PROFESSOR: Good afternoon. For the past couple of lectures, what I've been doing is using love and romance as the way to talk about broad issues like evolutionary psychology that could be talked about with a wide range of other examples. Love and romance just happen to provide a particularly good set of examples for that particular topic. I'm going to do the same thing now with attitude formation and the links between attitudes and behavior. But I'm going to switch from the love in relationships topic, to the topic of racist or prejudicial behavior and attitudes. Again, not because that's the only set of attitudes that are interesting or important, but because it makes for an interesting path through this material. So, when you read the book, get the topics discussed and rather more general terms. And, here, I'll discuss them in the specific terms of this particular problem.

I should note at the outset what it says about halfway down, I see the first page of the handout, which is, I'm going to do my best to give an explanation about why prejudiced attitudes are easy to come by and, are readily comprehensible in terms of psychological processes that we actually know something about. But explaining these things is not the same thing as excusing them. You can have a society that says, look, all else being equal, racial prejudice and particularly behaviors based on prejudices, based on gender race, national origin, religion, that those sort of biases are bad. And when they lead to behavior, we want to change that behavior. That's quite separate, not unrelated to, but it's separate from the question of how you would explain it psychologically. It's important to remember that explanation is not the same thing as excuse. Because I don't want people going out of the lecture saying, my psych professor explained that it's really, really easy to develop prejudicial attitudes, and that's OK. It's the that's OK part -- I also put on the handout a pair of quotes from the early days of the civil rights movement; one from Eisenhower saying that you can't legislate morality, and a response from Martin Luther King saying, well maybe not, but you can legislate the moral behavior. That's really the social policy point.

If you decide you don't like something that biology, psychology or whatever is pushing people towards, you simply have to do something to make it harder for them to go where that tendency might push them.

Well, in any case, what I'm going to do, is work a story about the development of prejudicial attitudes that's got four factors here listed at the top. And, I work through each of them. I hope you can see how they're tied together to make prejudice a very available option to us. Why this sort of thing happens to us with some regularity. The first factor that I list there is ethinocentrism. That's the tendency to think that your group is the best group. We're number one kind of thing. If you want to come up with, say, an evolutionary psych argument for why this might happen, it's kind of trivial. If you've got some notion that you want to get your genes into the next generation, then you might as well favor the people who are more closely related to you. So you should be more favorable to humans than to mice. You should be likely to be more favorable to people within your group, and even more so to people within your family, out of this sort of fairly straightforward application of evolutionary theory for example.

The more interesting aspect psychologically is how easy it is to get ethinocentric effects. So to get the we're number one effect. Interesting evidence for this comes from these minimal attachment experiments that called, minimal group affiliation experiments. There's a lot of them. Let me describe a couple to you.

So here's an experiment. You come into the lab, and we're going to do an assessment of your taste in abstract art. And, we're going to show you a bunch of abstract pictures. And, you're going to say how much you like it on a scale of one to seven or something. And, then the feedback you're going to get from this at the end is that you like the work of Paul Clay, one abstract artist, better than you like the work of Wassily Kandinsky, another abstract artist. Or it might be the other way around. I don't even remember, by the way, whether they were actually using real Clay and Kandinsky pictures. But you're going to get told that you're in the Clay group or Kandinsky group. This is not a group assignment where there is a lot at stake. Let us suppose that you're in the Clay group. You've been assigned to the Clay group. In the second part of the experiment, you're playing some sort of a game. I don't quite remember what the story was. And, it ends up with you being able to operate under one of two pay- off rules. In one rule, you're going to give the other group one buck, and your going to get two bucks. That's one possibility. The other possibility is a rule that gets you three bucks and give them four bucks. So, this is you. This is them. You've got a choice between option A, and option B. Clearly, the rational choice, from your vantage point is option B, because you get three bucks, and not two bucks. But, in fact, there's a bias towards picking this option. Why is that?

Well, if you get two bucks, you getting more then them. Here, I'm going to get more than I would have got there, but these guys are going to get a whole bunch more. Why should I give those Kandinsky lovers a bunch of stuff? Maybe it turns out that after you figure that these Kandinsky sorts really are a different sort of scummy kind of person, who you really ought to stick it to. Well, in fact, there's no difference between these groups. The group assignment is completely random. So any such distinction that you had made in your own mind is meaningless but maybe you didn't know that. So we can re-run this whole experiment by saying get rid of this silly cover story. We're going to flip a coin. You're in group B. B. A. A. B. B. B. B. Look guys, it's random. There is nothing differentiating your group from the other group. Guess what? You get the same result. I think it does get a little bit weaker, but not much. So, just the act of being in a group, causes you to be inclined to favor that group. You are also inclined to think that group membership is diagnostic.

