RESEARCH FOR SOCIAL IMPACT

THE ROLE OF TRUTH IN THE PURSUIT OF MORAL ENDS: A PHILOSOPHICAL AND EMPIRICAL INVESTIGATION

LYING IN THE NGO WORLD

January 2015

Contents

The Role of Truth in the Pursuit of Moral Ends: a Philosophical and Empirical Investigation

Maya Dank, Nathan Morton, Gilad Tanay & Yuval Ziv

Executive Summary

Paraphrasing Tolstoy, all true statements are the same, but each lie is different after its own fashion. For example, lying may be done for personal gain, to benefit others and sometimes both. This raises the question, when, if ever, is it morally permissible to lie for the sake of good causes? These questions are of great practical importance for NGO professionals, in particular, due to the fact that NGOs typically aim to promote good causes, and because they are often put in the position where their lies may influence many people and many lives.

The present study investigates the question of lying in the NGO world. The investigation advanced in two different yet complementary directions: first, it addressed the empirical question, to what extent are NGO professionals prone to lying for the sake of good causes. Second, it addressed the ethical question, when is it morally permissible to lie for the sake of good causes in the context of the NGO world.

In the ethical analysis the conclusion was reached in two stages. First, we discussed the conditions under which persons in general are either justified or have a moral obligation to deceive in order to bring about a moral good. Then the analysis extended to the particular special obligations a person may have in her professional role as an NGO executive, acting not as an individual but on behalf of her organization.

Taking this approach revealed an interesting finding: it seemed that there must be much more stringent restrictions on lying in the pursuit of moral ends for NGO executives compared to the average person lying to promote good causes.

However, this was ***not*** what we found **in the empirical study** conducted among 295 American and British NGO executives and professionals from the private sector, which examined what people actually believe in. In fact, it was discovered that both groups believed lying for moral

causes in the NGO world to be more permissible than in other contexts, even when the benefit to others was similar. Therefore, this belief seems to reflect a general social convention which clashes with the ethical recommendation.

Preface

Paraphrasing Tolstoy's famous quote, all true statements are the same, but each lie is different after its own fashion. For example, lying may be done for personal gain, to benefit others and sometimes both. This raises the question, when, if ever, is it morally permissible to lie for the sake of good causes? These questions are of great practical importance for NGO professionals, in particular, due to the fact that NGOs typically aim to promote good causes, and because they are often put in the position where their lies may influence many people and many lives.

Thus, the present study investigates the prevalence of lying in the NGO world. The investigation advances in two different yet complementary directions: first, it asks how NGO professionals behave in regards to lying (or more precisely, it asks about their attitudes towards the act of lying) in cases in which lying would promote good causes, meaning it is expected to benefit others. The second direction is a philosophical one: it asks the moral question when lying to promote good causes in the NGO world is ethically permissible. Accordingly, the report consists of two parts: an ethical (philosophical) part and an empirical (psychological) one.

Part One: The Role of Truth in the Pursuit of Moral Ends: a Philosophical Investigation

What are the general conditions under which it is permissible to lie in the NGO sector?

I. Introduction

The first component of the report is a philosophical analysis, and it aims to discuss the normative issue of *how NGO executives ought to act* (from a moral point of view): when are they morally justified in lying to promote some greater moral cause? The answer clearly will not be "never," so we will need a general theoretical analysis to discover a principled moral guideline.

We will arrive at our answer in two stages. First, we investigate the conditions under which persons in general are either justified or have a moral obligation to deceive in order to bring about a moral good. We start from here because these moral reasons certainly must figure in our analysis of well-intentioned lying, as they will apply in general. Determining the conditions under which any particular person might be justified in lying for good causes is what we will call the *general question* of our analysis.

The second half of our analysis then extends to the particular special obligations a person may have in the professional role of an NGO executive, acting not as an individual but on behalf of their organization. In theory, there may be additional reasons deriving from specific facts about their professional role that militate against well-intentioned lying, which would further constrain the executive's behavior. Or, there may be facts that *abrogate* the moral duties that the executives would normally be subject to when not acting in a professional capacity.[1] Theoretically, then, these could narrow or widen the conditions for moral permissibility that we discover in the first stage. We will call this our *specific question*. In the course of finding a

[1] To use an analogy from law: in virtue of their role as defenders of public safety, police officers have a wider justification in using lethal force (while *acting as an agent of the law*), while ordinary citizens have narrower legal duties.

persuasive answer to each question, we aim to formulate a tentative but specific, reasonable, and principled set of conditions under which the executive of an NGO is morally justified in telling a lie to achieve a moral end.

II. Methodology

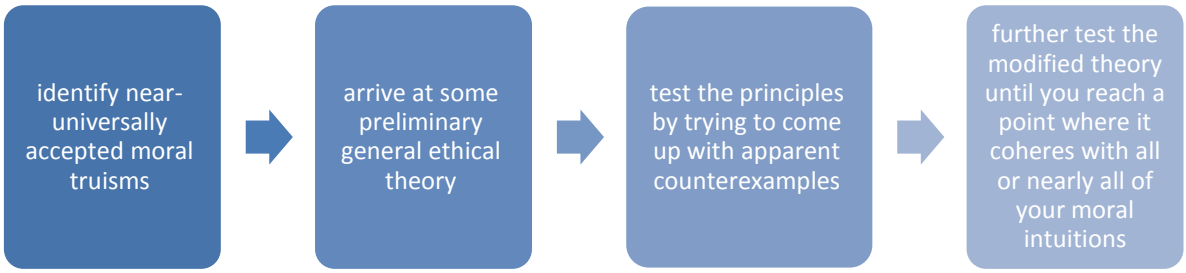
We will proceed by first answering the general question of when lying for good causes might ever be morally justifiable, and then go back to fine-tune our analysis in order to answer the specific question about the NGO world. Because of the relative generality of the first question, we will use two philosophical methods commonly employed in applied ethics: the methods of *reflective equilibrium* and that of *overlapping consensus*.

The method of *reflective equilibrium* is more "bottom up" in the sense that it gives strong weight to our particular moral intuitions. We start by identifying near-universally accepted moral truisms that just "seem" to be paradigmatic, *obvious* instances of ethical and unethical conduct: examples are "it is morally good to volunteer a large amount of one's spare time to feed the hungry," "it is permissible to intentionally cause another person to feel transient pain – however excruciating – if such pain is necessary to administer critical medical treatment," and "it is always wrong to physically or sexually abuse young children."

Second, we abstract from all these particular judgments to arrive at some preliminary general *ethical theory*. We try to give some explanation of *why* these particular types of actions are right or wrong, case-by-case, and try to identify a list of salient features or considerations common to all cases. We then generalize from these features/considerations to form some *general* principles, which are intended to apply to all possible actions. These principles at once analyze and explain why particular actions are morally good, bad, or neutral, and are also to be used as a *practical guide* in forming all of our moral judgments. For example, one morally relevant feature common to all of the particular judgments above is the potential for suffering or happiness that our actions bring to other people. A fitting general principle (especially taking the second example into account) we could formulate might be: *an action is right when the well-being it would bring to others is greater than any suffering (in the long-run, over people's lifetimes) it would cause, and an action is wrong when it has the opposite effect.*

However, the principles we arrive at by this second stage are not the endpoint. In a third stage, we *test* our principles by trying to come up with apparent counterexamples, where applying the theory yields conclusions that misalign with our intuitive judgments. So, we couldn't yet rest with the above principle as the final word on ethics until we subject it to severe criticism. Consider the following classic counterexample to this type of ethical system: you run a hospital, and five of your patients need different kinds of organ transplants immediately, or they will die. For argument's sake, suppose that you *know* that you could kill a patient with five healthy organs and harvest them to save the five, without anyone ever finding out. According to the general principle above, if you *can* do this, you ought to – after all, the *best* action is to minimize suffering and maximize well-being, and assuming human lives are equally valuable, all else being equal, *not* harvesting the organs is equivalent to choosing to save only one person out of six when one *could* instead save the other five. If one can accept the consequence that killing people to harvest their organs *can be* morally permissible, then the principle is fine as it stands. If however it seems "just wrong" that any theory of morality could sanction such actions (a reaction many will likely have), then perhaps instead this is strong evidence that there is something wrong with the principle.

