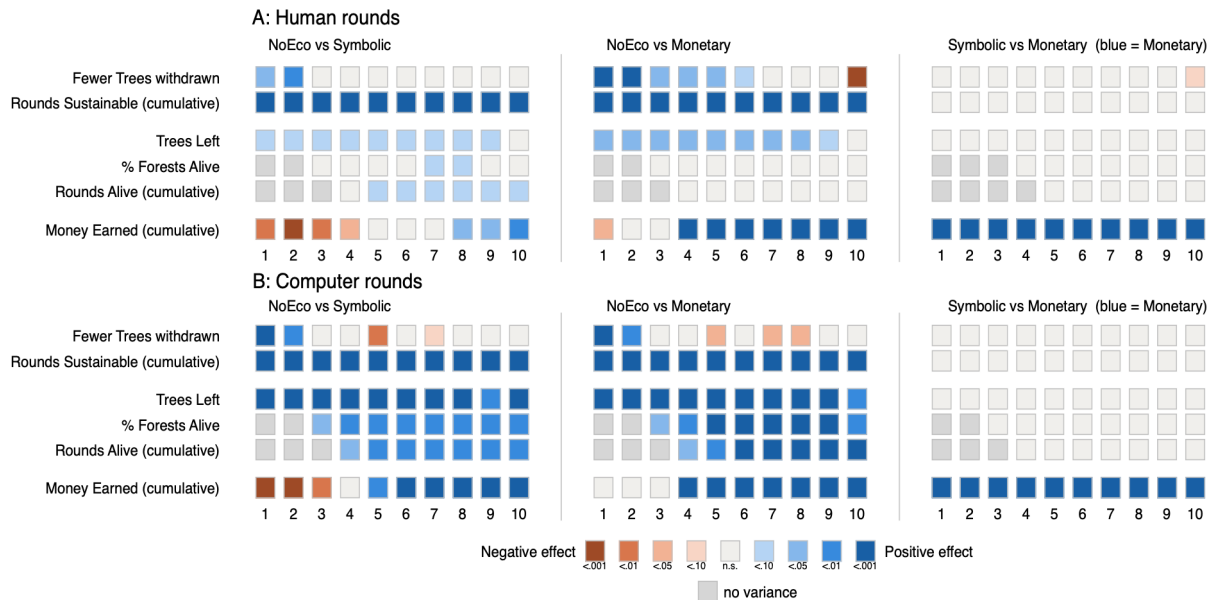


Fig. S3 Adaptation of Figure S1 with p -values adjusted with the Benjamini–Hochberg correction applied within each “strip”, i.e. across rounds within a given outcome and contrast. The correction family is the number of tests run in that strip, so strips containing no-variance cells are corrected over fewer than ten comparisons (in % forests alive over 8, in cumulative rounds alive over 7).

2. Effect of ecolabels on the number of trees withdrawn

For all of the regression analyses, we account for non-independence of repeated observations by estimating robust standard errors clustered at the individual level (for analyses involving repeated measures across rounds) or at the group level (for analyses aggregated by group), which is explained in the respective table descriptions.

Table S2. Consistent Effects of Ecolabels across different round inclusions

Predictor	Round = 1	Round ≤ 3	Round ≤ 7	Round ≤ 10
Symbolic EcoLabel (vs. NoEco)	-1.988** (0.772)	-1.425* (0.574)	-0.903* (0.382)	-0.544† (0.281)
Monetary EcoLabel (vs. NoEco)	-2.605*** (0.741)	-1.607** (0.548)	-1.060** (0.384)	-0.430 (0.302)
Constant	12.713*** (0.668)	12.492*** (0.404)	12.021*** (0.279)	11.945*** (0.234)
<i>F</i>	6.07**	4.84**	4.29*	1.87
<i>R</i> ²	0.072	0.03	0.013	0.003
N (Subjects)	244 (244)	732 (244)	1584 (244)	2076 (244)

Table S2. OLS Regression Results for Trees Withdrawn across rounds. Each column reports a separate regression estimating the effect of ecolabels for a different subset of rounds (1, 1-3, 1-7, 1-10). The dependent variable is number of trees withdrawn. Each separate OLS regression with two-way clustered robust standard errors by subject and group. F-statistics test the joint significance of Ecolabel predictors. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

3. Ecolabels exert similar effects on behavior across game versions

To test whether the effects of ecolabels differ between the human and computer-versions of the game, we estimated linear regression models including interaction terms between ecolabels and game version for each dependent variable. As shown in Table S3, none of the interaction terms are statistically significant (all $p > 0.1$), indicating overall similar results across the two game type conditions.

Table S3: Effects of ecolabels across game versions: OLS regressions with ecolabel by game version interactions

Predictor	% Depleted Forests	Round 1 Withdrawn	Rounds Sustainable	Avg. Trees Left	Total Trees Withdrawn	Avg. Rounds Alive
Intercept	0.65*** (0.05)	12.71*** (0.45)	3.71*** (0.36)	6.20*** (1.81)	80.05*** (3.12)	7.45*** (0.27)
Symbolic Ecolabel	-0.10 (0.08)	-1.99** (0.64)	3.09*** (0.51)	4.95† (2.55)	17.05*** (4.41)	1.70*** (0.38)
Monetary Ecolabel	-0.13† (0.08)	-2.61*** (0.63)	3.16*** (0.51)	5.75* (2.52)	14.62*** (4.35)	1.45*** (0.38)
Computer Rounds	0.16* (0.08)	1.29* (0.64)	0.19 (0.51)	-0.95 (2.55)	-20.11*** (4.41)	-1.13** (0.38)
Symbolic × Computer	-0.16 (0.11)	-0.75 (0.97)	-0.50 (0.73)	-1.43 (3.61)	0.13 (6.23)	-0.13 (0.53)
Monetary × Computer	-0.09 (0.11)	-0.85 (0.90)	-0.49 (0.72)	-4.10 (3.57)	5.74 (6.16)	0.33 (0.53)
<i>F</i>	4.01**	11.05***	17.07***	2.34*	18.81***	15.01***
<i>R</i> ²	0.04	0.10	0.15	0.02	0.16	0.13
N(Subjects)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)

Table S3. OLS regression estimates including interaction term for key outcome measures. Coefficients are unstandardized OLS estimates. Standard errors are in parentheses and clustered by subject. “Symbolic” and “Monetary” Ecolabel are dummy variables (No Ecolabel = baseline). “Computer Rounds” = 1 for computerized game version. Interaction terms test for treatment heterogeneity. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Despite only marginal differences in first round withdrawal, the individual groups in round 1 in the computer version (including the 3 computerized participants) withdrew on average significantly more trees than groups (4 human participants) in the human rounds ($m_{comp}=51.91$, $SD=4.33$, $m_{human}=44.66$, $SD=9.73$, $t(243)=-13.27$, $p<0.01$). This explains part of the level differences found between the game versions.

