

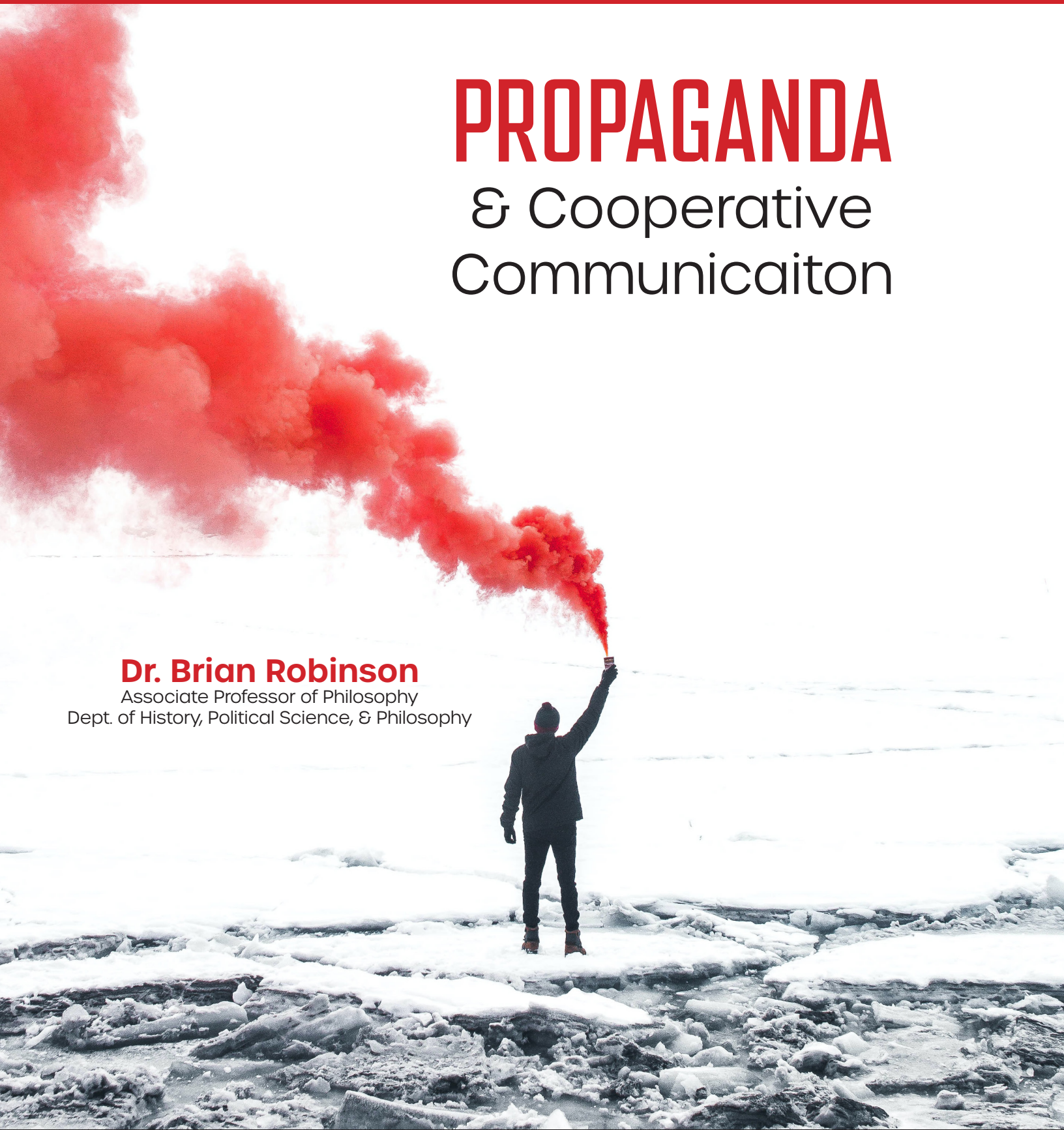
TEXAS A&M UNIVERSITY-KINGSVILLE®

# 42nd Annual Faculty Lecture

## PROPAGANDA & Cooperative Communication

**Dr. Brian Robinson**

Associate Professor of Philosophy  
Dept. of History, Political Science, & Philosophy





**Dr. Brian Robinson** has been at Texas A&M University-Kingsville since 2016. He is currently associate professor of philosophy and assistant director of the Center for Faculty Excellence. He also is currently working with researchers from other institutions on an external grant from the National Science Foundation entitled Investigating the Philosophical Commitments of University Scientists across Disciplines. He is the principal investigator and received over \$90,000 of a \$572,712 grant.

Robinson received the University Faculty/Staff Research Award in 2024, the Dean's Award for Outstanding Scholarly and Creative Production twice from the College of Arts and Sciences and the Summer Research Development Award. His research focuses on philosophy of language, moral psychology, and philosophy of science.

Robinson is a member of the American Philosophical Association, the Toolbox Dialogue Initiative, the Southern Society for Philosophy and Psychology and the International Association for the Philosophy of Humor.

He received his bachelor's degree as a University Scholar Major with emphasis in classics from Baylor University; a master's degree in divinity from Princeton Theological Seminary; a master's of arts degree in philosophy from University of Colorado Boulder; and a master of philosophy degree and a doctoral degree, both in philosophy from The Graduate Center, City University of New York.

## Abstract

Propaganda is a morally dubious form of communication, one that conjures the specters of manipulation and irrationality. Defining it, however, is challenging. For many of us, we simply know it when we see it. Existing accounts of propaganda typically aim to legislate the definition of the term so that it can be objectively determined what utterances are propaganda, in hopes of naming and shaming such speech. In this talk, I present a different approach culminating in a different sort of diagnostic definition. People disagree about what counts as propaganda. Instead of determining who is right, my account diagnoses the reason for these differences by looking for the more fundamental debate underlying disagreement about propaganda. I anchor this definition in Gricean pragmatics, in particular the idea that communication requires cooperation. In defending that view, I introduce the notion of non-consensual cooperation, which occurs when individuals unwittingly contribute to a speaker's communicative goal due to automatic cognitive processes. Propaganda exploits such cooperation or the audience's misplaced trust in a presumed speaker to misdirect trust in a social agent or institution. In the end, propaganda is not defined by its content or its circumventing of rationality, but rather by manipulating or exploiting trust.

## 1 Introduction

This talk will focus on propaganda. What does it do? I am going to try to present a theory of propaganda based on an account of cooperative communication. With that said, I have an initial question to get us started. Consider Figure 1 and ask yourself this question. Is that image an example of propaganda?

---

<sup>1</sup> Before I begin, I would like to thank the University and the committee that made the selection for this opportunity to give the 42nd Annual University Faculty Lecture. This is, I believe, the first time a philosopher has been afforded this honor and opportunity, and I appreciate it, and especially that it is the forty-second one that a philosopher gets to give some thoughts and questions and answers.