Let me describe a different experiment actually, as the easiest way to describe this. Suppose I put up a bunch of dots, and we're above the subitizing range here, obviously. OK. Quickly. How many dots are they there? I don't know. You could guess. You get lucky, and be right on. But you'd either be above or below, so you're going to do this for awhile. Guess how many dots there are. And, we are going to declare that you are an overestimater or an underestimater. So, that's the group assignment in this case. Now the interesting thing in this experiment is that you've done it with another person. And, the possibilities are that it could be two guys, both of them are overestimaters. It could be two guys, one of them an overestimater, one of them an underestimater. Similarly for females right? They could both be underestimaters and the critical conditions are. It could be a male, and a, female who, are both overestimaters both labelled as overestimaters or both labelled as underestimaters. And the critical condition is male overestimater, female underestimater or the other way around. But the important point here is that the male is identified as one, the female is identified as the other. So you've got people in all these various groups. I don't know. Somebody can figure out how many groups there must be for the full design here. Got people in all those groups. And, now you ask a question, at the end of the whole experiment, do you think there's a sex difference, a systematic sex difference between men and women on this task? What do you think the sex difference is, if there's a sex difference? All these groups, all the same sex groups, on average, I don't think there's a difference. Some people say, I think women are better. Some people think men are overestimaters or whatever. But, there's no systematic bias here. Nothing systematic happens here.

Here something systematic happens, in this particular version of it. You would declare that males as a group are overestimaters and females has a group are underestimaters. What's the evidence for that? One of each. Clearly not meaningful. There's no statistical reliability that you could glean from this. It could be that brown haired people or blond haired people or something are different. But you know that there's a group difference here, because males and females are definitely different groups. As soon as you've got evidence that there's a difference across this group, you are willing to start making assumptions that difference applies to the group as a whole. You see how that works? More or less? Somebody's nodding their head. That's encouraging.

Ok. So that's factor one-- that you're inclined to see your group as number one.

And this is a point we'll come back to, which is that you're inclined to see groups as having properties that pertain to that group. And, you'll jump to that quickly. Now the tendency to think that group identification tells you something about properties of the group, that's known as stereotyping. If I know that you are part of this group, I believe I know something about you. Period. When we talk about stereotyping, we tend to think of it in negative terms. That is a fairly self evident consequence of factor one mixed with factor two. If you are inclined to think that your group is number one, and you're inclined to think that group identity tells you something, it follows that being a member of another group, means you're of a group that is, at best number two. Ain't number one, that's my group. So you in this group that I've identified, are in the some other lesser group on this scale. That is fairly obvious. What's a little less obvious, and is, at least worth mentioning, is that stereotypes are not just descriptions of things that are common that to that population.

Here's the silly example. Let us consider the Asian women stereotype. Is it part of the Asian woman, is bipedality the Asian woman stereotype? No. Nobody sits around and says all those Asian women have two feet. That's stupid. Nobody's going to say anything like that, because everybody's got two feet. What's important in a stereotype is, stereotypes are different scores in the sense. It's not necessarily accurate ones you understand. They can be completely bogus. But what defines a stereotype, is what somebody thinks differentiates one group from the population as a whole.

So I put some data down here from one big study of stereotypes. Please note that these are not facts. I mean they are facts in the sense that they are data. They're not true facts about, in this case, the German population. What they are is what this particular group of subjects

reported believing about different groups. In this case, the Germans. I excerpted it from a huge study. The factors, efficient, extremely nationalistic, scientific- minded, and pleasure-loving. You will note that the largest single category, the highest value for the German population, in this collection of data points, is pleasure- loving, but that's not part of what would be considered the stereotypes here, because it's not higher then the population as a whole. This particular group of subjects asserted that 82% of people in the world would be pleasure-loving, and a mere 72% of the Germans would be. Therefore, pleasure- lovingness would not be considered part of that stereotype because it it's not something that distinguishes this view of Germans from the view of people as a whole. And, so, efficient, yes. Extremely nationalistic yes. And interesting is something like scientifically- minded would be considered part of the stereotype, even though it's not even held to be a majority. This perception doesn't say that a majority of Germans are scientifically- minded, but more Germans are scientifically- minded according to this view then the population as a whole. Again, I have no idea what the true data would be for the German population. I don't know if they think about science at all. But, the perception is that the stereotype would include efficient, nationalistic, and scientific, and not pleasure- minded, because of this issue of a differential relationship to the perceived baseline. To the population as a whole.

One of the factors contributing to the power of stereotyping is, what's known as the out group homogeneaity effect. So, the in group is you. You're in some group. The out group are other people. The out group homogeneaity effect is the tendency to think that all of them are kind of alike. You don't think that about you in the same way. From the last lecture, the nerdy high school freshman group. If that was my in group, I would know that they aren't all alike, because I'm one of them, and I know that I'm different from all those other nerdy high school freshman. But, the jocks, man, they're all alike. It's an effect, an ignorance effect.

Now, how does this play into bias? How does this ignorance factor play into not just thinking that all those people are alike, but the thinking less of those people? If you ask about what you know about groups that you don't interact with much, where do you got your information about them? You get your information from the news. What gets you on to the news? The fact that you're a good student, and you love your mother? No. But, if you decide to go off and commit an armed robbery, or something like that, that might get you on the news. And, if you are a member of a distinctive group, of some sort, that's not my group, I'm going to say, hey, look, I've got a data point about, -- we can keep picking on the Asian women -- I've got a data point about Asian women. That one committed an armed robbery. I know about Asian women now.

