# Altruism

Let’s start by recognizing a simple (and wonderful) fact…

People give.

~2% of GDP are donated to charity

2-4% of hours worked are volunteered

And they give in all sorts of other ways, too…

~50% vote in national elections

How many of you recycle?

Would help an old lady cross street?

Give a stranger directions

Pay more for sustainable agriculture?

Etc.

Giving helps others but costs the giver

Why do we do it?

Perhaps we want to look good to our friends?

Perhaps we want tax breaks?

Perhaps…

Don’t be so cynical!  At least some of us actually care…

But it turns out our caring has some pretty puzzling features

First puzzle…

An estimated 3.14–3.59 million have HIV/AIDS worldwide

.17 million die annually

Source: World Health Organization

How does this make you feel?

Let’s look at those statistics again…

An estimated 31.4–35.9 million have HIV/AIDS worldwide

1.7 million die annually

Source: World Health Organization

Does that make a difference in how you feel?

How come we genuinely care about curing HIV/AIDS but are insensitive to the magnitude of the problem?

# Another puzzle…

Consider the following canonical experiment

Subjects randomly paired and one subject is randomly chosen to move first

That subject is given $20 and asked whether she would like to give half ($10) or just $2 to her partner

76% give half

This is called *The Dictator Game*

Result replicates and generalizes… has been run all over the world, with high stakes, etc.

See: C. Camerer, Behavioral Game Theory (2003) pg. 56 on

Why does this happen?

For some of us, we feel guilty or are afraid we will be seen or…

But some of us *actually* care

But for those of us who care, this next result is kinda weird…

Subjects randomly paired and one subject is randomly chosen to move first

That subject is given $10 and asked how much she’d like to allocate to her partner

Then, subject is asked if she’d like to pay $1 to “exit” before her partner is informed of the game’s rules

28% exited

This is weird…

They could have gotten $10, given $0 and been better off

It's even weirder when we see that…

# Most of them had planned on sharing

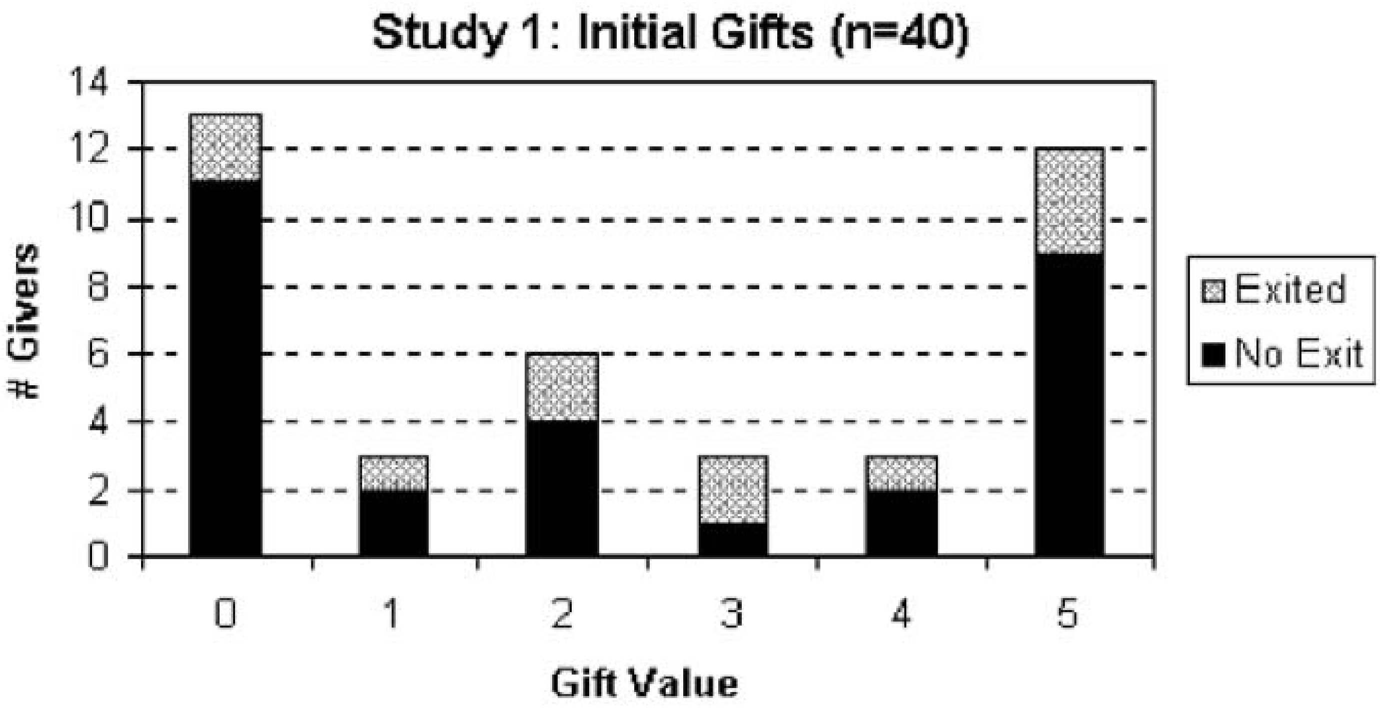



Fig. 1. Histogram of initial gifts in study 1.

If these people care enough to first promise to share, why would they choose to exit?

They care, but weirdly

What if you played with your grandfather

You may or may not share the $10.  But I doubt you would exit

You care, but differently from how you care about your grandfather.  Why?

Another puzzle…

Let’s look at voting…

Image removed due to copyright restrictions.

View the 2008 General Election Turnout Rates by County.

All over the country people vote.  Why?

Some of us vote because:

Our friends would yell at us otherwise
It makes us feel like we were part of something
We like getting free stickers

And some of us just care about the outcome of the election

But for those of us who care about the outcome, let's look at the chances of effecting it…

In California or Massachusetts, the odds of swinging an election are 1 in 50 million

If you spend an hour on voting, and you earn $10/hour…

Would you pay $500 million to have Kerry over Bush?

What if you can influence others' votes?

If you can influence 10 others, but that still implies a willingness to pay that is greater than most of our lifetime incomes

What about swing states…

Map removed due to copyright restrictions. See Figure 1 in the article "What is the Probability Your Vote Will Make a Difference."

Lighter colored states are those where your vote matters more
Source: ANDREW GELMAN, NATE SILVER and AARON EDLIN 2012

Even there…

"A single vote (or, for that matter, a swing of 100 or 1,000 votes) was most likely to matter in New Mexico, New Hampshire, Virginia, and Colorado, where your vote had an approximate 1 in 10 million chance of determining the national election outcome. Alaska, Hawaii, and the District of Columbia are not shown on the map, but the estimated probability of a single vote being decisive was nearly zero in those locations."

Here are some other things that might happen with 1 in 10 million odds:

You will have identical quadruplets

You will become president

And you are 10 times as likely to get struck by lightning or die flesh-eating bacteria than you are to change an election result

Voters care enough to vote, and they care about Kerry winning, but they can't be voting because they want Kerry to win

(In contrast, I buy chocolate, and I like eating chocolate. But… I buy chocolate because I like eating chocolate.)

Why is voting different from chocolate?

# Last puzzle…

# Why Habitat for Humanity is needed

All good reasons to provide housing to the poor

But…

Clearly, having a bunch of

kids with college degrees
who’ve never built a house
fly halfway across the world

is not the most efficient way to build new homes!

If we really wanted to maximize our impact, we’d spend those hours working and donate the money to hire local contractors

What we learned from these puzzles…

Of course, people care about HIV/AIDS, other lab subjects, voting, and access to affordable housing

They just care in weird ways

Can we characterize the ways in which caring is weird?

Can we understand when people will care, however weirdly?

And can we use this understanding to get people to care more? Or, given that they care, can you get them to have a bigger impact?

To understand caring and increase it, we need to introduce a new tool…

Repeated prisoner's dilemma

Let’s go back and review the prisoner’s dilemma

Let’s use the following payoffs…

| | C | D |
|---|---|---|
| **C** | b-c, b-c | -c, b |
| **D** | b, -c | 0, 0 |

b>0 is the benefit if partner cooperates

c>0 is the cost of cooperation

44

For example…

You and another student both have problem sets due next week

The other student can pay a cost c to improve your grade by b

And you can do the same for him

What is the Nash equilibrium?

| | C | D |
|---|---|---|
| C | b-c, b-c | -c, b |
| D | b, -c | 0, 0 |

b>0 is the benefit if partner cooperates

c>0 is the cost of cooperation

This is bad news for your grade on the problem set

Now that you remember the prisoners' dilemma, we're ready to talk about repeated games

The idea is simple…

In a repeated game, we simply repeat this game with probability δ

For example…

You and the other student don’t know if you will be assigned another problem set. You both assess that there is some probability that this happens, say .6

In this case, $\delta = .6$. This is the probability that we interact again in the future

Let’s modify this example slightly…

Imagine that we know we’ll get another problem set, but…

The other student can’t always tell if you actually put in effort into solving his problem set, or were thinking about your own problem set instead

Even if you do put in a genuine effort, you won’t always figure his problem set out.  Suppose there is a .6 chance that you will be able to detect whether I genuinely put in effort

What is $\delta$ in this case?

$\delta$ is still .6. But now it's the probability that you observed whether I cooperate

In general, δ is interpreted as:

the probability the players interact again,
and their actions in this round are observed,

and remembered,

etc.

Now that we know what δ represents, let's think about players' actions a bit…

When the game repeats, strategies are potentially a lot more complicated as they specify actions in every round

Example strategies:

> Cooperate in only the second round
>
> Cooperate in rounds 1, 2, 3, 5, 7, 11, …

Crucially, you can also condition your behavior on others' behavior

Let's see this in our example…

An example strategy which conditions on the other student's past actions is:

Help in round 1

Then, if the other student helped you in the last round, help in this round. Otherwise, don't help in this or any future rounds

This is called "grim trigger"

Finally, we're ready to think about Nash equilibria of this repeated game

Let's start with our example…

Suppose you play the strategy:

Help in rounds 1, 2, 3, 5, 7, 11, …

And imagine your partner's strategy involves helping in some arbitrary round k.
Should he deviate and defect in round k?

If k is 1, 2, 3, 5, 7, 11…

He gets b-c if he helps and b if he doesn't. So he's better off deviating

If k falls in one of the other rounds…

He gets –c if he doesn't and 0 if he doesn't. So he's better off deviating

This is not a Nash equilibrium

What *is* a Nash equilibrium?

For starters, if both of you:

Don’t help in any round

Then neither of you can do any better by deviating

So this is a Nash equilibrium

# What if you both play “grim trigger”?

Now, as long as you're sufficiently likely to be assigned problem sets again, and you value your grades sufficiently…

If you deviate, the future benefits you lose are greater than the current gains…

So neither can benefit from deviating and this is a Nash equilibrium

Notice that in this equilibrium, you and your partner always help

Let’s generalize this…

What has to be true about $\delta$, c, and b for the players to cooperate in equilibrium

What properties will their strategies have?

We identify two key properties

We'll sketch a proof of these in a minute, and you'll prove them formally in your HWs

Here they are…

# Properties of Cooperative Equilibria

(1) Cooperation can only occur when $\delta > c/b$

(2) When a player cooperates, the other player must "reciprocate" by being more cooperative in the future

Sketch of proof for condition (1):

Cooperation can only occur when $\delta > c/b$

Let's start by arguing one direction…

… that cooperation *can* occur when $\delta > c/b$

Suppose both players play grim trigger

Players won't deviate if the future benefits lost are greater than the current gains

Future benefits are greater when $\delta$ is high

Current gains are low when c is low, relative to b

When you'll do the math, you'll find that players lose from deviating exactly when $\delta > c/b$

Now we need to show the other direction…

… that cooperation can *only* occur when $\delta > c/b$

Start by recognizing that the punishment in “grim trigger” is the most severe possible because if your partner ever defects, with 100% probability you’ll never cooperate again

Thus, grim trigger makes the foregone future benefits from deviating greatest

So, if grim trigger can’t deter defection, nothing can

Before we prove condition (2), let’s consider some alternatives to grim trigger

Punish for 7 periods

Punish for 1 period (we call this “tit-for-tat”)

Notice that in both of these strategies, players condition cooperation on partners' past behavior

Condition (2) says that any strategy that supports cooperation MUST have some punishment for defection

I.e. strategies like the following cannot be equilibrium:

Cooperate in rounds 1, 2, 3, 5, 7, 11, …

We'll learn why from the proof

Proof of condition (2):

When a player cooperates, the other player must "reciprocate" by being more cooperative in the future

We do a proof by contradiction. Simply assume that condition (2) is false.

That means at least one player, say player 2, is playing a strategy where she cooperates, but her cooperation in future rounds doesn't depend on whether player 1 cooperates now

Suppose 1 deviates in some period. Then she gains c, and player 2 doesn't change a thing, so she loses nothing.

Obviously, she is better off and this could not be a Nash equilibrium. We've reached a contradiction.

So far, showed that cooperation can be supported in equilibrium

And described this equilibrium

However, cooperation isn’t the only equilibrium

In fact, all equilibria can be “invaded”

For example:

- TFT is an equilibrium
- AllC does equally as well, so might spread
- And if there are enough AllC, AllD does better than TFT, and will spread

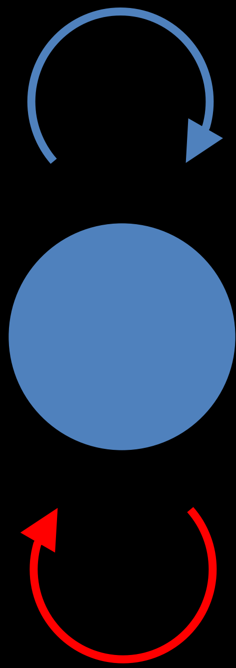
So it is not obvious which strategies will emerge in an evolution or learning process

Which we now investigate

Note for evolutionary process, we need to restrict strategies and make assumptions about how mutations work.