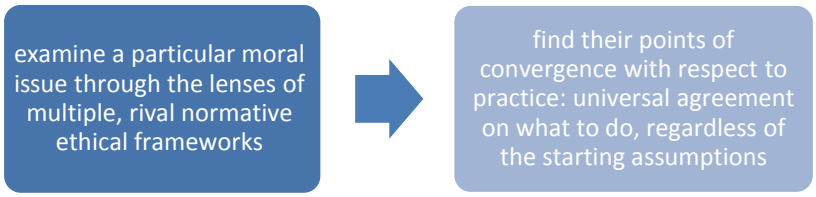
In general, if a moral theory conflicts with our intuitions, we have two choices: either we modify the theory in light of the "evidence" of our moral intuitions, or else we reject some of our initial intuitive moral beliefs, and hopefully use the theory to explain why these beliefs were mistaken. In the example we have used, the latter counterexample ought to move most people to modify the theory somehow. If we change the theory, the process then continues indefinitely – we have to further test the modified theory until we reach a point where it coheres with all or nearly all of our moral intuitions: one's concrete moral judgments and one's ethical theory are in harmony or have "reached equilibrium."



Method I: Reflective Equilibrium

In contrast, *overlapping consensus* is a "top down" approach, in the sense that we begin with general moral theories to derive conclusions about the rightness or wrongness of particular actions (e.g., whether lying to promote good causes might ever be morally acceptable). The method of overlapping consensus recommends that we examine a particular moral issue through the lenses of multiple, rival normative ethical frameworks, with fundamentally different assumptions, to find their *points of convergence* with respect to practice, regardless of whatever ultimate reasons we may give for their justification.

After all, as we will see, reasonable disagreement on ultimate questions of right and wrong is possible. Our analysis will therefore avoid favoring any one narrow moral perspective for this very reason – someone who does not accept our starting assumptions will not be persuaded of our conclusions before the inquiry even begins. *Overlapping consensus* is highly useful for finding universal or near-universal agreement on *what to do*, even if the moral justifications for the same conclusion diverge. We will draw on the three major traditions that dominate contemporary normative ethics. Our aim is to show from that no matter *which* of these perspectives one adopts, similar conclusions follow about the general conditions under which lying for moral ends is permissible.



Method II: Overlapping Consensus

III. The General Case: Lying for Good Moral Reasons

Reflective Equilibrium

If it is ever the case that there are some circumstances wherein lying for some "greater" good is morally justifiable or even morally required, then such a principle will likely lie between two diametrically opposed injunctions:

Extreme Restrictive Principle (ERP): It is never morally permissible to lie in order to achieve some better outcome.

Extreme Liberal Principle (ELP): It is always morally permissible to lie, whenever it makes things better, i.e., whenever the sum total of benefits or moral goods that result from the lie outweigh the sum total of harms the lie causes.

Problems with ERP

Upon considering the practical recommendations of each principle, it is evident that no reasonable and moral person could desire to put either principle into practice. Against the **ERP**, we can clearly conceive of circumstances wherein our moral common sense tells us that lying is permissible, or even obligatory. Consider the following examples:

The Battered Woman Scenario:

A battered woman has fled from her abusive husband, fearing for her life. She's hiding in your basement. Her husband shows up at your door unannounced, brandishing a gun in a drunken stupor, demanding to know where she is. He threatens to also kill you if you don't tell him where she is.

The Undercover Detectives Scenario:

A feud between two organized crime rings is responsible for an epidemic of crime in your city; the murder rate is up more than 300% from last year. Every week innocent people are injured or killed in the crossfire. It would be very difficult and much slower for the police to build a case without the use of undercover detective work. The very nature of undercover work would however require that the agents systematically deceive the gang members as to their identity and motives in order to be effective.

Regarding both scenarios, it is safe to say that virtually no reasonable and moral person would think that one ought to tell the truth. In fact, in the *Battered Woman* scenario, most of us would hold a truth-teller partially responsible for any harm that befalls the woman – morally speaking, if not legally. It seems there is a *moral obligation,* in this case, to lie.

Problems with ELP

Nor can the **ELP** pass muster. Consider another scenario:

The Con Artist Scenario:

A crafty con artist who is moderately wealthy defrauds a rich elderly widow, who is terminally ill, of her personal fortune. The widow originally intended to leave her estate to her well-off daughter. Since the daughter is richer than the con artist, the benefit he derives from inheriting the estate outweighs the loss for her.

Under the **ELP**, the con artist acts morally when he cheats the widow out of her money. After all, the lie results in an "improved" state of affairs where one person's significant gain in happiness outweighs a lesser loss for another person. However, we will venture to say that most reasonable and morally competent people will recognize that it is very *immoral* to con rich widows on their deathbeds out of their money, even if it is true that objectively speaking, the sum total of human happiness is increased. Just like the "murderer at the door" scenario above, this seems to be a fundamental intuition that falsifies any moral theory that would contradict it. So already, the **ELP** is refuted.

Perhaps the problem with this example lies in the fact that the con artist alone *personally benefits* from his own actions, and furthermore, he does not *need* the money. We might be

more sympathetic if, say, the con artist were to donate all (or even half) the money to hunger relief efforts, or perhaps if he would use the money to afford expensive medical treatments he needs to survive a chronic, life-threatening medical condition. Perhaps then we could tweak the **ELP** by limiting its application to *altruistic* lies, lies that aren't self-serving, unless (maybe) the self-serving nature of the lie is such that it thwarts some significant harm that the liar would otherwise befall. However, we believe even this restriction does not render the **ELP** very plausible. Consider the two following cases:

The Fixing the Test Scores Scenario:

The students of a high school in a middle-class neighborhood are performing above average on standardized tests. However, since there are monetary incentives for improved test performance, the school stands to gain a modest sum of money that could be used to build a new computer lab, or upgrade athletic equipment, or increase funding for more extracurricular activities, etc. – all of which would marginally improve outcomes for the students. Since the students are already performing well, test scores would have to dramatically improve to get the monetary reward. The principal comes up with a foolproof method to massively alter student's answers on tests, thus bringing up apparent performance to the requisite level without a significant risk of getting caught. He knows his teachers would be willing to falsify the answers and keep quiet. No one but the students would benefit from the improved funding, and the principal and cohort teachers will feel no guilt about fixing the test scores. Furthermore, the total pool of funds under the new policy is not capped at a fixed level such that the reward is "zero-sum" – other schools would not lose the opportunity to earn the same incentives (the school would not be "stealing" money from schools that honestly earn the incentives, especially underperforming, low-income ones).[2]

The Charity Scenario:

A woman hatches a plan to increase her charitable giving: every weekend she will panhandle on a street corner, wearing ragged clothes and holding a sign that says "homeless and hungry – please help." Every cent she collects will go to an organization working for hunger relief in the developing world.

[2] Currently in the US, there is a wave of scandals involving just this sort of behavior (www.nytimes.com).

