EXECUTIVE SUMMARY

How can we get people to behave themselves? Experts often assume that humans are selfish creatures who respond only to punishments and rewards, and who can’t be trusted to do a good job or refrain from lying, cheating and stealing unless given the right “incentives.” Yet every day we see people behaving ethically and unselfishly—few of us mug the elderly or steal the paper from our neighbor’s yard, and many of us help strangers. We nevertheless overlook the good aspects of our own natures and fixate on the bad things people do and how we can stop them.

This focus on bad behavior obscures the reality, and importance, of goodness, leading us to neglect the crucial role our better impulses could play in shaping society. Evidence from behavioral science and experimental gaming demonstrates that unselfish prosocial behavior (sacrificing to follow ethical rules, or to help or avoid harming others) is far more common and important than generally recognized. Under the right conditions, the vast majority of people act as if they have a conscience that causes them to act ethically and look out for others’ interests.

This paper unpacks how these empirical findings can be used to develop a “Jekyll/Hyde” model of how human behavior shifts predictably from purely selfish to prosocial, depending on certain social cues. The model can help us both understand how legal rules and social norms work, and help us use them more effectively. It also suggests that, rather than leaning on the power of greed to channel human behavior, policymakers often might do better to focus on and promote the force of conscience—the cheapest and most effective police force one could ask for.
How Good Laws Make Good People

What’s the best way to get people to behave themselves? Legal and policy experts often assume people are basically selfish creatures who respond only to punishments and rewards, and who can’t be trusted to do a good job or refrain from lying, cheating and stealing unless given the right “incentives.” Are CEOs neglecting their firms? Tie their pay to share price with stock grants and options. Are America’s children failing to learn their ABCs? Give teachers bonus pay if they raise test scores, and fire them if they don’t. Are Medicare expenses increasing too quickly? Use “pay for performance” schemes that give doctors and hospitals a direct financial motive for keeping health care costs down.

This emphasis on “incentives” and “accountability” relies on a *homo economicus* model of purely selfish human behavior that was developed for theoretical economics, but has since spread to be embraced by policymakers, business leaders, and experts in a wide range of fields from political science to philosophy. Today, it’s hard to find a serious discussion of the possibility that we might encourage or discourage particular behaviors by appealing not to selfishness, but instead to the force of conscience. Many modern experts would snicker at the very idea. Conscience is viewed as the province of religious leaders and populist politicians, not lawyers, businessmen, or regulators.

Blind to Our Own Goodness

This is odd, for every day we see people behaving ethically and unselfishly--few of us mug the elderly or steal the paper from our neighbor’s yard, and many of us go out of our way to help strangers. Our very language reveals our preoccupation with moral assessments. Just as the Inuit have many nouns for snow, English has a multitude of words to describe unselfish, conscience-driven behavior, including: virtuous; kind; fair; agreeable; honest; ethical; trustworthy; decent; upright; faithful; altruistic; humane; loyal; charitable; selfless; principled; conscientious; cooperative; generous; considerate; caring; and compassionate. Most tellingly, another simple word often used to describe unself ish behavior is “good.”

Policymakers and business leaders nevertheless usually overlook the unselfish, “prosocial” side of human nature, fixating instead on selfish misbehavior and how to stop it. This fixation may stem in part from certain biases in perception. For a surprising variety of reasons, including our psychological biases, the structure of our language and our society, and the way we select and train experts in law, economics, and business, we tend not to “see” ethical and unselfish behavior, even when it happens under our very noses. Americans watched their television screens aghast when hundreds of New Orleans residents began looting in the lawless wake of Hurricane Katrina. Few stopped to marvel at the miracle of the thousands of New Orleans residents who were not looting.
This collective blindness to our own capacity to act conscientiously—or, as behavioral scientists might put it, our capacity to act prosocially—can lead us to overlook the reality, and importance, of goodness, leading us to neglect the crucial role our better impulses could play in shaping society. Rather than leaning on the power of greed and selfishness to channel human behavior, our laws and policies might often do better to focus on and promote the force of conscience—the cheapest and most effective police force one could ask for.