4. Main results from human rounds (Fig. 2) replicate in computer rounds

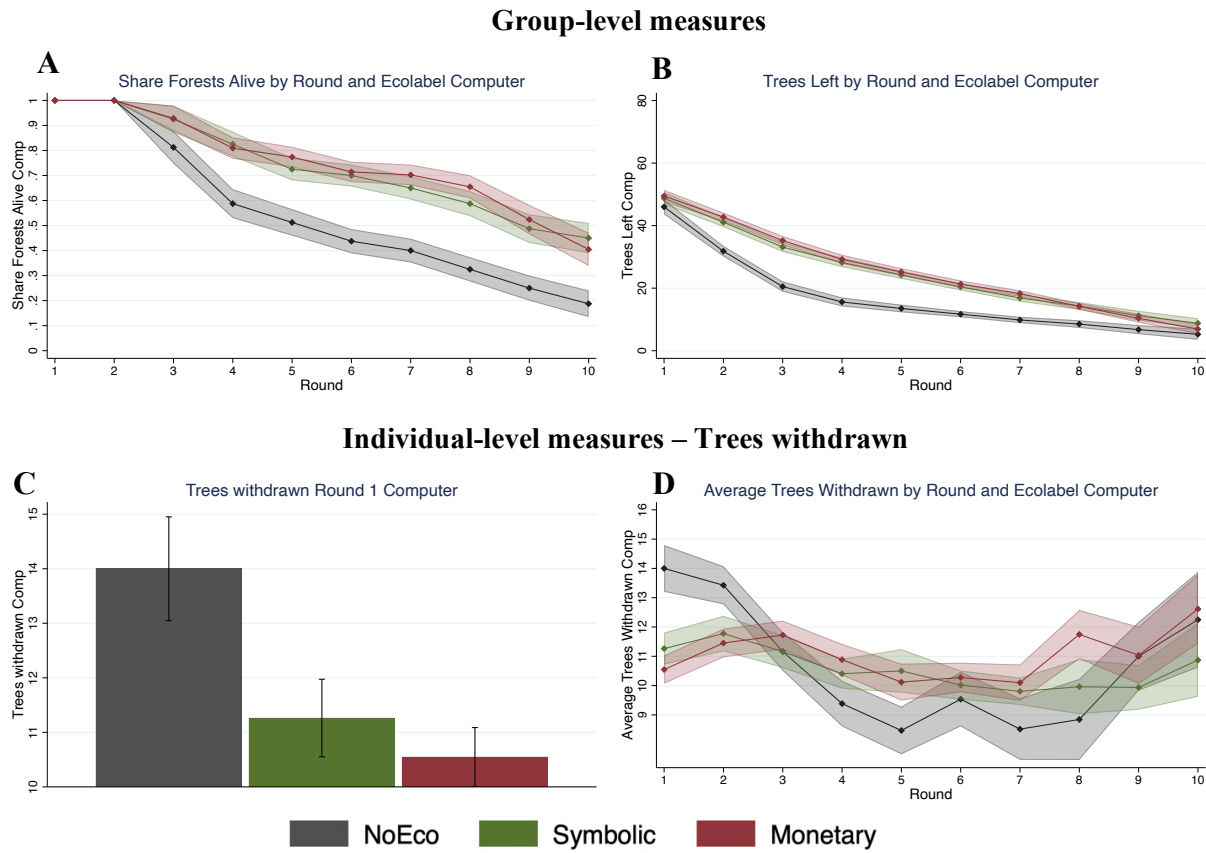


Fig. S3. Analyses of Computer Rounds following the approach reported for Fig. 2 in the main text. Top panels: Group-level measures. (A) Share of forests alive averaged across groups per round and ecolabel condition in computer rounds. (B) Number of trees left averaged across groups per round and ecolabel condition in computer rounds. Bottom panels: Individual-level measures. (C) Average trees withdrawn in computer round 1. (D) Average number of trees withdrawn per round and ecolabel condition in computer rounds. A, B and D use Cousineau-Morey within-subjects confidence intervals, shaded areas indicate 90% confidence intervals.

5. Benchmark adherence and sustainability standard across the game

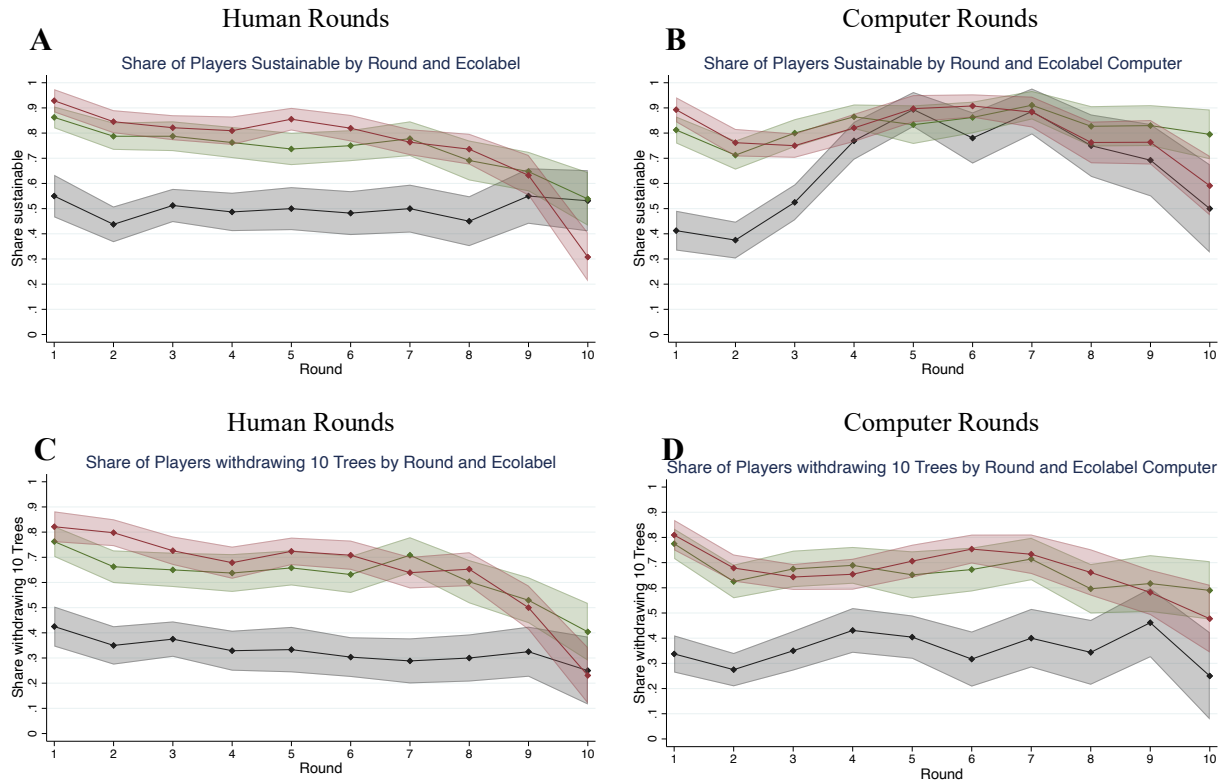


Fig. S4. Sustainable behavior in human and computer versions. A,B: Share of players behaving sustainably (withdrawing 10 trees or fewer) conditional on forest survival by round and ecolabel condition. (A) Human version. (B) Computer version. C,D: Share of players withdrawing exactly 10 Trees conditional on forest survival by round and ecolabel condition. (C) Human version. (D) Computer version. All graphs are for remaining players, i.e. players whose forest is depleted are excluded. All figures show Cousineau-Morey within-subjects 90% confidence intervals.