Figure 1 – A potential propaganda poster of Texas A&M University-Kingsville President Vela in front of the university bell tower with spray painted “Viva Vela!”

Now, consider why you have the view you do. If you think it is propaganda, what is it in this image that signals propaganda to you? If it is propaganda, what are you picking up on here that tells you it is propaganda? Alternatively, if you think this is not propaganda, what is your notion of what propaganda is that this image fails to meet?

I want to start by ruling out a few options. This image—if we assume it is propaganda—is not propaganda for any of the following reasons. These are answers we should rule out at the beginning. First, if it’s propaganda, it’s not propaganda because of its use of propagandistic tropes or styles. Admittedly, there are various stylistic elements that are common in propaganda images, such as the flared sun rays from behind a larger-than-life figure. That is a common motif in propaganda, as is the sort of spray paint technique. That said, that is not why this image is propaganda. There are two reasons for that. One is that there is some propaganda that you are supposed to recognize is propaganda, while some propaganda is not meant to be recognized as such. For the latter, the propagandist believes that it will be less effective if the audience realizes what the propagandist is doing. That is not always the case. Sometimes propaganda comes from the Ministry of Propaganda, but not always. So, the mere presence or absence of propagandistic styles, tropes, or motifs cannot be the feature that makes something propaganda or not. The second reason is that for there to be tropes, motifs, or styles that are indicative of propaganda, propaganda had to have come first. This image does have styles that historically were common in propaganda. That is why they are now associated with propaganda, because they already have been used over and over again in propaganda in the first place.

Second, this image is also not propaganda (again, assuming it is) because of the content, i.e., what is being asserted by it, what its message is. In fact, one might even claim that there is not a message being asserted. There is no semantic content here, as we would say in philosophy.

Third, if Figure 1 is propaganda, it is not because it presents some false information. This image does not present any false information. There is nothing incorrect or false about it. Vela is the president. That is the bell tower of the university. There's nothing false here. You can have propaganda that is true.

With these preliminaries out of the way, we can proceed with the task at hand. My goal is to present an account of propaganda that is based on the idea of cooperative communication. Propaganda is a form of communication. So, to get a sense of what it is, we need to step back for a bit and talk about how communication works, what meaning is, and how it is conveyed. To that end, here is the plan for accomplishing that goal. In section 2, I explain speech acts. Then in section 3, I focus on cooperation, which will lead us to cooperative communication in section 4. Finally, in section 5, I conclude by talking about the nature of propaganda itself.

## 2 Speech Acts

Set propaganda aside for a little bit. Let's talk about some building blocks that we need before we can get to a full theory of it. We start with speech acts. This is a term that philosophers have been using for the last 75 years or so, and by that we mean how to do things with words (Austin, 1962; Harris et al., 2018; Searle, 1969, 1976). Words do things, and they do different things. Or more accurately, we do many different things with words. Philosophers of language have been analyzing and cataloguing the different things we do with words, each of which is distinguished as a specific speech act. One common kind of speech act is an assertion. This is one we do a lot. For example, suppose I say, "Today is Tuesday," "The sky is purple," or "The moon is made of cheese." These are all assertions. Assertions are true or false. That is the point of them, to assert a proposition with a truth value, typically with the intent to get others to believe the proposition asserted.

Philosophers of language have spent a great deal of time considering the nature of assertions and their meaning, and how people interpret those assertions to understand the meaning. But this is hardly the only thing we do with language. We use words in lots of different ways. We make other kinds of speech acts. For example, we promise, pronounce, propose, pardon, prognosticate. And these are just examples of some speech acts that start with 'p'. Each of these are a different kind of speech acts. What speech act theorists do is try to identify individual speech acts and then give an account of what each of those speech acts are. What does an account of a speech act look like? An analysis of a particular speech act—in other words, a definition of it—answers two questions. First, what does that speech act do? Second, how does it do that?

Let me give you an example. Pronouncements are a kind of speech act. A pronouncement changes how things are. In other words, they create facts. They do not describe facts; they create them. How do they

do it? Well, by some convention the speaker of a pronouncement has the capacity to create certain facts by saying them in certain circumstances. For example, a baseball umpire yells, "Strike 3!" It does not matter whether the ball was in the strike zone or not, at least not in terms of the truth of what was uttered. That now is strike 3, regardless of whether the ball was in the strike zone or not. It is strike 3 because the umpire declared it to be so. The umpire is in this right kind of circumstance that they by convention have the capacity to create the fact through their utterance. If I in the stands yell, "Ball 4!" it does not matter. As a mere spectator, I cannot create the fact like the umpire can. As another example, a judge standing before a couple in the right circumstances, says "I now pronounce you husband and wife." The judge is not describing something by saying this. The judge creates the fact that they are now husband and wife. That's the nature of how pronouncements work.

Now, here is the claim that I want to make for the paper. Propaganda is a kind of speech act. Therefore, we must answer the same two questions. What does it do and how does it do that?

Before we can get to those questions, there are some preliminaries in terms of speech acts that must be addressed. We must consider the matter of how we should evaluate a definition of a speech act. How do we tell if it's a good one or not? We cannot perform some experiments to tell us. And there can be some good ones and could be some bad ones. So how do we tell?

To address this issue, we must consider what kind of definition is being offered. I propose that there are four different kinds, and the criteria for evaluating each is different. The first kind of definition is what I call a stipulative definition. In other words, "by 'x', I mean..." The speaker stipulates what a term means when they use it. The **stipulated definition** applies to no one else but that speaker. Here, there are not really a lot of limitations. A speaker to some degree can mean whatever they want. I could say, "By 'white', I mean black, and by 'black', I mean white" (Carroll, 1897). A speaker could do that. It is a little weird. It makes things more complicated for their audience, but they can use those words in whatever way they want so long as they are consistent, and especially if they tell us ahead of time. So, there is not really a great deal of evaluation here beyond efficacy. That said, we need not worry too much about this kind of definition. None of those are going to be the ones we are going to deal with when it comes to propaganda.

We can also have a **descriptive definition** or account of a speech act. In other words, "By 'x' everyone means...." Here the definition is trying to describe common patterns of usage. This is arguably what lexicographers in dictionaries do. They try to give descriptive definitions of words that match how people typically use those word, and when they use them different ways, then they give multiple definitions. The evaluation of a descriptive definition is based on the correspondence between the definition and actual usage. The evaluation requires some empirical checking in the world.

