The philosophical problem common to both Plato and Rawls was how to form a just society. Plato's solution was to institute a sustainable authoritarian state with the help of a "philosopher king". John Rawls' more modern idea was to build a social consensus around the form of the just society, by imagining an initial bargaining position, where, each participant, under a "veil of ignorance", has "forgotten" their own socio-economic status. The idea being, that by abstracting out socio-economic status, the participants in this imaginary constitutional convention are more likely to agree to principles of equality and justice for all, that, just by coincidence, would resemble the modern welfare state.

As a thought experiment, I suppose that is a fine thing to do, but I think the key to understanding what makes a just society is understanding the difference between humans and all other animals; and, (spoiler alert!) that difference has to do with our ability to create and maintain normative systems like morality, language, and truth.

We can think of human society as a kind of kluge - a contraption built in a haphazard way by using whatever bits and pieces of things are immediately at hand; over the long haul the environment often intervenes, creating inequalities, and we come up with further modifications in order to continually deal with the centrifugal pressures threatening to pull us apart. Looking back, we can see that the development of all human institutions - kinship, moral systems, language, myths, religion, government, money, legal systems, and educational systems - all show this gradual and haphazard growth process.

What frames it all is that every element of human culture comes from our primal ability to agree to form and follow rules of behaviour, where we also expect others to do the same. We can call this framing, "normativity"

In our closest animal cousins, the primates, there is no evidence for a shared system of rules and meanings that can override dominance. In the Darwinian state of nature, individuals have no incentive to share information with others unless it strictly benefits them to do so. Without a normative system in place already, language would probably never have developed. A normative system overrides self-interest and encourages altruism by successfully punishing cheaters. It is much more likely that a communication system such as language, with shared meanings, rules, and detachable units, arose after we first established an initial normative system. I describe this initial normative system here . In this essay I want to demonstrate how language depends on the additional normativity of truth to get off the ground.

In a fascinating book called The Handicap Principle, Israeli zoologist Amotz Zahavi, points out that unlike animal vocalization, which is tightly linked to an animal's abilities and physical state, "human language has no component that guarantees its reliability and prevents cheating." Language is a cheap egalitarian way to get messages across, unlike animal vocalization, where, as Zahavi puts it, reliability is hard to fake. (The smaller the lion, the more feeble the roar; the bigger the lion, the louder the roar.)

In animal communication, the signal is closely tied to the animal's physical state because the more effective the signal is in establishing and maintaining dominance, the more likely that animal will be reproductively successful. That is why truth is not needed in animal communication. "All these signals amplify the ability of the observer to spot superiority or defects in the animals that carry them." Weaker or inferior animals are not able to fake these signals because they are somehow deficient in the physical characteristics that are needed to produce the reliable signal.

Thus, I argue, the need for truth comes into the picture with the first appearance of language. Because we share, we humans need truth,

whereas non-human animals don't. Language is fundamentally based on sharing. It involves shared meanings, shared rules, and detachable and manipulable symbols that can be combined in numerous ways to construct novel sentences with unanticipated meanings. But unlike animal vocalization, the ease of communication with language makes it correspondingly easy to deceive others. Think of the kluge metaphor: by inventing language we opened a pandora's box of deception and misconception, and, in order to preserve reliability our ancestors had to add on an new regulatory system and today we are all still intimately involved in this system; we call it - "truth".

When we communicate we also share a universal commitment to tell the truth and counter lies and misinformation. With the development of language, humans took over the task of ensuring reliability from a largely unconscious nature.

Imagine a universal team sport, a game that everyone in human society is part of, a game where once you learn how to play it, you are in for life. That is what truth is. Truth isn't a thing, a property, or a relation. Truth is a system of regulating behaviour - a normative system.

In one sense truth isn't a game, because we can't opt out of playing without opting out of society. The way truth works is as if it were a referee that everybody, together, unconsciously imagines - a shared understanding of an idealized correspondence between our beliefs or utterances and an imagined, mind-independent, objective reality. Those who sometimes break these rules deliberately, can be called liars; they receive warnings and can be penalized for continuing to lie. Those who do not share this collective understanding of truth telling as the default mode of communication, and, always have to consciously pretend to follow the "truth referee", are called psychopaths, and, once discovered by the rest of us, they are kicked out of the game. This is an essential part of maintaining any human society, because when we don't recognize or do anything about

psychopaths, the pool of trust is in danger of being emptied and it becomes “game-over” for all of us.