One way to do so is “finite state automata.” Let me explain

*All C*



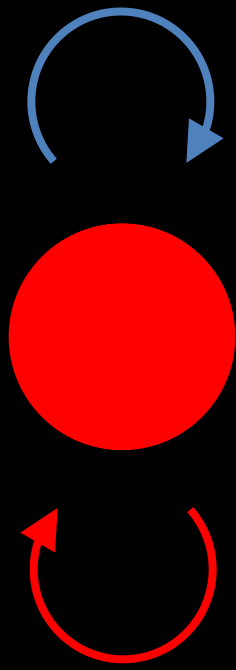
Left most circle represents "state" an individual starts at

Color of circle represents action taken in that state

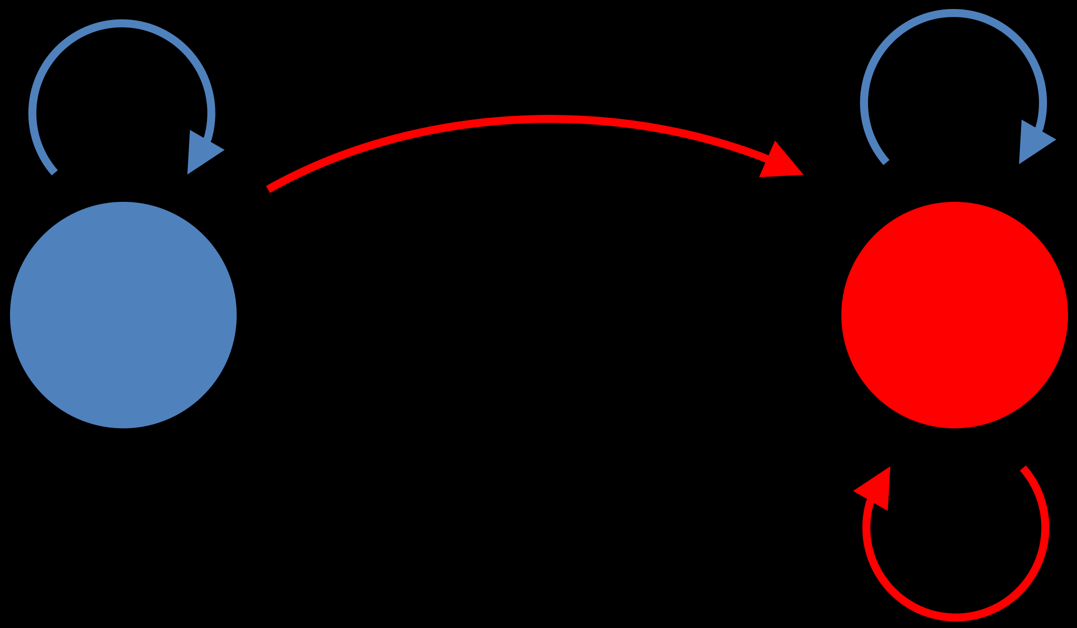
Blue arrow represents state "transition" to when Other player cooperates

Van Veelen et al. 2012

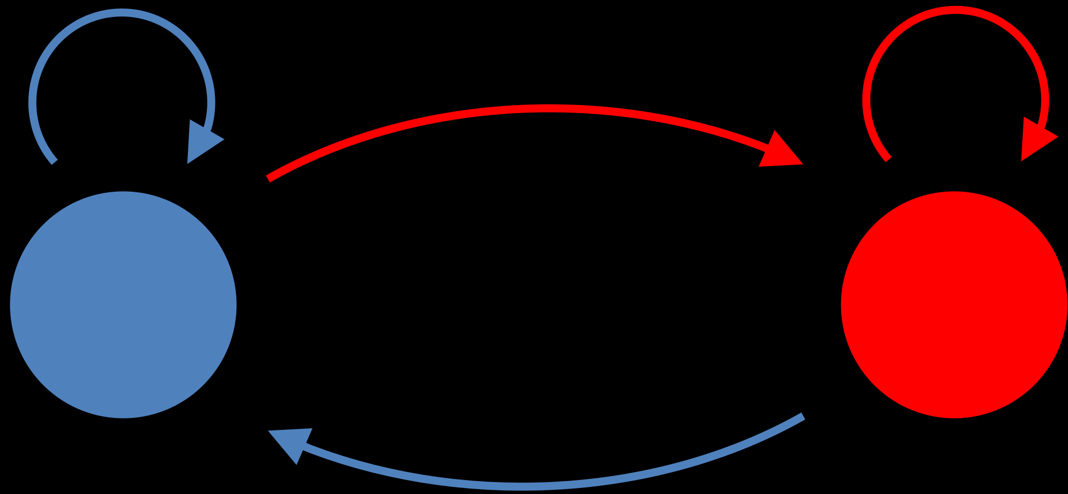
*All D*



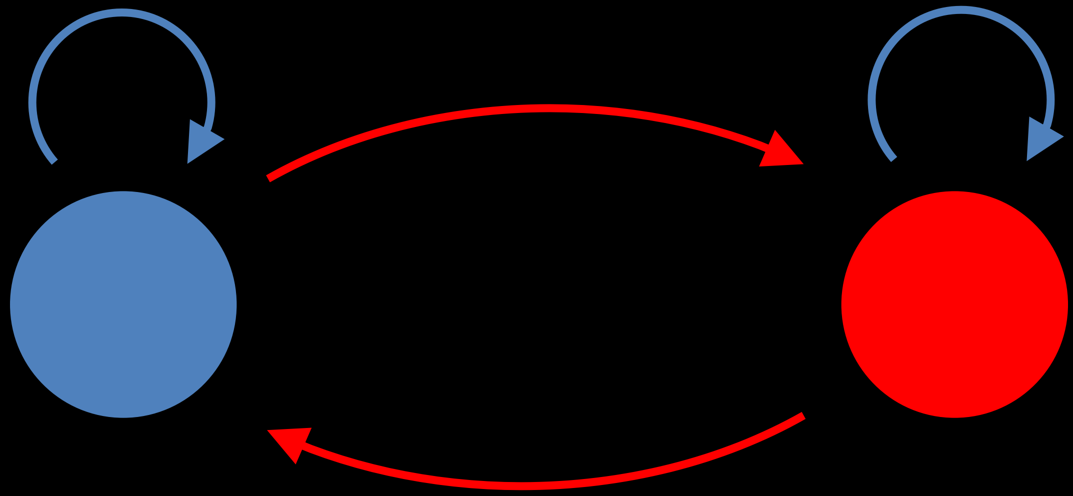
# *Grim Trigger*



*Tit-for-Tat*



# *Win stay loose shift*



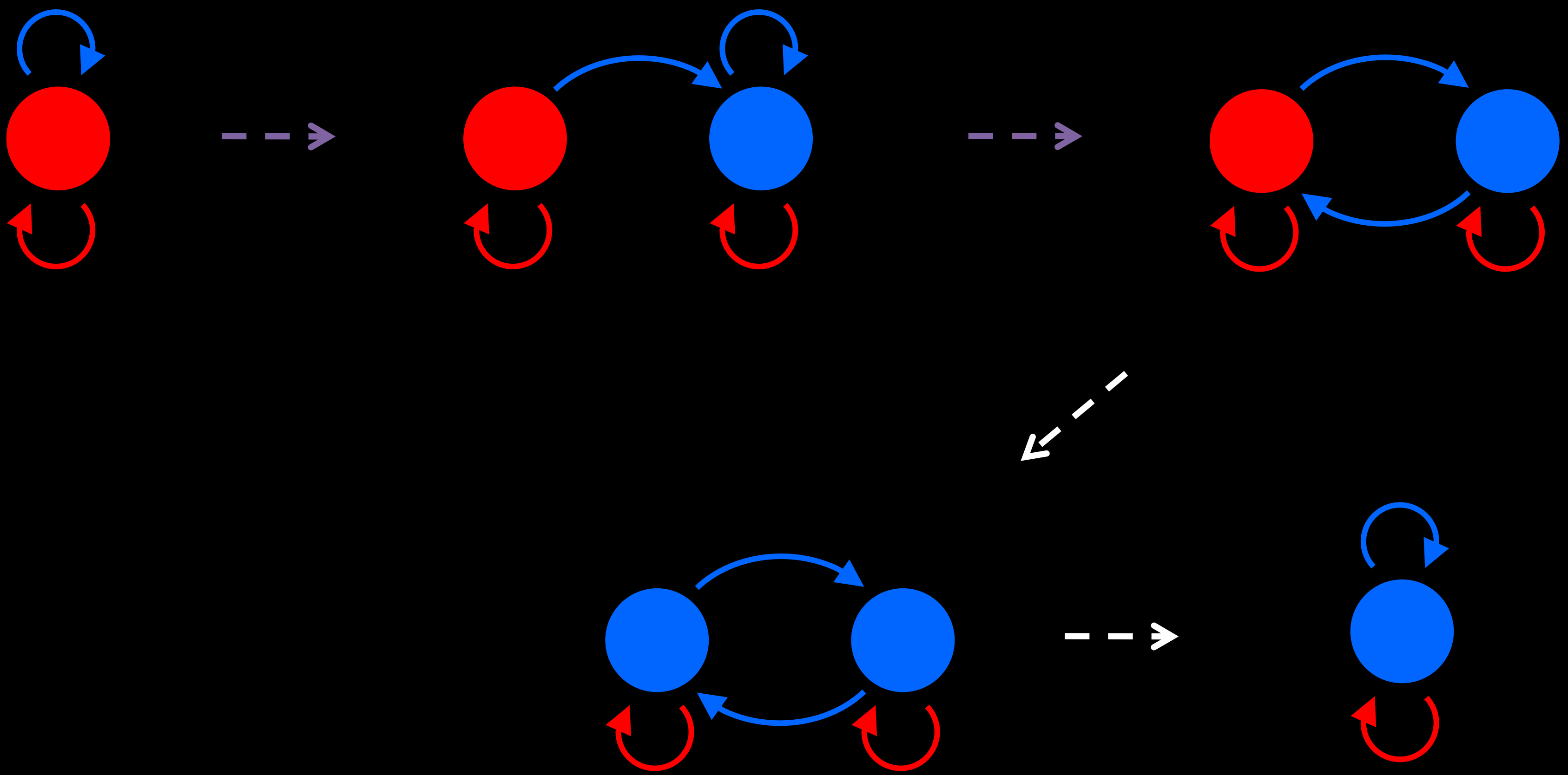
# Finite automata: way to go!

Any finite automaton can be reached from any other finite automaton by a finite number of mutations

The set of strategies is uncountably infinite

The set of finite automata is dense in the set of all strategies

# Mutations



Start with All D, run Wright-Fisher. Characterize the amount of cooperation played in the long run.