A third kind of definition is what I call a **legislative definition**. In other words, "By 'x' everyone should mean..." Here we do not care whether or not people actually use the term in the way that the definition is proposing. Someone giving a legislative definition is claiming that there is a correct way to use this word. It might side with one usage pattern over another. Alternatively, it might introduce a novel definition and assert everyone should adopt it. The motivation behind legislative definitions is that if everyone adopts this correct meaning, then things will be better in some way. Perhaps a miscommunication will be eliminated or a conceptual confusion

resolved. This correct definition might even allow for solving some larger social problem. We can then evaluate a legislative definition on whether it accomplishes that goal. But we cannot evaluate a legislative definition based on whether it accurately captures how everyone does use the word. That is a very different thing. It could be that no one uses a term in the way a legislative definition is saying should be. That fact, however, neither invalidates nor refutes a legislative definition.

Many (if not all) legislative definitions are examples of what I call the **Hobbesian program** of linguistic legislation. I give it this title because I can trace back this trend at least to Thomas Hobbes, the English philosopher and his great work of political theory the *Leviathan* (1651). In it he argues over the definitions of many terms, including terms ‘tyranny’ and ‘liberty’. He argues that the meanings of these terms have been usurped. He claims Aristotle and other democratically minded fools have defined or used these terms in ways that have been politically harmful. Hobbes tries to reclaim these terms from how they have been usurped, because these usurpers and their faulty definitions have caused a lot of trouble. If we can do this—if we can correctly use the word in the way that Hobbes is saying we should—we will fix somethings that are wrong in the world.

So, a Hobbesian linguistic legislation tries to say a word is being used wrongly, and because it is being used incorrectly, something wrong is happening in the world. And if we could just get the definition right, we will fix some problems in the world. In other words, we will at least remove some hindrance to improving things for people in the world by getting the definition right. Later on, I will argue that this program of linguistic legislation is exactly what is at work with most definitions of propaganda that are out there currently. As you may already detect, I will argue that approach is problematic. It is not working. It is not the right way to proceed, at least for propaganda.

Lastly, there is what I call a **diagnostic definition**. For now, I will leave this category unexplicated, giving an example of it toward the end of this talk. The short version is that a diagnostic definition defines a term or speech act in a way that both captures what the term means and also why people attribute the term differently in particular cases.

The point of this review of kinds of definitions is to foreshadow what we will encounter shortly. By and large, most accounts of propaganda that have been on offer have been legislative, and I think there are problems with that. I will offer a diagnostic definition instead.

### 3 Cooperation

Let’s now turn to talk about cooperation. What does it mean to cooperate?

I want to start with a preliminary distinction between cooperation and cooperative acts. These, I believe, are different, and distinguishing them will help us. They are too often confused and run together. A **cooperative act** is when one agent acts in a manner necessary for contributing to a cooperative goal. **Cooperation**, then, is, when all, or at least enough, agents act cooperatively. To clarify the distinction, we

can use a bit of game theory. If you are familiar with game theory, Table 1 should look familiar to you. If not, then this Table 1 needs a little explanation.

	Coop. Act	Exploit
Coop. Act	7,7	1, 10
Exploit	10, 1	3, 3

*Table 1* – A representation of the Prisoner's Dilemma

Imagine we have two agents. We will call them Agent Column and Agent Row. Each of them can perform one of two actions, which they will perform simultaneously. They can either perform a cooperative act or they can perform an exploitative act. If both perform the exploitative act, then they both get a utility (or payout) of 3, which is depicted in the bottom-right box. (Here, more is better, so higher numbers are better outcomes for each agent.) If, however, Column performs the cooperative act, but Row still opts for the exploitative act, then Row gets 10, while Column gets 1. Of course, if their actions are reversed, then the outcomes are reversed. Finally, if they both act cooperatively, they both get 7.

Cooperative acts are actions performed by individual agents. If both agents act cooperatively, they get the cooperation as the outcome. Cooperation is the outcome of sufficient numbers of cooperative acts. If one of the agents refuses to act cooperatively, then there is no cooperation.

Another preliminary to stipulate about the nature of cooperation is that cooperation is what we in philosophy would call a three-place relation (at least). In other words, when we talk about relations in philosophy, we are asking how many things there are that stand in a particular relationship to one another. For example, you might say trust is a two-place relationship: one person trusts a second person.

### **Trust (Person1, Person2)**

Friendship is another two-place relationship. Person A and Person B are friends.

### **Friend (Person A, Person B)**

Prima facie, cooperation may look like a two-place relationship. For example, Jack and Jill cooperate.

### **Cooperate (Jack, Jill)**

---

<sup>2</sup> In cases with much larger groups, not everyone has to act cooperatively for cooperation to be the outcome, so long as enough agents cooperate; the non-cooperative agents are then free riders on the beneficial cooperative outcome.

But that would be incorrect. It significantly under-describes the situation, such that we cannot assess the truth value of an assertion of a cooperation as a two-place relation. In other words, to claim that Jack and Jill cooperated is neither true nor false; it is an incomplete proposition.

I contend that cooperation is a three-place relation (at least). There are at least 3 things that must be there for cooperation to exist. Cooperation involves the agents cooperating and the end to which they are cooperating. For example, a fuller, a more accurate description would be to say Jack and Jill cooperated to fetch a pail of water.

### **Cooperate (Jack, Jill, fetch pail of water)**

One must specify who was it and what goal(s) were they acting towards. Only then do we have a complete proposition. Only then can we assess the truth value of claims of cooperation. So, it is not enough to say that you and I cooperated. You have to say you and I cooperated to play a game (for example).

The point of emphasizing this is to say that cooperation is always directed to some goal. If there is no goal, there is no cooperation. With that said, I want to present a minimalist account of cooperation. What is the minimum that must be the case for cooperation to exist?

Cooperation requires that there be multiple agents acting (either collectively or individually) in a manner to achieve a shared goal. That is the bare minimum of what must be the case for cooperation to exist.

This account excludes several things that often are present in others' accounts of cooperation and sometimes even thought to be required for cooperation. First, it does not require that all agents perform the same action. You can do one thing while I perform a completely different action. But the combination of those actions achieves our shared goal. Sometimes cooperation requires everyone to do the same thing—pushing in the same direction, as it were—but not always. Football teams, for example, have eleven players performing different actions to achieve the same goal.

Second, this account does not require that agents act at the same time. Again, that is common, but hardly required. Two co-authors can cooperate to write a paper by taking turns.

Third, it does not require coordination between the agents. We do not have to meet up beforehand and plan out how we will cooperate.