6. Coordinating effect of ecolabels and predictive power of early behavior replicate in computer rounds

6.1 Initial coordination in computer rounds

The coordinating effect of ecolabels observed in human rounds replicates in computer rounds. In the computer rounds, compared to the NoEco condition, participants in both ecolabel treatments withdraw significantly fewer trees in the first round (Fig. S3C), and a higher share of them adhere to the sustainable behavior of withdrawing 10 or fewer trees (Supplementary Results 5, Fig. S4) or exactly 10 trees (Fig. S4D, NoEco=33.75%, Symbolic=77.5% and Monetary=80.95%; Kruskal-Wallis Test: $\chi^2(2)_{ties}=48.57$, $p<0.001$; $MWU_{Symbolic \text{ vs } Monetary} p>0.3$, $MWU_{NoEco \text{ vs } Symbolic} p<0.001$, $MWU_{NoEco \text{ vs } Monetary} p<0.001$). The variances in the number of trees withdrawn in the first round are also significantly smaller in the two ecolabel treatments than in the NoEco condition (variance ratio test: Symbolic vs. NoEco: $F(79,79)=1.79$, $p_{two-sided}=0.01$; Monetary vs. NoEco: $F(79,83)=2.96$, $p_{two-sided}<0.001$).

6.2 Early withdrawal behavior predicts long-term outcomes

We test now whether round 1 behavior predicts behavior in later rounds, which is reflective of the effect of coordination to maintain similar behaviors across groups. Table S4 shows the regression results of game outcomes simply based on Round 1 behavior for both human and the computer version, while the main text refers to Table S5. In the human version, every additional tree withdrawn in Round 1 corresponds to surviving 0.17 rounds fewer (simple regression, $R^2=0.10$, $t=-3.90$; Table S4 column (1)). This translates to an increase in the odds of forest depletion by 12% for every additional tree withdrawn in Round 1 (logistic regression, $OR=1.12$ $R^2=0.036$, $z=2.70$; Table S4 column (2)). Participants that behave sustainably by cutting down 10 or fewer trees survive on average 8.86 rounds. These results replicate well in computer rounds: every additional tree withdrawn correlates with surviving 0.39 rounds fewer (OLS regression $R^2=0.38$, $t=-12.51$; Table S4 column (3)). This in turn translates to an increase in the odds of depletion by 53.2% for every additional tree withdrawn in Round 1 (logistic regression, $OR = 1.532$, $R^2=0.196$, $z=4.56$; Table S4 column (4)). Participants cutting down 10 trees survive on average 8.55 rounds in the computer rounds.

Table S4. Predictive Regression: Rounds Alive and Forest Depletion at end of the game on Round 1 withdrawal

	Human		Computer	
	(1) OLS Rounds Alive b/se	(2) Log Forest Depletes OR/se	(3) OLS Rounds Alive b/se	(4) Log Forest Depletes OR/se
Round 1 withdrawal	-0.17*** (0.04)	1.12** (0.04)	-0.39*** (0.03)	1.53*** (0.14)
Constant	10.39*** (0.39)	0.37** (0.15)	12.08*** (0.39)	0.02*** (.02)
R^2	0.1	0.04	0.38	0.20
Observations	244	244	244	244

Table S4. Predictive Regressions of number of rounds alive and likelihood of forest depletion on number of trees withdrawn in round 1. Coefficients are presented with standard errors in parentheses. OLS regressions estimate the number of rounds survived; logistic regressions estimate the odds of depletion, reported are odds ratios. These regressions correspond to any level of trees whereas in the main text we report the values for results above 10 trees withdrawn. Standard errors are clustered at the group level for human rounds. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table S5. Predictive Regression: Rounds Alive and Forest Depletion at end of the game on whether Round 1 withdrawal was above 10

	Human		Computer	
	(1) OLS Rounds Alive b/se	(2) Log Forest Depletes OR/se	(3) OLS Rounds Alive b/se	(4) Log Forest Depletes OR/se
Round 1 withdrawal above 10	-0.196*** (0.04)	1.11* (0.06)	-0.44*** (0.04)	1.61*** (0.20)
Constant	10.8*** (0.47)	0.4† (0.20)	12.78*** (0.46)	0.01*** (0.01)
R^2	0.10	0.02	0.38	0.17
Observations	244	244	244	244

Table S5. Predictive Regressions of number of rounds alive and likelihood of forest depletion on number of trees withdrawn above 10 in round 1. Coefficients are presented with standard errors in parentheses. OLS regressions estimate the number of rounds survived; logistic regressions estimate the odds of depletion, reported are odds ratios. These regressions correspond to level of trees withdrawn above 10 trees withdrawn i.e. 10 trees as baseline, with everyone withdrawing up to and including 10 trees coded as baseline. Standard errors are clustered at the group level for human rounds. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table S6. Predictive Regression: Rounds Alive and Forest Depletion at end of the game on the number of sustainable group members in Round 1

	Human	
	(1) OLS Rounds Alive b/se	(2) Log Forest Depletes OR/se
Number of sustainable group members in Round 1	1.06*** (0.28)	0.52† (0.20)
Constant	5.18*** (1.04)	11.25† (14.99)
R^2	0.24	0.06
Observations	244	244
Clusters/Groups	61	61

Table S6. Predictive Regressions of number of rounds alive and likelihood of forest depletion on number of sustainable group members in round 1. Coefficients are presented with standard errors in parentheses. OLS regressions estimate the number of rounds survived; logistic regressions estimate the odds of depletion, reported are odd ratios. Standard errors are clustered at the group level. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

7. Conditionality

7.1. Conditionality on number of sustainable group members

Table S7 shows the estimates of round-specific logistic regressions predicting whether participants behave sustainably, based on the number of group members who behaved sustainably in the previous round. The model includes a lagged peer behavior term and the analysis in the table is conducted for each round separately. Round 1 is excluded as there cannot be any lagged information. Sustainable behavior is a binary outcome (with 1 indicating 10 or fewer trees withdrawn). Peer sustainability refers to the number of other group members (0–3) who were sustainable in the previous round. We additionally estimated a mixed-effects logistic regression pooling all rounds (Rounds 2–10), with random intercepts for participants to account for repeated decisions and robust standard errors clustered at the individual level and an additional time trend as predictor (Round). Results indicate that each additional sustainable group member in the previous round increases the log-odds of behaving sustainably in the current round by 0.31 (SE = 0.12, $z = 2.61$, $p = .009$), corresponding to a 37% increase in odds (OR = 1.37, 95%-CI[1.08, 1.73]).

Table S7 Dependent Variable: Sustainable behavior in current round

Round	Odds Ratio	Std. Err.	$z_{uncorrected}$	$p_{uncorrected}$	$p_{corrected}$	N
1	-	-	-	-	-	-
2	1.720	0.293	3.18	0.001	0.005	244
3	1.678	0.278	3.13	0.002	0.005	244
4	1.693	0.265	3.37	0.001	0.005	240
5	1.581	0.293	2.47	0.014	0.025	212
6	1.259	0.234	1.24	0.216	0.278	204
7	0.982	0.185	-0.10	0.924	0.963	196
8	1.009	0.195	0.05	0.963	0.963	180
9	0.782	0.138	-1.40	0.162	0.243	176
10	0.568	0.130	-2.47	0.014	0.025	136

Table S7. Effect of lagged sustainable peer behavior on individual sustainability across rounds.