They all commit armed robbery. Well, you don't do anything quite that bald and stupid. But that's the sense in which you are willing to color the entire group on the basis of whatever information you have about one member of it. That's another version of this effect. Is likely to lead to negative assessments of the out group, because the information you get about groups that you don't interact with is skewed towards the negative. The stuff that's going to make the news about a group is going to be typically the negative effect.

So where are the strongest stereotypes in the American population -- well actually this is a somewhat old study at this point -- but when you assess, you can use various questionnaire assessments to assess how strongly a population holds stereotypic views of another population. The heavily stereotyped views held by an American population -- this is now about 15 years ago-- were held of Turks, Arabs to a lesser degree of the Japanese, groups that were not heavily represented in the general Americans mix. Less heavy stereotypes for groups with large immigrant populations in this country, because you were more likely to know some people in that group, and that makes the stereotypes less firm, less strong.

One of the reasons that these stereotypes matter is because along with being willing to build them quickly and easily, we are also inclined to think that the attributes that we put on other groups are causal. At least in others. This is what's known as the fundamental attribution error.

Let me explain a little. But I sent a note to a couple of my social psych colleagues yesterday, as I was thinking about this lecture, asking who was it who named the fundamental attribution error, because that's a great thing to be able to do. To be able to call what you work on fundamental. And, it turns out to be Nisbett, N I S B E T T, from the University of Michigan if you want to track that down. But, in any case, let me explain what this is. There are two broad ways of thinking about personality. We all have some notion that we've got a personality. And, thinking about the attributes that make up that personality can be divided into two broad categories that map onto the usual nature/ nurture kind of arguments in the field. There are trait theories that are typically on the more nature side of it. There are fundamental attributes of personality maybe coming from a genetic origin. You are who you are because you have these traits. The alternative from the more nurtured side of things, the environmental side of it is, is a situationalist account that says you are who you are because of where you are. Trait theory, you are here now because you where born smart, and hard working, and studious, or you basically have consistent over time, hardworking, studious attributes to you. The situationlist account says you're sitting here right now emitting student behavior because

you're in a student kind of environment. If we put you on a farm, you would not be sitting there in the middle of the field with a notebook taking notes about the cow. That's not what you'd be doing. In that situation, you'd be doing farm kind of stuff.

Like all such debates in the field, the truth is going to lie somewhere in between. There's going to be bits of both. You're not going to get any mileage out of arguing strictly one or strictly the other. What you think about this, what you think about the balance here, is important for the policy purposes, for instance. Why did this guy commit a crime? We know he committed a crime, because we just convicted him of committing this crime. Why did he commit a crime? Is it because he's a criminal sort of person? Fundamentally, dishonest, nasty, kind of person. Or, is it because he was in a situation that promoted criminal behavior?

Why does that matter? Well if you're in this mode, you may send him off to prison in both cases. In this case, you're likely to think of prison as a place where you put bad people to keep them out of the way for a length of time that's appropriate to whatever bad thing they did. But, it's basically as a punishment. This is the personality theory, that would lead you to call your prison a correctional institution. Because you think that you could correct this person. If you think he's fundamentally a bad person, the trick is to make sure he can't do that anymore. If you think it's because of the situation, that somehow forced to him into or pushed him into criminal behavior, you want to fix that. Modern prison philosophy is neither all the way here, or all the way there. But, the balance is really a personality theory question.

The fundamental attribution error is a tendency to hold to a more trait theoretic position when you're talking about other people than when you're talking about yourself. Why is it in error? Well it logically can't be the case that you are the product largely of the situation and that they are a product of their invariant parts. That over the population as a whole isn't going to isn't going to hold up.

Why did this guy rob the bank? He robbed the bank, because he's a criminal. Why did I rob the bank? I robbed the bank, because I was hungry, and the door was unlocked, and I didn't really rob the bank, I just kind of picked up the money that was lying on the floor, and anyway, it was other guy. Much more likely to give us situational account. Why did you get a bad grade on the midterms. Suppose you got a bad grade on the midterm. Why did you get a bad grade on the midterm. Well, the story was really lame, and it distracted me. And I didn't get enough sleep. And the course wasn't a big priority for me. That's why I got it bad grade. Your TA says-- well your TA is a good person, and doesn't say this-- but your TA looks at the exam and

says, why did they get a bad grade on the exam? They're stupid! That would be the fundamental attribution error. Why did your TA get a bad grade? I didn't get enough sleep and stuff like that. We're more inclined to give situationalist accounts of our own behavior and more inclined to get trait accounts of other people's behavior.

