

MITOCW | 20. Malleability and Inaccessibility of Preferences

[SQUEAKING]

[RUSTLING]

[CLICKING]

FRANK SCHILLBACH Let me get started on lecture 20. So I'm going to finish up with what we discussed last time, which is lecture 19. And then we're going to talk about lecture 20, which is about malleability and inaccessibility of preferences.

For now, we're going to talk for a little bit more about defaults and frames, and nudges in particular. So what we left it off last time was we talked about default effects. We talked about that in particular in retirement savings. Defaults can be sort of setting essentially in one default option. That is like, what happens if you do nothing in a retirement savings account.

Setting a certain default option can have very powerful effects in affecting people's behavior in a domain that's really important in this case, savings, and in a domain where traditional economic tools have not really gone very far, as in matching contributions, or any sort of like types of financial education, et cetera.

It's just not doing very much. It's expensive to do. And it's hard to change behavior, while using defaults is really powerful, and can cause large behavior change or large changes in outcomes even years later.

Let me talk a little bit about what is the optimal default decision regime. And that's getting into tricky territory. We talked about this last time a little bit. We'll talk about this again in the last lecture, on lecture 23, when we talk about policy.

So when do you think about active choice versus default, so what is active choice? Active choice to remind you is like somebody has asked at the beginning, or directly asked, you have to make a choice. And I'm just sort of forcing you to do it, more or less.

[INAUDIBLE] the company hires somebody and in the hiring package, there is just a form that you need to fill out. And otherwise you cannot really start [INAUDIBLE] company.

Sometimes that's not really enforceable, but effectively, essentially just everybody makes some choice. That's essentially active choice versus default, which is just like, if you do nothing, you either get zero retirement savings, or some positive amounts.

Now, if people are really different, then essentially, active choice is a great thing. Because if people really have different preferences, then we shouldn't just push them into one direction or the other. They should choose. Everybody chooses on their own what's best for them.

In contrast, if you think most people are the same anyway, and in addition, people don't really know what's best for them, well then it might be better to just provide one default [INAUDIBLE] most people that's best for most people. And then most people will stick to anyway, because these defaults tend to stick. And if people don't know what they want or what's good for them, then sort of letting them choose actively might make things potentially worse anyway.

And so that's sort of the fundamental trade off when people have come down on different sides of this. And it depends on your view, kind of, about sophistication and so on of customers who are [INAUDIBLE].

And one big and important issue with defaults and with the nudges, in general, which I'm going to talk about to you in a second is, we kind of want to make sure that we don't make people worse off in some ways by pushing people in certain directions.

For example, one thing you might say is, well, why not set the default to like 5% or 10% of retirement savings? But then you worry about people potentially over-saving in their retirement savings accounts. And then essentially they have lots of credit card debt, potentially, which might be even larger. Sorry, the interest rate on the credit card debt might be really large. And people get in trouble. Or people then withdraw prematurely, and so on, which is [INAUDIBLE] potentially worse for them.

We talked very briefly about the smart plan, which again, I'm going to get back to in lecture 23 as well.

Now I also talked briefly about sort of default effects also are quite powerful in other settings. For example, in organ donations, in some places, organ donations are explicit consent, which is like you have to opt in, and otherwise you will not be an organ donor.

In other countries-- and this is from a while ago-- in other countries, things are different, where there is what's called presumed consent, which is like you have to opt out. And otherwise, if you do not opt out, you're presumed to consenting to organ [INAUDIBLE]. We see essentially huge differences in otherwise quite similar countries, which sort of shows you that the default here can have really large effects.

There's some other examples of default effects in particular when it comes to voter registration, or green energy choices. In some cases, in particular when it comes to voter registration, it's actually quite hard to come with arguments against setting a default in a certain way. Like in some sense, getting voters to be registered when you get say a driver's license or the like seems just very straightforward and-- or when they turn 18, that seems like a very straightforward thing to do.

And there, in some ways, the defaults can really sort of remove some barriers and inefficiencies where we think we should agree that everybody should be able to vote if they would like to, with some exceptions, perhaps.

And so, if you can default people into getting registered to vote, that seems like a good thing, unambiguously. So there sort of default effects can be really, really powerful.

Now, let's talk a little bit about-- not just again, we're going to talk about this again in lecture 23, when it comes, about policy, but let me sort of define the issue and sort of be clear about that.

So you might have sort of come across the very famous book by Richard Thaler, who recently won the Nobel Prize in economics, and Cass Sunstein, called *Nudge*.

I'll give you the definition by Cass Sunstein, who says a nudge is a feature of the social environment that affects people's choices without imposing coercion of any kind of material incentive. So that-- default is sort of like a clear case of that, where essentially you just say, OK, if you don't do anything, I'm going to just pick a choice for you.

But you can choose, however, in whichever way you want. And there is essentially no material incentive. But I'm changing essentially your choice environment or what people would call sometimes the choice architecture.

And there's other versions of that. There's simplification, information disclosure. Again, these are all things that are usually available. People can figure it out on their own. But we make it easier for people.

And one of the big mantras [INAUDIBLE] about like, what's the main thing we should do to improve people's behavior? Dick Thaler would often say, make it easy for people. Essentially, people are often-- and he said that about himself-- people are lazy. And they make simple choices.

And if you make it easy for them to make the right choice, they will be more likely to make good choices for them. There's some other things like warnings, reminders.

Notice for reminders, for example, again, you already have the information that somebody gave it to you previously. So really it's only like if you have memory issues, that might be useful for you. There's no material incentives or [INAUDIBLE].

Use of social norms. That would be things like I'm telling you all of your friends did x. Now would you like to do x or y? Or I can send you letters and say your neighbors used so much energy, and you use more energy than your neighbor. And then I put like a sad face next to it, or something. And then you might feel bad about it because of social norms, and sort of reduce your energy usage, which people have in fact done.

You can also like increase ease and convenience. This is exactly what like [INAUDIBLE] convenience. That's exactly what Dick Thaler would say.

Those are things like, for example, I can put Snickers bars next to the counter at the cashier. Or I can put apples next to it. And people are more likely to buy apples when it's easy or simple or when they're reminded of eating apples.

And they're more likely to eat Snickers bars essentially if [INAUDIBLE] to them. To the extent that you want people to make what one might think of as good choices, making it easy for them to make the good choice without sort of restricting anybody's choice sets or coercing them in any way is considered a nudge.

And then there's also some things about framing of choices as in gains versus losses, in terms of their contracts. Any questions on this?

OK. So then let me give you, briefly, sort of-- and again, we're going to get back to all of this in lecture 23. But let me briefly sort of give you a couple of examples of nudges in some cases where it's pretty clear that some nudges are a good idea.

So one thing that we talked about quite a bit already when it comes to procrastination or self-control problems present bias, or [losing focus ?] is the health domain, and what people and in particular, individuals and society often have aligned goals, where individuals often want behavioral change. They want to improve their diets, increase physical activity, stop smoking, get vaccinated, use less energy, and so on.