Logistic regression estimating the effect of the number of sustainable peers in the previous round on the likelihood of sustainable behavior by the individual in each round of the game. Reported are the odds ratios, standard errors, z-values, uncorrected and Benjamini–Hochberg–corrected p-values. Round 1 is omitted as there is no lagged peer behavior. Sample size (N) indicates the number of individuals included in each round. Significant Odds Ratios are bold.

Table S8 shows the estimates of round-specific logistic regressions predicting whether participants who have behaved sustainably in all previous rounds are more likely to continue doing so when more group members also behave sustainably. The analysis is conducted

separately for each round (Rounds 2–10); Round 1 is excluded as it lacks lag information. Sustainable behavior is a binary outcome (where 1 means that 10 or fewer trees withdrawn).

To assess the overall effect across the game, we additionally estimated a pooled logistic regression including all rounds (Rounds 2–10) with robust standard errors clustered at the individual level and a time-trend predictor (Round). A mixed-effects specification was not used here because the model failed to converge (non-concave likelihood surface), likely due to the reduced within-subject variance. The simpler clustered logit thus provides a stable and conservative estimate of the pooled relationship. Results show that each additional sustainable group member in the previous round increases the log-odds of continuing sustainable behavior by 0.67 (SE = 0.12, $z = 5.84$, $p < .001$), corresponding to a 96% increase in odds (OR = 1.96, 95%-CI[1.54, 2.47]). The negative coefficient for Round ($b = -0.09$, $p = .001$) indicates that even consistently sustainable players are somewhat less likely to maintain sustainable behavior as the game progresses.

Table S8 Dependent Variable: Continue sustainable behavior in current round

Round	Odds Ratio	Std. Err.	$z_{uncorrected}$	$p_{uncorrected}$	$p_{corrected}$	N
1	-	-	-	-	-	-
2	2.262	0.417	4.42	0.000	0.000	244
3	2.184	0.345	4.94	0.000	0.000	244
4	2.635	0.437	5.84	0.000	0.000	240
5	2.415	0.394	5.40	0.000	0.000	212
6	1.931	0.295	4.31	0.000	0.000	204
7	1.778	0.280	3.65	0.000	0.000	196
8	1.801	0.294	3.60	0.000	0.000	180
9	1.517	0.235	2.69	0.007	0.008	176
10	1.125	0.205	0.65	0.517	0.517	136

Table S8. Effect of continued sustainable peer behavior on individual's continuation of sustainable behavior across rounds. Logistic regression estimating the effect of an additional group member's sustainable behavior in all the previous rounds on the likelihood of continued sustainable behavior in the current round. Reported are the odds ratios, standard errors, z-values, uncorrected and Benjamini–Hochberg–corrected p-values. Round 1 is omitted due to missing lagged peer behavior. Sample size (N) indicates the number of individuals included in each round. Significant Odds Ratios are bold.

Fig. S5 illustrates how sustainable behavior at the beginning of the game depends on both peer behavior and the forest's ecological state. Panel A shows that the share of sustainable players in Round 2 increases with the number of sustainable group members in Round 1, indicating positive conditional cooperation from the beginning of the game. Panel B (also Figure 4C in

main text) separates this relationship by players' initial behavior and reveals the heterogeneity underlying aggregate dynamics: participants who behaved sustainably in Round 1 become substantially more likely to maintain sustainable behavior when surrounded by other sustainable peers, whereas initially unsustainable players tend to remain unsustainable regardless of group composition or rather become even more unsustainable. Fig. S5C shows that sustainability in Round 2 also depends on the forest size, players in groups with more trees left after Round 1 (i.e., healthier forests) are more likely to behave sustainably. Together these results provide additional evidence for the condition cooperation discussed in the main text.

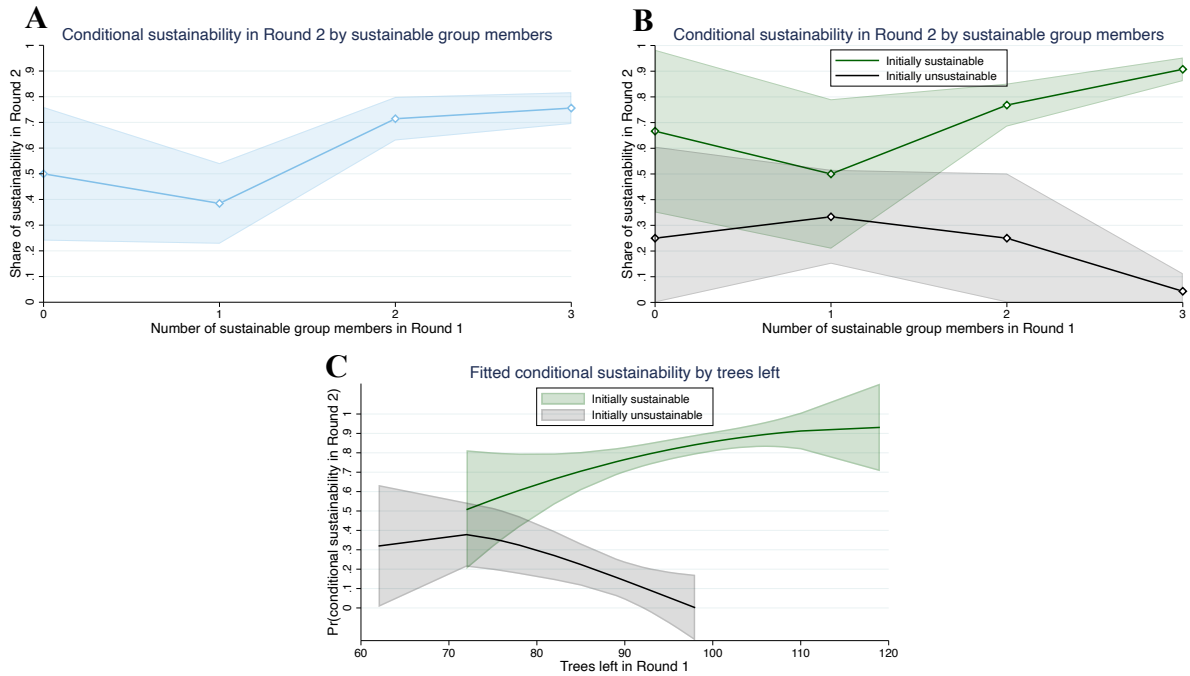


Fig. S5. Conditional sustainability by group behavior and forest state. (A) Share of sustainable players in Round 2 conditional on the number of sustainable group members in Round 1 (aggregated across all conditions). (B) Same relation as in (A), separated for players who were initially sustainable vs. initially unsustainable in Round 1. (C) Fitted probability of sustainability conditional on the number of trees left in Round 1, again separated for players who were initially sustainable vs. initially unsustainable. Shaded areas indicate 90% confidence intervals.