Fourth—and more controversially—I want to stipulate the minimalist account cooperation does not require mutual benefit. Cooperation often will involve mutual benefit, but it does not have to. For example, you and I could cooperate on saving your life. Say there is some sort of danger, and the goal that we both have is to keep you alive. I do that by sacrificing myself. That is not to my benefit, so there is no mutual benefit here, but we did cooperate. There was a shared goal. Even if it does not benefit me, it is still my goal. In other words, so long as the goal is shared, cooperation does not require that the benefit is equal or even mutual. Mutual benefit is common when we cooperate. A more robust theory of cooperation may require it (Axelrod, 1984). But my minimalist account is not unique in excluding it (Tuomela, 1993). I recognize that this is a large claim that needs some further argument, but for the sake of time, I'm going to just assume this point.

Now let's talk briefly about the nature of cooperation and competition. I claim that these are not incompatible, because cooperation is a three-place relationship. Whether we are cooperating depends on what is in that third spot, i.e., what it is we're cooperating on. In other words, both of the following statements can be true *at the same time*:

- A.     **You and I cooperate to play a game of chess.**
- B.     **You and I did *not* cooperate to win a game of chess.**

To have a game of chess (the goal) requires that we both sit down, we move our pieces in turn, we obey the rules of how chess pieces should be moved, etc. All this requires cooperation. It is a shared goal that takes cooperative action to achieve. Not cooperating to have a game of chess would look like me throwing the table and pieces (because I was losing) or violating the rules. So, you and I both perform cooperative actions to achieve this shared goal of playing chess.

Nevertheless, it is not the case that you and I cooperated to win the game of chess. I am acting in a way to win the game of chess, and you are acting in a way to win the game of chess. We are not cooperating on the same end. We do not share the goal that I win the game. My winning requires you losing, and vice versa. So, here we have a case where we both cooperate and compete at the same time, not to the same end, but to different ends. Sometimes, having a competition first requires cooperation. This is the nature of most games. So, cooperation and competition are not incompatible, at least not once we specify the goals.

Next, let's talk about cooperation and trust. Cooperative acts are fostered by trust in the other agents. If I am going to act cooperatively, I am hoping and often trusting that you will too, because if you do not cooperate, then I will not achieve my goal. If, however, you act cooperatively also, then not only do we achieve our shared goal, I now also have further reason to trust you in the future. That agents achieve the cooperative outcome further fosters trust among agents to encourage subsequent cooperative acts (Axelrod, 1984). The point is trust and cooperation create this sort of virtuous feedback loop. The more trust you have, the more cooperation you can get, and the more there's cooperation, the more trust.

This may seem obvious, and in many ways it is. I emphasize it because it will be important later on. Additionally, we cannot fail to note that trust and cooperative acts can be exploited. You may know I trust you and will act cooperatively. Resting in that knowledge, you choose to act selflessly and not cooperatively (what is often referred to as choosing the DEFECT option in a Prisoner's Dilemma), exploiting my trust and cooperation to your own gain. Again, this point is fairly obvious, but it is going to be key to the analysis of propaganda that I am building toward. Trust and an individual's cooperative act can be exploited. Now, I want to demonstrate a particular form of this kind of exploitation. Below, you will see an individual word, one word at a time. Please say aloud the color of each word. Do not read the word aloud, say the color it is printed in.



**RED**

**PURPLE**

**GREEN**

**BLACK**

Good! That was not too hard. The color of the word was also the word written. Now let's do a few more.





**BLUE**

**ORANGE**

**YELLOW**

**GREEN**



Those were harder. Why? As you surely noticed, unlike the first set, for this second set the word that was written did not match the color it was written in. Your first impulse was to read aloud the word, and you had to fight against that impulse to say the name of the color it was printed in. Now this is called the Stroop Effect (Scarpina & Tagini, 2017; Stroop, 1935). It has been widely studied for a long time in psychology. The Stroop Effect is an example of an automatic cognitive process that you must fight against. These are processes that happen quickly and basically automatically in our heads. They are hard to shut down. They are hard to circumvent. They will happen then we can react to them. But there is not a lot we can do to stop them. In this case, reading of the word itself happened faster than finding the name of the color the word is printed in, because reading words, once you are literate in a language, happens automatically. It is very hard, if not impossible, to not read a word in a language that you are familiar with. Try it sometime. Look at a printed word in English and try *not* to read it. Hear a spoken word in English and try to *not* understand what that word is and what it means. You cannot. If you see it, you read it.

The thing about these automatic cognitive processes is that there are many of them at work in our minds all the time. The Stroop Effect is just one example. For our present purposes, one key point about these automatic cognitive processes is that at least some of these automatic cognitive processes are goal-directed actions. For the most part, I do think it's fair to say at least, some of these cognitive processes that we perform are goal directed, and that goal directedness is part of what defines that process as what it is. In the case of the Stroop Effect, your goal for the automatic cognitive process involved is to understand. To be fluent in a language or to be literate in a language is to have the goal of understanding other speakers of that language.

For where we are headed, considering this goal is critical. The reason is that by keeping in mind the goal-directed nature of automatic cognitive processes, we can define a new form of cooperation, one that has so far not fully been noticed in the literature. I contend that there is what I call *non-consensual cooperation*. Admittedly, that name sounds oxymoronic. How could you have cooperation that is non-consensual? To play a game of chess, you and I each have to consent to playing the game. This is true for the vast majority of cases of cooperation. Typically, a cooperative act is an act that one consents to perform. So, cooperation is usually consensual.

---

<sup>3</sup> I think that in all likelihood all actions are defined in terms of the goals they're trying to achieve. But we won't worry about that claim for today.

That said, however, we have just seen that there are some actions that are quick and automatic, like language processing. When you are driving and see a billboard, you do not first consent to read and then consensually perform the action of reading. You just read it. So, this is one of a small range of cases I want to claim that involve non-consensual cooperation. You cooperate, you act cooperatively, and cooperation is achieved, even though you did not consent.

This non-consensual cooperation, I argue, consists of performing an automatic cognitive process that contributes to cooperatively achieving someone else's goal. Cases like this are both exploitative and cooperative. The exploiter, in other words, gets you to perform a simple automatic cognitive process or action. The goal of that action is not set by you. It is set by the exploiter. They put you in a situation where you cannot help but perform one of these simple actions, and to do it in a way that contributes to the goal that they set. So, you acted with a goal at least briefly. That also was someone else's. Thus, when you perform that act, you cooperatively contribute to the achieving of that goal, even though you did not choose to. In other words, you briefly take on the goal of the exploiter.

Now, I cannot emphasize this next point enough. This only applies to these automatic cognitive processes. These are very quick, very automatic processes again. I cannot get you to non-consensually cooperate to play a game of chess. I cannot get you to non-cooperatively consent to deliberative, non-automatic actions. To be *absolutely clear*, I do not claim and in fact strongly deny that any sexual act can be a case of non-consensual cooperation. I cannot get you to non-consensually cooperate with a lot of the actions that we normally focus on. But for this small group of automatic cognitive processes and things like that, non-consensual cooperation is possible, at least on the minimalist account of cooperation I have presented so far.