And often, there are societal costs of these behaviors, or lack of those behaviors. Sometimes there's externalities from smoking or getting or not getting vaccinated. Sometimes [INAUDIBLE] health care costs and the like from obesity and so on, where it's bad for everybody if the population is sick.

So these goals are now aligned. And the social planner might want people to improve their diet, or the government might want that, or might want to reduce smoking and so on. But individuals, as we saw previously, often fail to follow through.

And sort of often, education and information interventions are often ineffective. Even price interventions are often ineffective, or in the sense of like increasing prices, or incentivizing people, it doesn't always work. Plus, it's quite expensive to do.

So one natural question you might want to ask is, can we use some nudges to align intentions and actions?

And the reason why this is really popular, in particular, among governments, is that it's really cheap to do this. So it's actually sending some reminders or the like or providing information in one way or the other, is essentially free. And if governments don't have a lot of money, that's an easy thing to do as opposed to paying people, which is quite expensive in many cases.

So here's an example of a free intervention which is about flu shot communications. This is a study by Katy Milkman and co-authors from 2011.

So here, the goal of the intervention is to get people to sign up, or to actually get flu shots. So the control group got an informational mailing. I'll show you this in a second. The treatment group got the same, plus an encouragement to make a date plan. And the treatment group got the same plus make a date plan, and a time plan. Let me show you what this looks like.

So here's the control condition. This is just an informational mailing. It's saying here, the company-- this is a company trying to do this-- is holding a free flu shot clinic. And companies very much want their workers to get flu shots, because that reduces essentially sort of sick days. If everybody gets their flu shot, not only are they healthier and it's probably good for them. They're like less at risk of dying and infecting each other, or just being sick. But it's also good for the company, because the company now gets workers to actually show up more, and have fewer sick days.

So here's sort of the control condition. It looks like a perfectly reasonable letter, where you say, here is these days, on which you can sign up. And you are informed about the dates and times of the workplace flu clinic.

Now, the treatment condition is the date plan condition, which essentially is asking people to-- or invite people to choose a concrete date for getting a flu vaccine. And the rest of the information [? mailer ?] is exactly the same.

So essentially, that's just asking people, pick a day of the week and a month, or a month and a day, and a day of week, and just write it down. And presumably, that also then encourages people to put it into their calendar or the like and sort of acts as some form of a reminder.

Notice, there's no incentive here. There's no financial incentives. There's no coercion and so on. It's just saying why don't you just pick a day right now. Just have a look, check your calendar, and just figure out. Make a plan for this.

And the second condition is then make a date and a time plan condition, which is to say, why don't you also pick a time, which is even more concrete. That would be Monday, October 26, at like 9:00 AM or whenever I'm going to go. Presumably, that also encourages people to put it into their calendar.

Now what you then get in this experiment-- it is a very simple experiment. You get, when you just send a letter, 33% actually get the flu shot. When you also send a date plan, 34%. And when you send a date plan and a time plan, it's 37%.

Now, these effects are a little bit underwhelming in the sense they're kind of small. It's like 1.6 percentage points at 4.2 percentage points. And in relative terms that's not huge effect.

But notice what's beautiful about here is that the costs are essentially zero. So in some sense, if you value additional flu shot adherence, if you value more people getting flu shots, then a 4.2 percentage points increase, which is about like a 10% or whatever, 11%, 12% or something increase relative to the control group, that's quite a bit of an increase for something that's essentially free. So the cost effectiveness of this is, of course, very, very high. Because it's essentially free, or maybe it cost somebody a few minutes to sort of write this down, or design the sheet.

But really it costs exactly the same to send the sheet here on the left versus the sheet here on the right. So that's kind of an example of a win-win kind of situation where we make things essentially more effective. It's not really hurting anybody, and it's cheap and free to do. And so everybody is happy. And it's sort of uncontroversial in the sense of nobody would argue that this is a bad thing.

Of course, some people might sort of have some concerns about the flu shots themselves, which I don't think there's any actual evidence against that as far as I know. But that's sort of an example where we say, look, this is very cheap to do. So presumably you should sort of pick the right mailing, or the right design, if I'm mailing, to encourage people to do desirable things, at least from the company's perspective, while also preserving people's freedom, not spending any money, but also not encouraging anybody. Any questions on this?

OK. So here's another example of a similar intervention, which is signing up for FAFSA. This is a paper by Bettinger et al. in 2009, where essentially, people were provided free additional assistance in completing and filing applications for college financial aid. This is FAFSA, the financial aid for college.

And it really increased college enrollment-- not only FAFSA completion, but also college enrollment. And there's sort of different conditions. One is control versus just providing people information only. Here's [INAUDIBLE] information of like you could get some FAFSA help and so on.

But it seems to be really what you need is some additional assistance of completing and filing this application. Just somebody essentially walking you through the filing has a huge impact.

There's a bit of a question of, is that really just a nudge, or is it like a more powerful intervention. But it's a really minor intervention where some people just need some level of help in filling out this form. And it has pretty large effects, if you look at the FAFSA [INAUDIBLE] completion of these forms, reasonably large effect on college enrollment.

And you might sort of think this is a hugely important choice, and sort of if you get people to make that really big change, that's pretty remarkable for doing some very simple things. [INAUDIBLE] help me complete and fill out a simple form, something that you could have done anyway.

So, and here again, so this is a relatively cheap thing to do. And it's essentially equivalent to the impact of several thousands of dollars of education subsidy. So if you wanted to essentially pay people to go to college, [INAUDIBLE] subsidize that, which you might want to do for other reasons anyway, one key issue with that is that not only is it perhaps not particularly effective, but also you're going to pay a lot of for marginal people.

You're going to pay a lot of people who would have gone to college anyway. You're going to pay the subsidy. So if you have a low budget, that's really tough to do. And sort of then there's not a lot of-- the cost effectiveness tends to be very low.

Because you have essentially a low impact, but like really high costs. And you can read about this more in the link that I had.

But essentially the bottom line here is, again, you can do very small changes that can have pretty large effects. There's other types of examples that government use, or where governments use-- not just-- another example would be sending out letters for people to pay their taxes, which essentially are like reminders, or often appealing to people's sort of prosociality in various ways. But the government sends you letters. And there's different ways in which you can write these letters.

You could say you should pay your taxes, or everybody's paying their taxes, or good citizens pay their taxes, and so on. And people-- there's a whole industry of nudged units where people then try to figure out what's the best way of doing that.

And by doing that in different ways, that can have a pretty large effect on people's behaviors. And the governments sort of love this, because it's a very cheap thing to do. And it can have large impact and change behavior quite a bit. In this case, in the tax case, increase revenue by quite a bit without costing the government very much.

Because often, [INAUDIBLE] send some letters anyway. And they send some more in some less effective ways of doing that.

When you think about the intervention that I have here on the slide, the FAFSA one is essentially just-- the effects almost surely will be persistent in the sense of like, this is a one shot choice, whether you send your kid to college or not.