The **ELP** tells us that the principal either *ought to* cheat, or at least that cheating would not be morally wrong. Weighing the consequences, the benefits of the deception are very minor gains in the education students receive. *No* harm results, since by stipulation no one will ever find out, and the funds are not taken away from any other school that could use the money. And yet, we intuitively feel that the action is wrong somehow owing to *the deception itself*, that the principal and teachers are engaged in a massive fraud, even *if* it is done with noble intent. However, the **ELP** has us only weigh benefits and harms, and it is hard to precisely say what this "harm" consists in, in terms of human happiness or welfare. What seems to drive this judgment is a sense that the *relative triviality* of the benefit does not justify lying, even in the absence of harms. Perhaps this is because the act of lying intrinsically has negative moral weight regardless of consequences, or perhaps this is because lying *always* has negative consequences that the **ELP** does not take into account (We will examine this question later). For now, we observe that even a very liberal account of when lies are morally justified must impose some minimum "threshold" on the net moral value that results from a lie, below which lies are morally forbidden. Lies that improve the world in relatively inconsequential ways seem to still be morally wrong.

Likewise, the **ELP** will condone panhandling to collect money for charity. If no one learns of what the woman is doing, those deceived are not harmed – perhaps they even feel good about being charitable, and so long as they would not otherwise give the money to an even greater cause or use it to meet their own basic needs, hungry people benefit. But this sort of behavior *still* seems morally wrong. Again, most people will feel that there is something about the deception itself that corrupts or undermines all the good that the woman achieves by contributing the money. Unlike the cheating principal, this woman might actually bring about a fairly significant amount of good through her deception if she is able to collect enough money. What seems perverse about *this* scenario is that the lie is totally *unnecessary*. That is, there are (roughly) equally effective non-deceptive means to achieve the same moral end of feeding the hungry. Instead of lying to collect money, the woman could collect the money honestly, say, by spending her weekend volunteering for charity, soliciting donations from individuals *and* local businesses and philanthropic organizations. The **ELP** commits us to saying that *any* lie that brings about a moral good is morally permissible. But when there are ways to achieve the same noble end *without* lying, the lie is gratuitous and therefore again seems wrong if avoidable. Even if no one is harmed by the well-intentioned lie (unlike the "widow" case above), there seems to be

something about lying *itself* such that we ought to *at least* use it as a tactic of last resort when there are other options available.

In light of these counterexamples, both principles need to be revised: we need to find a tenable middle ground. The **ERP** is too demanding because *sometimes* lies are necessary to protect life and limb and to rectify great injustices. The **ERP** is too permissive because not all better results are worth lying for, nor is lying an equally acceptable means of bettering the world when there are other ways to do so. We can take the insights from each counterexample and use them to weaken each principle so that they are no longer so vulnerable. Here is a first pass:

Moderate Restrictive Principle (MRP)*: Lie only when there is no alternative action you could possibly take to prevent a great harm from befalling others or to rectify some gross miscarriage of justice.*[3]

Moderate Liberal Principle (MLP)*: Lie only when doing so is the most effective means (by a significant margin, relative to non-deceptive means) to bring about a substantially better outcome.*[4]

These principles are *much more* reasonable principles that *approach* a middle ground from opposite sides of the spectrum. Of these two principles, which should we opt for? Using the method of reflective equilibrium, we can test each principle against "morally grey" ethical dilemmas where the right answers are not as clear (in contrast to the above counterexamples). These are situations where our intuitions pull us in opposite directions – we recognize some morally compelling reasons to lie and other reasons to tell the truth. Here are three such cases:

---

[3] Some examples of "great harms": losses of life, debilitating injury or illness, and deprivation of basic needs. Some examples of "gross miscarriage of justice": deprivation of fundamental human rights, uncompensated restitution for past mistreatment, and systemic inequality of treatment under the law.

[4] An example of a "significant" margin of effectiveness over other alternatives: it is not morally permissible to deceive someone to secure food for 10,001 needy families when one could secure food for 10,000 families *without* lying. Examples of "substantially" better outcomes are results like permanent increases to the general welfare of the people *most* in need of help. Providing an ice-cream cone to every child in your town is *not* a goal worth using deceptive practices according to the Moderate Liberal Principle, but securing adequate housing for ten homeless people might be worth the lie.

The Adulterous Partner Scenario*:*

A close friend confides in you that he cheated on his wife on a single occasion. He is not having an affair, and he sincerely regrets the lapse of judgment – he loves his wife. However, he fears that if he tells her, it will ruin the marriage and break up his family. His wife is suspicious and asks you if he has ever cheated. Since she trusts you, lying would probably settle the matter permanently, sparing her from heartbreak and your friend from losing his family.

The Witness for the Prosecution Scenario:

You are the star witness for the prosecution in the trial of a man charged with murdering his ex-wife. Near the approximate time of the murder, you saw the man wearing blood-soaked clothes, fleeing from the crime scene, in a car registered in his name. After agreeing to testify, you became privy to a battery of forensic evidence overwhelmingly indicating that the man was set-up by a third party. However, the unscrupulous District Attorney successfully lobbied the judge to have this evidence sealed in a corrupt back-room deal. The evidence is such that you are now convinced that the accused is innocent, but you are also convinced he was in fact at the scene of the crime (as part of the set-up). If you take the stand and answer the prosecution's questions truthfully, the man will likely go to jail. However, if you instead lie and testify that the man wasn't present (you could say, under oath, that you misidentified him in your statements to the police), the prosecution's case has a good chance of falling apart, setting the man free.

The Supercomputer Scenario:

One hundred years from now, the military of a large, industrially advanced civilization has designed a top-secret supercomputer that can calculate the outcomes of policy decisions with very high precision. The nation is in the midst of a severe economic crisis; inflation and unemployment are rampant. The democratically elected president uses the supercomputer to calculate with 99% certainty that mobilizing the economy for a deep space exploration project would lead to a quick and long-term economic recovery. However, there is a significant problem: the supercomputer calculates that the space exploration project is so technologically infeasible that actually conducting it will be an absolute disaster for the economy and lead to a loss of a great many lives; in other words, it is the mobilization, the massive preparation effort, that will lead to the sustained recovery and not the project itself. The president runs hundreds of other

possible proposals for economic recovery through the supercomputer, but each plan would either lead to weaker recoveries, take a much longer time, or could not win the democratic support of the public and elected leaders. The president (with a small core of deeply trusted advisors) tasks the supercomputer with devising an elaborate plan to deceive the populace and elected representatives into thinking such a project will be successful, to fabricate evidence to make it seem as though an active space exploration project is ongoing once the mobilization begins, and to tie up all "loose ends" such that the plot is never discovered. The computer calculates (again with 99% certainty) that the plot will not be exposed within the next 50 years.

If we follow the recommendations of the **MLP**, then we ought to lie in all three cases, for lying is the most promising and immediate way to promote a *significant* moral good: a marriage and family ties are saved; an innocent person narrowly avoids wrongful incarceration; a nation is pulled out of a deep depression. Every lie helps people without causing *harm* to anyone. To people who are resiliently committed to the proposition that the only thing that *matters*, morally speaking, are relative states of actual human wellbeing – this will make perfect sense.

Independently of this observation, in the third example it is less clear that it is just *obviously correct* for a president to lie to millions of people even if the deception is for their benefit (a caveat: although this is intended as intuitive "evidence" against the **MLP**, we fully recognize that people will feel the force of this moral intuition to varying degrees, and people who think it *is* obviously correct will of course be those who are *most* convinced of the **MLP**). The lie causes no obvious "harm," defined in the sense of relative states of human wellbeing. However, we submit that many people who would ultimately approve of the president's deception can at least *feel* the force of a countervailing intuitive moral reason against the president's lie: that it is wrong to massively manipulate public opinion, especially when one is a trusted leader, vested with political power deriving from public consensus.[5] For many, these intuitions won't *outweigh* the moral reasons to lie, but they will feel some ambivalence. We will not speculate on whether most or even a significant majority of moral agents would approve or disapprove of the president's deception. What we *do* think is that very many people will intuitively feel as though *it would be morally optimal* if the president did *not* have to lie – that the lying is a "necessary

[5] Note that this is another example of the intuition that something is wrong about the act of lying itself, either through some indirect harm or some moral reason that doesn't reduce to changes in wellbeing.

evil," to use common parlance, even if the benefits are *unquestionably* worth the deception. This fact is hard to explain if the lies cause no apparent actual harm apart from the outcomes they lead to.