Truth works because we believe in it and respect it as an impartial referee. It’s a beautiful thing just like a well-played game is a beautiful thing. Even though it’s a fallible system that somewhat belies our faith in it, the fact that it takes all the participants, their dedication, and their commitment to the truth to make it possible, also makes it work better.

We can adhere to telling the truth and come to value and defend it when we all expect everyone else to do the same. Furthermore, we can have strongly felt judgements about liars which will serve as motivation to help each one of us to be part of the collective enforcement of moral and epistemological norms. Every culture has collective ways of punishing lying and immorality, from shaming, inducing guilt, and ridiculing, to more serious sanctions like shunning, and expulsion. In many cases, moral emotions are the indispensable motivators for detection and enforcement of cheating.

I’m going to argue that games, normative systems, and language all share some of the non-Darwinian qualities of a Common Pool Resource, and if we can understand what a common pool resource is, then we can understand the basics of what human behaviour is. I put it to you, dear reader, that the concept of a common pool resource, or CPR for short, developed by the Nobel Prize Winning American Institutional Economist, Elinor Ostrom, is a key concept for understanding normative systems. I maintain that it is the basic underlying substructure of all human behaviour and what really distinguishes us from the rest of the animals.

Where common property is on a small-community-scale, everyone needs to be the eyes, ears, and bodies on the ground, in order to detect and prevent overfishing, hogging water from a reservoir, overgrazing, or any other overuse of communal resources; and, in a common pool resource,

each and every member both follows the rules and enforces the rules. Being on the same “team”, in effect is a group identity that goes with being a part-owner of a communal resource. The double function of adherence/enforcement exists in all CPR’s and normative systems. As Ostrom reports, the most stable and workable CPR’s are the ones where commitment to follow the rules is at the same time a commitment to enforce the rules. It’s when this double commitment is absent that you get the so-called “tragedy of the commons”, a situation where the commons is degraded by over-use.

Ostrom also found that common pool resources that survived over generations all demonstrated a powerful sense of collective identity amongst the CPR owners. We can see how this works by thinking about how introducing teams to a sport energizes the game. “Team identity” - identifying with team players,wearing the same colours, sharing similar tasks and objectives, feeling strong emotional bonds with teammates - is a powerful motivator that makes each player give it their all.

A game is played through when the players respect the rules and abide by the referees calls. In the same way a common pool resource is maintained because its common owners believe in and abide by its rules, individually and collectively enforcing the rules at all times.

Truth works in the same way as do norms and common pool resources. It works because everyone believes in it, everyone commits to it, and everyone judges that those who don’t are morally deficient. This explains why lying is more complex than telling the truth. Truthfulness is presupposed in almost all conversations; if truth is part of the background, then it is lying that requires the extra effort. Sure enough, lying can be detected by a machine, because it takes extra psychic and physical energy to pull off a lie, whereas telling the truth is simply our default mode of communication.

There is a philosophical “theory” of truth called “deflationism”, which gets its appeal by presupposing this point, claiming that “truth” is nothing more than a logical device, when it takes for granted the fact that it is already assumed to be the default mode of communication before we even utter a word. A real theory of truth should explain this fact, not take it for granted.

In a team sport such as hockey, when a player breaks a rule, he or she is called out and penalized by the referee. All the players know the rules and abide by the referee’s calls, or they don’t get to play. In contrast, and this is an extremely important point, there is no real physical referee in normative systems, yet we seem to function pretty well most of the time by internalizing the rules and checking ourselves against everyone else. All humans have the amazing ability to “internalize” rules - to impartially follow and enforce rules by unconsciously imagining some proxy for a referee, like an “impartial observer” or the “rules of grammar”.

Humans are different from all other animals because we have normative systems like morality and truth. These systems run on shared understandings and common expectations. When trust fails, when expectations fail, normative systems fail. Like a common pool of resources, they must be maintained by frequent checking for rule-breakers, and by procedures for punishing or ultimately, expelling them. And, normative systems share both with self-organized systems and common pool resources, the facts that everyone participates and no one is in charge. The crucial difference between normative and non-normative systems like human conventions, is that normative systems like morality don’t support self-interest with positive reinforcement; normative systems work to yoke self-interest to the collective interest. Normative systems, like truth, are fallible, improvable, and they are not based on Darwinian self-interest. And that, in a nutshell, is what creates the basic foundation of a just society.