And once you increase the college enrollment, which might have some other issues, but once you do that, there's going to be persistent effects. Because once kids go to college, that has effects on their behaviors and so on.

Now when you're saying something else, which is some other decisions, [INAUDIBLE] just like about paying taxes where every year you get this letter from the government that sort of tells you you should pay your taxes, maybe the first time it works. Maybe the second time as well. But at some point, you're like, whatever. I don't care.

That's a great question. And I don't know whether we have lots of evidence or not. I think the governments would sort of-- I think there's some research on this that I just am not sure I'm aware of.

I think governments or people would say, well, in some sense, we don't care that much about it. Of course we care. But it doesn't change the fact that in the first year-- so we know that in the first year, it works pretty well. And there's some things that work better than others.

Many governments send some letters or some forms of mailers anyway. So what you want to do is kind of like make sure you optimize, at least the first or second time you contact people. And that might have some persistent effects by themselves, even if you don't re-contact them.

Because once you start paying your taxes, maybe then it becomes a habit, and people do it anyway. But even if that's not persistent, governments would say, well, if that gets people to pay taxes more at least once or twice, that's worth sort of the money spent. [INAUDIBLE] somebody essentially needs to design this letters. There's like, templates of what you can use.

And that's essentially very cheap to do and has pretty large effects at least in the short run. I think [INAUDIBLE] correct, which is sometimes it's like, some effects-- it's just information versus some things about making you feel bad, or social norms. I think the information things-- if you think people just forget stuff-- for example, if people forget to get their flu shot and so on, you should probably send reminders and so on and get them to do that.

If there's sort of these social norms, or sort of [INAUDIBLE] makes you feel bad in some ways, maybe those things are maybe more likely to potentially at least go away once you do it like 17 times, because at some point, people are like, yeah, whatever, the same letter I got always. And I haven't paid my taxes last time. I'm not going to pay them now either.

It's additional assistance in completing and filing applications. So essentially, it's not-- so the financial aid is always available. It's just who takes advantage of financial aid depends on how easy you make it to people, and whether you support them, provide some support in completing and filing applications.

There's also [INAUDIBLE] with that where essentially, once you do your taxes, essentially, it already prefills you the forms, which is even easier to do. But essentially, this falls under the category of making things easy for people. And that can sort of change behaviors quite a bit.

But the key part in the nudges is-- and let me get back again to the definition-- is that there's no financial incentives. So this part here is without imposing any coercion, or without paying any kind of material incentives, or at least, only very, very small incentives. There's a bit of a question how to define that.

But once I pay you \$1,000 to go to college or whatever, subsidize you, that's not a nudge anymore. Or once I'm taking away certain options from you, [INAUDIBLE] eliminating choices from your choice set, again, that's not a nudge. That's something else. That's changing your choice set.

This one is-- nudges are explicitly about keeping people's choices fully free and available, but making it easier for them to make a certain choice. Presumably, a choice that's desired in some ways by the social planner, or about people's plans for their future.

So anyway, there's a large sort of set of examples of different nudges and so on. And they can be fairly effective. Of course, not for everybody and not for the fraction of people who are swayed by these nudges is not huge, in part because if you think about it, essentially people need to be marginal in some ways.

If people just really, really don't want to go to college because it's too expensive for them, or whatever for other reasons, or if people really want to go to college anyway, the nudge will not change that behavior. Whether somebody fills out the form or not doesn't really matter.

But if somebody is sort of like, ah, maybe I should send the kid to college or not, but now [INAUDIBLE] form and then they forget about it and so on and so forth, for that kind of people, there will be potentially large effects.

But sort of then by definition, in some ways, the fraction of people who are actually marginal will not be huge. But again, since the costs tend to be extremely low, it tends to be very much cost effective to do so.

Now again, sort of previewing a little bit what comes in the policy lecture is, so minor interventions can have large effects. But nudges can often achieve sort of-- in some cases, nudges can achieve unambiguous improvements.

Like, if you send people reminders and they forget to get the flu shot, and now they do the flu shot, or if people don't take their medication, and you remind them, or provide information, or some other forms [INAUDIBLE] [? tell ?] people to take their medication, it's pretty clear that that's making people better off.

But there's also a bunch of challenges and other situations which is kind of like-- in some cases, it's not at all obvious. Like, which nudge to choose-- are we making everybody better off? Are some people made worse off?

For example, should everybody save for retirement? Are we pushing people to save too much? Should everybody go to college? Maybe in some ways, for some people, it's more suitable than for others to go to college. Often it's a huge financial expense for the family.

And if then the job prospects are not necessarily better compared to not going to college, then it's not clear that one should do that.

There's also some evidence or some concerns about nudges making people feel bad. So you get these letters that says you're destroying the environment, and [INAUDIBLE] all these sad faces, and you're worse than your neighbors, people might just feel bad about it. And we should put some weight on that.

If you send these letters to 1,000 households, and 50 households or 20 households change their behavior in certain ways, but 300 households feel bad about it because they feel uncomfortable getting these nudges, or you get all these phone calls and spam about the environment from sending too many letters, we should put some weight on that.

And then it's like, then there's some trade offs. It's not like entirely free. And then there are some costs and benefits of that.

There's also some questions about which self should we respect. We talked about this a lot before, which is if one self wants to be very virtuous and exercise a lot and so on, and the other one wants to sit on the couch and watch TV, or if one self wants to smoke and the other one does not, it's not obvious that we should respect the long run self compared to the short run self that wants to just enjoy themselves, and so on.

And sort of now there's tricky issues [INAUDIBLE] are we actually making people better off or not. And we'll get back to these issues in the last lecture, which talks about policy. Any questions on this lecture?

This is actually lecture 20 about malleability and inaccessibility of preferences. This is, in a way, a more radical deviation from neoclassical economics. Because so far, we have always said, people know their preferences. People have certain beliefs. Their beliefs might be wrong, and so on.

But crucially, people knew what they were doing, and they were deliberately making certain choices. And whether their preferences were perhaps including present bias, or social preferences, a reference dependence or the like-- but these preferences were given and fixed. And people knew what those preferences are.

Now we're going to deviate from that and think about [INAUDIBLE] evidence that people might not actually A-- these preferences might be malleable. And B, people might not even understand why they want what they want. And sort of the preferences are sort of quite mysterious to people. And it's quite easy in some ways to manipulate people in some ways, without them even understanding it, knowing about it.

So I tell you about some very fascinating psychology work, which is called-- the paper is called "Telling more than we can know" from Nisbett and Wilson, which is a classic paper in social psychology. And then we're going to talk a little bit about willingness to pay, and then the paper by Ariely et al on coherent arbitrariness, which you were supposed to read for today.

OK. So let's step back a little bit, and talk about the Nisbett and Wilson paper. So this is-- so many questions-- you might have any questions about cognitive processes underlying our choices, evaluations, judgments, and behaviors.

You might wonder, why do you like a certain person versus not? What do you like about them? And why do you like this person and not another person? And why you're friends with this person versus somebody else? And why do you like Math versus English, and so on and so forth?