## 4 Cooperative Communication

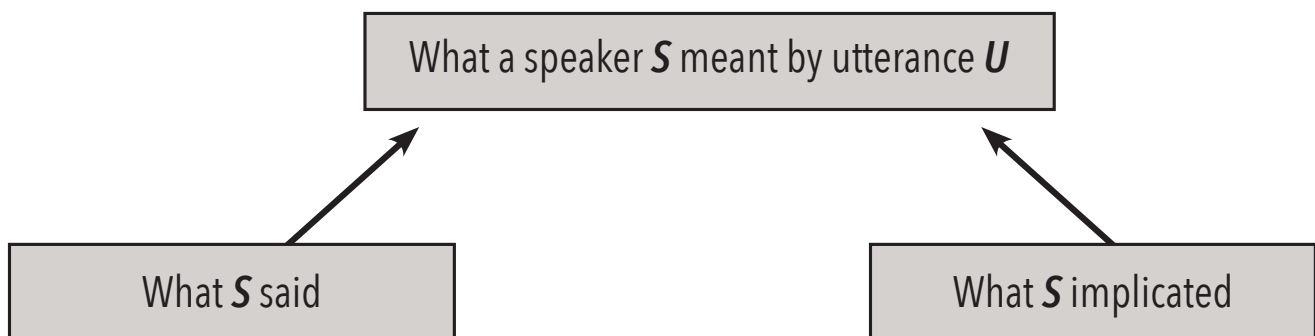
With the basics of cooperation settled, we can now turn to cooperative communication. Here, I will not construct a theory, but rather lean heavily upon the existing theory of the philosopher Paul Grice. This theory presents something called the Cooperative Principle (Grice, 1975), which has been highly influential in a lot of philosophy of language and linguistics since the 1960s. The Cooperative Principle states, "Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of

---

<sup>4</sup> One might contend that you do consent by either of the following means. First, your consent could stem from the fact that you consented to having your eyes open in a venue where words might be placed on signs. Your consent could stem from consenting to learning to read. I find neither argument compelling. First, a sign could be sprung on you unexpectedly, but you would still read it without consent. Plus, your eyes are open to drive, not to read billboards. Second, most people learn to read so young, they did not consent to it. Plus, no one consents to being fluent in their first language, and the same effect works for hearing words.

the talk exchange in which you are engaged” (Grice, 1975, p. 45). In other words, when it is your turn to say something, say something that contributes to the purpose or direction of that conversation as is required at that step. Be cooperative. This is supposed to state not only how people should communicate, but also (*ceteris paribus*) how people do communicate. Conversations are not just a collection of random statements, but they build off of one another. One person says something, then the next person says something that cooperates with it. That is how communication works—it is the logic of conversation—Grice argues.

Grice’s cooperative principle is the cornerstone of his larger theory of meaning, in particular making sense of what a speaker (whom we will call *S*) meant by an utterance *U*. In other words, we take what *S* said, and we want to figure out what *S* meant. They often are not the same thing. To do that we have to not just take what they said, but also what they implicated. This is a technical term that Grice introduced to cover things like, implied, suggested, made as if to say, etc. What he means is often a great deal more to what we mean than what we say. In fact, sometimes what we mean can be the exact opposite of what we say, and that is what Grice is trying to get at.



For example, imagine we have two speakers in the following conversation:

- A: “I’m out of gas.” (Said while A is standing next to their car on the side of the road)  
B: “There’s a truck stop around the corner.”

Taken purely at the level of what is said, B’s statement is uncooperative. It *says* nothing about where to get gas (the topic of the talk exchange, as set by A). That is, unless we also take B to have implicated that the truck stop is open, and that they likely have gasoline available for purchase that A can get and haul back to A’s car. What B *meant* can be worked out by asking, based on what B said, what must be have implicated in order to cooperate with the purpose of the talk exchange. We must assume, in other words, that B was being cooperative.

For now, we will skip over the rules (or maxims, as he calls them) that Grice proposes are needed to cooperate conversationally. It is enough to say that speakers follow these rules, at least at the level of what is implicated. That is how they are cooperating. That is how we can work out what speakers mean.<sup>5</sup>

This notion of working out what is implicated is important for Grice's theory of cooperative communication and where we are heading on propaganda. While what is implicated can be worked out by referring to Grice's Cooperative Principle and four conversational maxims, that is not what audiences typically do when they assess meaning in situ, on the fly in real life. Typically, speakers just get what is being implicated. It's pretty clear. Grice's analysis is a post hoc method to explain how audiences work out speakers' meanings.

How it actually works in real time is different. Typically, audiences understand speakers' meanings (including what they implicate) quickly, even automatically. This automaticity is important, because it means that we automatically parse and process utterances. We cooperate non-consciously, not only with understanding what is said, but also what is implicated (at least usually). Sometimes that cooperation is even non-consensual because these cooperative acts of communication can be exploited.

Here are a couple of relatively benign examples. In *Duck Soup*, Groucho Marx says, "One morning I shot an elephant in my pajamas." [Pause] "How he got my pajamas, I don't know." Granted, it is a joke, but one that works by exploiting our cooperation, in the case how we parse the syntax of the first sentence. Our automatic, cooperative parsing of the first sentence is to take him to have meant that he was in his pajamas. The other possible way to parse that sentence is that the elephant was in his pajamas when he shot the elephant. That one does not occur to us as quickly or automatically. Groucho knows this. He exploits this cooperation to playfully spring upon us the alternate way to parse that first sentence when he utters the punchline. He exploits our cooperation for a humorous purpose, which we presumably will consent to after the fact.

As another example, the comedian Mitch Hedberg said, "I haven't slept for 10 days... because that would be too long." Like Marx, he is playing. He is exploiting how we automatically parsing 'for' to mean "in" (as in, "in 10 days") instead of the literal 'for', meaning a duration of time. He is playing with English language and exploiting the fact that we tend to parse things one way, but they could be parsed another.

These are two examples of non-consensual cooperation in communication, albeit rather benign ones. Of course, we can easily find worse examples of seemingly completely uncooperative conversations. Imagine we have a police officer interrogating a criminal. The cop says, "Where were you last night at 1 AM?" The criminal retorts, "I don't know... the moon." This looks like an uncooperative conversation. In fact, examples like this have been typically thrown against Grice's theory as a form of refutation (Asher & Lascarides, 2013; Sperber & Wilson, 1995). The argument goes that the interlocutors are not cooperating, but they clearly understand one another well. Since Grice's theory says understanding requires cooperation, Grice must be wrong. Or so his critics argue.