## 3.2 Levels of cooperation for $\alpha = 0$ and varying continuation probability

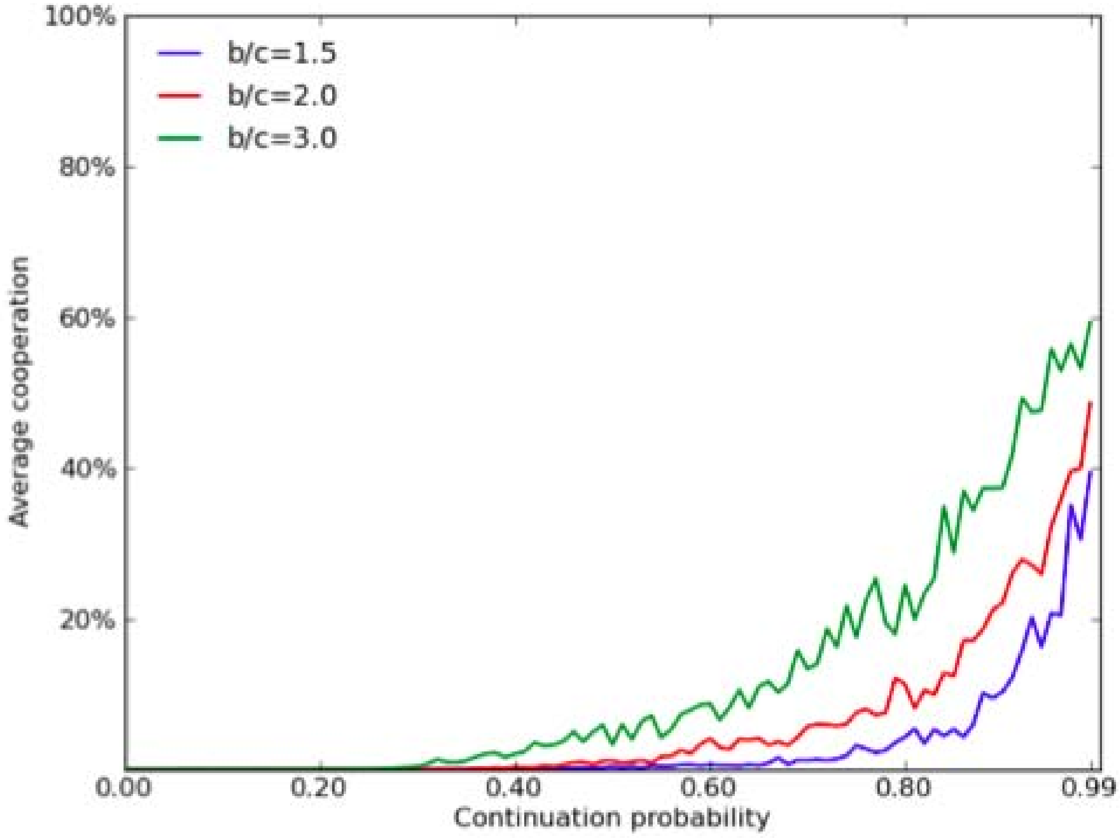
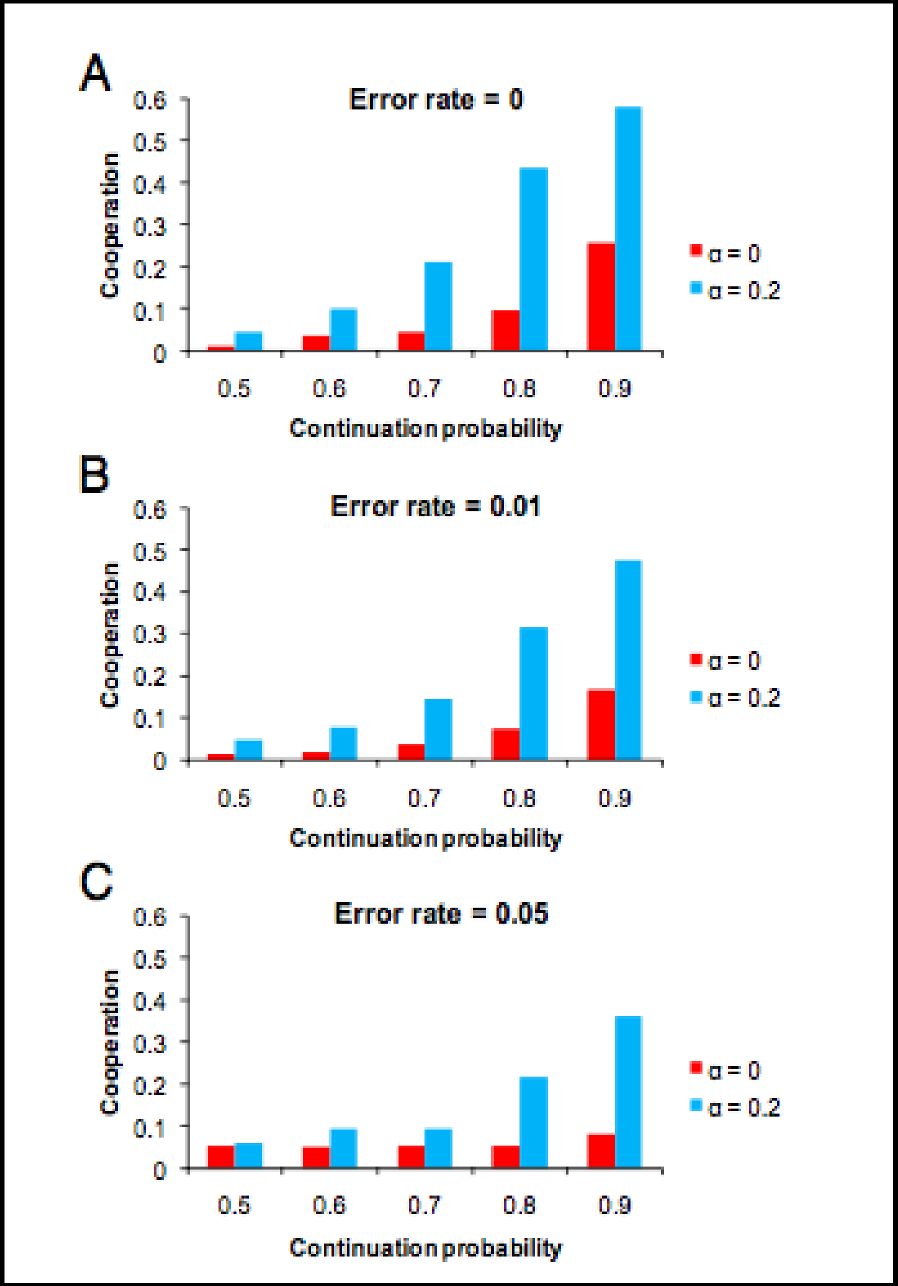



**Fig. S5. Levels of cooperation for different *b/c* ratios and continuation probabilities**

- Cooperation increases with delta and b/c

This is true even when we add “errors” (sometimes players accidentally play C when mean D and vice versa)

(ignore the blue)





Source: Veelen, Matthijs, Julian Garcia, et al. "Direct Reciprocity in Structured Populations." *PNAS* 109, no. 25 (2012): 9929-334. 

Many other simulations have been run with other strategy spaces and mutation rules.

e.g. in readings will see:

1) “genetic algorithms”

(a few “loci,” each with several “alleles” coded for what do in first 3 periods, and what do as function of past 3 periods)

2) “axelrod’s experiment”

(best game theorists submitted strategies)

What does this have to do with public goods contributions?

(e.g. giving to charity, voting, carbon emissions…)

Our results on repeated prisoner's dilemma EASILY extend to a public goods setting

Just need:

- Everyone chooses whether or not to contribute to the public good
  - Costs self, benefits all, socially optimal for all to contribute but dominant strategy not to.
- Everyone has a chance to "punish" those who have not contributed
  - Punish could be defecting E.g. I refuse to help those with homework who don't recycle
  - Or "costly punishing"
    - I punch in face anyone who doesn't recycle.
- Crucially, everyone must have a chance to "punish" everyone who didn't punish properly
  - E.g. I won't work with anyone who works with a nonrecycler (ostracism)
- And again, and again, and again…

Contribution to the public good can then be sustained, if

1) $\delta$ sufficiently high
2) more likely to be punished if don’t contribute or if don’t punish when “should”

(You will do proof in homework. We haven’t seen any simulations on this. Final project?)

Before we go on to evidence, some comments…

Notice what properties equilibria don't have:

1) Doesn't depend on (privately known) effectiveness of action

2) Doesn't (necessarily) depend on level of public good

Finally, we're ready for the evidence that this model works

We'll show evidence that reciprocity:

- Exists in animals
- Emerges "spontaneously" in humans

Then, we'll show evidence that altruism in humans is consistent with reciprocity

- More altruism when c/b is lower and when $\delta$ is higher
- People avoid being observed defecting
- Altruists not sensitive to effectiveness of gift
- People sometimes motivated to avoid learning effectiveness

Let’s start by looking at the evidence that reciprocal altruism exists in animals

Here is our favorite…

Image courtesy of kjelljoran on Flickr. CC BY-NC-SA

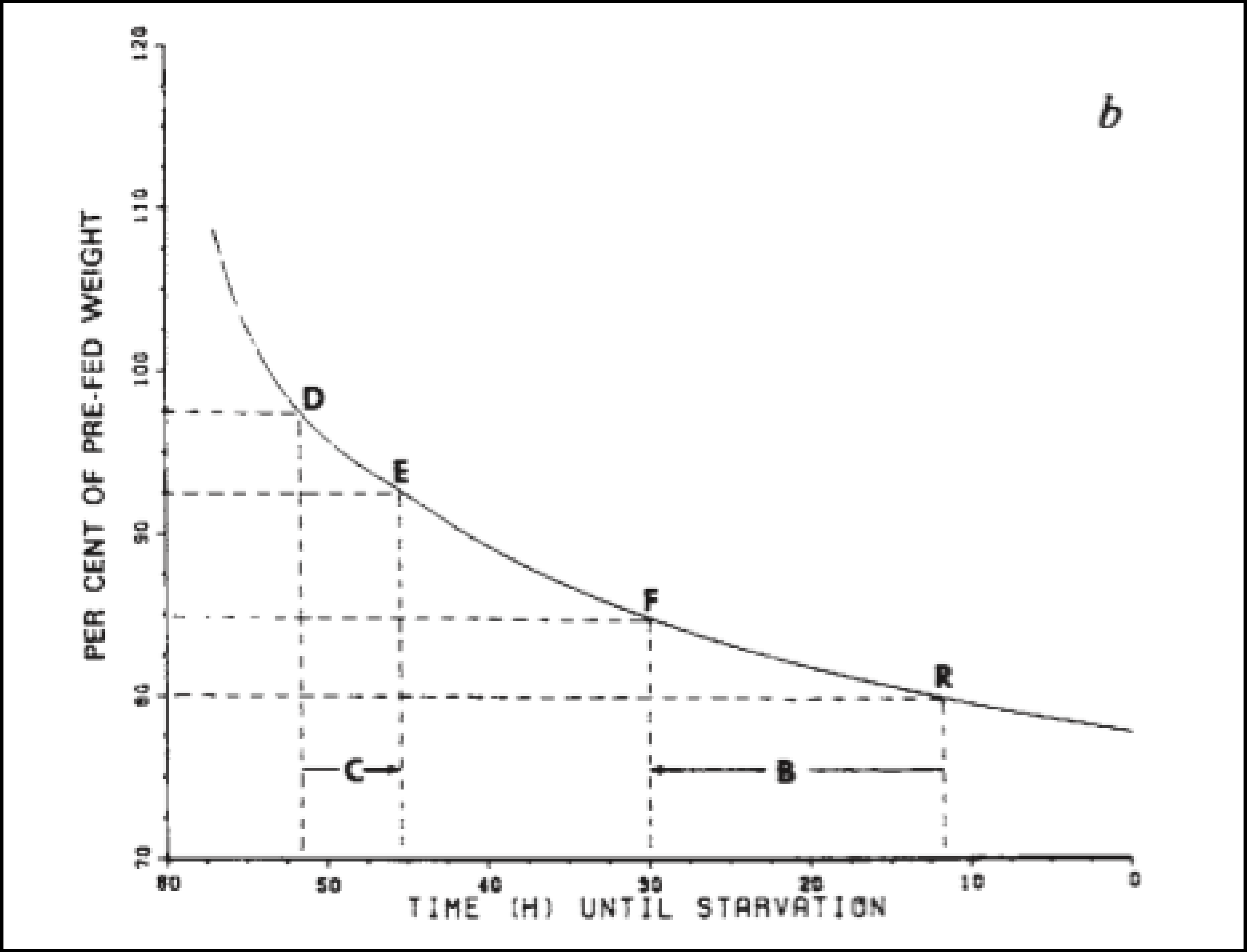
And our second favorite:

# Vampire bats

Vampire bats live off the blood of large farm animals

Sometimes, they forage unsuccessfully and remain hungry

If they remain hungry for too long, they die






Source: Wilkinson (1984)

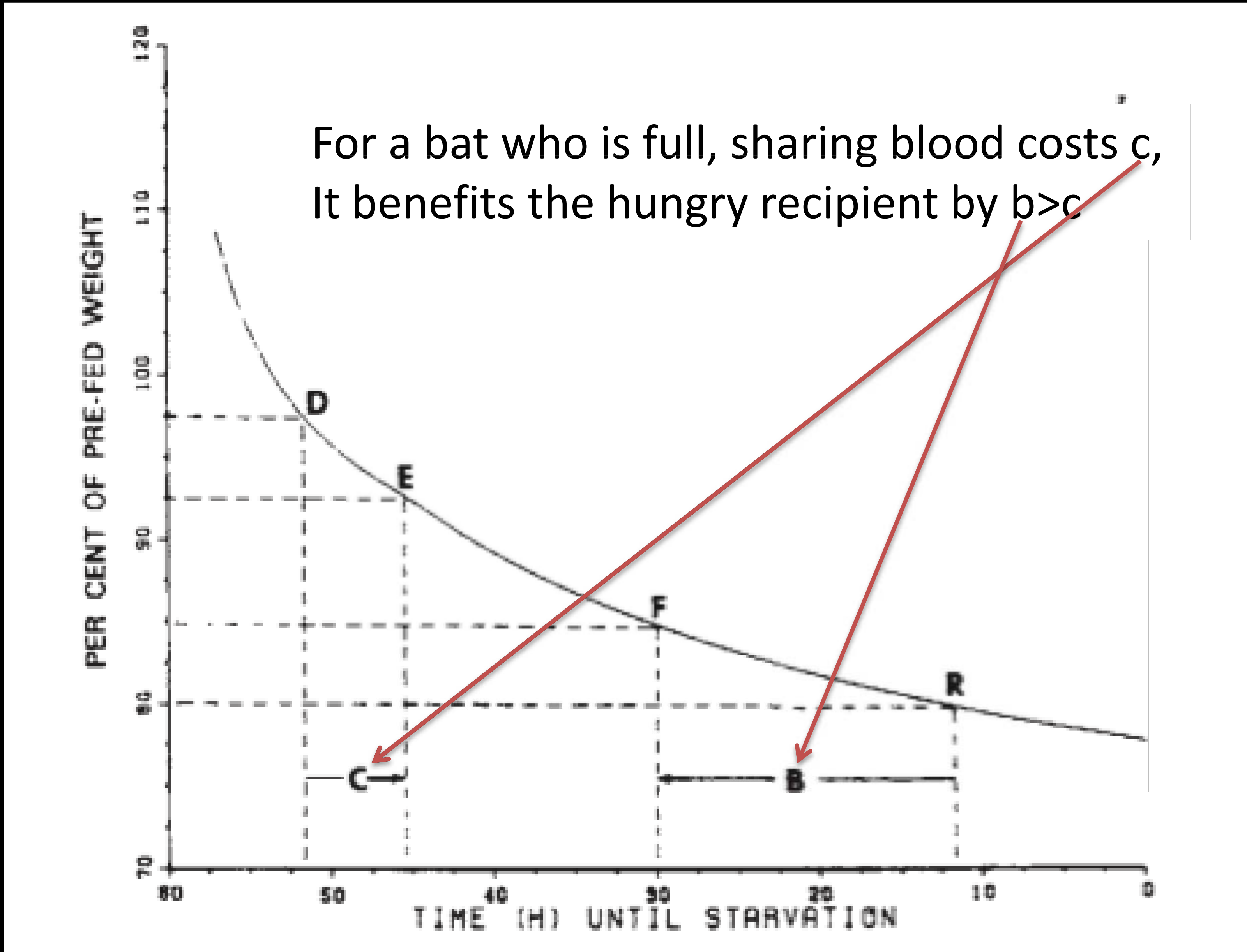
It turns out bats can share blood through regurgitation

Frequently, bats who successfully foraged for blood will share with those who came back with an empty stomach

They do this even with bats they're not related to

Why would they give up this valuable nutrition?