All right, so let's see where this has gotten us to. We're inclined to put people into groups. We're inclined to assign attributes to those groups. We're likely to assign more negative attributes to groups than we ought to because our information about groups that we don't know is being skewed in that direction, and we're likely to attribute behavior to what we now perceive correctly or incorrectly as the traits of the group. And, so you can see how you're going to end up with a story that is a pretty negative account of some out group or other.

Oh by the way, there's a very interesting wrinkle on the fundamental attribution error, or at least there's certainly used to be, I do not know whether this is still the case in the population as a whole, and I certainly don't know, though I would be very interested to know, whether it's true at MIT. Let's try this as an experiment, and see how your intuition goes -- OK we can do the trait versus situational thing, and we're going to have a math test. You're interpreting your own score on a math test. That score can be good or bad, as you may know. There are explanations of why the test for was good. A trait theory would be, I'm a genius. A situationalist theory would be, I'm lucky. And, if the test is bad, you can have an assessment that it says I'm dumb, or you going to have an assessment that said, that the test is unfair. The interesting bit is that one gender is more likely to say this. And the other gender is more likely to say this. One gender is more likely to say this. And the other gender's more likely to say that. So each of these cells can be associated with a gender. So I got a good grade on the test. I'm a genius. Who are we talking about?

AUDIENCE: Men.

PROFESSOR: All right. And, it follows that this must be the female cell. I got a bad grade on the test. I'm dumb.

AUDIENCE: Female.

PROFESSOR: And, that is the historical finding. These are studies that came out in the early days of the women's movement. And, I don't know about whether or not it's still pertains. But the disturbing finding at the time was that males were inclined to give trait theoretic answers for the good stuff, I'm good, I'm bright, I'm gorgeous. And, females were likely to say I'm lucky and

it's all makeup, or something like that. And on the other side, on the bad side, the females were likely to give trait theoretic answers. I'm dumb and ugly and depressed and its terrible. And the guys were likely to say I'm brilliant, I'm gorgeous, et cetera, and the test was really kind unfair and my teacher hated me and stuff. It would be interesting to know whether that was still true. We might as well take a poll. How many people think that if we actually collected data, we'd find that something like that was still true? How many think that we would find it has gone away? I have no idea if there's new data on that. The basic point is that with that modulation possibly we tend to see situational explanations of our own behavior, and trait explanations of other people's behavior.

Now let's take a look at this last factor, which I'm calling the role of the ignorance in person perception. And see how that can lead to what looks like a biased outcome, perhaps even if you didn't have these other factors. This ignorance factor's now going to interact with these other factors to make biased outcomes. It's quite easy to come by. So, let's do a version of the classic physics joke, about assume the horse is a sphere. We're going to over simplify the situation.

The issue here is who are you going to be friends with? Well, first of all, we have to go back to the good, ernest high school discussion about making friends with people. And, does it matter if they're wearing the latest designer whatever? And, the answer, of course, is no, because you shouldn't judge a book by its cover. Good. Nice cliche. And, it is, of course, true. In the first assume the horse is a sphere over simplification, the problem, let us assume that the set of people with whom you might be friends in the world, is the set of let's just make it all MIT undergraduates. And we know we don't want to judge books by covers, so, therefore, what you're going to do is, set up in depth interviews with everybody. And decide who's going to be your friend on the basis of that. No. That's not going to work. Well, all right, the other alternative is, don't want to judge books by their cover, therefore I won't talk to anybody ever again. That's not going to work either. So it is self- evident again, various bits of this are self-evident, you've gotta make snap decisions on the basis of imperfect information. It doesn't mean that you have to make your decisions on the basis of whether or not they're wearing designer whatevers, of course. But you're necessarily going to have to make a first cut through the population on the basis of essentially superficial information. Well, what's that going to do? Let us assume that in the world, in an act of massive, further over simplification, there are bad people and are good people. And, your job is to divide the world into those two categories.

So you're going to make an assessment. You're going to perform an act of person perception, and divide the world into bad people, and good people. We've got a nice simple two buy two design here. Ideally you want everybody to be in those two populations, and in those two cells. Even if you could do in depth interviews with everybody, it's not clear you'd never make a mistake, but clearly, if you're just going to be basing your decisions on relatively superficial information, there are going to be errors. I labeled these Type 1 and Type 2, which is actually jargon from signal detection land, but don't worry about that. It just, in this case, give us a chance to ask the question. Which of these type of errors is worse? If you had a choice about which error to make, how many people, given the choice between, let's be clear about this, you've got a good person, and you declare that good person to be bad, or you've got a bad person, and you declare him to be good. You got a choice between which of those errors you're going to make. How many people vote for Type 1? How many people vote for Type 2? OK. So a Type 2 person. Why do you you prefer the Type 2 error? Any Type 2 person?

AUDIENCE: [INAUDIBLE]

PROFESSOR: You're missing out on good people. That's the nice person answer. You don't want to inadvertently tar some nice, good person with the label of being bad. How about a Type 1 person?

AUDIENCE: [INAUDIBLE]

PROFESSOR: There's a lot of people who can be your friends. You don't need all of them. And you want to get to those bad people. How come? Anybody else?