This example is deliberately intended to prime intuitions *against* lying, and towards the reasonableness of the **MRP**, which may seem too rigid in the first two examples. The **MRP** identifies what is wrong about the president's lie, in that there are other available means to affect an economic recovery, though not the *optimal* one by a significant margin, thus making the lie run afoul of the "necessary means" restriction. The president ought to choose the best non-deceptive option, even though this helps his citizens less than is actually within his power to help. We think it is definitely possible that many people would approve of lying in the first two cases but *not* in the third, siding with the practical recommendations of the **MLP** *and* the **MRP** depending on the particular case.

Turing quickly to the **MRP**, the principle forbids lying in the 'Adulterous Partner' scenario, since although divorce and family breakup is very *painful*, it is not a "grave" harm or "gross" miscarriage of justice (e.g., on par with starvation or human rights violations), and lying is not your only choice – you could try to change the subject, or refuse to answer the question, which would certainly raise suspicions without necessarily *confirming* them. In the "wrongly accused" scenario, as long as it is the case that one could employ non-deceptive but less effective means to try to rectify the injustice, one should not lie – which is not to say one should tell the truth. One could make a personal sacrifice and refuse to take the stand, risking punitive retaliation for contempt of court from a corrupt legal system. Or, one could uphold her oath, but publically aver the man's innocence outside of court and work towards his freedom post-conviction by drawing attention to the unknown evidence of his innocence.

We will also point out that like the third example it is not at all *obvious* to everyone that it is moral to lie in these first two cases as well. The **MRP**'s recommendations may seem overly stringent and demanding, but not *wrong*. We can prime these intuitions by very slightly altering the first two examples or by adding additional information. In the first example, suppose instead that you are the *woman's* most trusted friend, and the partner who cheated is a mere acquaintance. Even if those who would ultimately lie to the woman to spare her emotional harm, as with the third example, for many the lie will feel like a "necessary evil," even if it is better for her not to know. Many people will be inclined to say that on some level, the lie is a

deep *betrayal*, given your deep friendship – although again this goes *beyond* considerations of relative wellbeing if the lie is never found out. In the second example, suppose that the man is homeless and suffers from serious health problems. Possibly, going to prison might *improve* his wellbeing in terms of providing better shelter and medical care than charities or social welfare programs could offer. Perhaps he would ultimately be happier in the long-term if he went to prison (even if he has a *present* desire not to go to prison). Weighing the consequences, according to the **MLP** you ought to *refrain* from lying, but for what can seem to be perverse reasons. Although going to prison is better for him, many will feel as though this analysis leaves out some important moral feature, namely, the value of the man's *liberty* above and beyond one's own wellbeing. For those with a strong intuition that freedom is worth having even if it leaves one worse off, the **MLP** gets the moral explanation wrong. The **MRP** restrictions are designed in part so that it can accommodate these "additional" putative moral reasons.

As we have stressed, neither of these considerations are ultimately decisive in favor of either principle. We expect fully reasonable and morally competent people will disagree over "the" right course of action, and accordingly, which principle to follow in a given case. Crucially, it is entirely conceivable that someone could accept the recommendations of the **MLP** in, say, the first two examples, but *not* for the third example. These are indications that the optimal, *more nuanced* maxim regulating when lying is morally permissible is probably less restrictive than the **MRP**, but more stringent than the **MLP**. The right "answer" in any given situation will likely lie somewhere in between the practical recommendations of either of these two principles. We will leave open the question of whether there *is* some "perfect" principle that could tell us, for any given case, if a lie is morally permitted or required in the pursuit of some greater moral good. Any such principle *we* could offer would in any case be highly contentious and probably incredibly complex.

In conclusion, reflective equilibrium can only get us so far. Although we have found two reasonable and plausible principles, we have not yet identified any reasons to prefer one to the other, at least most of the time, nor have we found a way to weaken them further towards a point of near-convergence. Matters of great practical significance are at stake: if we *always* follow the recommendations of the **MRP**, we will pass up opportunities to foster significant amounts of wellbeing and alleviate or eliminate deep suffering. If we *always* follow the **MLP**, then we ought to lie in *every* such opportunity, which means that lies will occur very often. A

hardline proponent of the **MLP** could quite understandably complain, "if the lie harms no one, why does it matter if lying will occur very frequently?" We have located some *intuitive* reasons to think there's something wrong about the very act of lying, even when the lie goes unexposed. But we have no "deeper" theoretical understanding of the underlying reasons *why* we have this intuition. As we will contend, this intuition is well grounded, and it indicates that we ought to *incline* towards the **MRP**, that the best moral principle probably lies closer towards *not* lying as a means to realize a better state of affairs, in most cases. The method of overlapping consensus will help us to actually explain why we ought to ascribe so high a moral premium on honesty, such that we ought to be very cautious about using deception as a means to a moral end.

Towards Overlapping Consensus

Using the method of overlapping consensus, we can derive principles that specify the conditions under which it is acceptable to lie for moral causes using several different kinds of moral justification. To the extent that the principles thus derived agree on these conditions (in the ideal case, if they are near-identical), then we can rest secure in the knowledge that there are many different justifications that we could give to justify the same ethical maxim – no matter what your theory is, you ought to accept such a principle.

Most ethical theories fall into one of three major camps: they either take a *consequentialist*, a *deontological*, or a *virtue-based* approach in determining the nature of right and wrong, and its consequences for how we ought to live. We cannot possibly survey all such particular versions of these ethical theories, but we can group them according to this classification scheme, which captures in a very general sense the way in which a theory analyzes the nature of moral goodness and right action.

Deontological Theories

Deontological approaches to ethics share in common the view that actions are *intrinsically* obligatory, permissible, or forbidden. This means that the action "by itself" has a positive, negative, or neutral moral status, *unconditionally*, or without the need to consider any fact beyond the nature of the action alone, crucially *leaving out the consequences of that action*. Deontological theories are therefore often described as *duty-based* ethical frameworks: to live an ethical life is to follow moral precepts such as "honor your promises," "do not murder," and so forth – regardless of particular circumstances that might tempt us to violate the rule, most

notably, when following the rule would lead to undesirable outcomes, or when violating the rule would lead to desirable outcomes.

There are two broad categories of secular[6] deontological ethics: *Kantianism* and *contractualism*. They agree that moral conduct is a matter of following universal and *impartial* rules of conduct that are discoverable by reason, however, they diverge over the source from which these duties derive, and *how* we discover them. Kantian ethicists follow Immanuel Kant's understanding of ethics as a set of principles that we commit to obey, while expecting the same commitment from others, forming a kind of universal code of moral conduct. The Kantian formula for discovering impartial moral imperatives requires that we only act in ways that we could *universally will* that everyone follows. A similar idea should be familiar to many readers in the form of the "Golden Rule": "treat others as you would want to be treated." These are not the same notions, but the root idea is that behaving morally at bottom requires treating people *fairly*, with equal and impartial consideration of their interests.[7] From this it follows that we owe other people a certain degree of dignity and respect. As Kant would say, we cannot treat other people *merely* as a means to getting what we want, but as "*ends in themselves*" – that is, we must respect other persons' intrinsic value by not infringing on their autonomy, just as we would desire to be treated ourselves.[8]

Contractualist deontological theorists shift emphasis from the reason of the individual to a *social* justification of moral rules. Very briefly, according to contractualists, ethics is (or ought to be)

---

[6] Strictly speaking, deontological ethics refers to any system of thought that characterizes morality as a set of duties to engage in or refrain from particular actions simply because these actions are deemed obligatory, permissible, or forbidden, without appeal to anything other than the type of action that it is. Therefore, *religious* codes of conduct (e.g., the Ten Commandments, or the Five Pillars of Islam) derived from the teachings of sacred texts likewise fall under the umbrella of deontological systems insofar as these are unconditional imperatives.