We can model the decision to share blood as a prisoner's dilemma…




Reprinted by permission from Macmillan Publishers Ltd: Nature.

Source: Wilkinson, Gerald S. "Reciprocal Food Sharing in the Vampire Bat." *Nature* 308 (1984).

To see that sharing blood is driven by reciprocity, let’s check that our two conditions are met

(1): Cooperation can only occur when $\delta > c/b$

Bats live in stable broods for many years:

“Seven out of nine adult females present in one group in 1980 were still roosting together in 1982; in another group, one pair of females marked… in 1970 roosted in the same area in 1981.”

So $\delta$ is high

(2) When a player cooperates, the other player must "reciprocate" by being more cooperative in the future

Researcher randomly selected bats and prevented them from feeding

"Each trial in which the donor fed a starved bat was compared with the subsequent trial in which the donor was starved.

Starved bats… reciprocated the donation significantly more often than expected had exchanges occurred randomly"

Now let’s look at humans

We’ll start by showing reciprocal cooperation can emerge “spontaneously”

That is, we can go from an equilibrium where everyone is defecting to one in which everyone is cooperating

It turns out this can happen even in the places you'd least expect it…

… like in the trenches of WWI

In WWI, the two sides discovered that it was easy to defend and virtually impossible to overtake an area defended by a trench

The two sides tried to attack by going around each other’s trenches, but each made the trench longer and longer

Until it extended all the way to the sea…

At this point, they reached a stalemate which lasted for about 3½ years

But what happens when the same people sit across from each other for long periods of time?

δ increases, and…

Now, the British didn't seem so keen on killing Germans…

"…German soldiers [were] walking about within rifle range behind their own line.  Our men appeared to take no notice."

“At seven it [the artillery shell] came—so regularly that you could set your watch by it.”

And the Germans weren't so keen on killing the British…

"So regular were they [the Germans] in their choice of targets, times of shooting, and number of rounds fired, that, after being in the line on or two days, Colonel Jones had discovered their system..."

"Suddenly a salvo arrived but did no damage. Naturally, both sides got down and our men started swearing at the Germans, when all at once a brave German got on to his parapet and shouted out, "We are very sorry about that; we hope no one was hurt.  It is not our fault, it is that damned Prussian artillery."

Cooperation emerges!

How can we tell it emerged because of repeated PD?

Look for evidence that our two conditions were met:

- We already know that $\delta$ was high
- Was there reciprocity?

"If the British shelled the Germans, the Germans replied, and the damage was equal: if the Germans bombed an advanced piece of trench and killed five Englishmen an answering fuselage killed five Germans"

"The really nasty things are rifle grenades… The can kill as many as eight or nine men if they do fall into a trench… But we never use ours unless the Germans get particularly noisy, as on their system of retaliation three for every one of ours comes back."

Here is some more evidence…

"During the periods of mutual restraint, the enemy soldiers took pains to show each other that they could indeed retaliate if necessary. For example, German snipers showed their prowess to the British by aiming at spots on the walls of cottages and firing until they had cut a hole."

Source: Axelrod The Evolution of Cooperation

Amazing, right?

Note:

We are NOT saying this is the only source of altruism

Also note:

Need not be aware

Might “feel good” to give, or “feel guilty” if don’t, or believe its “right thing to do”

(We are offering explanation for WHY we feel/think this way!)

Let’s switch gears

Next, we’ll look at a bunch of evidence for condition (1): Cooperation can only occur when $\delta > c/b$

Let’s start in the lab…

Researchers had subjects play a repeated prisoners' dilemma

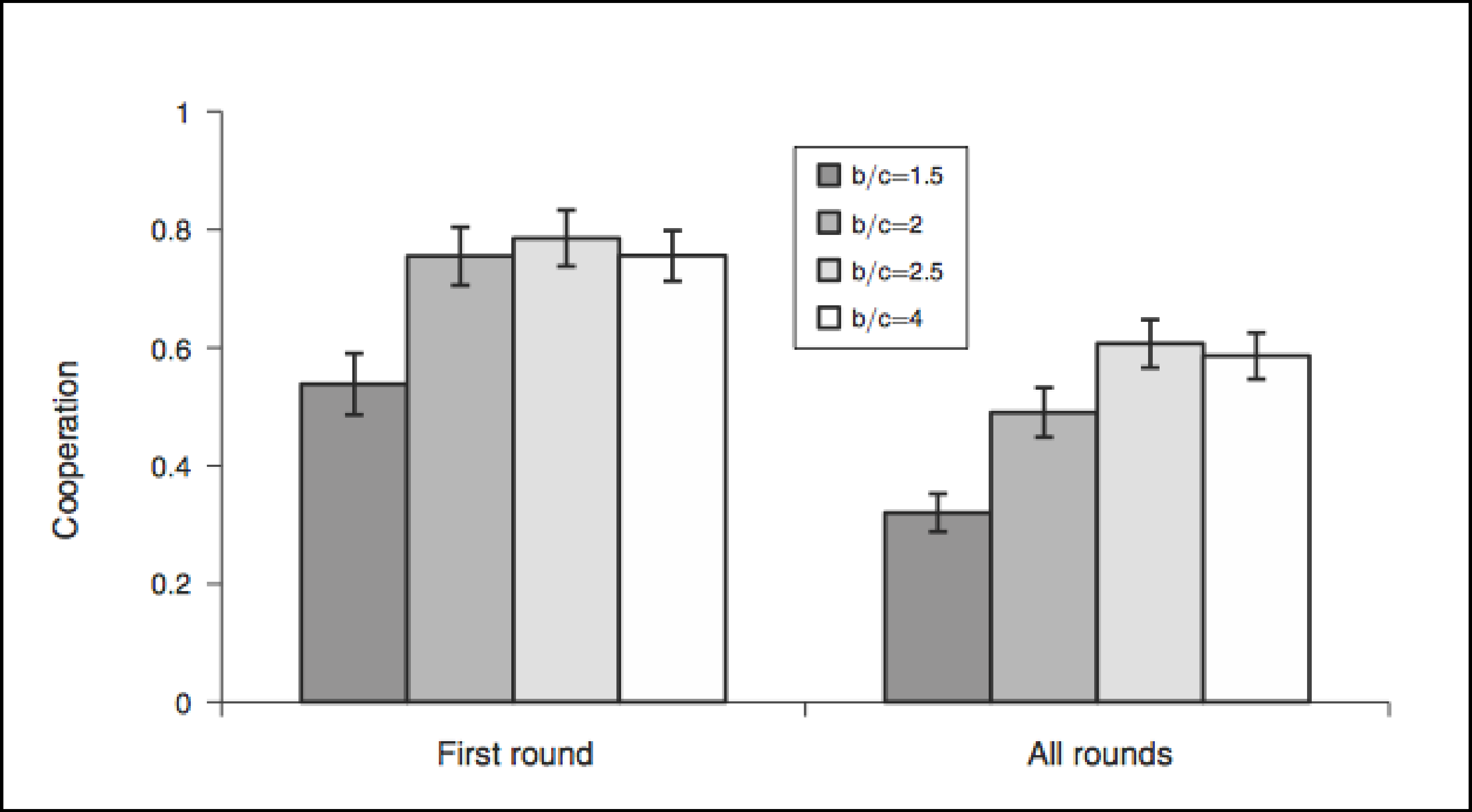
Set $\delta=7/8$

Vary b/c: 1.5, 2, 2.5, and 4

Cooperation supported

Cooperation NOT supported

# Cooperation was significantly lower when it was not supported in equilibrium:






Source: Fudenberg et al. 2012

OK, that was c/b

What happens when we vary δ?

For now, let’s stay in the lab

Researchers had subjects play a few rounds of *public good game*

In the control group, they stopped there

There were three different treatment groups in which they interacted again in a punishment or reward round

Subjects contributed a lot more when they knew the punishment or reward round was coming

So that’s what happens when we increase the likelihood of future interactions

But that’s not the only interpretation of δ

Subjects in a computer lab were assigned to play dictator game with a randomly selected partner

Researchers varied δ by very subtly manipulating whether subjects thought they were being observed

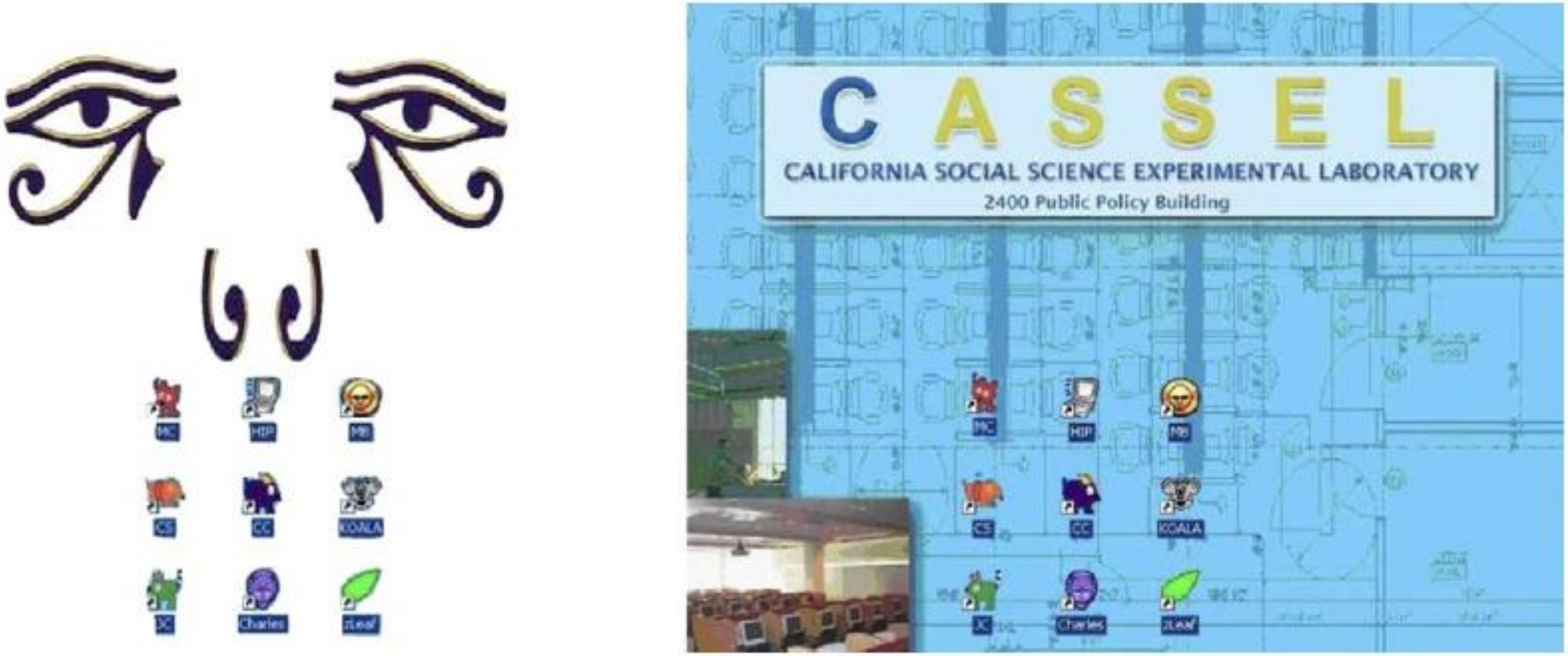



Fig. 1. Eyespots (left) and control (right) desktop displays.