8. Types and conditionality

8.1. General Conditionality and distributions of Types

In this section, we classify players into behavioral types and inspect how the distribution of such types relates to outcomes in the game. Following previous research^{3,4} and our preregistration, we consider three types: Unconditional Cooperators (UC) are players who maintain sustainable behavior (withdraw ≤ 10 trees) throughout the entire experiment⁵. Unconditional Defectors (UD) are those withdrawing more than 10 trees in each round. The remaining players, that vary their behavior across the game are classified as Conditional Cooperators (CC).

Across all conditions, 24.18% of players are classified as UC, 6.56% as UD, and 69.26% as CC. The distribution of types in each treatment is shown in Fig. S6. The distribution differs across conditions. In the *NoEco* condition, fewer players are classified as UC ($m_{NoEco}=11.25\%$ vs $m_{Symbolic}=37.5\%$ & $m_{Monetary}=23.81\%$; $\chi^2(2) = 15.04$, $p = 0.001$) and more as UD ($\chi^2(2) = 10.13$, $p=0.006$), compared to the ecolabel conditions. The share of conditional cooperators differs marginally across conditions ($75.0\% / 60.0\% / 72.6\%$; $\chi^2(2) = 4.91$, $p = .086$).

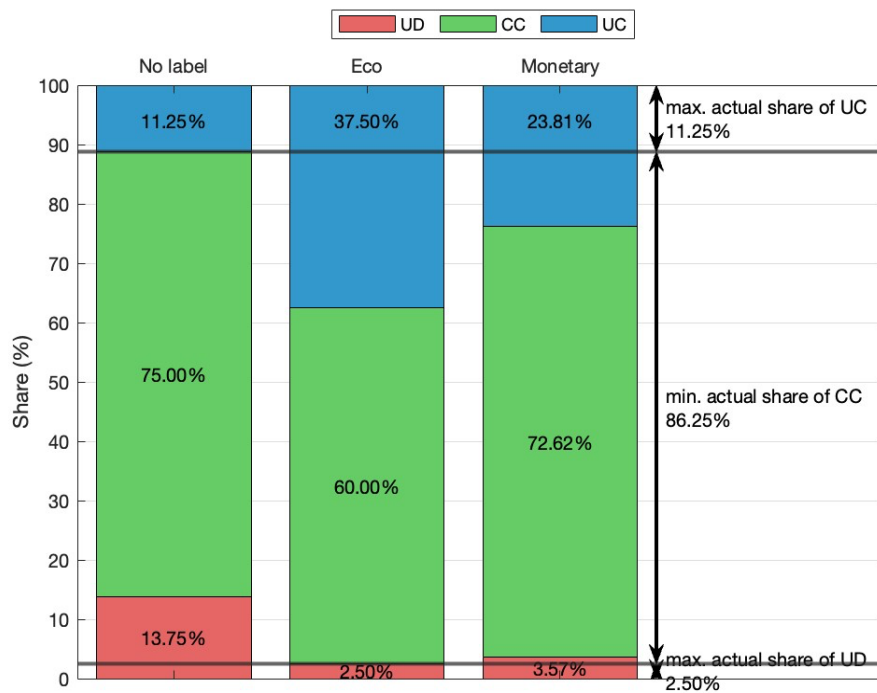


Fig. S6. Distribution of behavioral types observed in the game by ecolabel condition. Each bar represents all players in one condition. The bars are divided into the shares of Unconditional Cooperators (UC), Conditional Cooperators (CC) and Unconditional Defectors (UD).

Within and across groups, the distribution of types relates to forest survival and other outcome measures. Regressing group compositions on outcomes, shows that groups with at least one UD type player survive an average of 2.9 rounds fewer (mean rounds alive= 6.29, $t=4.16$, $p<0.01$) compared to groups without. On the other hand, groups with at least one UC type player survive longer by 1.72 rounds (mean rounds alive= 9.128, $t=2.63$, $p<0.01$) than groups without.

Table S9: Mean number of rounds alive per group composition (types) and ecolabel condition

Mean Rounds Alive (<i>Std. Error</i>) N	Number in Group	No Eco (3)	Symbolic (4)	Monetary (5)	Average (6)	Average when ≥ 1 in group
Unconditional Cooperators (UC)	0	7 (.91) 44	8 (2) 12	7.75 (1.01) 32	7.41 (.62) 88	9.128 (0.23)
	1	8 (.71) 36	9 (0.63) 24	9.57 (.20) 28	8.77 (.36) 88	
	2	-	9.78 (0.15) 36	9.6 (.) 20	9.71 (.16) 56	
	3	-	8.5 (1.5) 8	10 (-) 4	9 (1) 12	
Conditional Cooperators (CC)	0	-	7 (0) 4	-	7 (-) 4	8.533 (0.29)
	1	4 (-) 4	9.5 (.5) 8	10 (-) 4	8.25 (1.436) 16	
	2	6.75 (1.25) 16	9.88 (.13) 32	9.5 (.34) 24	9.06 (.41) 72	
	3	7.67 (0.88) 36	9 (.63) 24	8.25 (.94) 32	8.22 (.49) 92	
	4	8.167 (1.17) 24	8 (2) 12	9 (.82) 24	8.47 (.65) 60	
Unconditional Defectors (UD)	0	8.545 (.67) 44	9.278 (.39) 72	9.44 (.29) 72	9.17 (.24) 188	6.285 (0.67)
	1	6.7 (1.02) 28	8 (1) 8	5.67 (1.67) 12	6.67 (.72) 48	
	2	4 (-) 8	-	-	4 (-) 8	

Table S9. Mean number of rounds alive per group composition and ecolabel condition. The left-most column shows the three behavioral types. Columns 3-6 show the number of rounds alive averaged across all groups that contain the specified number of players of the corresponding type, by ecolabel condition. The last column shows the average conditional on at least one player of that type (UC, CC or UD) is present in a group. Each cell contains the mean number of rounds alive in bold, the standard error in parentheses and italic, and the number of observations, i.e. the number of participants in a group with the corresponding group composition. For example, in the first row of the NoEco condition, 44 participants (11 groups, each with 4 participants) are part of a group with no Unconditional Cooperators (UC), while this holds for 12 participants for Symbolic and 32 participants in Monetary. Errors in the table are unclustered, while the regressions above use clustered errors at group level. N indicated as individual subjects.

A challenge in accurately identifying the players who cooperate conditionally in our experiment is that their observed behavior may be indistinguishable from that of unconditional cooperators or defectors: the players might constantly cooperate or defect as they conditionally respond to others' constant cooperation or defection. To address the potential underestimation of conditional cooperators' prevalence due to this issue, we reason that the lowest observed shares of UC and UD across the three treatments are the upper bounds for the actual shares of unconditional cooperators and defectors. Fig. S6 shows that these actual shares are quite small, and at least 86.25% of participants are conditional cooperators. We conclude that most participants are conditional cooperators, but ecolabels may affect their behavioral expression.