---

<sup>5</sup> It is worth nothing that while what is implicated can be worked out by referring to Grice's Cooperative Principle and four conversational maxims, that is not what audiences typically do when they assess meaning in situ, on the fly in real life. Typically, speakers just get what's being implicated. It's pretty clear. Grice's analysis is a post hoc method to explain how audiences work out speakers' meanings.

<sup>6</sup> That is, unless you are a linguist or philosopher of language with Gricean leanings like me. I regularly work out what was meant in conversation like Grice suggests. But I—and those like me—are admittedly a bit weird.

If my account of cooperation—especially non-consensual cooperation—is right, then we can actually make sense of this kind of conversation as cooperative. Remember, cooperation and competition are not incompatible. There is some cooperation going on in this conversation between the cop and the criminal. There is a back and forth, a dance between the cop and the criminal; they both agree to engage in and cooperate toward this exchange. If the criminal (we are assuming for the sake of argument here that the criminal is in fact guilty) were completely uncooperative, the criminal would say nothing at all. The cooperation here is not in terms of the cop's objective of catching the criminal. The criminal clearly is not cooperating with that. But the criminal is cooperating with the goal of having an interrogation. It is on the basis of that level of cooperation that they understand one another. They are both cooperatively contributing to having an interrogation. Again, if the criminal were completely uncooperative, the criminal would say nothing at all. So, this is still a form of cooperative communication. This kind of conversation does not refute a Gricean approach to meaning and understanding and communication.

An even more challenging example, however, is the following. To invoke stereotypes, imagine a group of construction workers, and a woman walks past them on the street. One of the workers yells, "Hey! Where are you going, beautiful?!" The woman says nothing in response. Therefore, this exchange is seemingly an even less cooperative exchange than with the cop and the criminal. Nevertheless, she does understand what was said and meant. My claim here is that this is an example of non-consensual cooperation in communication. She did not get a chance to consent to being catcalled. However, because language processing is so automatic that when you see or hear language in a language that you are fluent in, you cannot, but you cannot help but process it. You will pick up on complex implicatures that take a while to work out, but at the base level of what is said (and to an extent, even a little bit of what is implicated) you cannot help but get it. You cooperate non-consensually. The cat caller set the goal of harassing this woman, and she is unable (and unwilling) to consent to that goal. However, she is not able to not cooperate with it either.

This notion of non-consensual cooperation in communication is central to the notion of propaganda I want to develop, which we can now finally turn to.

## 5 Propaganda

First, remember our two questions: what does propaganda do and how does it do that? To get started on this question, we must consider one underlying question. Is all propaganda morally wrong, i.e., impermissible? How we answer that question is going to be critical to our analysis of what propaganda is.

To highlight the contrast, consider two examples, the first of which is Figure 2. In Figure 2, we have old Uncle Sam. "I want you for the US Army. Enlist now!" It is iconic. It was part of the US Army's recruitment drive in the First World War.



Figure 2 – Iconic Uncle Sam recruitment poster from the First World War by James Montgomery Flagg (Public Domain)

Now consider a similar example of a recruitment poster, only this time it features Darth Vader from Star Wars front and center. It is a recruitment poster for the Galactic Empire. Behind Vader is a host of Imperial soldiers, such as Storm Troopers. In big letters, it proclaims, “ENLIST TODAY! He can’t do it alone!” The message is similar, but the morality is likely very different. Vader, who killed so many and slaughtered children, is the epitome of evil.

With these two images in mind, I ask the reader: Are they both morally wrong (i.e., impermissible)? To determine whether they are both propaganda, we also have to think about this question. How we answer it matters to our analysis, in other words, what we are providing an analysis of.

---

<sup>7</sup> The artist Cliff Chiang produced this image in 2010, and it was reproduced in Pablo Hildago’s *Star Wars Propaganda: A History of Persuasive Art in the Galaxy* (2016). [https://starwars.fandom.com/wiki/He\\_Can%27t\\_Do\\_It\\_Alone?file=Enlist\\_Today.jpg](https://starwars.fandom.com/wiki/He_Can%27t_Do_It_Alone?file=Enlist_Today.jpg)

Suppose, for example, you think that only the Darth Vader poster is wrong. Vader is on the dark side, and he did kill lots of people, including children. In that case, what we are going to be giving a theory of is negative propaganda—impermissible propaganda—not all propaganda. If, however, you think that all propaganda—even if it is for a good cause—is impermissible, then that will mean what we end up with is all propaganda. To simplify things somewhat, I am going to focus on what we might call here negative propaganda, impermissible propaganda, taking the former option.<sup>8</sup> I will omit the argument for that position out of expediency.

Now let's look at a couple of theories of propaganda. First, Edward Herman and Noam Chomsky (1988) focused on the political economy of the mass media and how propaganda manufactures consent. Their objective was not to develop a philosophical account of propaganda that answers our questions. It is a different sort of analysis of propaganda that they offer. Nevertheless, we can construct something of a theory for them that this propaganda model that they create. On their view, what propaganda does is manufacture consent to serve the interests of the elite. It does that by having information pass through these 5 filters that they identify that the elite control.

Next is Jason Stanley's theory of propaganda (Stanley, 2015), which has gotten much more attention of late. On his view, propaganda is a contribution to political discourse that is presented as the embodiment of certain ideals and either supports or undermines the realization of those ideals. That is what it does. In terms of how it does that, Stanley contends propaganda does that by means of emotional or non-rational means.

It is worth noting that for Stanley, not all propaganda is wrong or impermissible. Some propaganda can support a good ideal. That said, it is still propaganda because of how it does that. Stanley is not a huge fan of this supporting propaganda, but it is at least somewhat permissible when the ideal is morally good. So, in terms of our two questions, there are two different (though related) things propaganda can do, either support or undermine some ideal. For Stanley, the more important feature of propaganda is how it does the supporting or undermining. It does it irrationally by circumventing rationality. This emphasis on rationality is key for Stanley's analysis, and is common in other accounts of propaganda. For Stanley, this appeal to rationality is explicitly an appeal to an idealized Kantian notion of rationality of reason from the enlightenment, where reason is opposed to emotion. For reasons I will not go into here, I find this notion of rationality unsustainable and potentially even ridiculous. So, I think it is problematic at best to define propaganda in terms of non-rationality. At the very least, this approach is often going to radically over-identify cases of ration of propaganda, because so much of our political discourse is not a pure appeal to reason.

Though these two definitions of propaganda were rather different, they have two things are common between them. First, they are both examples of the Hobbesian program of linguistic legislation. Both Stanley on the one hand and Herman and Chomsky on the other are trying to claim that there is a problem, namely that propaganda (or at least certain kinds of propaganda) are bad and we have not done a good job of identifying propaganda. The term, they claim, is ill-defined or has been usurped. Thus, if we can correctly and precisely say what propaganda is and how it works, we can spot this evil speech that causes so much trouble. We can, in other words, name it and shame it, thereby diminishing its power in society.