Source: Haley and Fessler 2005

Also…

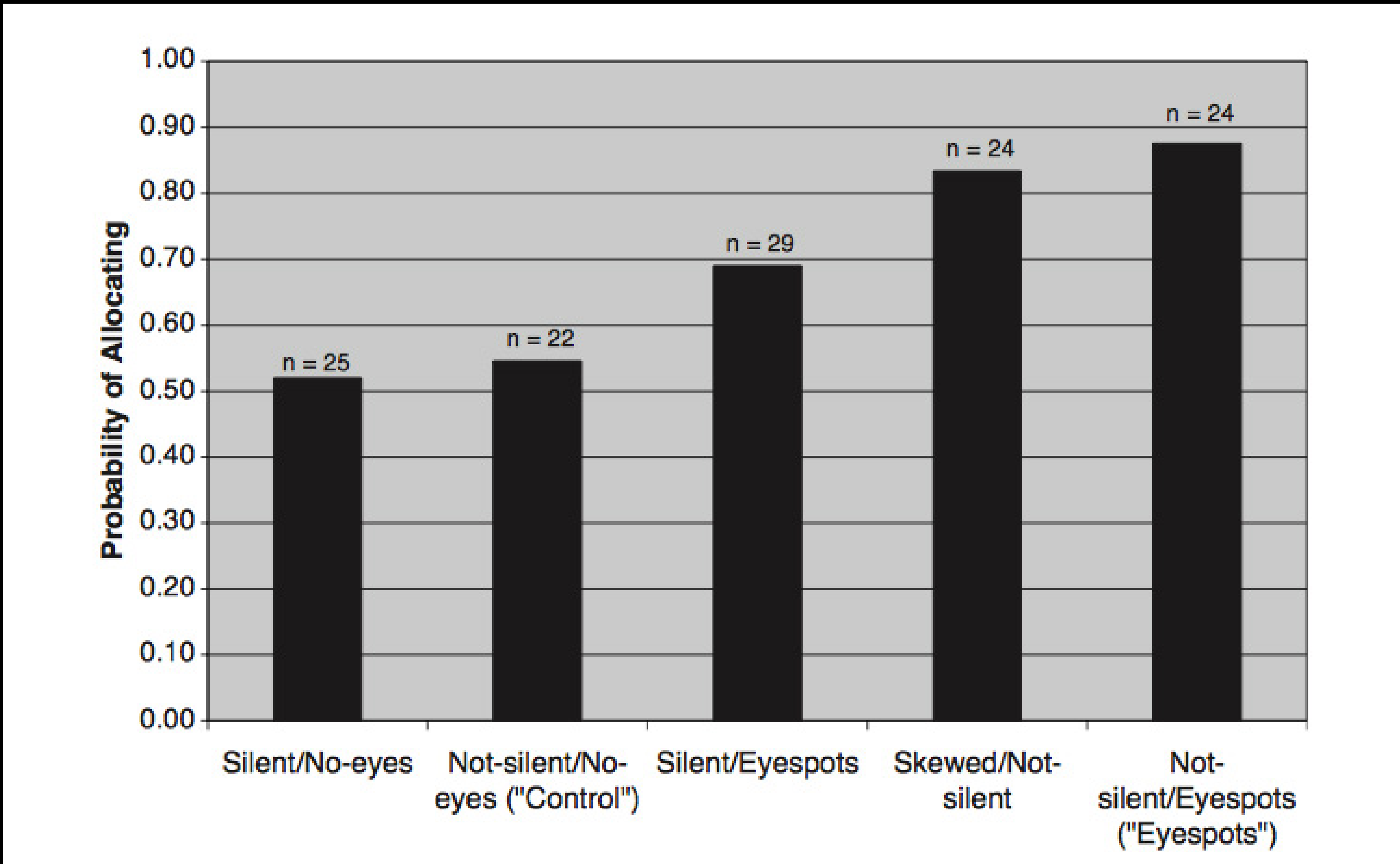
Had another subtle cue of observability: whether subjects wore ear-muffs

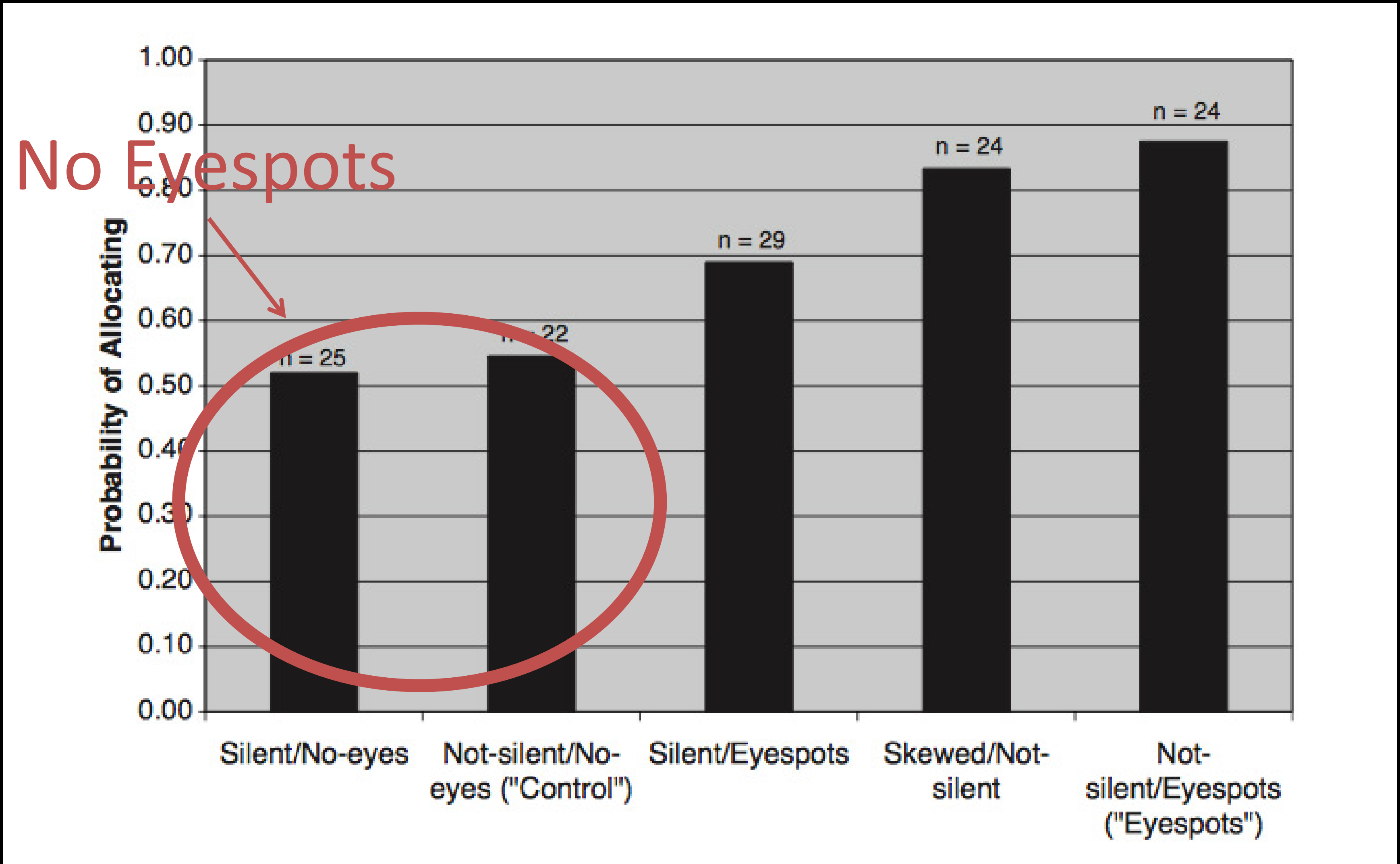
Had a second type of eye-spots, called “skewed”

What happened to giving by dictators?

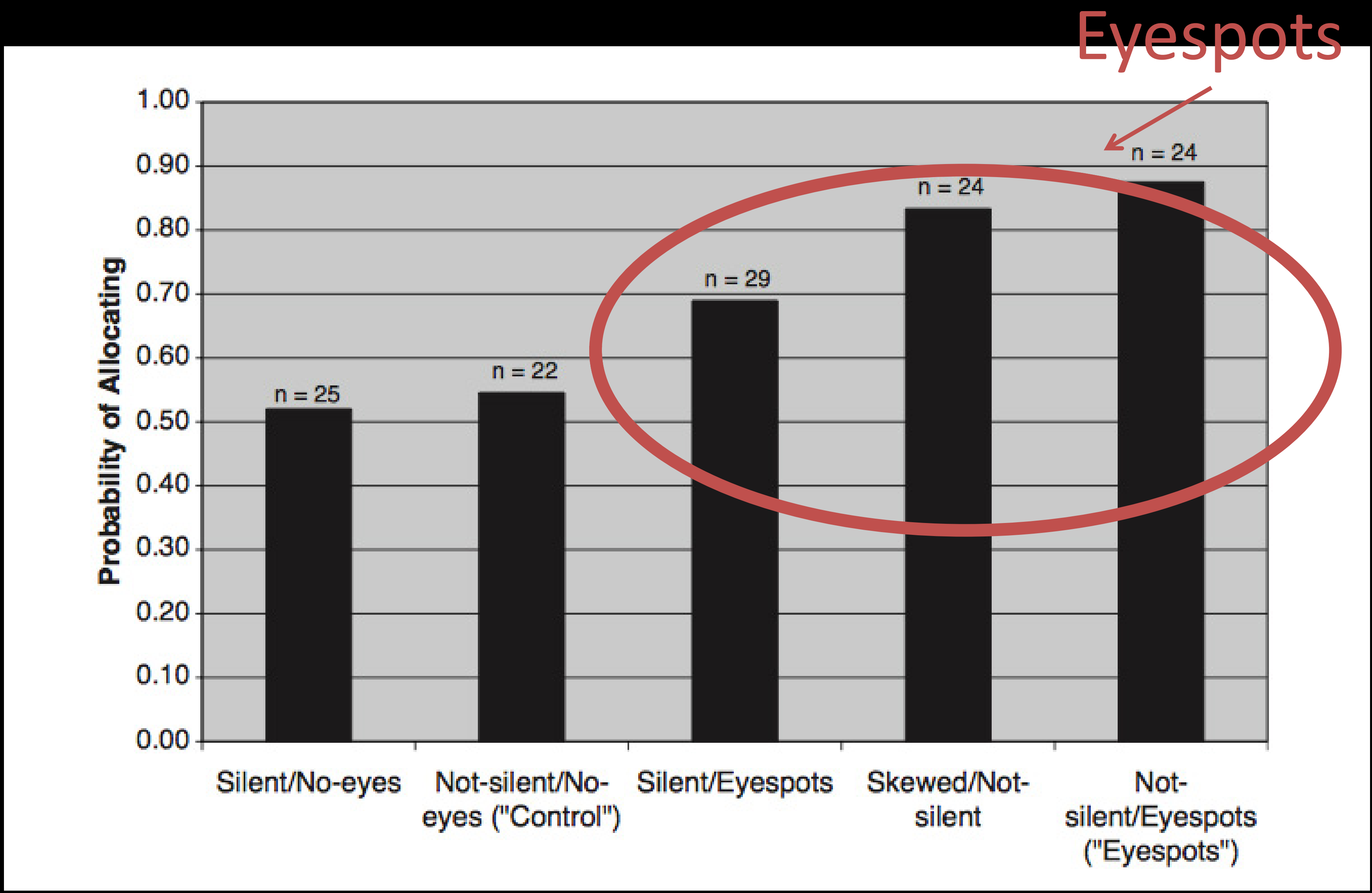
Even subtle cues of observability can increase cooperation

That's powerful!


Probability of Allocating
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
n = 25
n = 22
n = 29
n = 24
n = 24
Silent/No-eyes
Not-silent/No-eyes ("Control")
Silent/Eyespots
Skewed/Not-silent
Not-silent/Eyespots ("Eyespots")




Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.



Eyespots
Probability of Allocating
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
n = 25
n = 22
n = 29
n = 24
n = 24
Silent/No-eyes
Not-silent/No-eyes ("Control")
Silent/Eyespots
Skewed/Not-silent
Not-silent/Eyespots ("Eyespots")

There is *a lot* of other evidence that observablity increases altruism in the lab

Does it work outside the lab?

We ran a large scale field experiment with a real public good to find out…

We collaborated with CA electric utility, PG&E, to market their SmartAC blackout prevention program

In CA, super-peaks on hot days can threaten grid stability

SmartAC helps by targeting central ACs

PG&E traditionally:

- Mailed customers informational fliers
- Offered them $25 to participate
- Asked them to call a hotline to sign up
- Achieved sign-up rates of ~2%

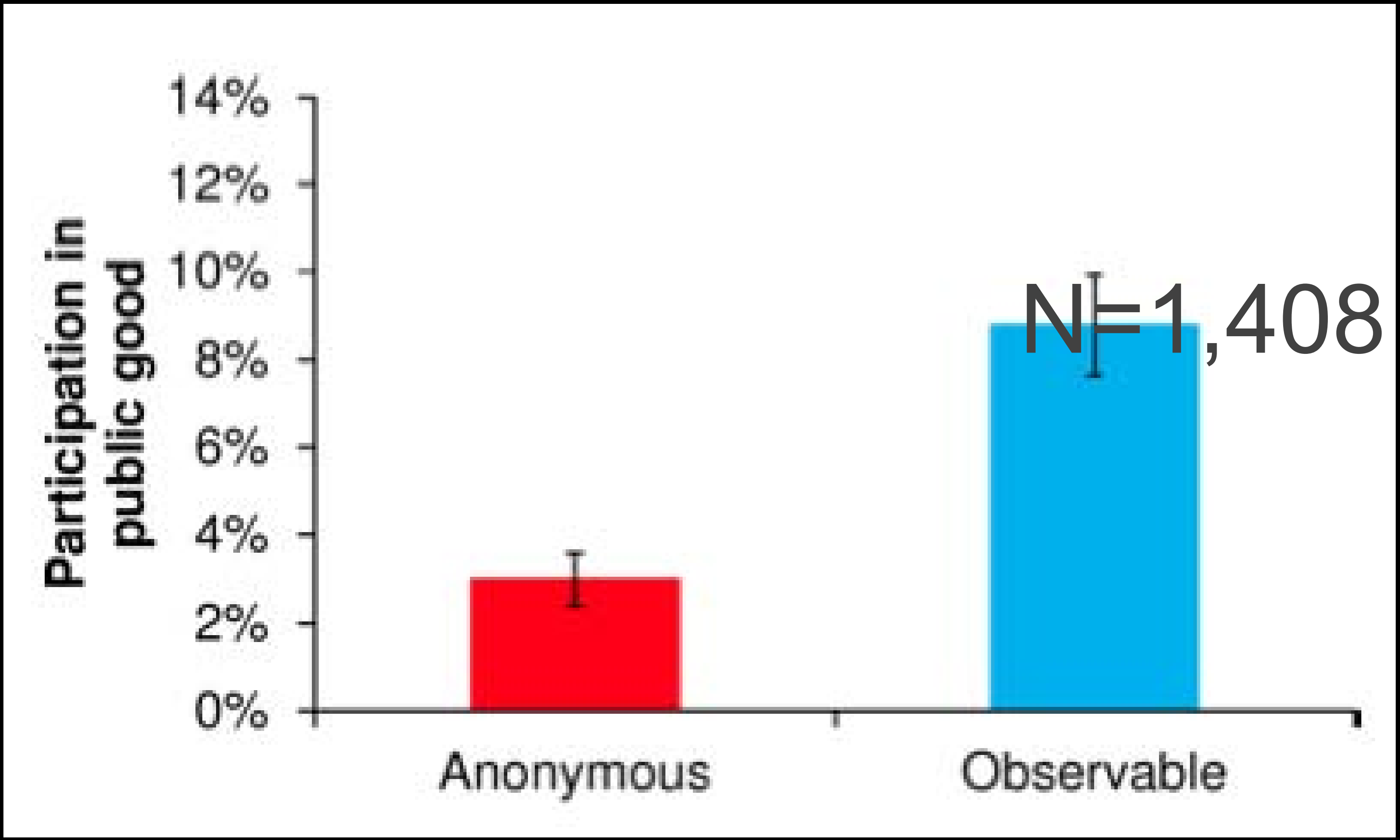
We

Marketed SmartAC to residents of homeowners associations (HOAs) in Santa Clara County

Asked customers to sign up on publicly posted sheets

Varied observability by varying the sheets

Here are the results…





Source: Yoeli, Erez, Moshe Hoffman, David G. Rand, and Martin A. Nowak. "Powering Up with Indirect Reciprocity in a Large-Scale Feld Experiment." Proceedings of the National Academy of Sciences of the United States of America. 2013. 