How did you solve a certain problem once you were sort of asked to solve some problems? How did you come up with a solution? Why did you take this job, or why did you take this class, versus another class?

And you might sort of say, well, I like this one and not that one. And I really like sort of like computer science, versus math versus economics.

But when you then ask [INAUDIBLE] yourself for others, why exactly do you like that one, versus something else? You'll realize quickly that people often don't quite have a good answer to those questions.

And Nisbett and Wilson's fairly provocative paper says, essentially, we have no idea where these preferences are coming from. And we're just making stuff up.

And so let me give you some examples of that. The first example is what's called Maier's two-string problem from 1931. This is [INAUDIBLE] one of these classic psychology experiments.

The experiments works as follows. There's two cords hung from the ceiling of a lab with many objects, such as poles, ring stands, clamps, pliers, and extension cords.

Subjects are now asked-- are told that a task is to tie the two ends of the cords together. So there's two cords hanging on the side of each side of the room. And subjects are supposed to tie these two cords together.

The problem is that the cords are placed far apart from each other, such that subjects can't, while holding onto one cord, reach the other. So you can't just take one cord and then sort of like try to reach the other, because it's too far away.

And so there's all these different objects in the room. And subjects are sort of trying to figure out how to do this.

And so then they usually come up with one or two solutions that are not really solutions, because it doesn't really work. So they can get [INAUDIBLE] like extension cord, they use the pliers, and so on, and the clamp, and some ring or whatever. And it just doesn't work.

And then they sort of try it and keep trying, but it doesn't really work out. And then they're told to do it another way.

And then Maier, the experimenter, walks into the room at some point, or walks around, and essentially, accidentally-- this is sort of like on purpose of course-- accidentally puts some of the cords in motion. And subjects then very quickly afterwards figure out the solution within the next 45 seconds.

Of course, the reasoning here is, why they figure it out is because now the cords are put in motion. And subjects sort of figure out how to do this.

And what you're supposed to do is essentially put the cords in motion. And once you put them in motion, you can essentially reach them both and then put it together.

I have a video here. I'm a little traumatized the last time where it didn't work out. So I'm not sure [INAUDIBLE]. Now I don't even see it. So maybe let's just skip it. But you can watch the video of that. It's the first two and a half minutes video. [INAUDIBLE] see anything. But there's a video that sort of demonstrates this task that you can sort of just skip this.

So now then afterwards, people are asked, well, how did you come up with the idea of using a pendulum? And people then have all these like explanations. Well, it just dawned on me. It was the only thing left. I thought about all these different things. And at the end, I came up with this other solution, which is using these strings as a pendulum.

Or I just realized the cord would swing [INAUDIBLE] the weight on it. So what the solution at the end is like, you put the plier on one of the cords and then swing it. And then you sort of use the other one. And then you could put them together.

There's one Harvard psychology faculty subject, in fact, had the following explanation, which is, having exhausted everything else, the next thing was to swing it. I thought of the situation of swinging across a river. I had an imagery of monkeys swinging from trees. This imagery appeared simultaneously with a solution. The idea appeared complete.

Now of course, the problem here is that-- the problem here is that's all bullshit. Because the reason why people come up with a solution is because, as I told you, Maier was walking in and accidentally putting the cords in motion.

And there's differences in timing. Sometimes they would come in earlier, sometimes later. And every time the experimenter walks in, people would sort of accidentally then figure it out, and then come up with these elaborate explanations.

And the reason why we know that this is because of the experimenter are coming in is because if the experimenter doesn't accidentally come in, and puts the cords in motion, then essentially, people are not figuring out.

So we know the true causal effect here is the experimenter coming in and suggesting a solution to people. But people then make up all these stories, including like the monkeys swinging from trees, how they came up with a solution, which is clearly not how they came up with the solution. The reason why they came up with the solution is because Maier touched the strings.

OK. Now there's a second example here, or three examples in total. So let me tell you about example number-- or let me ask first, is there any questions about this study? By the way, you should lower your hands, if you asked [INAUDIBLE] previously, like [? Brian ?] and [INAUDIBLE]. Unless you have have additional questions.

Let me tell you the second study now. This is a study by Latane and Darley, which is, I told you this a little bit about [INAUDIBLE] these are the impacts of bystanders and witnesses on helping behaviors.

I told you about the story of the Good Samaritan before. This is sort of a similar study. And this is essentially a situation that's, again, created by social psychologists. It's not a real situation, but people are meant to think that it's real.

And so here, the more people over here, someone in another room, having what sounds like an epileptic seizure, the lower the probability that any given individual will rush to help. And you get similar results for individuals' reaction to dangerous-looking smoke coming out of the ceiling of a room. So like there's people in the experiment. And they have smoke coming out in a room.

And the more people that are there, the less likely is any given person, perhaps because they're freeriding is to actually do something about it. And this looks kind of very dangerous. But nobody does anything, because essentially other people are also not doing anything.

Now when you then ask people afterwards, you know, what influenced your choices? And essentially, when you ask them directly and tactfully and bluntly, you always get the same answer, that essentially, they think their behavior had not been influenced by the other people present.

They think you're like-- for whatever reason, they didn't do anything, or maybe they thought it wasn't so bad, or there wasn't really danger, or whatever. But of course, these are randomized experiments. We know that some subjects were influenced by the presence of other people, because precisely, the experiment was creating that variation, and it affected people's behavior. [INAUDIBLE] people essentially just don't understand that, and sort of then come up with other explanations of why they do or did what they, in fact, did.

And number three is sort of very similar. These are called erroneous reports about position effects. These are studies where certain items are positioned in different ways, to get people to-- when people were evaluating them.

So passerbys were asked to evaluate some clothing. And they were asked about the quality and the preferences-- the quality of certain goods and people's preferences between different goods.

And the way this experiment was set up-- and this is sort of essentially like a marketing experiment if you want-- essentially, there's a pronounced left-to-right position effect. The rightmost object was heavily over-chosen in these experiments.

And that's in some sense irrelevant why that's the case. But essentially, from previous experiments, we know that essentially the rightmost object is most likely to be chosen, perhaps because people look at it first. [INAUDIBLE] look at it last and so on, it doesn't matter.

What's important for our purposes is that the rightmost object is most heavily chosen. People like that object the best in these experiments.

And then there's going to be randomization in the positioning of experiments. And people are then asked, well, why did you choose this sweater versus another?

And no subject whatsoever ever mentions the position of the article in the room, and virtually all subjects denied it was, when asked directly about the possible effect of the position of the article. And people come up with all these explanations. It's like, green is really always my [INAUDIBLE] favorite color. And this sweater is really fluffy. And this and that. And I really like it for reason x and y.

Of course that's true for some people, but we know that at least some people must be swayed by these position effects, because we set up an experiment to precisely do that. And then that essentially swayed behavior in a very predictable way.

So essentially, there are these determinants on people's preferences of behavior, that people do not seem to understand.