8.2. Conditionality by Label

Section 8.1 showed that the shares of unconditional cooperators and unconditional defectors differ across conditions. Table S10 reports logistic regressions of sustainable behavior in the current round on the number of sustainable group members in the previous round, using human rounds 2–10 and dropping rounds after the forest died. Across conditions, the association is positive (Table S10 Model 1, Fig. 4, OR=1.28, 95%-CI[1.08, 1.51], $p=0.004$), yet this association is not uniform and shows a significant negative effect in the NoEco condition. Table S10 Model 2 shows that the interaction with the label conditions drives the overall positive associations, and that only in the label conditions participants exhibit clear conditional cooperation rather than free riding.

The contrast is clearest among participants who were already sustainable in round 1, for whom conditional cooperation rather than free riding is the relevant response (Model 3; Figure 4C). The slope on peer sustainability is negative under NoEco (OR=0.71, 95%-CI[0.52, 0.97], $p=0.029$) and significantly more positive under both ecolabels (Symbolic, OR=1.83, 95%-CI[1.14, 2.93], $p=0.012$; Monetary OR ratio = 2.11, 95%-CI[1.29, 3.47], $p=0.003$; joint $\chi^2(2)=10.73$, $p=0.005$). Restricting further to the early rounds, where labels are most often present (Model 4, rounds 2–4), the within-condition slopes are significant in both ecolabel conditions (Symbolic, OR=2.14, 95%-CI[1.14, 4.00], $p=0.018$; Monetary, OR=2.57, 95%-CI[1.36, 4.86], $p=0.004$) while NoEco remains flat (OR=0.88, 95%-CI[0.56, 1.38], $p=0.577$; joint $\chi^2(2)=9.36$, $p=0.009$).

Table S10. Conditional sustainability by ecolabel condition

Predictor	(1) All Rounds 2–10	(2) All Rounds 2–10	(3) Sust. in R1 Rounds 2–10	(4) Sust. in R1 Rounds 2–4	(5) All Rounds 2–10	(6) All Only Round 2
Sustainable group members (t–1)	1.277** (0.110)	0.779* (0.090)	0.707* (0.113)	0.881 (0.200)	0.684* (0.107)	0.463* (0.154)
Symbolic (vs. NoEco)		2.428† (1.114)	0.981 (0.508)	0.733 (0.600)		
Monetary (vs. NoEco)		0.902 (0.431)	0.576 (0.313)	0.374 (0.286)		
Symbolic × Sust. group members		1.155 (0.243)	1.832* (0.441)	2.425* (0.952)		
Monetary × Sust. group members		1.829** (0.389)	2.111** (0.534)	2.914** (1.156)		
First Round Sustainable					1.656 (0.647)	0.944 (0.836)
First R. Sust. × Sust. group members					1.962*** (0.366)	4.706*** (1.952)
Constant	1.207 (0.231)	1.412 (0.320)	2.330** (0.725)	1.877 (0.763)	0.944 (0.298)	0.857 (0.556)
χ^2		8.11*	10.73**	9.36**	87.31***	64.71***
N (Subjects)	1,832 (244)	1,832 (244)	1,501 (191)	570 (191)	1,832 (244)	244(244)

Table S10. Conditional Sustainability by EcoLabel Condition. Logistic regressions predicting sustainable behavior in the current round (1 = 10 or fewer trees withdrawn) on the number of other group members who behaved sustainably in the previous round (0–3). Entries are odds ratios with standard errors in parentheses, clustered by subject. Human rounds only; round 1 is excluded because it has no lag, and rounds after the forest died are excluded. Columns (3) and (4) restrict the sample to participants who behaved sustainably in round 1. (5) and (6) show the interaction between number of sustainable group members and initial sustainable behavior. NoEco is the reference condition. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

9. Correlations with Environmental Attitudes

9.1 Environmental attitude interaction and behavior

Table S11. Environmental attitude regressions with ecolabel condition interactions on forest outcomes.

Predictor	Rounds Alive		Round 1 Withdrawn		Trees Left		Total Trees Withdrawn	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Env. Attitude	0.50** (0.18)	-0.03 (0.25)	-0.77* (0.37)	-0.55 (0.62)	2.30** (0.85)	1.99 (1.28)	4.38* (1.84)	-0.44 (2.61)
Comp. Rounds	-1.05*** (0.18)	-1.05*** (0.18)	0.75*** (0.21)	0.75*** (0.21)	-2.83* (1.25)	-2.83* (1.26)	-18.09*** (2.52)	-18.09*** (2.53)
Sym (vs. NoEco)		1.57*** (0.30)		-2.20** (4.10)		3.61† (2.08)		16.61*** (3.20)
Mon (vs. NoEco)		1.61*** (0.33)		-2.95*** (0.63)		3.40 (2.12)		17.43*** (3.59)
Symbolic x Attitude		1.21** (0.39)		-0.52 (0.91)		2.52 (2.27)		9.76** (3.77)
Monetary x Attitude		0.16 (0.45)		0.72 (0.81)		-2.41 (2.19)		1.64 (4.63)
Constant	8.51*** (0.14)	7.41*** (0.26)	11.16*** (0.26)	12.88*** (0.55)	9.80*** (1.14)	7.49*** (4.89)	90.67*** (2.13)	78.96*** (3.38)
<i>F</i>	21.11	16.84	8.25	6.96	6.52	3.69	30.83	22.31
<i>R</i> ²	0.063	0.162	0.025	0.113	0.018	0.035	0.100	0.174
N(Subjects)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)

Table S11. Environmental attitude regressions with ecolabel condition interactions on forest outcomes.

Separate OLS regressions predicting total trees withdrawn, rounds alive, trees withdrawn in Round 1, and trees left alive at Round 10. Results replicate for Rounds Sustainable and for depletion probability in computer rounds. For each outcome variable we (1) show the general association with environmental attitude and then (2) the interaction effects across ecolabels. For (2) Environmental attitude and ecolabel condition (No Eco, Symbolic, Monetary) were included as main effects and interactions. To increase power for interactions (see^{6,7}), the combined dataset was used that includes human and computer rounds; a factor for game type was included in regressions. Environmental attitudes were mean centered. Robust standard errors clustered at the participant level are shown in parentheses. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

9.2 Environmental attitude and Rounds Alive

Tables S12 and S13 are further exploring the main results reported in Table S11. Table S12 explores the robustness of the main effect of environmental attitude. We explore associations split by game version for (1) individual environmental attitude and (2) group average environmental attitude. Results hold for other measures discussed above (Table S11). Table S13 reflects this follow-up, split by ecolabel conditions, with results staying only significant in the symbolic ecolabel condition.