Source: Yoeli, Hoffman, Rand, and Nowak 2013

# How does this compare to the $25 incentive?


Increasing observability has the same effect as paying

That’s the main result

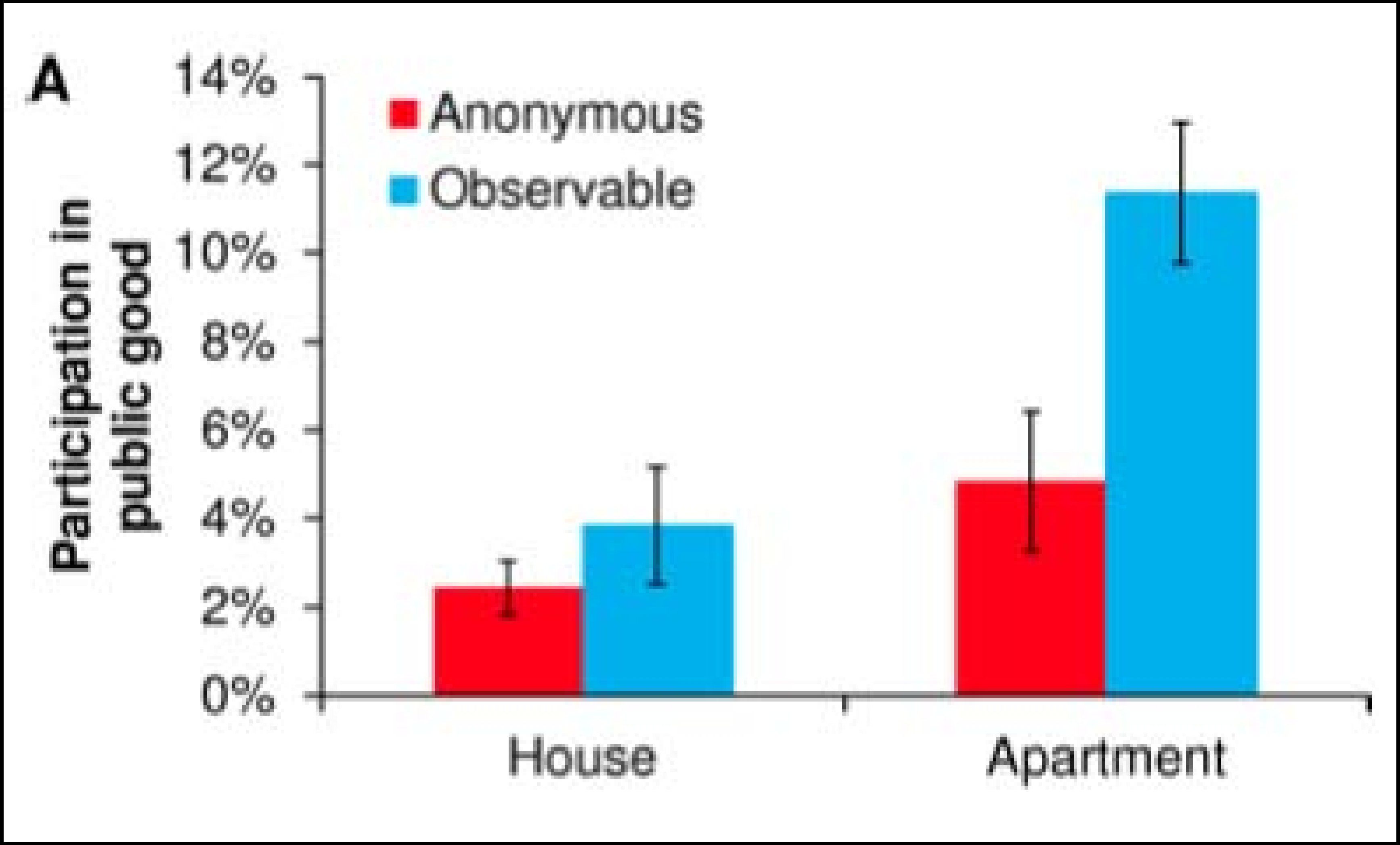
The theory gives additional testable predictions…

When eliminate observability, reduce $\delta$ to “zero”

The higher $\delta$ was in the first place, the bigger the difference observability will make

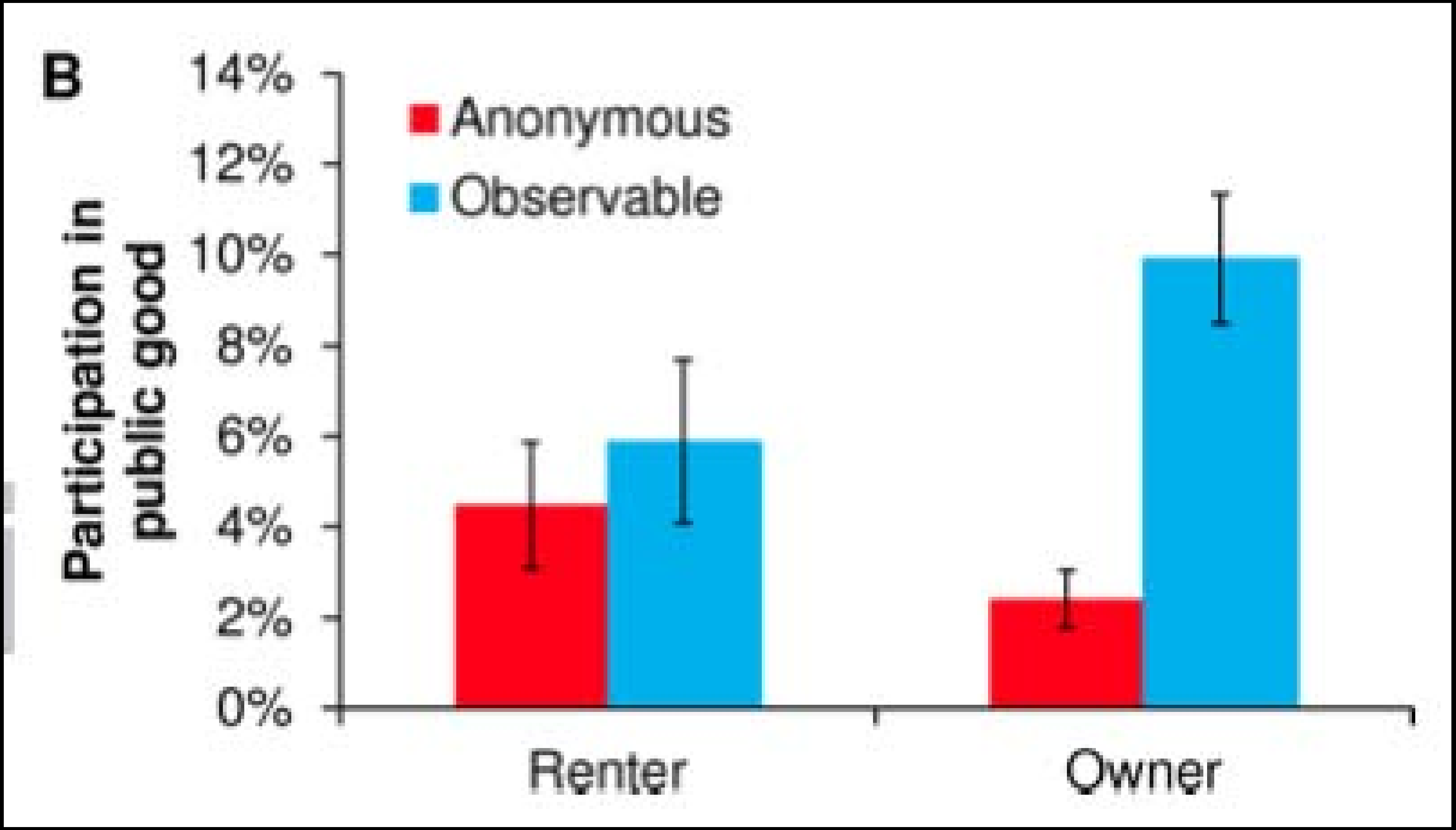
We can test this by looking for situations with higher $\delta$:

- Apartment buildings vs. Townhomes
- Owners vs. renters





Source: Yoeli, Erez, Moshe Hoffman, David G. Rand, and Martin A. Nowak. “Powering Up with Indirect Reciprocity in a Large-Scale Feld Experiment.” Proceedings of the National Academy of Sciences of the United States of America. 2013. 





Source: Yoeli, Erez, Moshe Hoffman, David G. Rand, and Martin A. Nowak. “Powering Up with Indirect Reciprocity in a Large-Scale Feld Experiment.” Proceedings of the National Academy of Sciences of the United States of America. 2013. 

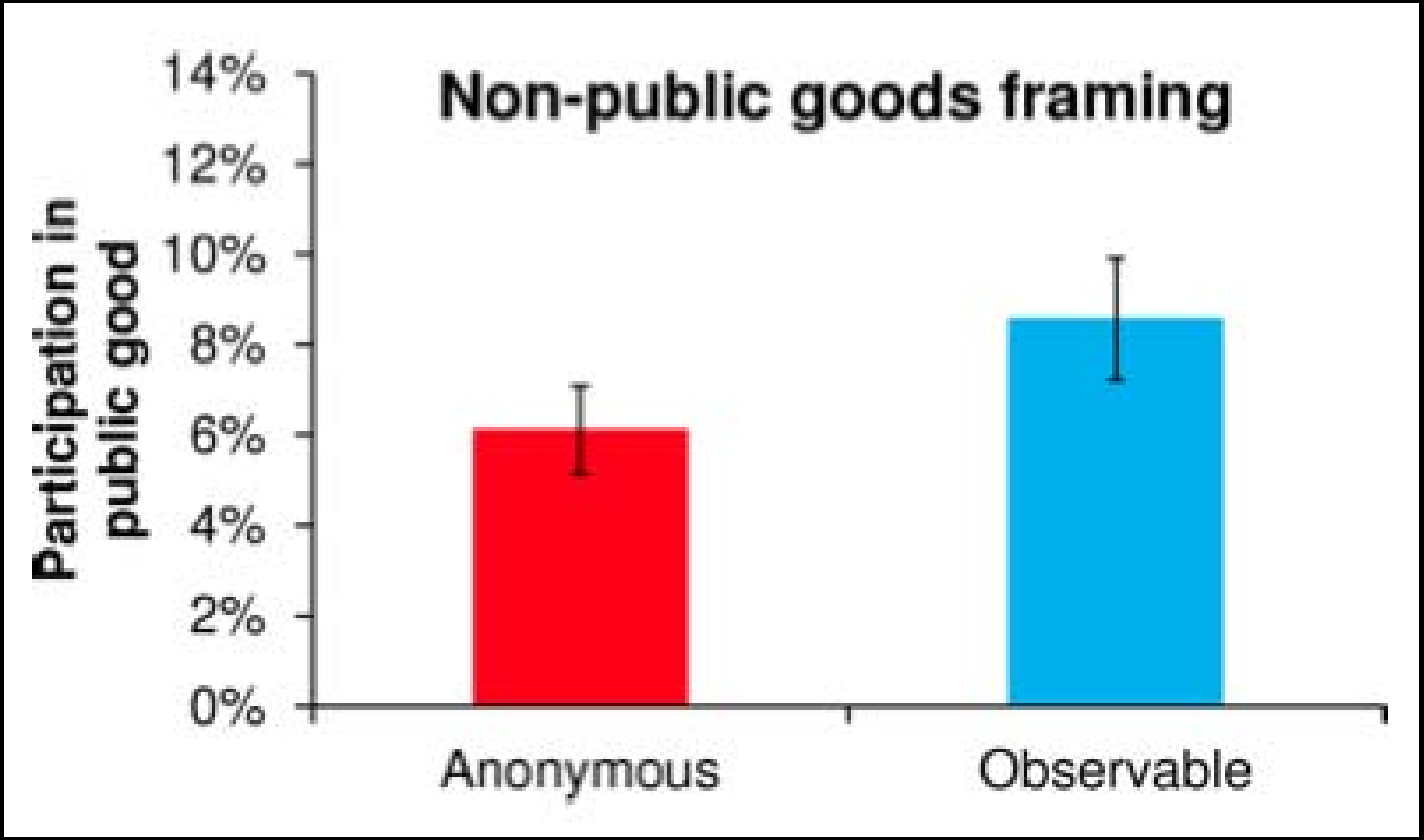
# The Experiment

We also double check that the effect is specific to public goods

Add additional subjects

Replicate experiment, except strip public good language from the marketing materials

# Results:





Source: Yoeli, Erez, Moshe Hoffman, David G. Rand, and Martin A. Nowak. “Powering Up with Indirect Reciprocity in a Large-Scale Feld Experiment.” Proceedings of the National Academy of Sciences of the United States of America. 2013. 