---

<sup>8</sup> For those that are oppositely inclined, taking all propaganda to be impermissible, everything that I say below can be recast to accommodate that view with only minor alterations.

I have a problem with this approach, because I simply lack faith that if we could settle on the definition of what propaganda is and that by naming and shaming, we will lessen propaganda's existence or its impact. It seems dubious at best that we could point to a particular political utterance and say, "This fits the technical definition propaganda!" and expect anyone to be thereby freed from the utterance's propagandistic power. People have been trying to define propaganda for a while now, and these attempts at Hobbesian linguistic legislation do not appear to have done anything to lessen the impact of propaganda over time.

A second point to note about both accounts of propaganda is that they are both focused on the semantic content of the utterances. In other words, they focus on what is being asserted (or what question is asked or what command is given) either directly or indirectly by propaganda. Underlying that assumption of semantic content is the notion that what propaganda is trying to do is get you to believe something, get you to consent to something, get you to do something, or get you to value something. I also think this approach to propaganda is flawed. The problem is that there is plenty of propaganda that lacks any semantic content.

So, I what do I want to say about propaganda? Propaganda is about manipulating trust, not necessarily belief. Consider a speaker who, whipping a mass audience into a frenzy, says, "Scientists are the enemy of the people!" This has content. There is a proposition that is being asserted, and there is an ideal being appealed to here. But what the speaker is primarily trying to do is get the audience to stop trusting scientists. If people's trust in scientists is eroded, then the propagandist does not need to argue against all or even most of what scientists say. They will not be believed or listened to from the outset.

Consider again Figure 1, which we can assume for a moment to be an example of propaganda. I made it, and even I could not tell you what it is asserting. It is not trying to get you to believe anything. It is trying to manipulate your trust, this time trying to get you to trust someone more.

Why then is propaganda fundamentally about trust and not belief? In short, because belief is secondary and not necessary. A propagandist does not need to control what you believe if they can control whom you trust. The desired beliefs and actions will follow naturally as a consequence of the manipulated trust. If I can manipulate your trust, I can then manipulate your beliefs. If I can manipulate your trust, I can then manipulate your actions.

Now I can finally give you my full account of propaganda. Propaganda is:

- An utterance to a mass audience, [Mass condition]
- by which the speaker intends to misdirect the audience's trust in a social agent process [Misdirect condition]
- either by [Exploit condition]
  - o exploiting the audience's trust in the presumed speaker or
  - o by means of non-consensual cooperation.

---

<sup>9</sup> Granted, there is not one settled definition, but there is considerable agreement across definitions about most cases.

Let me step through this definition briefly. The mass condition stipulates that propaganda must be before a mass audience. I cannot whisper propaganda to you and to you alone (and never repeat it). It would not be propaganda then. Propaganda must be uttered to a big group of people (or repeated so that the cumulative audience is big). How big is big? That is a question I will set aside for now, as it likely depends in particular who the social agent or process is that is the target of the propaganda. The necessary size of the audience for propaganda about the president of Texas A&M University-Kingsville is surely smaller than for propaganda about the President of the United States.

The misdirect condition focuses on what propaganda is trying to do. Propaganda is about misdirecting trust in someone or something or some process in society. The social agent that can be individuals, groups of people, institutions, or anything of that sort that classifies a social agent or a process by which something gets done in the social world.

The misdirect condition is inherently normative. It assumes that for any social agent, there is a correct amount of trust that one should have in them.<sup>10</sup> The goal of propaganda is to manipulate the audience's trust to some other point, to get people to trust the target social agent more or less than they should. (To be clear, the target social agent can be the speaker or someone else.) Some propaganda is designed to get you to trust someone less than you should. Other propaganda is meant to get you to trust someone more than you should. If, however, you currently do not trust someone as much as you ought to, then an utterance trying to raise your trust to the right level is not propaganda. I will come back to the issue of how much trust you should have in a moment.

It is not just what the utterance tries to do, however. It also matters how. This is where the exploit condition comes in. In order for an utterance to count as propaganda, it must try to exploit your trust in one of two ways. The first is by means of nonconsensual cooperation, which I discussed above. Some propaganda works by conveying some sort of communicative message about a social agent's trustworthiness that you cannot help but process. In such cases, there is no consent. You, as the audience, see the propaganda poster or here the propaganda message and automatically process it.<sup>11</sup> That process may then lower or raise your trust in the target social agent past how much trust you should have in that agent. Alternatively, you may consent to it. The misdirection in trust, however, is achieved because you trust the speaker (or whom you take the speaker to be, in cases where it is unclear) more than you should.<sup>12</sup>

---

<sup>10</sup>It could also be a range instead of a fixed point on a scale of trustworthiness. No social agent is completely trustworthy. Were there such a social agent, then they should be believed automatically without question or further investigation every time. However, since we are all liable to be mistaken or to say things unclearly (to say nothing of lying or other forms of deception), no one is completely trustworthy. A social agent's trustworthiness is a function of their tendency to hold false beliefs, to utter things unclearly, to lie, to bullshit (Frankfurt, 2005), to deceive, etc.

<sup>11</sup>This automaticity is a key part of why, when it comes to propaganda, simpler is better. After the automatic processing, the more you intentionally consider the message, the less likely you may be to have your trust manipulated.

By way of an example, consider again Figure 1, the poster with President Vela on it. If I plastered that all over campus one night, it would look like the speaker is President Vela or the university. People around campus would not know I did it, and so they would presume the speaker is the university. In that case, whether or not the poster was seen as propaganda depends on how much trust they already had in the university. In other words, whether or not someone sees an utterance as propaganda depends on how much trust the place in the target social agent in the first place, and so whether they see trust as being *misdirected*.

This point demonstrates that my definition is what I am calling a diagnostic definition. I am not trying to settle whether *particular* utterances are propaganda by means of my definition of propaganda. People will disagree on cases. For Figure 1, I am guessing that some of you think it is propaganda, and some of you think it is not. What my definition is trying to do is diagnose why people disagree. Attributions of propaganda, I am claiming, will hinge upon the following. First, whether you think all propaganda has a negative connotation or not, which we are more or less setting aside. Second and more importantly, it depends on whether the intended shift in trust is regarded as a misdirection. Figure 1 was trying to shift your trust. In the imaginary scenario where I put it all over campus, I am trying to build trust in President Vela and the University. If you think that is appropriate and that trust is warranted, then you will not see this as a misdirection of trust. In that case, Figure 1 is not propaganda. Alternatively, if that trust is not warranted, then you the poster as a misdirection of trust. Then, Figure 1 is propaganda. This prior assessment of trustworthiness is the main feature around which attributions of propaganda will hinge.<sup>13</sup>

To wrap things up, I offer a prediction. Everything so far has been an argument to plausibility. My account of propaganda may be intuitive, at least to some, but there was no decisively logical argument offered. Empirical evidence could, however, support my overall account. To that end, I will make a prediction. If I am on the right track, here is what we should find if we were to do some experimental checking.