OK. So let me summarize what did we learn. So there are many instances in which subjects have no idea why they choose what they choose. And then people-- [INAUDIBLE] can read this more in this paper, which I think is a beautiful paper-- people appear to make up stories that are based on their a priori, implicit causal theories.

What I mean by that, essentially, they have some theory of why they do certain things. And when you ask them, then, OK, well, why did you do what you just did? People will sort of come up with some ways of justifying [INAUDIBLE] in some ways their behavior, based on some theories that they had about themselves.

And sort of-- I always like fluffy sweaters, and therefore you choose the sweater. Of course, it happens to be the only [INAUDIBLE] sweater if it's [INAUDIBLE] on the right versus on the left. But people then sort of have their ways in which they then explain what they do based on essentially things that they sort of essentially make up. Any questions on this before I move on?

And I encourage you to read the paper, more for fun than for anything. It makes you kind of wonder quite a lot in why you do what you do, and why you prefer certain things versus others.

So let's briefly talk about two experimental design tools, which will be useful for the coherent arbitrariness paper. One is what's called the strategy method, which you already came across a little bit. And the second one is what's called the BDM, the Becker-DeGroot-Marschak procedure for eliciting valuations. I'll try to be quick.

So we're often interested in behavior in rare contingencies. So often, we may ask the question, how would people behave in many different contingencies? Some of them are often quite rare, that don't happen very often.

And why do we care about such contingencies? Well, sometimes we just care about them like [INAUDIBLE] it's inherently important how people behave or prevent certain contingencies, disasters, earthquakes, droughts, et cetera. We care a lot about these things, even if they're sort of rarely happening. So we kind of want to know how people behave in certain situations.

But events and rare contingencies also can affect, on top of that, events in likely contingencies. Here's a simple example.

If your roommates think you'll punch them in the face if they borrow your stuff without asking them, they will not do it. Of course, hopefully, you are not punching your roommates, and I very much do not want to encourage you to do so.

But here, the key part here is that punching is rare but important, precisely because you might do so in case they misbehave, they will not borrow your stuff or steal your stuff in the first place. So essentially, sort of these [? equilibrium ?] or rare contingencies might be quite important, not because we think they actually happen very often, but they sort of discipline behavior in other cases.

Now, what's [? known ?] as strategy method? Well, the strategy method helps you to elicit behavior in many potentially rare circumstances by asking subjects what they would do with a choice implemented if the circumstances arises.

That is to say, I'm asking you-- for many different cases, I'm going to implement one of those cases. And that's going to allow me to be [INAUDIBLE] [? compatible ?] in various ways.

So since the decision does count if the contingency occurs, subjects have an incentive to choose correctly for each contingency. That is to say, I could say, suppose this is what we're doing in class in some sense. We talked about social preferences. I said, you know, I'm asking you to make a choice. I'm going to only pick one of you to implement that choice. And you have the incentives now to answer truthfully, because it could always be the case that your choice will be implemented.

And so that allows then the experimenter, me in that case, to generate a lot of data from one simple experiment, where essentially, you can give a subject many different decisions, and then say, only one of your choices will actually be implemented for sure. Or you can ask many people the same question and say, [INAUDIBLE] only one of you, or one of your choices will be implemented.

And so then that's essentially incentive-compatible, because it can always be that your choice is the one, or one of the choices is the one that counts. And when you look at sort of experimental evidence, it seems to suggest that the strategy methods are asking, essentially, if I can ask you-- so there's experiments where I can ask you 100 different questions, and only one of them will be implemented, versus a different group would be asked-- a randomized group would be asked only one question. It turns out that people actually answer in experiments these questions pretty similarly in both of these cases.

So that's to say the strategy method seems to elicit individuals' true preferences pretty well. Sorry. [INAUDIBLE] messed up. But I think that's fine.

So what's now the Becker-DeGroot-Marschak procedure? It's essentially a version of that, which looks something like this, which is, people are asked-- [INAUDIBLE] so what the goal here is to understand people's willingness to pay for a good.

So subjects are told that a price for the good will be randomly selected. And the price goes from like \$0.50 to like, say, \$10. For each of these prices, the person is asked, do you prefer buy versus not-buy in these options.

And so then, I ask you to fill out the entire form. Tell me, for each of these prices, what do you want to do? If the price is \$3, what do you want? If the price is \$1, what do you want? If the price is \$10, what do you want?

And then I'm going to afterwards then pick a price and say, OK. Now the price is \$3.50. And now I'm going to just look at like, what did you actually say, and whatever your choice was then for that specific price is going to be implemented. Is that clear?

Now, just to be clear what the problem is that we're trying to overcome here-- suppose I'm trying to sell you a mug. Or suppose I have this mug here. I'm saying, would you like to buy this mug? It's a very beautiful mug with lots of trees on them. I'd like to sell it to you.

And I'm eliciting your willingness to pay. What's the problem that arises from just asking you this directly, without the BDM procedure? What problem-- what is the BDM procedure sort of helping with?

So if you go to a marketplace, or try to bargain with me on this mug, which again, it's a beautiful mug-- you might want to shade your valuation. You might be willing to pay \$5, but you're going to say \$2, because you are hoping to get a bargain and get [INAUDIBLE] a cheap price from me. Right?

And the key part here is that you are essentially hoping that your willingness to pay whatever you are offering me will essentially change whatever price I'm offering the mug to you. Notice that that's not the case here in the BDM procedure.

In the BDM procedure, whatever you say is independent of the actual price that's implemented. I'm saying, essentially, I'm going to randomly pick one of those 10 or 20 prices here. And you're going to tell me, for each of those prices, what you want, either buy or not buy. And your choices are-- since I'm randomizing afterwards, which price selection you're going to pick, your choices are irrelevant for which of the actual prices are selected.

And so now essentially get around this issue about people shading or essentially underreporting their willingness to pay. Precisely that's what makes it incentive-compatible. And that's essentially what's called the BDM mechanism. That's what many economists use in many, many experiments.

There's another version of what's called the BDM which is a more straightforward version of it. Notice here you have to ask 20 questions. Ask would you like to pay-- would you like to buy it for \$0.50, \$1.00, \$1.50, \$2.00, and so on and so forth.

Another version of this is, I'm asking you straight up what's your willingness to pay. And the bid is then compared to a price determined by a random number generator. I'm just saying, like, tell me what your willingness to pay is.

Then I'm going to do like a random number generator. If the subject's bid is greater than the price that's generated by the random number generator, he or she pays the price, not the announced willingness to pay, and receives the item being auctioned.

If the subject's bid is lower than the price, he or she pays nothing and receives nothing. And so now here again, the final price the person must pay is independent of what the person indicated as your willingness to pay, which essentially solves the incentive compatibility issue.

Think of like this version that I have on this line as just a more efficient way of eliciting the question that's on the previous slide. This is essentially exactly the same thing, except for that it's a more efficient thing to do.

The problem with this version, often, is comprehension. People get confused. Or people tend to still try to bargain and give like lower willingness to pay. So sometimes, this version works better because it's much easier, much transparent, when we say, here's a price, do want to buy it or not, versus and so on.