N=1,005

Increasing observability made a difference in a real-world public good

Worked much better than cash incentive

Let’s stay in the field…

Theory predicts that if people are going to defect, they'd rather not be seen doing it

And if they have the opportunity not to be seen, will sometimes take it and defect

Here is evidence that they do…

# Researchers teamed up with the Salvation Army to fundraise outside a supermarket

# There were three doors


Figure 1: The Store Studied. Doors 1 and 2 were the main entrances, while Door 3 was the side "recycling" door.

Researchers varied whether volunteers fundraised in front of one door or two

And watched what happened to traffic out of the store…

## Table 2: Store Traffic

| | Asking Condition | | | | |
|---|---|---|---|---|---|
| | **Silent Opportunity** | | **Direct Ask** | | |
| | Doors with Solicitors | | Doors with Solicitors | | |
| | Door 1 (Opp1 ) | Doors 1&2 (Opp1&2) | Door 1 (Ask1) | Doors 1&2 (Ask1&2) | Total |
| **Gross Traffic** | | | | | |
| Door 1 | 2,563 | 2,508 | 1,728 | 1,918 | 8,717 |
| Door 2 | 2,284 | 2,174 | 2,321 | 2,166 | 8,945 |
| Total Passings | 4,847 | 4,682 | 4,049 | 4,084 | 17,662 |
| **Imputed Traffic, Preferred Specification, Opp1 Total as Baseline** | | | | | |
| Total Passings | 4,847 | 4,847 | 4,847 | 4,847 | 19,388 |
| Door 3 Increase | 0 | 181 | 798 | 763 | 1,742 |
| **Traffic As Percent of Total** | | | | | |
| Door 1 | 52.8% | 51.7% | 35.6% | 39.5% | |
| Door 2 | 47.1% | 44.8% | 48.1% | 44.6% | |
| Door 2 Increase | 2.3% | - | 3.5% | - | |
| Door 3 Increase | 0% | 3.7% | 16.4% | 15.7% | |

The number of people leaving the store mysteriously fell

They were leaving through door 3!

People went out an inconvenient side door to avoid the Salvation Army volunteers, so as not be seen defecting

This doesn’t have to be conscious, if internalize punishment through emotions such as guilt

To see this, let’s think back to our puzzle from the beginning of class…

# Promised donations by subjects who paid to exit from a dictator game

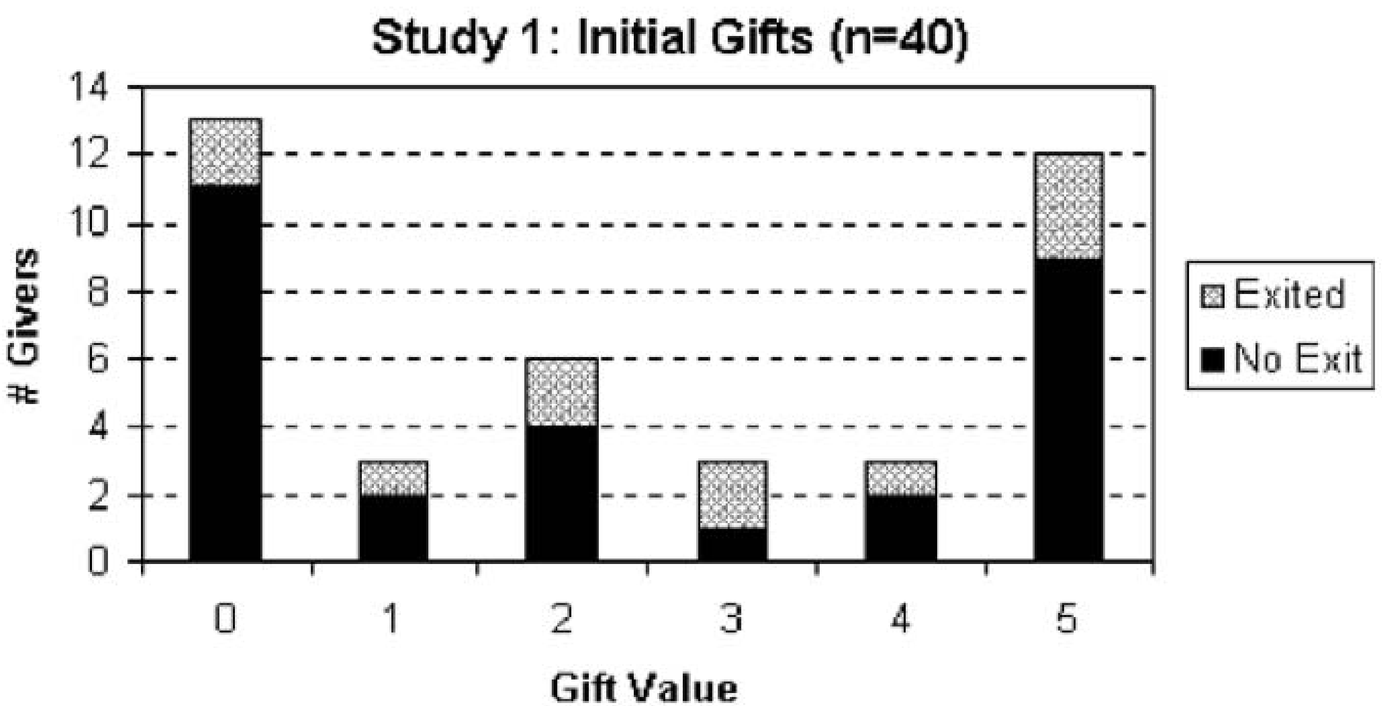



Fig. 1. Histogram of initial gifts in study 1.

There, too, subjects were willing to pay a cost to avoid being seen defecting

Next…

… evidence that altruists are not sensitive to effectiveness of gift

Reminder…

The theory predicts others focus on our actions, not their effects

So we focus on actions, not their effects

Here is some evidence for this…

Researchers teamed up with a political organization to raise funds

Solicited funds via a mailer

Offered a “match” which varied

## MATCHING GRANT
# NOW IS THE TIME TO GIVE!

Troubled by the continued erosion of our constitutional rights, a concerned member has offered a matching grant [4 treatments: of $25,000; of $50,000; of $100,000; blank] to encourage [redacted] at this time. To avoid losing the fight to defend our [redacted] this member has announced the following match: [3 treatments: $1; $2; $3] for every dollar you give. So, for every [3 treatments: HPC*1.00; HPC*1.25; HPC*1.50] you give, [redacted] will actually receive [$x]. Let's not lose this match -- please give today!

Giving is a lot more effective when the match is 3x than 1x

When the match is 3x, a $1 contributes $4

When it is 1x, it is only $2

Did givers respond?

# Not really…

| | Control | Treatment | 1:1 | Ratio 2:1 | 3:1 |
|---|---|---|---|---|---|
| Implied price of $1 of public good: | 1.00 | 0.36 | 0.50 | 0.33 | 0.25 |
| PANEL A | (1) | (2) | (3) | (4) | (5) |
| Response rate | 0.018 | 0.022 | 0.021 | 0.023 | 0.023 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Dollars given, unconditional | 0.813 | 0.967 | 0.937 | 1.026 | 0.938 |
| | (0.063) | (0.049) | (0.089) | (0.089) | (0.077) |
| Dollars given, conditional on giving | 45.540 | 43.872 | 45.143 | 45.337 | 41.252 |
| | (2.397) | (1.549) | (3.099) | (2.725) | (2.222) |
| Dollars raised per letter, not including match | 0.81 | 0.97 | 0.94 | 1.03 | 0.94 |
| Dollars raised per letter, including match | 0.81 | 2.90 | 1.87 | 3.08 | 3.75 |
| Observations | 16,687 | 33,396 | 11,133 | 11,134 | 11,129 |

# Not really…

| | Control | Treatment | 1:1 | Ratio 2:1 | 3:1 |
|---|---|---|---|---|---|
| Implied price of $1 of public good: | 1.00 | 0.36 | 0.50 | 0.33 | 0.25 |
| PANEL A | (1) | (2) | (3) | (4) | (5) |
| Response rate | 0.018 | 0.022 | 0.021 | 0.023 | 0.023 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Dollars given, unconditional | 0.813 | 0.967 | 0.937 | 1.026 | 0.938 |
| | (0.063) | (0.049) | (0.089) | (0.089) | (0.077) |
| Dollars given, conditional on giving | 45.540 | 43.872 | 45.143 | 45.337 | 41.252 |
| | (2.397) | (1.549) | (3.099) | (2.725) | (2.222) |
| Dollars raised per letter, not including match | 0.81 | 0.97 | 0.94 | 1.03 | 0.94 |
| Dollars raised per letter, including match | 0.81 | 2.90 | 1.87 | 3.08 | 3.75 |
| Observations | 16,687 | 33,396 | 11,133 | 11,134 | 11,129 |

# Not really…

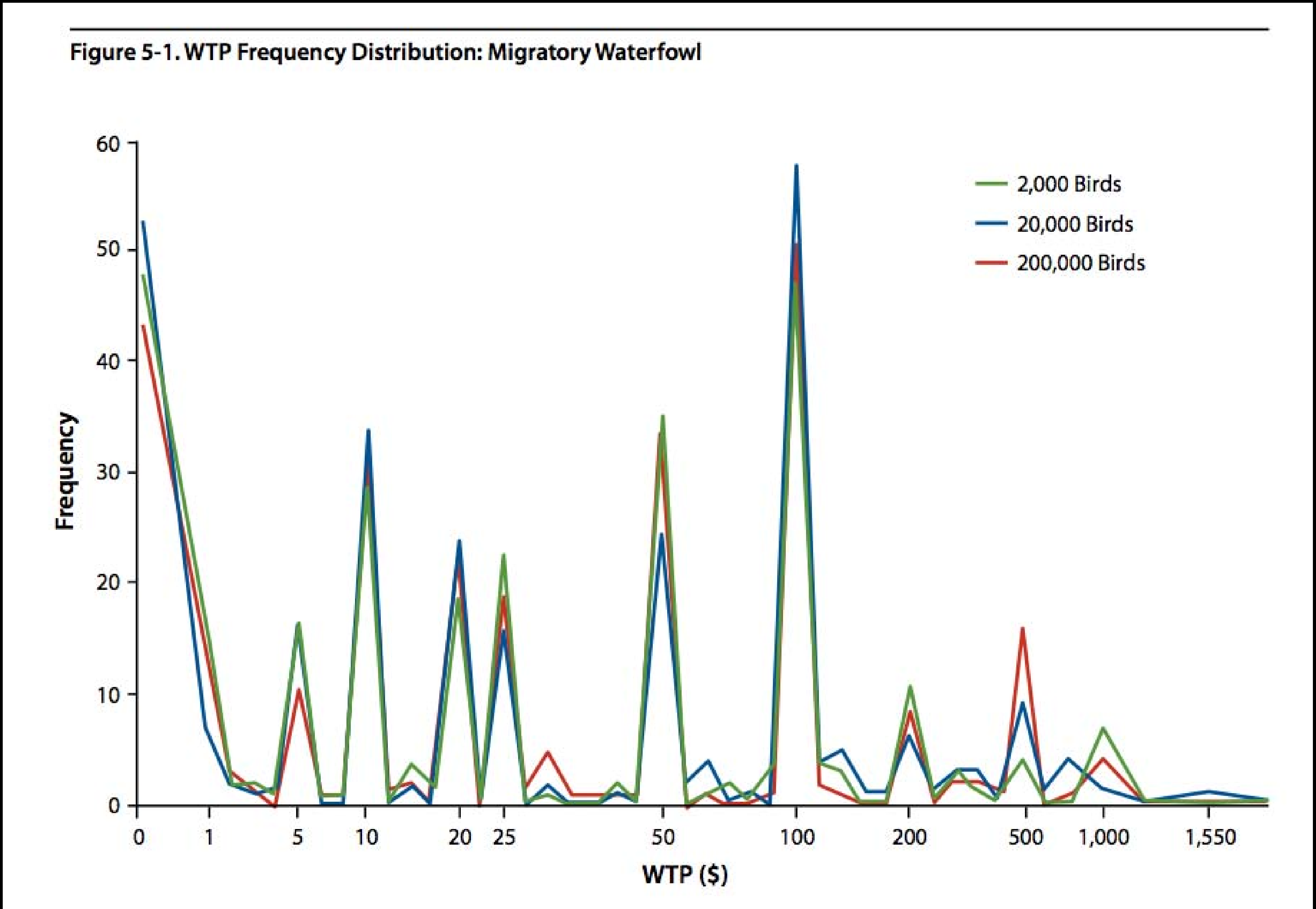
| | | | Ratio | | |
|---|---|---|---|---|---|
| | Control | Treatment | 1:1 | 2:1 | 3:1 |
| Implied price of $1 of public good: | 1.00 | 0.36 | 0.50 | 0.33 | 0.25 |
| PANEL A | (1) | (2) | (3) | (4) | (5) |
| Response rate | 0.018 | 0.022 | 0.021 | 0.023 | 0.023 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Dollars given, unconditional | 0.813 | 0.967 | 0.937 | 1.026 | 0.938 |
| | (0.063) | (0.049) | (0.089) | (0.089) | (0.077) |
| Dollars given, conditional on giving | 45.540 | 43.872 | 45.143 | 45.337 | 41.252 |
| | (2.397) | (1.549) | (3.099) | (2.725) | (2.222) |
| Dollars raised per letter, not including match | 0.81 | 0.97 | 0.94 | 1.03 | 0.94 |
| Dollars raised per letter, including match | 0.81 | 2.90 | 1.87 | 3.08 | 3.75 |
| Observations | 16,687 | 33,396 | 11,133 | 11,134 | 11,129 |