AUDIENCE: It could be harmful.

PROFESSOR: Yeah. it could be dangerous. If we dichotomise this into good and really bad, nasty, dangerous people, then that intuition becomes a little clearer that you whatever you else you do, you want to keep these people away. This is applied signal detection theory. Usually you do signal detection theory in visual perception land or something. But this is what the next page of the handout has on it. But here's where this is coming from. Let us suppose, again, for the sake of vast over- simplification, that there are on a scale of goodness, that the only two types of people in the world. There are good people and bad people. And, you want to pick the good people and reject the bad people. But the difficulty is that you can't successfully see this, because your information is lousy. The effect of your information being lousy is that what you see is a distribution of goodness and badness, something like that.

By the way, if you were doing this in vision land, this would be one light and another light. Can you tell the difference between a dim light and a bright light or something like that. By the time it goes through your nervous system, rather than being always looking exactly like this, or always looking exactly like this, sometimes the bright light looks a little dim, sometimes the dim light looks a little bright. And how do you decide which one you've seen? So, you got a person in front of you. How can you decide whether they're good or bad? Well, the best you can do, is put a criterion in there somewhere. So let's just divide it. If I do that, that's OK. That means I'm going to declare everybody on this side to be good, and everybody on this side to be bad. And, OK, so I'm declaring all of these people, who are, in fact, good to be good. The difficulty, the sad thing, is the good people who I'm declaring to be bad, so the Type 1 errors are here. These are good people who I declared to be bad. That's too bad. OK. Here, on this side are all the bad people who I declared to be bad. That's exactly what I wanted to do. But these guys, these are the Type 2 errors. These are bad people who beat my criterion level, and I said they're good.

Now, you should be able to figure out, looking at a picture like this, there's no way to eliminate error. If this is the situation of the stimuli I have to deal with, there's no way to eliminate error. All I can do is apportion error. So, if I decide that the Type 2 errors are the dangerous errors that I need to avoid, then I'm going to move my criterion over. This is the second picture on the handout. If I move my criterion over so that I reduce my Type 2 errors to just these few, let's say, now the result is that I've massively increased my Type 1 errors. I'm now declaring all these lovely people to be people I don't want to make friends with. It's sad, but that's the way it go. Because, as the gentleman back there said, yeah. There are a lot of people here, so I've got plenty of people to be friends with. And these guys will just have to deal with the fact that they're not my friend. That's just the way it is. But, I'm not letting any of these mean, nasty, rotten people in, except for these guys. Most of these guys, I'm going to avoid. OK. Now, look what happens when you deal with an out group. When you deal with a group other than your own. If the argument is that part of what makes an out group the out group is the fact that you know less about them. The way to express that in these sort of signal detection terms, is as an increase in the noise. An increase in the spread of these distributions. So what that's going to end up looking like is you still have the good people and the bad people. But now your perception is less accurate. OK. That'll do. So, they now just overlap more, because we just don't know as much about these people. You're not as good at -- you want a trivial example of this?

Let's take wolves, there's an out group you know, the ones with the big sharp, you know, grandma what big teeth you've got, kind of wolves. Maybe there's a good wolf out there somewhere. A nice wolf. You know the kind of wolf that we're supposed to have adopted back in antiquity to make into dogs eventually. But you meet a wolf on the street, and you don't know much about him. Where should you draw your threshold before bringing him home to play with your six year old? You're going to draw your threshold out here somewhere, right? I don't care if I reject the one nice wolf. It's just really risky to bring wolves home. And, that's because you're just really, really ignorant about wolves, you just don't know much about them. Maybe if you knew wolves better, you'd know who the nice ones were.

All right. With human populations it's obviously much less dramatic than that. But you don't know much about these other people, so the distributions theoretically overlap more. You still only want to make very few errors where you let bad people in next you. So that's going to cause you to move that threshold still further over in this direction. Not because you don't like these people. Understand that there's no explicit bias going on here. You're just being cautious in this in this story. You can get explicit bias out of those the first three factors. But here, this factor has no explicit bias in it at all. Just ignorance. So now you say oh good I'm only letting this percentage of really bad people and now obviously there's a little problem here. Your Type 1 where you reject good people. Now you have rejected almost the entire population of this other group. You know that you're not biased in your heart of hearts, because you can still say, as it says on the handouts, this little tale of the distribution some of my best friends are X. Whatever that out group is. Some of my best friends are white, black, Christian, Jewish, whatever the other out group is you're dealing with. This signal detection story will get you there with some people in the group who are fine, because they beat your threshold, and the vast bulk of the rest of it, who disappear because you're applying the same caution to an out group, that you were applying to the group that you knew something about. So, that's how ignorance can end up being a factor in producing what looks like biased behavior. If you compare these two, you'd have to say I'm biased against this group. Because 100 of these people, I'm only letting five of them be my friend. 100 of these people I'm letting 60 of them be my friend. That's a biased outcome from no explicit bias.