While here it's people are sort of [INAUDIBLE] because they're used to, essentially, bargain, they tend to sort of under-report at least sometimes. Any questions on this?

So now I'm also going to skip this video. It's a fun video. Who knows the story of Tom Sawyer and the fence? I think it's in the readings, no?

Tom Sawyer was misbehaving. He was essentially punished to paint the fence. He really didn't want to do it. But then his friend comes along, and essentially, he sort of tricks his friend into thinking this is really like such a fun task, and it ends up being essentially, the friend even paying him for painting the fence, and he not having to do it.

And the point of the story, of course, is that people's willingness to pay is malleable in the sense, like, even to the extent that people don't even know whether they're willing to pay for something, or you have to pay them something to do it.

And so essentially, people's preferences are inherently malleable by the way they're being marketed, or the way they appear to them, or being sold to them. You can watch the video in the slides if you download it. For whatever reason this is not working on Zoom, at least for now.

So stepping back a little bit-- so overall, so far we talked about two key components of individual decision making. Utility functions-- what people want and what they care about, and beliefs-- how people perceive themselves and patterns in the world.

And so understanding these, of course, is important because that determines people's choices. Now so far, we have always sort of pretended and said people's preferences and beliefs are always sharply-- people are always sharply aware of what they want and believe that costs [INAUDIBLE].

For example, a homeowner might have reference-dependent preferences, but they know what their preferences are. There might be some issues about how the reference point is determined, but the preference is the function. The functional form is always fixed, and it's known.

A person might have the wrong theory of the world, but she always has some beliefs in mind, what she uses to make those choices

Or a smoker might act suboptimally, but he always has a fully specific strategy in mind for all his current and future decisions. So people-- the smoker might sort of have some issues with present bias, or some wrong beliefs in some ways or the other. But the utility function was always assumed to be fixed and known.

Now we're going to deviate from that. If you think about it like, what are your preferences, and what do you like and what do you not like-- in many cases, we actually don't know.

And in many cases, people sort of just make things up as they go along. So that's to say, like, people are essentially making some choices. And this one of the examples that I showed you from earlier-- the three examples that I showed you first, where people are just induced to make certain choices one way or the other, depending on their choice environment, depending on their social influence.

In Tom Sawyer's case, what he tells his friend. His friend is now willing to pay to paint a fence, which actually is a punishment. So people's preferences seem inherently malleable.

And so that's what this paper about coherent arbitrariness is very much about. And so if you think about it, in some ways, almost all of economics, and any class that you have taken, the professor-- or in any book, people will write down the utility function as if that's some truth that we know about the world. And this is what the utility function looks like.

But if you think about it, in many cases, actually, we don't know what is the utility function [INAUDIBLE]. If somebody asked [INAUDIBLE] particular about new goods, who knows what our preferences actually are.

So for example, if you say, you're trying to buy a monitor, and there's like a 30 inch monitor versus a 24 inch monitor, who knows how much you're willing to pay for that. Maybe \$50, maybe \$100, maybe \$20, maybe nothing. It's very hard to sort of figure this out.

And sort of when asked to make these kinds of decisions, people tend to sort of construct their preferences on this stuff. They essentially make stuff up.

And so because of that, because people are sort of fundamentally in some ways unsure about their preferences, this construction of their valuation is very much easily manipulated by, often, cues that should be really irrelevant.

And so one example that I showed you a little bit-- I think in the second class, or from the survey, is what's called anchoring. This is also in the Ariely paper, which essentially is, what they were doing is, they were asking people about their willingness to [INAUDIBLE] their Social Security number. And then afterwards, their willingness to pay for different things, about like, wine, design books, chocolates, and so on.

Notice that these are somewhat unusual things that people wouldn't necessarily buy every day. And essentially what you see is people who are in the highest quintile of their Social Security number-- this is like year number five-- have much higher willingness to pay. And this is also all incentive-compatible [INAUDIBLE] methods essentially of using some form of BDM methods.

People's willingness to pay is way higher when their Social Security number-- the last digits of their Social Security number are higher.

Now of course, that shouldn't be the case. Your Social Security number has nothing to do with your tastes for wine. Because that's [INAUDIBLE] explicitly random. Yet, people are easily manipulable. And the difference in sort of valuations is like, [? huge. ?]

If you happen to have a low Social Security number, or the last digit are low, then you are willing to pay \$11.73. If it's high, you're willing to pay like three times or even more as much. So the huge differences show that people are pretty easily manipulable.

Now notice that these questions that I asked here are sort of somewhat unusual sort of items. This is not asking like how much are you willing to pay for food at the food truck, or whatever stuff that you do every day, or for a pizza.

Because there, people kind of know already how much they're paying anyway. And if I just ask you something else that's different from the market price, people probably would not want to be paying like \$30 for it, even if their Social Security number tends to be quite high.

Now, Ariely et al go further with this. In some sense they have sort of more radical deviations.

So they elicit people's willingness to accept. And this is how much they have to be paid to endure an unpleasant sound for a different length of time. Now, why did they pick like an unpleasant sound?

In part, because they could provide people like a sample of it. In part because people don't have any experience with that. It's completely unclear how much you are supposed to be willing to pay-- are willing to accept for listening to a sound. So there's no price or market price for it. And people sort of had to sort of in some ways rely on their own preferences for that.

And then they know it's very easy to change the quantity. You can do like 10 seconds, 30 seconds, 50 seconds. It's very easy to manipulate that.

Now what is the procedure? Subjects we're listening to a 30 second sample of the noise. Then there were answering whether they hypothetically would be willing to listen to the noise for another 30 seconds for x cents. And then they were asked their willingness to accept for 10, 30 and 60 seconds afterwards.

Notice that number two is only a hypothetical question, and is essentially entirely irrelevant for number three. Sorry. This is supposed to say number three, not number one.

So number two here is like completely irrelevant. This should really have no effect whatsoever in affecting your willingness to accept for the 10, 30 and 60 seconds. Again, this is just a hypothetical question that really does not matter, and is not going to be implemented at all.

And then the experimental variation here is that x was varied across subjects. Now, why would x matter? What's going on here? What's the experiment trying to do?

People are very unsure about what their valuation is. So now what they do is, essentially, they use the x -- the amount that's offered as like an anchor. The same way as in, usually in markets, when you go to a store and look at how much do certain things cost, and you don't know the quality of the underlying items, often people sort of try to infer quality or lack of quality from the price.

And if we sort of say, if I'm telling you I'm going to pay you like \$100 for this, it must be really, really painful. So people try to sort of infer in some ways something from that. Now notice that this is explicitly hypothetical.

[SOUND]

And so on. And so really it shouldn't matter. If people were sure about like-- I just played you the sound. So you shouldn't be able to tell me how much you like it or dislike it.

So this should really be irrelevant. But people seem to sort of essentially just not really know what's appropriate, or how much they're willing to accept or not to do that.