Table S12a Number of rounds alive on environmental attitudes

	(1) OLS Human Rounds Alive b/se	(2) OLS Human Rounds Alive b/se	(3) OLS Computer Rounds Alive b/se
Individual Environmental Attitude	0.44* (0.21)		0.55* (0.24)
Group avg. environmental Attitude		1.76* (0.84)	
Constant	8.51*** (0.28)	8.51*** (0.27)	7.46*** (0.17)
R^2	0.02	0.08	0.02
Observations	244	244	244

Table S12. Regressions of number of rounds alive on environmental attitudes. Coefficients are presented with standard errors in parentheses. OLS regressions estimate the number of rounds alive. (1) number of rounds alive in human rounds on individual environmental attitudes. (2) number of rounds alive in human rounds on group average environmental attitudes. (3) number of rounds alive on individual environmental attitudes in computer rounds. For (1) & (2) errors are clustered at the group level. All environmental attitudes are mean centered. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table S12b Number of Trees Left Alive after Round 10 on environmental attitudes

	(1) OLS Human Trees Left Alive b/se	(2) OLS Human Trees Left Alive b/se	(3) OLS Computer Trees Left Alive b/se
Individual Environmental Attitude	2.37* (1.16)		2.23 † (1.25)
Group avg. environmental Attitude		9.40* (4.65)	
Constant	7.46 (0.17)	9.80*** (2.29)	6.98*** (0.91)
R^2	0.01	0.04	0.01
Observations	244	244	244

Table S12. Regressions of number of Trees Left Alive after Round 10 on environmental attitudes. Coefficients are presented with standard errors in parentheses. OLS regressions estimate the number of trees left alive. (1) number of trees left alive in human rounds on individual environmental attitudes. (2) number of trees left alive in human rounds on group average environmental attitudes. (3) number of trees left alive on individual environmental attitudes in computer rounds. For (1) & (2) errors are clustered at the group level. All environmental attitudes are mean centered. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table S13 Number of rounds alive on environmental attitudes by Ecolabel

Ecolabel Condition	β (Individual Attitudes on Human Rounds Alive)	SE	p-value	N
No Ecolabel	0.07	0.29	0.804	80
Symbolic	0.90	0.40	0.037	80
Monetary	0.07	0.31	0.832	84

Ecolabel Condition	β (Group Attitudes on Human Rounds Alive)	SE	p-value	N
No Ecolabel	0.36	1.42	0.803	80
Symbolic	3.99	1.20	<0.01	80
Monetary	0.25	1.20	0.838	84

Ecolabel Condition	β (Individual Attitudes on Computer Rounds Alive)	SE	p-value	N
No Ecolabel	-0.13	0.37	0.73	80
Symbolic	1.47	0.43	<0.01	80
Monetary	0.19	0.42	0.66	84

Table S13. Regression predicting number of rounds the forest remained alive (out of 10). First block, based on individual environmental attitudes in human rounds; second block, based group-average environmental attitudes in human rounds; third block, based on individual environmental attitudes in computer rounds. Results are presented separately by ecolabel condition. Only the Symbolic ecolabel condition shows a significant positive association. Standard errors are clustered at the group level for the first two panels (human rounds).

10. Eye-Tracking results

To examine how ecolabels influence attention, information processing, and environmental decisions, we use eye-tracking for 100 participants while they made decisions throughout the experiment. We recorded eye gaze using Tobii EyeX eye-trackers (60 Hz frequency, with 1920x1080 pixel resolution 22" monitors and chinrests at 58 cm distance, see⁸ for equivalent methods). This additional dataset enables us to investigate whether the different ecolabel conditions shift the attentional focus and/or information search patterns during the decision stage of our experiment (see Figure S8). Particularly, we were interested whether participants acquired information differentially in the Symbolic compared to Monetary ecolabel conditions, as these could be indicative of the underlying motivational forces that drive behavior decisions in each ecolabel condition.

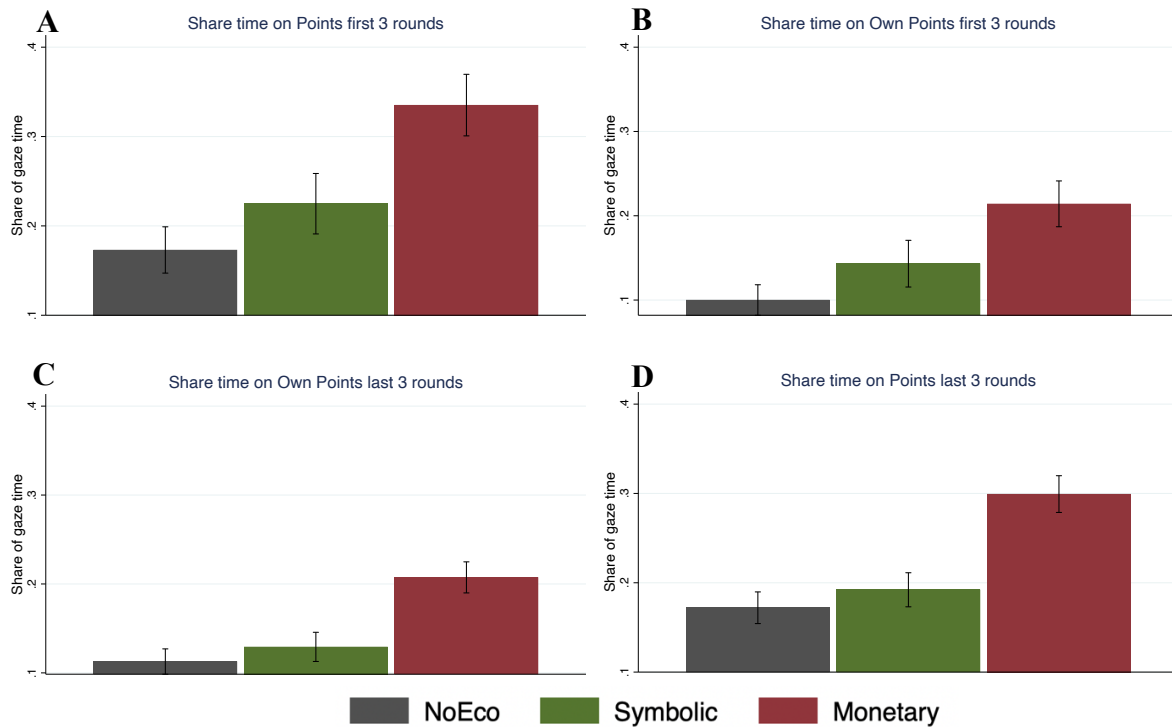


Fig. S7. Eye-tracking results for the allocation of attention by ecolabel condition as share of time focused on Points Earned. Panels A–B show the share of gaze time in *the first 3 rounds* directed to (A) Groups Total Points Earned (B) Own Total Points Earned. Panels C–D show the share of gaze time in *the last 3 rounds for each group (i.e. the last 3 rounds before they depleted the forest)* directed to (C) Groups Total Points Earned (D) Own Total Points Earned; in the eye-tracked sample (N=100; Tobii EyeX, 60 Hz). Bars reflect means across 3 rounds within the first three rounds, or the final 3 rounds before forest depletion (group-dependent) of the human rounds; error bars indicate 90% CIs.

Figure S7 shows results from the analyses of eye-tracking data, which includes all groups, i.e. before any of the groups deplete the forest (first 3 rounds) and the last 3 rounds for each group. We find that participants in the Monetary condition allocated a larger share of their attention

to payoffs compared to the Symbolic condition (points earned: $\beta_{\text{Round1-3}}=0.110$, $p_{\text{Round1-3}}=0.005$; $\beta_{\text{Last3Rounds}}=0.114$, $p_{\text{Last3Rounds}}=0.007$), especially to own earnings (own points: $\beta_{\text{Round1-3}}=0.070$, $p_{\text{Round1-3}}=0.015$; $\beta_{\text{Last3Rounds}}=0.065$, $p_{\text{Last3Rounds}}=0.113$). Differences between NoEco and Symbolic were small and not statistically significant (see Table S14). We regressed the share of fixation time on each ROI on ecolabel condition using OLS with standard errors clustered by participant. Results are robust to mixed-effects models with participant random intercepts.