[7] A related concept stemming from deontological thought that is likely familiar to the reader is that of a *human right,* which *cannot* be violated, unconditionally, even if for some manifestly greater good.

[8] This is not intended to imply that people have *unconditional* freedom to do whatever they want. In fact, it's quite the opposite: the duties that deontological ethics imposes upon us are typically a finite list of *restrictions* on how we may limit other's free choices ("*negative* duties") – but for example, if all reasonable people could agree to (say) taxation to provide for public goods, or compulsory vaccination to vouchsafe public health, these limitations on freedom would be morally justified.

concerned with how we treat other people and coordinate our collective behavior, hence ethical rules ought to be understood as a set of principles people would *consent to*, in an ideally rational and fair process of negotiating the principles we could all agree to obey. This is the fundamental idea underpinning the idea of a "social contract" – a fictitious but theoretically useful thought experiment, in which the laws of an ideal society are those to which *all* persons would consent to submit themselves, if negotiated in an ideally fair and rational process (see, for instance, Rawls' book "A Theory of Justice" for the most well-known contemporary articulation of such a theory). One can identify a common core here with Kantian approaches: moral behavior is about using a rational process to figure out the best system of rules that we ought to live by. Kantians see this process as something individuals can discover on their own by pure reason; contractualists emphasize the need for social negotiation. At bottom, morality is about impartiality, fairness, and respect for the autonomy and informed consent of all persons.

Consequentialist Theories

Consequentialist theories of ethics depart from the deontological perspectives precisely in denying that actions have an intrinsic value; whether an act is right or wrong depends on factors that are not inherent to the very nature of the act.[9] As the label suggests, consequentialist theories determine whether an act is right or wrong by looking to the consequences of the action. In extremely simple terms, an action is right if it produces more morally desirable outcomes than undesirable ones; conversely, an action is wrong if it produces more morally undesirable outcomes than desirable ones.[10] Specific consequentialist theories diverge over just

---

[9] To fill in the details on what it means for the "nature" of an action to be "intrinsically" wrong, consider a simple case like the act of stealing for personal gain. The deontological formula would have us reflect on whether a reasonable person could want to live in a world where people are perfectly free to take others' personal property by force for their own selfish benefit. Since (presumably) no one could desire such a world, stealing is wrong. The categorical wrongness of stealing follows just from *what it means to steal* – to take something from another person against her will.

[10] The astute reader might ask: desirable or undesirable, *according to whom*? That is, whose preferences do consequentialists take into account? I strongly desire chocolate peanut-butter ice cream and detest mint chocolate chip; you have precisely the opposite preferences. We have just enough money to buy only one gallon, so whose desires win out? Consequentialists agree with deontologists insofar as they understand morality to essentially involve *impartial consideration* of other people (e.g., I do not deserve special treatment over anyone else; no matter how much I may *enjoy* such a privilege, it would be *unfair*). The issues involved here over how to *weigh* and *balance* interests are very complex (for example, should we give a child's interest's just as much consideration as those of its parents? What about an animal's

which outcomes are desirable: common proposals include producing the greatest pleasure or happiness for the greatest number of people, or satisfying people's subjective preferences to the greatest extent, in sum total. Nevertheless, the common idea behind consequentialism is that actions only have a derivative moral significance in their effectiveness at bringing about morally desirable outcomes.

There is an important distinction between rule and act consequentialist theories. As the name suggests, rule consequentialists agree with deontologists that morality is a matter of following the proper set of rules. However, for a consequentialist a moral rule can only be justified if following that rule typically tends to bring about desirable consequences – not because the rule has any intrinsic binding force due to type of action that it is. An example would be "stealing is wrong, because a system of general respect for private property produces the most prosperity and happiness." By contrast, act consequentialism holds that individual actions alone can be moral or immoral depending on whether they produce the best consequences in a particular context. There are no general rules or principles; for instance, the act consequentialist will not always act on the common sense maxim "stealing is wrong." Theft can only be wrong in particular cases, where the theft will hurt more people than it helps. This is probably typically the case, so in most cases, the act consequentialist's behavior will look like the rule consequentialist's. But there will be events where the act consequentialist and rule consequentialist disagree over whether a particular act of theft is wrong, and these will occur exactly whenever the aggregated consequences of stealing are superior to those of not stealing. Suppose a rich person is about to spend $500 on a fancy dinner. According to a crude version of act consequentialism, one would be justified or perhaps even obligated to steal that money and buy food for a destitute family, whereas a rule consequentialist would respect the rule "stealing is (in non-exceptional cases) wrong," and this rule is founded on the recognition that a system of respect for private property is a very important good that is greatly harmed by each act of theft – so, in the final analysis, perhaps a more nuanced act consequentialism could incorporate this

---

preference not to be harmed or killed *versus* the sick person's interest to use the animal for medical research?). However, in this simple case, it seems clear that there's no morally relevant difference between us such that one of us should get their way. So, the only *fair* thing to do in this case is to buy another flavor we will both enjoy, even if no one gets their first pick. This accords with our common moral common sense.

harm into the final calculation of harms and goods, concluding that the theft on balance leaves the world worse off (this point will be important later when we discuss the case of lying for greater goods).

Virtue Theories

A final class of theories we will consider, virtue ethics, differs from deontology and consequentialism by eschewing actions, shifting perspective to the whole person and their stable character traits: the proper object of moral evaluation is the agent, not the individual act (i.e., people are more or less good, not any particular action they might perform).[11] What really matters, according to virtue ethics, is the character of a moral agent. Having a good character means having a disposition to perform actions that exemplify what Aristotle called 'the Virtues.' Giving an exhaustive list of virtuous character traits is no easy task (if it is even possible), but hopefully we can agree on some of the most important qualities of a good person: courage, self-determination, self-control, fairness, justice, benevolence, trustworthiness, kindness, and prudence are perhaps a good place to start. Possession of a particular virtue to a particular degree is demonstrated by the person's tendency to perform actions of a certain type. For example, a benevolent person will regularly donate to charities, help strangers, and forgive people without expectation of reward or even at great personal sacrifice. Although being a good person is a matter of having a kind of disposition, not every action will always manifest a virtue. A virtue theorist will say that we cannot judge a person's moral worth based on a single particular action; what matters is the pattern of behavior that exemplifies a good or bad character. Consequently, actions taken in isolation from a person's character are neither good nor bad, neither intrinsically (per deontology) nor derivatively (per consequentialism). Therefore, the categories of right action and wrong action are conceptually derivative: they reduce to what a person of good or bad character would do, given the totality of relevant information about the agent and the context in which she acts. It is a mistake to say that any

[11] This theory matches the way we sometimes think about morality, e.g., that there are people who are very morally "upright," even if they have an occasional moral failure. Similarly, there are people who in general lack a "moral compass," even if they sometimes do the right thing.

action is either right or wrong because it exemplifies or is contrary to the virtues.[12] This gets the explanation exactly backwards: a virtuous person (who is fair, benevolent, and trustworthy) is strongly predisposed not to act contrary to the virtues. E.g., stealing per se is not wrong; rather, good people typically do not steal, all things considered.