And x essentially is sort of then anchoring [INAUDIBLE]. OK. So now we get these somewhat messy graphs. Can somebody explain to me what this graph shows? What do we find?

These are [INAUDIBLE] certain subjects. And then there's like 10 seconds, 30 seconds, and 60 seconds here.

So people are always-- for each of these lines, on average, at least, willingness to accept goes up, in the way you expect, in the sense of, for more time, people are asking for more. That's very reasonable.

So 30 seconds, presumably, are worse than 10 seconds. And 60 seconds are worse than 10 seconds. But the levels seem to be completely arbitrary. Essentially, giving people a high anchor increases the levels by a lot for each duration. And giving people a low anchor decreases their willingness to pay at least a little bit compared to no anchor, and surely compared to the [? high anchor. ?]

That is to say-- and this is sort of where this term comes from, coherent arbitrariness, which is, essentially, people seem to be coherent in the sense of like, once you fix a certain level, based on that level, if you ask them, OK, if you tell me about 10 seconds and I ask you about 30 seconds, about 60 seconds, this demand curve or supply curve, if you want, looks pretty reasonable.

But the actual level to start with is completely arbitrary, because I can essentially manipulate you by quite a bit. Look at the differences in magnitudes. This is like 50 versus 30. That's almost like twice as much, depending on just this [INAUDIBLE] anchor.

And there's sort of different rounds in that that you can sort of do this one way or the other. And this is like essentially increasing or decreasing, starting in 10 seconds versus-- and then going up. People's willingness to accept goes up the longer it is. Or if you start with 60 seconds and go down, essentially, people's willingness to accept goes down the shorter it is.

So the direction is very much coherent. The direction of essentially this sort of change in terms of relative to duration is very much coherent. The level seems very much arbitrary.

And so this is where the title of the paper comes from. Arbitrariness, essentially, people's willingness to accept depended strongly on x . For x equals 50, the willingness to accept is like about \$0.59 on average.

For x equals 10, it's only \$0.40. But it's also coherent in very sensible ways. Their willingness to accept are highly sensitive to the duration in the-- very much in the expected direction. Longer is always perceived to be more painful.

OK. So how do we think about this? Well, so one, preferences can be influenced by relevant cues. For instance, [INAUDIBLE] arbitrary initial question, or the Social Security number, or whatever that's being elicited to start with.

But once people have stated a preference, related preferences, they're like, essentially, surrounding preferences are consistent in a sense of like, if you think listening to the sound is painful, asking you to listen to it twice as long will be more painful. Therefore, I have to pay you more to do that compared to what you said in the first place.

Any questions or is that clear? OK. So now there are some concerns about x might be viewed as a hint from the experimenters, how bad the sound is.

But you know, people just listened to the sound for 30 seconds to start with. Like, I gave you the sound for 30 seconds. I told you exactly what it is. So in some sense, that seems like not really a concern anyway.

But there's also-- the very end of this experiment, where x was generated by the last two digits of the subject's Social Security number, and explicitly so, and still there was a correlation between essentially x and subjects' willingness to accept, which really shouldn't be in any other-- in the absence of such anchoring effects.

Are stakes too low? I told you about like, \$0.30, and \$0.50. Maybe people just don't care, and so on. But there's also another experiment [INAUDIBLE] 10-fold stakes. And they got essentially the same results.

So this evidence now is very much consistent with the idea that subjects are searching for their preferences. They don't quite know their true willingness to accept for the sound, and this is essentially arbitrary.

But they know their willingness to accept should relate to each other in a coherent way. And in fact, once you fix the level in some ways, they are in fact coherent, or essentially they're sort of making sense. They're [INAUDIBLE] consistent with that, once you fix a certain level.

Now this is now getting back to Tom Sawyer. Now, the paper goes-- so this essentially is sort of saying, your level of willingness to accept for a certain good that's unpleasant is malleable. Essentially, I can manipulate you to ask for a lower or a higher price, depending on some irrelevant question that I ask you, for a given thing that is perceived to be unpleasant.

Similarly, I can manipulate you to pay more or less for a certain good that's perceived to be a good, something that you want. So the stuff on the Social Security number that I showed you previously, going back, these are all kind of things-- presumably, there's some use for these things.

Belgian chocolates are supposed to be delicious. Red wine, even if you don't like wine, you can sell it to somebody else or something. These are all things that seem to be worth willing to pay for something. And now essentially, given that you are willing to pay something, or some positive amount for it, I can manipulate now-- or the experimenters, in this case, can manipulate people into higher or lower willingness to pay.

And so that's true for positive or negative things. Now, what's amazing about the Tom Sawyer story, however, isn't the-- it's ain't that work? It's sort of like, that's the expression he sort of says towards the end.

The amazing part here is that not only is Tom Sawyer able to manipulate his friend's willingness to accept, or willingness to pay in one direction or-- increase or decrease it, but he's even able to flip it from something that he hates doing, that's essentially something that you really do not want to do, where essentially one has to pay somebody to do it, but instead, his friend is willing to pay Tom Sawyer for being able to do it, right?

The friend, now, instead of like having the friend to pay for it, you can say, I have to pay you \$10 to do it, the friend is very happy to say, Tom, I'm going to pay you some amount. So I think he gives him some apples or some candy or whatever, so that he can actually do it himself.

So essentially, what the manipulation here does is not only changing the level for something that's good or bad, but it's flipping the sign from willingness to accept to willingness to pay, which is a more radical deviation.

You might sort of think we know what's good for us and what's bad for us. But it seems to be that what Tom Sawyer is doing is manipulating sort of the social perception of the item or the activity, and that sort of flips essentially the sign of the item.

So then Ariely now does that in a very beautiful experiment in class. And I always have wanted to sort of replicate this, but I'm not sure I should. So what he does is, or at the time, this is with MBAs-- where he did a poetry reading from Walt Whitman's *Leaves of Grass*.

And there, he has an experiment, where half of the class is asked hypothetically-- again, hypothetically, whether they'd be willing to pay \$10 to listen to Ariely recite poetry for 10 minutes.

So he's like, OK, in class, some of you are able to listen to this. But it could be like-- I think it's outside of class, because otherwise [INAUDIBLE]. And are you willing to pay \$10 for attending this poetry reading?

And I think actually, Ariely is a very good sort of story reader. So it might actually be fun to listen to it. But some of you know to ask hypothetically, again, are you willing to pay \$10 for it?

The other half then are asked hypothetically are they're willing to accept \$10 to listen to Ariely recite the poetry for 10 minutes. That's to say, it sounds like it's pretty painful to listen to Ariely. He's going to pay you \$10 to do it. And are you willing to do that?

Notice that here is only hypothetical. These are all hypothetical choices that will not be implemented. So these hypothetical questions should really have no effect whatsoever on what's asked afterwards.

Because afterwards, people are in fact indicating their monetary valuations for one, three or six minutes of poetry reading.

Now you can already guess what's happening here is here, the first condition is essentially asking for people's willingness to pay. That's essentially selling the item as like a good, that's something that you really want. And you're sort of essentially encouraging people to say, well, you should be willing to pay something for it, even as much as like \$10 for 10 minutes.