**Experimental Gaming and the Science of Conscience**

Luckily, modern behavioral science offers policymakers a guide for how to put conscience to work. To a behavioral scientist, conscience might be better described as unselfish prosocial behavior, an objective approach does not require speculation on the internal motivations that sometimes lead people to behave ethically and sacrifice to help or avoid harming others. Extensive empirical evidence from behavioral economics, social psychology, and evolutionary biology proves that, far from being rare and quirky, unselfish prosocial behavior is not only common, but highly predictable—and easy to manipulate.

Over the past half-century, behavioral scientists have devised an ingenious parade of experiments to test what real people do when placed in situations where their material interests conflict with the interests of others. “Social dilemmas,” “ultimatum games,” “dictator games,” and “trust games” all test what human subjects actually do in various situations where they must choose between selfishness and prosociality. The results of such experiments demonstrate beyond reasonable dispute that, far from being rare and quirky, unselfish prosocial behavior is endemic. Researchers around the globe have run hundreds and perhaps thousands of experimental studies that consistently demonstrate that unselfish prosocial behavior is a real and very common phenomenon. Sometimes—in fact quite often—we sacrifice our own material payoffs in order to help or to avoid harming other people.

That possibility should interest anyone who lives among, cares about, or deals with other human beings. But it should especially interest those who study and care about law, regulation, public policy, and business management. Each of these fields deals with the central problem of getting people to behave in the fashion we think of as “conscientious”—to work harder than the minimum required, to pay taxes instead of cheating, to keep their commitments, to respect others’ rights and property, and to refrain from violence, theft, and mayhem.

**The Jekyll/Hyde Syndrome**

At the same time, the empirical fact that people sometimes act unselfishly is only useful if we have some idea of when, and why, this happens. What determines when we act selfishly, and when we show consideration for others’ welfare and
for following ethical rules?

Luckily, experimental gaming demonstrates not only that conscience (or at least conscientious behavior) exists, it also teaches a great deal about when and why conscience comes into play. In particular, the data demonstrates that while most people are willing to sacrifice for others, they are only willing to act unselfishly in certain conditions. We seem to be collectively afflicted with a “Jekyll/Hyde syndrome” that causes us to shift predictably between selfish and unselfish modes of behavior in response to certain social cues.

In particular, three social cues seem especially important to triggering unselfish prosocial behavior. The first is instructions from authority. As we have known since the days of Stanley Milgram’s famous experiments on obedience, in which subjects obeyed instructions to administer what they thought were potentially fatal shocks to another human being (really an actor pretending to be shocked), people tend to do what they are told to do. This instinct for obedience, it turns out, can also be employed for more prosocial purposes. When asked to do so, subjects in experimental games routinely act prosocially—even when it is personally costly for them to do so.

Perceptions of others’ behavior also play a critical role. We are herd animals who act nicely when we think others are nice, and nastily when we think others will be nasty. When experimental subjects are led to believe others will act prosocially, they become more likely to act prosocially themselves—again, even when they must sacrifice to do so.

Finally, people seem more inclined to behave unselfishly in experiments when they believe others will enjoy large gains, not small, from their unselfishness. We seem to be “intuitive utilitarians” who are willing to sacrifice more when we believe others will benefit more from our sacrifice.

By manipulating social variables like instructions from authority, beliefs about others’ behavior, and perceptions of benefits to others, researchers have been able to dramatically change the behavior of human subjects in experimental games. When the social cues favor prosociality, behavioral scientists can elicit universal or near-universal unselfishness. Conversely, when subjects are told to act selfishly, believe others would act selfishly, and believe selfishness is not too costly to others, they exhibit near-universal selfishness.

The Role of Personal Cost

Experimental gaming permits us to develop a relatively simple, three-factor model in which conscience is triggered primarily by the three social cues of instructions from authority, belief in others’ prosociality, and perceptions of benefits to others. However, saying that social context matters does not imply that personal costs don’t. People are far more capable of acting unselfishly than the homo economicus model admits. At the same time, the experimental evidence suggests that the supply of conscience is not unlimited. As the personal cost of
acting prosocially rises in an experiment, the incidence of prosocial behavior observed declines.