Applying the Overlapping Consensus Approach

Now let us apply these particular theories to the act of lying, and more narrowly, lying in the service of achieving good causes: what do the theories have to say about when it is ever permissible (if at all)? Do they agree? We will start with consequentialism, since it is likely to be the *most* permissible ethical framework along the spectrum of principles we considered above. If the only standard by which we judge its moral value is what the lie achieves in practice, then lying in the pursuit of objectively morally desirable ends is – at first glance, anyway – permissible so long as the lie does not have ultimately undesirable consequences.

Above, we distinguished between rule and act consequentialism, and we will start with the latter. An act consequentialist will predictably answer our question as follows: "lie, whenever doing so will increase the sum total of human wellbeing, and do not lie whenever doing so will decrease the sum total of human wellbeing." While this is (strictly speaking) a *principle*, it is completely insensitive to the nature of deception. It is just a derived instance of the only "rule" act consequentialists accept – always act so as to maximize wellbeing, and for that reason we cannot use it to say anything very informative. Fortunately, we submit that the act and rule consequentialist will *generally* agree on which lies are morally justified; though there will be rare exceptions, the *practical recommendations* of act consequentialism with respect to lying will converge with those of rule consequentialism. This owes to general facts about the nature of interpersonal communication, the only context in which *any* deception can take place. Once we analyze lying from the perspective of rule consequentialism, we will be better positioned to explain these reasons.

Since rule consequentialists must propose a more specific principle, they have to provide an analysis of when the specific act of *lying* is apt to cause more harm than help. Recall that

[12] For example, a virtue theorist could not say "to keep a promise is right, because it exemplifies trustworthiness, but breaking it is wrong because it runs contrary to trustworthiness." All we can say is that the trustworthy person reliably honors their promises.

throughout discussing the **ELP** we observed that we intuit – albeit with varying degrees of conviction – that acts of deception in some circumstances are "just wrong," in a way that is not *obviously* connected to relative states of aggregate human wellbeing. This required us to put some restrictions on the moral permissibility of lying in the **MLP**: the lie must achieve *significant* improvements in relative wellbeing, and should not be used if an equally effective *non-deceptive* means is available. Even the **MLP** ran into an apparent problem to the extent that it permits a president to systematically deceive an entire nation, so long as it leaves the populace better off than any other alternative action he could take.

Perhaps this is an indication that there are morally relevant factors we have failed to notice (as we alluded to above). In other words, we have *assumed* that the only negative consequences of a lie are the *actual* sanctions and loss of trust for getting caught, and potential future guilt. If one can get away with it with a clear conscience, well-intentioned lies are perfectly moral, even virtuous. The intuition that this leaves out the wrongness of *the lie itself* might point to there being *unavoidable* "hidden costs" to lying, costs in terms of human wellbeing, costs that we can "cash out" in consequentialist terms of harm to human welfare.

For a consequentialist, these costs must be "extrinsic" in the sense that lies are wrong because they have negative consequences, not because lying is wrong in some fundamental sense. However, as we have suggested, these harms may be so tightly associated with lying that no matter what one does, they are present even when a lie goes undetected and is perpetrated without compunction. We believe that there are *at least* two such hidden costs[13]: that (**a**) lies have cascading or "snowball" effects which tend to compound the initial lie and lead to further ones, and perhaps more significantly, (**b**) that lies, if commonly practiced, *always* undermine the fabric of social trust that makes civil society possible.

The cascading effect of lies is likely well known to anyone who has ever significantly distorted the truth. This has both an outward and an inward dimension. The outward dimension is that lies can tend to take on a "life of their own" beyond the initial mistruth. All too often a person may have to cover for the initial lie with a second one, which can lead to ever greater and

[13] There may be more hidden costs, perhaps many – we focus on these two because they are especially prevalent and harmful.

complex deceptions. (As the fairy in the movie *Pinocchio* observes, "a lie keeps growing and growing until it's as plain as the nose on your face"). As more lies pile up and one's version of events departs ever further from reality in complexity and misinformation, it becomes ever more likely that the truth will eventually come out, potentially causing much worse harm than if one had just confessed to the first lie.

The inward dimension of the snowball effect has to do with the liar's own respect for the truth. Lying can become habitual; with each "successful" (undetected) lie, one learns that lying can be an effective means to achieving one's goals – especially if one is good at not getting caught. This of course makes a person ever more prone to deceive. A lie that starts even with the best of intentions can normalize deception. Because there is nothing about well-intentioned lies that makes them easier to get away with, the liar begins to believe that the truth is a tool that can be manipulated for *whatever* means, altruistic or not. Consequently, if altruistic lying becomes habitual, this can easily turn into lying for selfish motives, which certainly can harm others.

Finally, lying produces false beliefs in other people and creates misinformation (when successful), and this is something that all rational people ought to find undesirable, or even actively oppose. For we have no choice but to rely on other people to acquire information that is instrumental in helping us satisfy our own desires: if we could not trust medical experts (for instance), we would have no reason to believe that we ought to wash our hands before handling food. Most truths can in some way help us get what we want, but knowledge is diffused among many people[14] – therefore, we are able to lead much happier, fulfilling, and safer lives, but *only because* we presuppose that others will not intentionally mislead us. Even the most rudimentary forms of social organization would not be possible without this background norm of *reciprocal honesty* – in any act of communication, there is an unstated shared understanding that I can trust that you are being sincere (that you believe what you say), and likewise, you can rely on me to be sincere with you. Language itself requires a shared, public understanding of what words mean and the proper contexts for their application in order for an utterance to convey a "default" message. If we did not presuppose that at least *most* of the time, *most* of the people

[14] When it comes to the sum total of human knowledge, there is a *division of labor*: no one person could possibly know for themselves all of the facts they will ever need to know in order to lead a meaningful life.

with whom we speak believed what they said, it would be impossible to distinguish a sentence's meaning from its very denial.

This fabric of social trust is supremely valuable, but it is damaged each time a lie is perpetrated. Even when a successful lie goes undetected, someone *still* ends up with a false belief that they may act on at some future point, potentially to their detriment. Furthermore, since each new lie erodes the *liar's* respect for the value of truth, the aforementioned cascading effect gradually undermines the liar's commitment to and dependence upon the societal bonds forged by trust, again, *even if the lie is never discovered*. On the other hand, liars *are* caught a fair amount of the time – and when this happens, especially in the case of significant lies that impact many people, this directly damages both the liar's future credibility, and the credulity of those who are deceived, further undermining people's natural and enculturated propensity to trust each other.

From a rule consequentialist perspective, lying *inevitably* damages a great public good: our ability to easily trust one another. Imagine if most people did accept something like the **ELP**: lying would become so rampant that it would be difficult to trust what anyone says. This would not only cause societal breakdown, it would undermine the very possibility of meaningful communication.

What rule consequentialist principle ought we to accept, then? The **MLP**, recall, placed restrictions on when the benefits of lying can outweigh the costs. The lie must produce significantly better results than all alternative actions (don't lie when you can effect roughly the same amount of good through non-deceptive means), and the lie must achieve something of "sufficient" moral worth (only marginal benefits are probably not worth the lie). If we add the significant negative consequences of deception into the "calculus" of harms and goods, lies then ought to be used even less frequently in the pursuit of moral good, only when the stakes are appreciably high. With the benefit of a theoretical underpinning, we submit that if we *fully appreciate* the actual "hidden" costs of lying, the **MLP**'s recommendations end up being pushed a lot closer in the direction of those of the **MRP** – perhaps not towards full convergence, but appreciably more in the direction of considerable restraint. We cannot naively assume that undetected lies have no bad actual consequences; harming the fabric of social trust should not be done lightly. To return finally to why the recommendations of act consequentialism and rule consequentialism converge in the special case of lying, given that communication is a system that critically depends upon default "rules" of truth telling, one ought to observe these rules

except of course for when the good consequences outweigh the bad. Of course, since damaging the public good of social trust has a high negative moral value, lying ought to be quite rare as well on act consequentialism. Therefore, the recommendations of act consequentialism are, for all practical purposes, subsumed under those of rule consequentialism, which inclines towards a restrictive view.