Table S14. Comparison of OLS regressions for fixation-time outcomes first 3 rounds and last 3 rounds for each group.

Predictor	Share Time on Points		Share Time on Own Points	
	First 3 Rounds	Last 3 Rounds	First 3 Rounds	Last 3 Rounds
No-Ecolabel	-0.052 (0.033)	0.005 (0.033)	-0.044* (0.022)	0.004 (0.033)
Monetary	0.110** (0.038)	0.114** (0.041)	0.070* (0.028)	0.065 (0.041)
Computer Rounds	-0.110*** (0.019)	-0.069** (0.022)	-0.060*** (0.017)	-0.026 (0.019)
Constant	0.280*** (0.028)	0.187*** (0.026)	0.173*** (0.018)	0.134*** (0.027)
R^2	0.099	0.052	0.076	0.021
N (Clusters)	594 (99)	594 (99)	534 (99)	517 (99)

Table S14. Comparison of OLS regressions for fixation-time on points / profit ROIs first 3 rounds and last 3 rounds for each group. OLS predicting the share of fixation time spent on different regions of interest (ROIs) by ecolabel condition. Symbolic ecolabel is the baseline. As preregistered, one participant was excluded due to insufficient data quality. Standard errors clustered by participant (N = 99). † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Fig. S8 shows a replica of the participant decision screen. Information order was varied to create four different treatments in which information was positioned: Information about the personal or group account was presented on the left or right side of the screen. Information about the environmental impact (trees withdrawn) or Monetary payoff (points earned) was presented at the top or bottom. This manipulation allowed us to test whether making information more salient through screen positioning (left vs. right, top vs. bottom) affected decisions and controlled for natural order effects.

<u>You</u>	<u>Round 2</u>	<u>Group (You + 3 others)</u>
10	Withdrawn Trees	50
12.5	Points	57.5
yes	Ecolabel	3
Trees left in Forest		90
How many trees do you want to withdraw		
You can withdraw between 0 and 20 trees		
Continue		

Fig. S8. Illustration of Decision Screen and Eye-tracking results. The participant decision screen, illustrating the type of information that participants could acquire before making a decision. Note that the location of the personal or group earnings, environmental impact, and payouts were varied across four conditions to counterbalance acquisition order (which typically starts in the top left) and potential saliency effects. Further, the NoEco did not include the Ecolabel row.

11. Location check

To assess whether pooling participants across locations affected the treatment effects, we tested condition-by-site interactions for the main dependent variables first-round withdrawals and forest survival separately for the human and computerized rounds. We found no evidence that the effect of ecolabel condition differed between sites for the human rounds for first round withdrawal, ($F(2,243)=0.08$, $p=0.924$), or the computerized rounds, ($F(2,238)=0.08$, $p=0.924$). Similarly, condition-by-site interactions were not significant for the number of rounds alive in either the human rounds, ($F(2,55)=0.94$, $p=0.396$), or the computerized rounds, ($F(2,238)=0.73$, $p=0.485$). The descriptive treatment patterns were also similar across sites, with one exception in rounds alive, providing no clear evidence that pooling across locations altered the estimated effects of the ecolabel treatments.

Table S15. Location interaction

Predictor	Rounds Alive		Trees Withdrawn Round 1		Trees Left		Total Trees Withdrawn	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Site (Konstanz)	0.379 (0.262)	-0.385 (0.456)	-0.701 (0.493)	-0.802 (0.988)	1.670 (1.623)	-0.245 (2.733)	3.372 (2.702)	-3.010 (4.939)
Comp. Rounds	-1.053*** (0.181)	-1.053*** (0.182)	0.750*** (0.206)	0.750*** (0.206)	-2.828* (1.254)	-2.828* (1.260)	-18.094*** (2.515)	-18.094*** (2.525)
Sym (vs. NoEco)		1.177** (0.427)		-2.521** (0.943)		4.583 (2.895)		12.740** (4.509)
Mon (vs. NoEco)		1.177** (0.435)		-3.031*** (0.821)		1.063 (2.732)		14.396** (4.823)
Symbolic x Konstanz		1.151* (0.576)		0.396 (1.254)		-0.865 (3.868)		10.932† (5.808)
Monetary x Konstanz		1.056† (0.612)		0.059 (1.167)		6.179 (3.898)		7.417 (6.585)
Constant	8.353*** (0.186)	7.568*** (0.336)	11.451*** (0.362)	13.302*** (0.743)	9.119*** (1.441)	7.237*** (1.891)	89.290*** (2.523)	80.245*** (4.337)
<i>Site x Label, F(2,243)</i>		2.23		0.07		1.93		1.78
<i>F</i>	18.11	16.31	7.52	6.36	3.40	2.49	27.24	22.96
<i>R</i> ²	0.0483	0.1481	0.0143	0.1069	0.0101	0.0325	0.0923	0.1692
N (Subjects)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)	488 (244)

Table S15. Location interaction. OLS regressions of each outcome on experimental site, with and without site $\times$ ecolabel interactions. Amsterdam is the reference site and NoEco the reference condition. Human and computerized rounds are pooled with a Comp. Rounds indicator, giving two observations per participant; standard errors in parentheses are clustered on the participant. Rounds Alive, Trees Left and Total Trees Withdrawn are measured at the end of the game (Round 10); Trees Withdrawn is the Round 1 withdrawal. Column (1) reports the site main effect only, column (2) adds the ecolabel main effects and the site $\times$ ecolabel interactions. F-statistics test the joint significance of all predictors. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

References in Supplementary Results

1. Royston, P. & Parmar, M. K. B. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Stat. Med.* **30**, 2409–2421 (2011).
2. Uno, H. *et al.* Moving Beyond the Hazard Ratio in Quantifying the Between-Group Difference in Survival Analysis. *J. Clin. Oncol.* **32**, 2380–2385 (2014).
3. Romero, J. & Rosokha, Y. Constructing strategies in the indefinitely repeated prisoner’s dilemma game. *Eur. Econ. Rev.* **104**, 185–219 (2018).
4. van Huyck, J., Battalio, R. & Beil, R. Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. *Am. Econ. Rev.* **80**, 234–48 (1990).
5. Axelrod, R. An Evolutionary Approach to Norms. *Am. Polit. Sci. Rev.* **80**, 1095–1111 (1986).
6. Leon, A. C. & Heo, M. Sample Sizes Required to Detect Interactions between Two Binary Fixed-Effects in a Mixed-Effects Linear Regression Model. *Comput. Stat. Data Anal.* **53**, 603–608 (2009).
7. McClelland, G. H. & Judd, C. M. Statistical difficulties of detecting interactions and moderator effects. *Psychol. Bull.* **114**, 376–390 (1993).
8. Fischbacher, U., Grammling, D., Hausfeld, J. & Zíka, V. Identity breeds inequality: Evidence from a laboratory experiment on redistribution. *J. Public Econ.* **222**, 104866 (2023).