These empirical results indicate that if we want people to be good, it’s essential not to give them strong motivations to be bad. Unlike Oscar Wilde, most of us can resist small temptations. It’s the big temptations that do us in.

An Example: The Perils of Ex Ante Incentives

As an example, consider the disturbing implications that conscience carries for the contemporary enthusiasm for trying to channel human behavior through financial incentives. This practice is particularly common in the business world, where federal tax law since 1993 has required corporations to tie executive pay in excess of $1 million to “objective” performance metrics.

Unfortunately, behavioral science predicts this approach may often be counterproductive. Unless corporations can somehow develop “complete” employment contracts that fully specify all duties and obligations under every possible set of circumstances, emphasizing ex ante incentives will often have the perverse and unintended effect of promoting opportunistic, even illegal, behavior. Consider how the widespread adoption of stock option plans to “incentivize” executives at Enron and Worldcom to raise stock prices had the unintended effect of incentivizing them instead to commit massive accounting frauds.

To see why this might happen, recall that unselfish prosocial behavior seems triggered by at least three important social influences: (1) instructions from authority; (2) beliefs about others’ selfishness or unselfishness; and (3) perceptions about the magnitude of the benefits to others from one’s unselfish actions. Emphasizing ex ante financial incentives undermines all three. This is because offering a material incentive to induce someone to do something inevitably sends the unspoken signal that selfish behavior is both expected and appropriate to the task at hand. It suggests that others in the same situation are behaving selfishly. Finally, it implies selfishness must somehow be beneficial. (Otherwise, why is it being rewarded?)

Incentive contracts can also create large temptations that kill off conscience. The investment banking industry, for example, is notorious for employing incentives schemes that allow its traders to reap rewards that may reach into the millions of dollars. As we have seen in recent years, in the effort to reap these rewards, Wall Street traders took on excessive risks that nearly brought down their firms and the wider economy. Similarly, mortgage brokers paid bonuses for loan volume approved millions of inappropriate and shaky subprime loans.

Unless done very carefully, focusing on extrinsic incentives can have the unfortunate side effect of “crowding out” internal incentives like trustworthiness, honor, and concern for others’ welfare. Emphasizing material incentives, it turns out, does more than just change incentives. At a very deep
level, it changes people. Relying too much on selfishness can become a self-fulfilling prophecy. By treating people as if they should care only about their own material rewards, we ensure that they do.

**Taking Conscience Seriously**

Why should contemporary legal and policy experts be eager to do the extra work needed to incorporate the idea of conscience into their analysis? The answer is simple: we can’t afford not to. Peace and prosperity depend on our human capacity for courtesy, consideration, and forbearance. Today, unselfish prosocial behavior is so deeply woven into the warp and woof of Western life it often goes unnoticed. People take cash out of ATM machines without hiring armed guards; beefy young men stand patiently in line behind frail senior citizens; drivers wait for red lights to turn green, even when the police are nowhere in sight. We take for granted the countless unselfish acts of cooperation and restraint that bind us together in a civil society, just as we take for granted the gravitational force that keeps us from floating out into space.

But just as we cannot live well without gravity, we may not be able to live well without conscience. The statistical evidence indicates that cultural habits of unselfish prosocial behavior are essential to both economic growth and psychological wellbeing. Evidence is also accumulating that unselfish prosocial behavior is on the decline in the United States. Just as environmental scientists have become concerned about many sources of scientific data that point to the possibility of global warming, some social scientists have become concerned about the growing evidence that points to the possibility of “conscience cooling.”

If Americans are indeed becoming collectively more selfish, unethical, and asocial—concerned only with their own material welfare, and not with the fates of their communities, nation, or future generations—this shift threatens both our happiness and our prosperity. We need to respect, and cultivate, conscience.

**Tell us what you think of this Issues in Governance Studies.**

**E-mail your comments to gscomments@brookings.edu**

This paper from the Brookings Institution has not been through a formal review process and should be considered a draft. Please contact the author for permission if you are interested in citing this paper or any portion of it. This paper is distributed in the expectation that it may elicit useful comments and is subject to subsequent revision. The views expressed in this piece are those of the author and should not be attributed to the staff, officers or trustees of the Brookings Institution.