Now this question of explicit versus either no bias or implicit bias is an interesting one. You don't necessarily have a clear idea of the biases that you may have. One of the more interesting and more disturbing bits -- there's a thing called an implicit attitude test, the IAT. If

you want to try this out on yourself, go to www.prejudice.com I think is the right site. That is one site. But if that fails go and find your way to the website mosuronbanashi at Harvard. She is one of the leading practitioners of this, and her website will link you to a place where you can try this out on yourself. As the website will tell you, before forewarned. You may find the results of this experiment to be disturbing to you. But, it's well worth trying out yourself. Now what is this experiment about? This experiment is, in effect, a version of a Stroop interference test. The classic stroop interference experiment is an experiment where what you do is you see a collection of words and your job is, whatever, the word says, it's just to tell me what color the ink is that the word is written in. So if I write cat in red ink, you say red. And, if I write dog, in blue ink, you say, dog. No. You say blue. That's a different interference. The problem is, that if I write red in blue ink, some people will simply make the mistake of saying red, and everybody on average, will be substantially slowed down, because of an inability to suppress that response. And if I do red in red ink they'll be speeded up. So, if the two sources conflict with each other, you're slowed. If the two sources agree with each other you're speeded. OK. So, here's what you do in an IAT experiment. What you do, is you tell people, I'm going to show you some words. And, if they're good words, they're nice words, you push this button. And, if they're nasty words, you push this button. OK. No problem. So nice boink. Evil. Boink. Pain. Boink. And so on. Not very tough. OK. Second task. I'm going to show you some pictures of people, if it's an old person, I want you to push one button, if it's a young person, I want you to push another button. So we put me up there, boink. Put you up there, boink. Now what we do, is we do mixed blocks. Where you're going to see words and pictures together. And, I'm going to tell you, OK, now if you see a nice word, or an old face, push this button, if you see a nasty word, or young face, push this button. It's just the tasks on top of each other. Whatever we do in this regard, you'll be slower, because now you have to keep two rules in mind at once, but what's striking, is it if you do nice and old, you're significantly slower than if you do nice and young. Its as if nice and old maps the same response. It causes a conflict that doesn't work for you. And nice and young does. As if you've got a biased in favor of young over old.

If we ask you to why is this an implicit attitude test if we give you an explicit attitude test that says what's your attitude towards young and old people. You may perfectly well report I love old people. I love young people. I love everybody. But, you'll still come up with this result. I deliberately did this one, because it doesn't tend to carry an awful lot emotional loading for people. But the reason this may be a disturbing test for people to take -- you should still go off and do it-- is that if I do nice and black, African American pictures and nasty and white,

regardless of your explicit report, of what you consider your bias to be, the white population in this country, will, on average, have slower reaction time to nice, black pairings then to a pairing of nice and white. Actually, one of the more interesting and depressing findings, is that doesn't even completely reverse with an African American population of subjects. In the African American population, the last time I checked on the data, the pairings were roughly equal. So, the African American population has presumably some ethnocentric biased in its own favor, an implicit bias in its own favor, but that's counter balanced, by some incorporation of overall bias against, and so they come out is roughly equal.

The debates and the literature about this line of work is this talking about implicit attitudes which is a notion that regardless of what we think, we're all racists or something like that. Or, is this just say that we have somehow incorporated that we know, at some level, the biases of the culture, as a whole, even though, we, ourselves, may not be biased. It is an interesting question beyond the scope of what I can talk about today, whether there is a serious difference between those two. But, the disturbing finding is, -- and you can do this -- it's not black, white, old, young, by no means is the limit on this game. Once you've got the methodology, you can do it on anything. So, shortly after September 11, they started doing these tests on opinions about nice -- and you don't need to do with faces either, so if you want to do Arab versus non- Arab, you can just do it with names. So, you do Abdul, and Mohammed and so on, and then you do Chris, and Jane. And, you find that in the American population, at the moment, nice and Mohammed is slower then nice and Robert. It's not surprising, but, it is, nevertheless, disturbing how easy it is to show these effects. They're robust. They show up across populations, and they show up fairly independent of what people report. It doesn't matter if you explicitly, and let's assume that people are reporting honestly. It doesn't matter if you explicitly have the bias. You can show something that looks like a bias with a test of this sort. Anyway, give it a try. It's interesting and disturbing.

And the interesting and disturbing department, I will continue in a minute or two talking about the link between attitudes and actual behavior. So, this is a good place for a short break.

So, look. Bias. We can presumably agree that bias isn't a good thing, but if it stays in the realm of private opinion, or if it stays in the realm of implicit opinion that you don't even know what you have yourself, it's not exactly front page news, but we also all know from reading the front page, that there are regrettably frequent occcasions where one group is willing to slaughter another group based on very little more, if anything more, than group identity. So, an

absolutely critical question, is how can people be moved from attitude to action? And, the disturbing aspect of what we know from experimental psych about this, is that it is surprisingly easy to have your behavior controlled by outside forces. The experimental work, of course, cannot get people to go out and slaughter each other. Nothing like that would be even faintly moral, but rather like the Clay and Kandinsky experiments. You can do experiments that show that it's surprisingly easy to manipulate the situation in ways that changes behavior in ways, that changes behavior in ways that look at least a little bit disturbing.