Next…

More evidence that people aren't sensitive to effectiveness of their action…

In a survey, researchers asked,

“Keeping these factors in mind, what is the **most** that your household would agree to pay **each year** in higher prices for wire-net covers to prevent about 2,000 (or 20,000 or 200,000 depending on the version) migratory waterfowl from dying each year in waste- oil holding ponds in the Central Flyway?“






Source: Desvousges et al., 1992

The theory offers a similar explanation to the puzzle from the beginning of last class…

An estimated 31.4–35.9 million have HIV/AIDS worldwide

1.7 million die annually

Source: World Health Organization

Since theory predicts we focus on action

We don’t need to know the magnitude of the problem

Last piece of evidence…

Just saw evidence that people don’t care about effectiveness

Sometimes do they not only not care, but they prefer not to know

Here’s the evidence…

Subjects played anonymous dictator games in the lab

# In the control group, played a typical variation of the game…

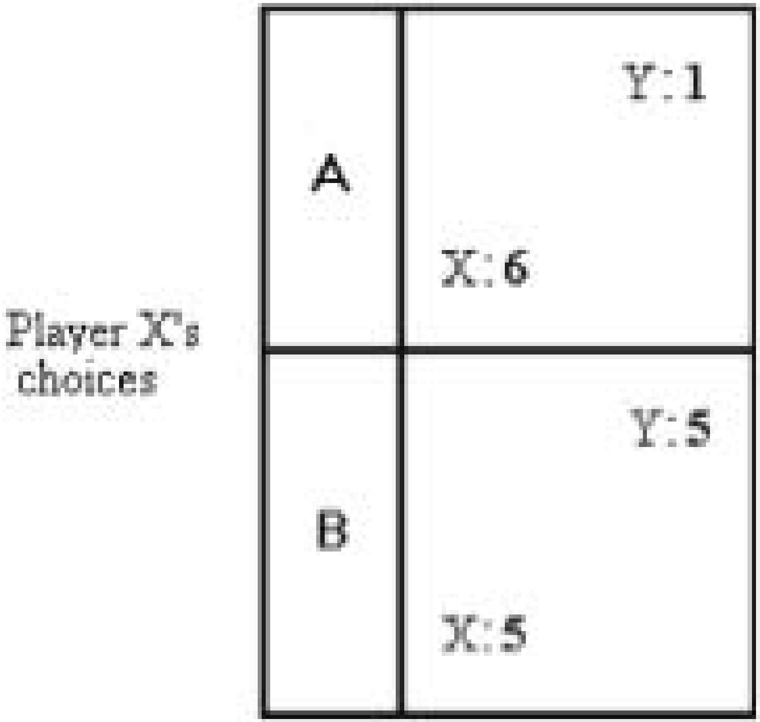



**Fig. 1** Interface for baseline treatment

# And, as usual, they were frequently altruistic…

**Table 1** Comparison of baseline and hidden information treatments

| Treatment | Proportion choosing "A" (unfair choice) | Proportion revealing true payoffs |
|---|---|---|
| Dictators' (Player X) choices | | |
| Baseline | 5/19 (26%) | |
| Hidden information (Matrix 1, baseline payoffs) | 10/16 (63%) | 8/16 (50%) |
| Hidden information (Matrix 2, alternate payoffs) | 13/16 (81%) | 10/16 (63%) |
| Recipients' (Player Y) hypothetical choices | | |
| Baseline | 0/19 (0%) | |
| Hidden information (Matrix 1, baseline payoffs) | 13/32 (41%) | |

# In the treatment group…

Payoffs were determined by a coin-flip

If game 1 chosen, dictator can ensure fairness

If game 2 chosen, dictator can ensure better for both

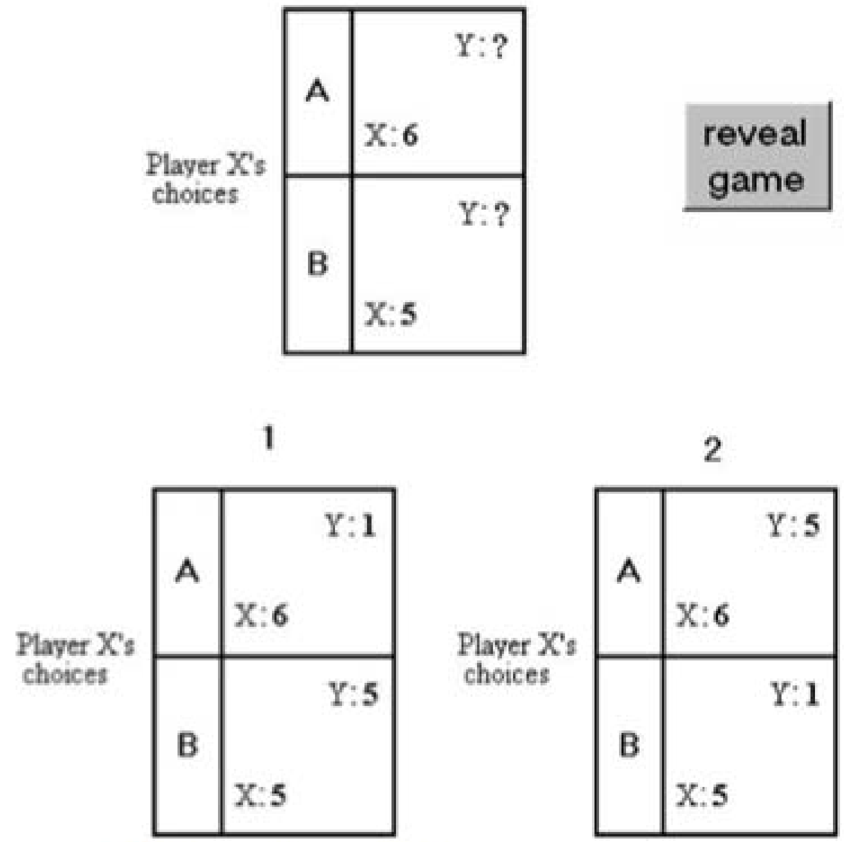



**Fig. 2** Interface for hidden information treatment

# In the treatment group…

Dictator given the costless option of revealing payoffs or remaining ignorant

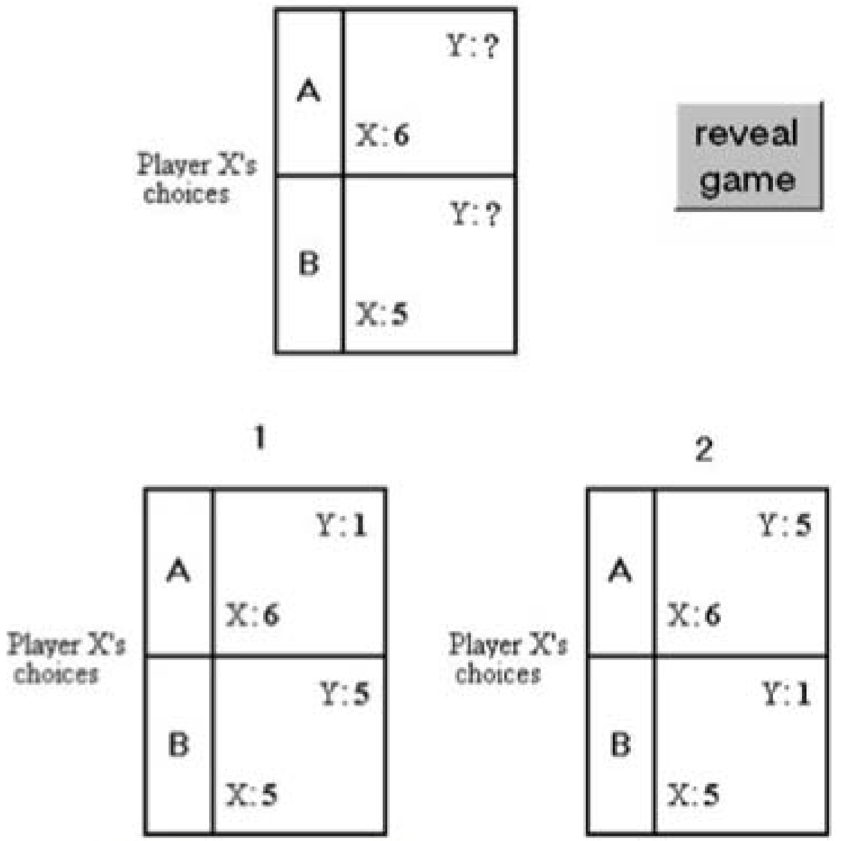



**Fig. 2** Interface for hidden information treatment

If the dictator cares about the receiver, should always reveal

Do they?

# No.

**Table 1** Comparison of baseline and hidden information treatments

| Treatment | Proportion choosing "A" (unfair choice) | Proportion revealing true payoffs |
|---|---|---|
| Dictators' (Player X) choices | | |
| Baseline | 5/19 (26%) | |
| Hidden information (Matrix 1, baseline payoffs) | 10/16 (63%) | 8/16 (50%) |
| Hidden information (Matrix 2, alternate payoffs) | 13/16 (81%) | 10/16 (63%) |
| Recipients' (Player Y) hypothetical choices | | |
| Baseline | 0/19 (0%) | |
| Hidden information (Matrix 1, baseline payoffs) | 13/32 (41%) | |

Remind you of anything?

- Why Habitat for Humanity is needed

Finally, we’re ready for policy implications

First one…

Increasing observability is feasible, cheap, and effective

Can be implemented in other settings to increase altruism

Second one…

People are not motivated to learn about efficiency

Making efficiency well-known will increase altruism

Magnitudes of effects from increasing observability are big. We expect the same for efficiency

Costs of implementing both are often low

Good place to start increasing altruism

Third policy implication is a bit of a political point…

Some political philosophers (e.g., Marx and Skinner) have argued that have urged social orders that rely on altruism

Our theory from class says you get altruism only under the right circumstances, and that it is of a particular type

Any political theory that wishes to exploit altruism must consider its limits, namely the two conditions we identified

E.g., cannot rely on anonymous cooperation without reciprocation (ehem, Marx)

Last piece of advice for both academics and practitioners…

When looking for ways to increase altruism, use the theory

How can we get people to be more prosocial?

In our 2-person analysis, we assumed that players remember their partner’s past actions

When they condition their strategies on partner’s past actions, they implicitly judge them

So we are effectively assuming players evolve or learn a way of judging someone’s past actions

I.e., a moral code

Now that players never play each other twice, it’s unrealistic to expect them to know partners’ past actions

But it’s perfectly realistic to assume they know others’ judgment of these past actions

We'll start with the following model (Ohtsuki and Iwasa 2006):

A population with a large number of players

Each period, players are randomly paired and play a prisoner's dilemma

With some probability, δ, the game continues

And the next period, the players are matched with someone else. They never play each other again

Each

They take one of two values: *good* and *bad*

We’ll explain where they come from in a moment

Players have strategies composed of two parts, (d, p)

Let's see what each part is…

Players choose their actions according to the rule p(i, j)

It depends on the player’s reputation, i, and their partner’s reputation, j

For example p(B,G) = C means a Bad dude who plays against a Good dude Cooperates

In addition, players have a rule for updating reputation *reputation dynamic* that assigns reputation, d(i, j, X)

It assigns a reputation based on the player's type, i, his partner's type, J, and the player's action X

For example, d(G, B, D) = G means, "Assign type G to a good player who defected against a bad player"

Note that the interpretation of δ is:

The probability the game is played again

And reputations are observed

And accurately conveyed

Etc.

We are ready to find Nash equilibria

Focus on symmetric Nash equilibria where everyone plays the same (d, p)

There are lots, but only a few cooperative ones

They have the following properties…

# Properties of Cooperative Equilibria

when more than 2 players

1) Cooperation can only occur when $\delta/(1+\delta) > c/b$

2) Reciprocity: p(G, G) = C, p(G, B) = D

Three others:

Justification of Punishment: d(G, B, D) = G

Identification: d(*, G, C) = G and d(*, G, D) = B

Apology: p(B, G) = C

So…

Even with more than 2 players, we get cooperation

And it has the same properties (plus a few)

OK, but we don’t always have diadic interactions

For example, all the most interesting puzzles from the beginning of class!

- Keeping oceans clean
- Preventing resistant antibiotics
- Voting
- Etc.

How do we model those?

Let’s introduce a new game called the *public goods game*

N players move simultaneously

Each player has two strategies, C and D

If play C, pay c to contribute b

If play D, pay nothing and contribute nothing

All contributions are shared equally by all players, whether they contributed or not

As in the prisoner's dilemma, Nash equilibrium is for everyone to play D

But it is socially optimal for everyone to play C

How

How do we incorporate this into a repeated environment?

Make the following modification to Ohtsuki and Isawa:

Simply assume that each period, with some probability, we play a public goods game

Otherwise, we play diadic relationships

We haven't proved it, but we're sure the results will be qualitatively the same

This would make a great final project!

OK, we’ve shown that cooperation is supported when there are two players

And showed you that it looks similar when there are more

And argued that the model could easily accommodate public goods and thus apply quite generally to altruism