And that'll [INAUDIBLE] encourage people, since they don't quite know, listening to your professor is like doing some weird poetry could be really great, or could be really bad, they don't quite know. But here, essentially, they're manipulated into paying for it, while here they're manipulated into being paid for it, or asking for money to be paid for it.

So now, what they then get is their willingness to pay, and their willingness to pay conditions when the first condition here, where they are given the hypothetical \$10 for 10 minutes.

Here, the willingness to pay goes up and is positive. And here, it's negative and goes down. So what we see essentially is that not only is it the case that when you make it essential, when you ask people how much are they willing to pay, people are in fact afterwards willing to pay for it.

But notice that everybody could have just said zero if they didn't want to pay. But in addition, once you sort of fix their willingness to pay, the demand curve is very much sensible in a sense. Once you're willing to pay some amount for one minute, you're going to be willing to pay more money for more minutes.

Essentially, once we define this good to be a good for you, or like this item or this activity to be like a good thing for you, people say more must be better, and sort of are willing to pay more for more minutes. On the other hand, once the item is determined as like a bad, in the sense of this is really a bad activity. Listening to your professor reciting poetry is really awful, and you have to be paid for that, then they sort of say, well one minute is bad, but like six minutes surely is really bad. And you have to pay me a lot more for six minutes than for one minute.

And that's essentially saying subjects don't know whether this reading is good or bad. But they do know, essentially, either way, more requires more money. And that's essentially exactly the coherent arbitrariness, where essentially, the level or even like the sign is really unclear. But once the sign is sort of fixed, people behave in pretty reasonable ways.

I actually don't remember. I should look it up-- surely in the paper, whether it's possible in this condition to actually have-- [INAUDIBLE] so the [INAUDIBLE] experiment would surely always ask you a whole range where you can always say, between say, minus six or minus 10, and plus 10. Now tell me your willingness to accept versus willingness to pay. So if it's positive, it's willingness to pay. If it's negative, it's willingness to accept.

I think that's what's-- so surely, that's how the experiment should have been done. I would like to think and hope that's also how it was actually done. I don't remember. This is quite a while ago. But it's really in the paper, so I don't know if anybody remembers.

But I think it's exactly right. In some sense, the experiment is very much sort of manipulating people in the way of saying like, well, so the prompt here really gets you into like, OK, you're supposed to pay for this. Now it's kind of a little bit of a weird thing to say I'm willing to accept.

Professor says, are you willing to pay \$10 for it? And then you say, well, how much are you willing to pay for like six minutes? And then you say, well you have to pay me like \$5, is a bit of a weird thing to say and the other way around.

So it's very much sort of like I think a set up to sort of generate the effects that they're looking for. I think the underlying-- so I think the underlying essence-- and there are some other experiments that show somewhat similar results-- the underlying essence is right that particularly for unknown goods, people's preferences are very much malleable. And people just don't know what they are.

And there's also some other things. For example, if you think about, for example, social activities, where maybe your friends like them, or they don't like them, and people are very much like-- and this is exactly the Tom Sawyer example, where people are very much sort of malleable.

And you can vastly shape what people want versus not in many cases. And for example, one example would be like, there's some very nice work by Loewenstein and coauthors that looks at education. And sort of, do students want to study versus not. And do you want to be a nerd versus not in class.

And so depending very much on your environment as a kid, when you grow up, if all of your friends are sort of essentially not working hard, and sort of want to be cool, and sort of do not want to study, and it's not a cool thing in your environment to study, people might just not want that.

But on the other hand, if you have like five nerd friends who are all working really hard, or like, if you're running around MIT, studying is a cool thing. And everybody sort of is working hard. And that becomes suddenly very much like a desirable activity.

So I think more generally, while this is a bit of a contrived experiment, surely people's preferences are very much malleable through their social environments, and what they think others think, and how they perceive it. And that can be [INAUDIBLE] shaped by-- manipulated in certain ways, and potentially affected. That could be very much like a policy angle if you wanted to change people's behavior profoundly without spending much money, in fact.

Let me-- I think that's mostly what I have to say. Let me summarize, and then tell you about next week for a second.

So we asked the question, whether people have stable preferences. And it seems to me that people don't have clear preferences for goods and experiences, and sort of construct their preferences on the spot.

Now, they're influenced by environmental cues that, in a way, that doesn't necessarily reflect the true utility from the good or experience. And that's very much also what companies are doing [INAUDIBLE]. And sometimes, some items are made really, really expensive. Like, why is this thing expensive? And somehow companies are able to create some fads or some way of making things desirable by just making them expensive and making people want them in that way.

That very much relies on the fact that people are in fact sort of malleable. And you can sort of introspect and think about what things in the world make you generally happy. What things do you really like? And what things are sort of more things that like, well, other people like them, and you kind of get sort of manipulated, or you get sort of like tricked into doing that.

And in some sense, and we'll talk about this next week a little bit about happiness-- it's important to sort of try and figure it out and try to sort of not perhaps be as much influenced by others, and rather figuring out what you generally and truly like.

So there's a nice series of experiments that sort of demonstrate this coherent arbitrariness, in very sort of like clean variation, and somewhat contrived context. So you might wonder then, does it matter in the real world?

Well it's perhaps less important in settings where people have experience. Again, like if I had done the same experiments with like pizza, you know how much you're willing to pay for pizza. You know the market price. And you know also, and so on.

And so there, probably, your willingness to pay and your preferences are very much like set. But in some other cases, in particular like new environments, [INAUDIBLE] sort of new-- when preferences are shaped-- for example, think about maybe back when you first started at MIT, sort of seeing lots of other students, when norms are shaped, people's preferences are very much malleable and can be influenced profoundly.

Now, there's not so much like actual field evidence in high stakes settings that are sort of this clean nature. Having said that, the examples of, for example, environmental-- sorry, educational choices is perhaps the most compelling one because that's really a high stakes setting that matters a lot. But people's preferences or their choices are very much malleable.

That's all I have to say on coherent arbitrariness. So next time, Monday, we're going to talk about poverty. Please read the paper by Mani et al. And then on Wednesday we're going to talk about happiness and mental health.

Now, I promised you a guest lecturer. And so you know, I was thinking about what's a good thing to do. And so you may have heard about this at UC Berkeley. They have llamas come to campus to make students happy, and destress them in some ways.

Now, of course, I can't bring any real llamas. But what I found, and you may have seen this on *The Daily Show*.

There's this thing called Goat To Meeting, where you can essentially get like a goat or a llama to come to your meeting for 10 minutes.

And I did ask for like a llama to arrive. But there's apparently a chance that it might be a goat or a cow. So we'll have a visitor at the end of the lecture. I think I asked for 2:20, or between 2:20 and 2:30. And which we'll have a guest lecturer that will hopefully either tell us about happiness or maybe make some of you, or at least myself, happy.

That's all I have for today. I'm happy to answer any questions that you might have.