One of the classics, back in the '50's, that gets this literature going, was done by Ash at Columbia at the time. I think that the picture is still in the book. A marvelous collection of male, Columbian nerds from the '50's. Anybody read the chapter yet, and happen to know if it's there? Here's the basic experiment, here you come into this experiment, and you're doing an experiment on visual perception. Ash shows you a card with three lines on it. Your job is to say which line is longer. Now, the odd thing about this, well, you're not an experimentalist, so why should you care? But it's a little odd, as an experiment to be doing this in a group. But it turns out, you're doing this in a group. And, so Ash holds up the card, and you say B, and you say B, and you say B, and everybody said B. Next card. I'm not going to bother changing my cards, but you get the basic idea. On the critical trial, up comes this card. He says, C. He says C. He says C. She says C. C. C. Now, we're up to -- I stuck with her because, she's got glasses. Because the nerdy Columbia guy has glasses on too. What does she do? Well, why did all these guys say C? Are some kind of morons? They're all confederates of the experimenter. The only real subject is this person. And, the question is, does she say C? The answer is about a third of the time in the original Ash experiment, the answer's yes, she says C. It is completely clear that even when she doesn't say C, she's uncomfortable. This is an experiment on peer pressure. And, it's perfectly clear that when everybody else is saying C, she's busy taking off her glasses, and checking them, and stuff like that, to see what's going on here. There's something wrong. What do you think manipulates -- the standard result is you got about a third of the people complying with the pressure. What reduces that compliance?

AUDIENCE: How [INAUDIBLE]

PROFESSOR: Yeah, sure. Presumably it's hard to get the result. But, OK if we keep the physical stimuli the same. Yes. AUDIENCE: [INAUDIBLE]

PROFESSOR: If somebody picks A, and it just looks noisy. I don't know if they ever did that particular

manipulation. That's an interesting question. That might change things.

AUDIENCE: If they have six or seven people, [INAUDIBLE]

PROFESSOR: It doesn't take any support. You probably know this from arguments with groups of people or something. It's hard to be the first person to voice the minority view. It's much easier to be the second person. I think I actually have the data from that. Yeah. One supporter. One supporter in the group drops compliance from one third to 1/12. So, big drop in the amount of compliance. The more people who say C, the more likely you are to comply. The smaller the group, the less likely you are to comply. But, the point is, that even in a matter as seemingly straight forward as which line is longer, you can feel that pressure from others.

The most famous experiment in this canon is an experiment done by Stanley Milgram. And, in that experiment, here's the set up. You come into the lab for a study on learning. The effects of punishment on learning. And, there are two of you. And, one of you is going to be the learner and one of you is going to be the teacher. OK. And we're going to decide this randomly. Flip a coin. You're going to be the teacher today. Now, in fact, this is not random. The subject is always the teacher. The learner is always a stooge of the experiementer and here's what happens. You're told you going to do some sort of a task, and your job, as the teacher, is to give the learner a shock every time they make a mistake. This was done back in the '60's with this great big hunk of electrical equipment, with a gazillion switches on there, running from 15 volts to, I think 450 volts, in 15 volt increments, with instructive little labels like mild, and then, up here somewhere is severe, and by the time, you get up here, it's labeled something like XXX. The rule is every time you make a mistake, you increase the voltage. Now, we will give you a 45 volt shock. You, the teacher, gets a 45 volt shock, just to see what it's like. And, a 45 volt shock from this apparatus is mildly unpleasant. It's nothing you'd want to sign up for. It's not going to kill ya, though the suggestion is that this might. And, that's the rule of the game. Now, before doing the experiment Milgram went and asked everybody under the sun, what the result would be. Well, the answer is that everybody under the sun will bail out pretty early here. Nobody's going to get them very massive shocks. He asked theologians, he asked psychologists, he ask people off the street, and everybody agreed this is not going to lead to much in the way of shocks. And, this made the result of the first experimental a little surprising. Everybody went all the way through to 450 volts. The entire population of subjects in the first study went through to 450 volts. Now, were they with a thrilled about this? No. It was also absolutely clear that the subjects were very uncomfortable about this. And, that they