Deontological prohibitions on well-intentioned lying actually converge to the **MRP** as well, and surprisingly, their justification strikingly resembles that of the rule consequentialists. If rule consequentialism initially *seemed* too permissive when it comes to lying for good causes (although this was in fact an illusion), deontological theories would at first glance have the opposite problem of appearing too strict. At first glance, a deontologist might say that we have a *duty to always be truthful*, because no reasonable person could want to live in a world where they are lied to.[15, 16] Lying in the interest of some greater good could never, ever be morally justified if our duties really are that strict. Doing so would appear to be tantamount to shrugging off our moral duties just to achieve a better outcome, but outcomes should not enter into our moral decision-making on a deontological view. This would seem to commit a deontologist to the **ERP** – never lie, under any circumstances. Kant ([1797] 1996) infamously bit this bullet, defending the view that it is wrong to lie to the murderer at the door who comes looking for his victim.

However, as with consequentialism, there is a missing factor we have not yet accounted for that explains this radically unintuitive result. The very nature of moral dilemmas (not just those that

[15] As Kant argues, it is not so clear that we can even coherently *conceive* of such a scenario, for reasons we have already discussed: how could *language* itself (or in fact any system of communication) function, if listeners have no more reason to believe that what a speaker is saying is true than that it is false? Deception, therefore, is only possible where there is already a background norm in place for speakers to tell the truth – as a *practical* norm of communication and coordinating our activities, not *merely* as a moral imperative.

[16] Contractarian versions of deontology also converge on the importance of a reciprocal norm of honesty, from a slightly different angle. If we are drafting a social contract, participants have a strong reason to encourage honesty and sanction lying, because deception undermines the basic trust that makes even the most rudimentary social order and coordination possible. Put simply, we need to be able to trust other people in order to live in a social world at all. No sane person could *really want* to live in a world where no one could be trusted. We have a reasonable expectation that others be truthful with us, therefore, consistency in our own moral behavior demands that we should be honest with others as well.

involve lying) involves different moral reasons for taking opposite courses of action. A consequentialist will view a moral dilemma as a high-stakes situation where each course of action might have strong payoffs as well as negative consequences, such that it is quite difficult to balance the two. A deontologist might view the *very same* dilemma as a situation where their duties *come into conflict* with one another. Yes, they admit, we have a duty not to lie, but we *also* have duties to help others and to rectify wrongdoing, and the moral dilemmas above pitted the duty to tell the truth against these duties: e.g., the eyewitness has a duty to be truthful (*especially* under sworn oath), but she also has a duty to protect the wrongly accused.[17]

One has to suitably modify a duty-based theory of ethics to tell us which rules to follow in situations where one cannot fulfill one moral duty without violating another. Necessarily, some duties must "trump" other ones. The deontologist can accept the **MRP**, because the *duty to help others* is not *so* strong as to always (or usually) trump other, more fundamental moral duties, including our duty to tell the truth. This is how lying even to help someone could still generally be wrong, depending on what's at stake; in lying, one violates a universal, reciprocal norm of truth telling which we all depend upon. In deontological language, we could *not* rationally will that people universally act on the **MLP**, even if it can greatly benefit others. However, protecting someone from great harm or injustice is a special and extreme case of helping others, and this duty is more demanding – such as to override the prohibition on lying where it is necessary to fulfill the duty. Arguably, under the conditions of the **MRP –** where one's life, fundamental needs, or rights are imperiled, and you are forced into a situation where lying is the *only* means to protect the person – we can see how a duty to help could override the duty to tell the truth. As much as we value the ability to trust others, we obviously value our own lives and basic wellbeing more (for what good is a norm that *sustains* human flourishing, if one cannot lead a life worth living in the first place?). Therefore there is a strong case to be made that we could will that people universally act on the **MRP**, or a principle closer to it than the **MLP**, sacrificing the truth only when there is no better way to protect human life, freedom, and dignity.

---

[17] Kant infamously concluded from his moral system that we have a "perfect duty" to be truthful – meaning a duty that is strictly inviolable, whereas we have an "imperfect duty" to help others – meaning we ought to do so *as much as practically possible*, without violating our perfect duties. So while Kant would disagree with this analysis, certainly *we* must depart from his analysis of the situation.

At this stage, we should pause to observe that the deontologist's *intrinsic duty* to tell the truth in non-exceptional scenarios, based on a universal norm of reciprocal honesty, seems virtually indistinguishable at least in its practical recommendations from the rule-consequentialist's *derivative good* of maintaining broad societal trust. Rule consequentialism and a "flexible" deontology able to handle conflicting duties look *extremely* similar in practice, at least in the case of well-intentioned lying. Not only do the two theories converge on the same principle, but they also tend to do so for the *same reason* (at a level of sufficient generality): the immense value of maintaining a norm of reciprocal honesty. Whether one views this norm as intrinsically valuable (because it is universally desirable) or only derivatively valuable (because of truth's instrumental value for getting what we want), seems to be a distinction without much of a practical difference.[18]

Although speculative, this may be an indication of a deeper convergence; contemporary philosopher Derek Parfit (2011) has famously argued that the "best versions" of both deontological theories – in Kantian and contractarian forms – and rule consequentialism tend to approximate the very same *fundamental and uniquely correct* moral theory. These theories "climb the same mountain" from different starting points (e.g., different conceptions of the good), but they wind up at the same "peak." Parfit's full argument for his conclusion is quite sophisticated and involves matters far beyond the scope of this paper. But the skeleton of his argument is as follows: according to the most plausible forms of rule-consequentialism, an act is morally permissible just in case it is "optimific" – it tends to "make things go best" from an impartial point of view. Deontologists determine moral rules by finding a set of principles that every ideally rational being could will everyone to obey (Kantianism), or rules that could command universal consensus from ideally rational negotiators (contractarianism). But since optimific principles are *just* those moral rules that would make things go best from an impartial point of view, ideally rational beings could only universalize/negotiate *exactly* the same

[18] In fact, there is no obvious contradiction or tension in saying that the norm is *both* intrinsically *and* also instrumentally valuable. Honesty could be a good thing *in itself* because of the universal desire that others be honest with us (even where the truth would make us less happy), *and also* because we have a strong *self-interest* in being able to trust others.

optimific principles that rule consequentialists promote, and *only* those principles.[19] So, the best forms of deontology end up endorsing the same practical recommendations of rule-consequentialism! Parfit calls this his "Triple Theory" – moral behavior requires that we conform to rules that are "optimific, uniquely universally willable, and not reasonably rejectable." We have seen that this certainly seems true *at least* in the case of lying: as we have argued, "things go best" in a world where all people are committed to a norm of reciprocal honesty, except in circumstances where deceit is *the only possible means* to avert a *great harm* or *injustice* (that is, the **MRP**).

Our final matter is that of virtue theory: do its practical recommendations also converge to the "deontological / rule-consequentialist" **MRP**? If so, we will have a very strong overlapping consensus indeed. However, one seductive pitfall we must be careful to avoid is that virtue theory cannot recommend any moral *principles* by its very nature. Because it is focused on *character* and not *actions*, there is no way to evaluate an action in isolation from the whole person's pattern of behavior. Of course, moral rules have to be universal and impartial, which further means that they cannot take into account the particular idiosyncrasies of any single individual. A person-centric ethic will not be able to give us any general guidelines as to when lying for good causes could possibly be moral.