Imagine we showed Figure 1 and other purported examples of propaganda to research participants. Then we asked them whether or not these particular examples were propaganda. Participants respond on a scale of 1 (absolutely not) to 5 (absolutely). But first, we also asked them how trustworthy they regard the targeted social agents (e.g., President Vela in the case of Figure 1), likewise on a five-point scale of 1 (very untrustworthy) to 5 (very trustworthy). My prediction is that should see a response pattern like Figure 4.

---

<sup>12</sup> It is a presumed speaker here, because in some cases (especially for visual propaganda) where it is not readily apparent who the speaker is. So audiences must make an assumption of the speaker's identity. That assumption then plays a role in how much their trust is misdirected.

<sup>13</sup> That does mean that effective propaganda will no longer be regarded as propaganda. If I can shift your trust—if I can misdirect it and get you to start trusting me more and more and more—you will not see what I am doing as propaganda. You will see it as appropriate, not as a misdirection of trust.

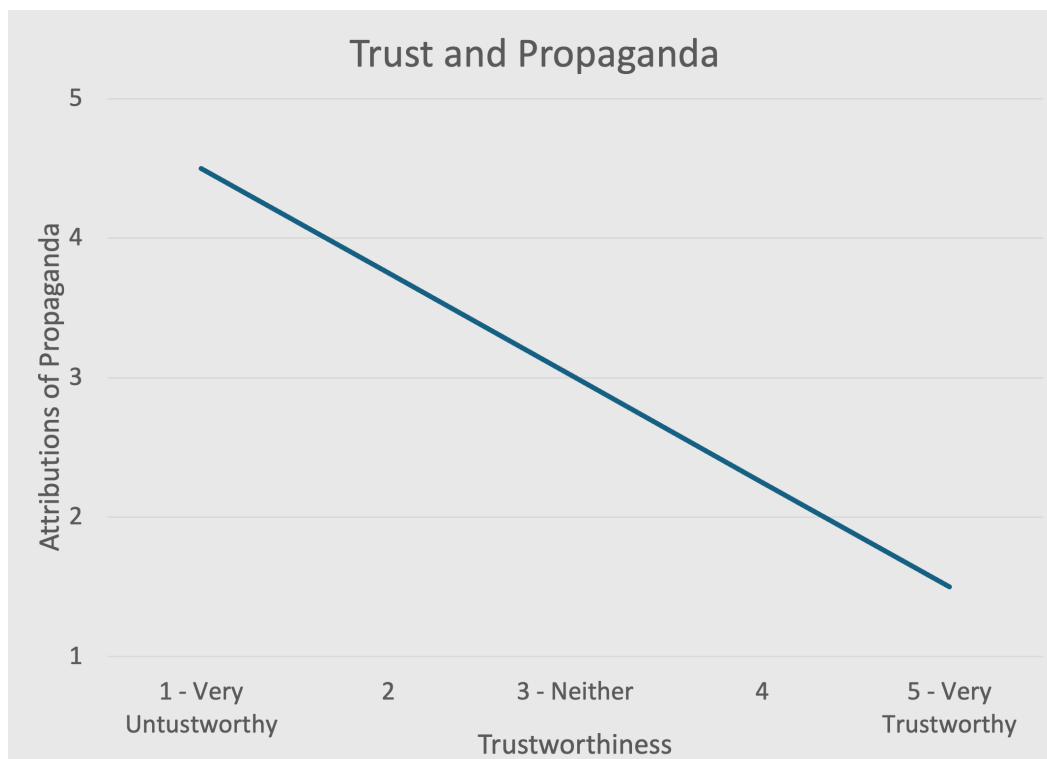


Figure 4 – Predicted response pattern when asking first about a social agent’s trustworthiness and then whether an utterance about or depiction of them is propaganda. The more people initially find someone untrustworthy, the more likely they are to attribute propaganda.

For those who say President Vela is very untrustworthy, they would be much more, much more likely to say that Figure 1 is propaganda. On the contrary, those who initially see President Vela as trustworthy will be much less likely to say Figure 1 is propaganda. That is the prediction that comes out of my diagnostic definition.

## 6 Conclusion

In conclusion, then, propaganda is about manipulating trust as a means of control. It is not first and foremost about belief, as commonly supposed. It is about manipulating your trust. It is not based fundamentally on irrationality. It is about exploiting trust or nonconsensual cooperation.

That said, while the debate over the definition of propaganda is intellectually and philosophically interesting, it is something of a red herring. Pointing this out is the core of what a diagnostic definition does here. My account of propaganda explains what propaganda does and how it does it. But more importantly, it points to a deeper debate. Arguments over whether particular utterances are propaganda obscure deeper debates about whom to trust and how much.

## Works Cited

- Asher, N., & Lascarides, A. (2013). Strategic conversation. *Semantics & Pragmatics*, 6(2), 1–62.  
<https://doi.org/10.3765/sp.6.2>
- Austin, J. L. (1962). *How to Do Things with Words*. The Clarendon Press.
- Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.
- Carroll, L. (1897). *Symbolic Logic*. Macmillan and Co.
- Frankfurt, H. (2005). *On Bullshit*. Princeton University Press.
- Grice, P. (1975). Logic and Conversation. In D. Davidson & G. Harman (Eds.), *The Logic of Grammar* (pp. 64–75). Dickenson Publishing Company, Inc.
- Harris, D. W., Fogal, D., & Moss, M. (2018). Speech Acts: *The Contemporary Theoretical Landscape*. In D. Fogal, D. W. Harris, & M. Moss (Eds.), *New Work on Speech Acts*. Oxford University Press.
- Herman, E. S., & Chomsky, N. (1988). *Manufacturing Consent: The Political Economy of the Mass Media*. Pantheon Books. 1988
- Hobbes, T. (1651/2008). *Leviathan* (J. C. A. Gaskin, Ed.) Oxford World Classics Edition. Oxford University Press.
- Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology*, 8.  
<https://doi.org/10.3389/fpsyg.2017.00557>
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/CBO9781139173438>
- Searle, J. R. (1976). A Classification of Illocutionary Acts. *Language in Society*, 5(1), 1–23. JSTOR.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Harvard University Press.
- Stanley, J. (2015). *How Propaganda Works*. Princeton University Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
- Tuomela, R. (1993). *What Is Cooperation?* *Erkenntnis* (1975–), 38(1), 87–101. JSTOR.



The Annual Faculty  
Lecture is sponsored by  
the Office of the President