questioned whether they should do this. And, Milgram had an absolutely stereotypical, he'd prepared, in advance, the response. Please continue, the experiment must go on. And, that was it. You were free to leave. Though he didn't explain this to you in great detail. But if you said, should I do this? He said, please continue. The experiment must go on. And people did. Now, he was a little surprised. In the original version, the alleged learner had been taken out and put in a different room. And Milgram figured, well, look. Maybe these people just didn't believe the set up story here. So, they moved the alleged learner to a position where he was visible and making noises about this. He's vigorously protesting as the voltage gets larger. At some point, up here, he says he's not going to respond anymore. So, the experiment's rigged, so that he keeps making mistakes as a result. No response is considered an error. And Milgram is there saying please continue, the experiment must go on. Oh. He also makes useful comments like, I have a heart condition, and stuff like that. It's pretty vivid stuff. So, what happens in that case? Well. Great. No longer do 100% of the subjects go all the way through to the end. Only 2/3 of them go all the way through with a subject who has stopped responding, and has announced for all you know you're killing this guy, perhaps. This is pretty disturbing kind of stuff. It was very disturbing to the the subjects. Actually, this is one of the experiments that produces the need for informed consent in experimental psychology. You didn't just go and hijack people off the street and say, you have to be in my experiment. These people were volunteers. But, there was no sort of consent process the way we would have today. And, the level of distress produced in these subjects was part of what drove the field to require informed consent. Why was the experiment done at all? This is an experiment done in the '60's, less than a generation after World War II. And, a question that had obsessed social psychology, since World War II, was how could the Nazi atrocities have happened? Who were the people who went and killed millions of other people? Not the soldiers, but the people who killed six million Jews, and I don't remember, a million and a half gypsies, and some large number of gays and so on. Who were these people? There was a theory out there that encapsulated in a book called *The Authoritarian Personality* which among other things, fed into stereotypes about the Germans which was that there was a certain type of person, who was willing -- the Nuremberg trials, the war crime trials after the war, had produced over and over again, the line, "I was just following orders". And, the notion was, well there are some people who are just good it that. They just follow orders. And, the rest of us were not like that. Milgram suspected that was not the case. Milgram suspected that the answer was that under the right situation, many people could be pushed into acts, that they would objectively think were impermissible. This is his effort to get at that question.

If this sounds like current events to you, every social psychologist with half a credential was on the news after the Abu Ghraib Prison Scandals earlier in the year, because it just sounded so much like this sort of problem again. The people who were charged, many of them responded with the response, that they were simply if not following orders, following the implicit instructions that they felt around them. You got to endless articles in the papers, saying, Oh, you know, so and so was just like everybody else back home. I don't understand how he, I don't understand how she, could end up in these pictures doing things that any reasonable person would say are unacceptable in a in a military prison situation. It's the same kind of question, different scale of magnitude, of course, to the Nazi atrocities. But it's the same question. How do people end up doing things like this?

Before attempting to answer, which I can see is going to run into my next lecture, let me tell you about a different experiment designed to get at the same question. This is an experiment where you think you're in a study, sort of a consumer relations kind of study, that you might run into it the shopping mall. They've set up the experiment at a shopping mall. And, the cover story is this. We're doing some research on community values. Because we're basically doing investigations for a legal case. Here's the situation. This guy was living with a woman to whom he's not married. His employer found out about it, and fired him. This has gone to court, and the court case hinges on community standards. If community standards are that living in sin, out of wedlock, is a bad, bad thing, then it's OK to fire him. Otherwise not. So we've gotta find out what the community standards are. Let's all have a discussion. So, I've told you the story. I've got the cameras rolling here. I, the experimenter, I'm going to step out and you guys discuss. OK. You guys discuss. That's great. Now, I come back and there's a group of ten of you or something discussing this. I come back in, and I say, that was great. Thank you very much. I know what you believe, because I've been watching this. But, I want you, you and you, please to argue from the point of view, that the guy should be fired. I know it's not what you really believe. But just argue for that. OK. Comes back in. OK. Now I want you, you, and you to join that argument arguing that the guy should be fired. Fine. And then in the penultimate step, is I want everybody to have a chance to look into the camera, and say why the guy should be fired. I know you don't believe. But, just give me a little speech as if you believed it. OK. Cool. OK. Last step. Here's this statement that says that I can use my videotape in any fashion I wish including submitting it as evidence in court. Will you please sign. I'll be back in a minute. I gotta go rinse a few things out. But you guys decide. The question is do people sign? I mean this is obviously basically asking you to perjure yourself if that wasn't what you believed. Now

this was a huge experimental design originally. They were crossing a number of people with gender and they originally had a plan for, well I can't remember, a gazillion groups. They called off the experiment early, because like the Ash experiment and like the Milgram experiment, this experiment produced extremely strong feelings in their subjects. They had gotten a form of informed consent, which I can't go into now, but even with that, it felt unethical to them to keep going. So, they had 33 groups. It was a busted design. But they had 33 groups. Of those 33 groups, of those 33 groups, how many did the Milgram thing, went all the way and everybody sign? Do you think? The answer is one. I think. One group got total obedience. 16 of the groups unanimous refusal to sign. Nine groups got majority refusal to sign. So, in this experiment, compliance was much, much lower.

And, the question that I'll take up next time, for the start of the next lecture, is what's the difference between the Milgram experiment -- let me say one last thing about the Milgram experiment. This is not an isolated result -- Milgram's at Yale. It's not just an isolated experiment that happens in New Haven in 1960. Replicated all over the place. Doesn't matter in America. Doesn't matter age, sex of the subjects. It replicates beautifully. Why does this work, and the other experiment didn't?