However, there is one very interesting point of contact between virtue theory and the convergent justification for the **MRP**. Virtue theory gives us a new perspective on the aforementioned phenomenon that dishonesty slowly corrupts one's respect for the truth, just as any other vice erodes one's inner moral character. The unique insight of the virtue theorist is that the virtues – honesty included – must be routinely practiced and *cultivated*, so that they become habitual and ingrained into the person's core identity.[20] The opposite is also true: if we routinely spurn virtue, we will end up cultivating bad moral habits. So virtue theory can at least

---

[19] This is one of the most contentious parts of Parfit's arguments, as it relies on substantive and controversial claims about what people have reasons to do (reasons which go beyond what people in fact actually subjectively desire), an argument that (as mentioned) reaches far beyond the scope of this paper.

[20] The virtue theorists remind us that it is often not *easy* to be a good person – it takes effort. Certainly this is true of honesty, as anyone who has ever been tempted to tell a convenient lie must know. And it also converges with our suggestion above that the **MRP** is in fact the right moral guideline, *even if* it is difficult to follow in practice because it is practically impossible to *always* do the right thing.

help us flesh out the earlier argument that dishonesty erodes the public's norm of reciprocal honesty, because deceitful behavior becomes a learned habit. From an impartial point of view, this is very undesirable – if people have few moral qualms about telling the *occasional* lie, lying will tend to become normalized. Conversely, because honesty is also a learned habit and because truth telling is sometimes more difficult and inconvenient than the occasional lie, continually practicing honest behavior is necessary to maintain the fabric of social trust.

Although virtue theory cannot yield any absolute moral principles, it does mesh well with our previous arguments and helps us understand even more why honesty is worth safeguarding. It, too, converges with the justification for not lying except in the extreme cases even if we cannot formulate a *principle*. Practically, we predict that someone who follows virtue theory's ideal recommendations will exemplify behavior that from an outside perspective looks as though the person is following the **MRP**.

If our line of reasoning is correct, we have seen that there is a *robust* overlapping consensus recommending a moral prohibition on lying in non-exceptional scenarios (defined by the practical recommendations of the **MRP**). This is a fortunate and significant result. In general, the only good causes that license deception are safeguarding others' or one's own life, liberty, or fundamental needs, and *only when* deception is the only way to do so. This is because following a norm of *reciprocal honesty* is so important to a well-lived life, and because there are good reasons to consider this norm both intrinsically and derivatively valuable – in fact, if Parfit is right that there is convergence at the level of *theory* as well as practice, the distinction between intrinsic and derivative value may not even matter. Lying both outwardly and inwardly corrupts this norm due to its cascading effects and through the spread of false information. And although virtue theory cannot *recommend* any principle such as the **MRP**, a focus on the *person* instead of *acts* still leads us to expect that leading a virtuous life will be virtually indistinguishable from following the **MRP**.

MRP

- Rule / Act Consequentialists
- Deontologists
- Virtue Theorists

As our conclusions are universal, they extend also to NGO executives simply in virtue of their being mere rational agents. As we will argue, however, there are special facts about the nature of the executives' professional role that place *additional* constraints on the **MRP**, further limiting its applicability while the executive is acting in a professional capacity. We turn now to our special question.

IV. The Special Case: Well-Intentioned Lying in the NGO Sector

Having found some stringent yet plausible criteria to guide us in determining when it is justifiable to lie in the pursuit of moral end, we turn to the specific matter of lying for good causes in the NGO sector. We will show that people in leadership roles at NGOs have to contend with uniquely challenging demands, which further *restrict* the conditions under which lying for good causes is morally permissible. Consequently, we will need to amend the application of the **MRP** within the context NGO sphere by adding new constraints.

The additional burdens placed upon NGO executives derive from the *interplay* of two salient features of NGOs: (**a**) their unique *mission* to promote good causes, and (**b**) constraints on how they may legitimately influence other people in support of their cause. For NGOs promoting humanitarian and charitable causes, well-intentioned lying is actually a matter of great *practical* significance, one that executives may have to frequently confront. There are good reasons to think that lying to promote good causes is much more often at least a tempting strategy for executives of such organizations, relative to their counterparts in the private sector or other non-charitable enterprises. There is an obvious reason for this: it is the *job* of these executives to promote good causes. Their daily work involves activities such as coordinating food relief, or promoting awareness of human rights abuses – work that often saves lives, helps people meet their basic needs, resist oppression, and so on: exactly the sort of *overriding considerations* which might permit deception. Like us, these executives must consider *opportunity costs* for

every decision they make: their resources are limited; both organizations and individuals can only expend so much time, effort, and money to achieve their planned goals. But *unlike* most other people, the relative costs and benefits NGO executives must consider can have profound, life-changing effects on the health and welfare of many people.

For example, an NGO focused on economic development may be able to fund the construction of either five schools or one hospital, but not both, even if there is an urgent need for both education and healthcare in the region. An executive at the NGO who sincerely wants to promote both objectives might do so most effectively by telling the right lies to the right donors, and then covers her tracks well. The rest of us do not frequently confront moral quandaries like the examples above in everyday life, where deception is a means to promote some apparent greater good – which is precisely why they may seem *to us* as mere hypotheticals, only designed to construct an "in-principle" moral theory. They may be useful in performing ethical analysis, but most of us can expect to confront such *momentous* decisions once or twice in a lifetime, if ever. Not so for the executive of an NGO: these types of difficult decisions could be very real choices executives are forced to make.

The second crucial feature of NGOs derives from the how effective and ethically managed NGOs are able to accomplish their humanitarian and charitable goals: by exercising their *moral authority* in the public sphere to rationally persuade others of the importance and righteousness of their cause. To explain what we mean by this, we will very briefly foray into Jürgen Habermas' *discourse ethics* and his theory of how individuals and institutions exercise power.

Habermas identifies the three dominant forces effective for exercising social control in today's increasingly technological, hyper-capitalistic, and globalized society: these are *violence*, *money*, and *moral authority*.

To use violence as a force for change is to coerce others into obeying your will, most directly through the use of *actual* physical force. But more often, violence takes the more subtle form of the mere threat of physical violence.[21] Violence always involves coercion, subtle or overt,

[21] E.g., a heavy detail of armed police can shut down or deter mass demonstrations, or a surveillance state can keep its populace in check by monitoring a person's communications, movements, activities, *etc*.

typically on the part of governments,[22] whether democratic or despotic, although sometimes individuals and other private interest groups (e.g., multinational corporations) with special access can influence more corrupt governments to use direct state power to protect their private interests.

Money, the second force, is a more "soft" but manipulative form of power, especially because of its relative invisibility. Money is most overtly used to exert social control through buying influence over those in power. Most insidiously for Habermas, money can corrupt free and open forms of social communication by infiltrating mass media. In an era where ownership of media outlets is increasingly concentrated in the hands of a few powerful and wealthy individuals, money has an increasingly disproportionate influence on public opinion. Corporate owned media outlets can exert control over which opinions and arguments are expressed, selectively portray and interpret information in a way that serves their interests, etc.

Finally, there is *moral authority*. The notion of moral authority is impossible to fully understand apart from Habermas' "theory of communicative action," which developed into his later idea of a "discourse theory." In extremely broad terms, discourse theory resembles a form of contractarian consensus-building process. Habermas argues that the only valid claims are those that can stand up to the scrutiny of critical reason. His notion of reason, however, is not individualistic but social-*discursive*: under ideal (and at least presently unrealizable) circumstances, he imagines that the whole of civil society is party to an ongoing voluntary, non-coercive, and critical discourse. In an ideal form of such discourse, everyone has an equal chance to advance their point of view, no claim is immune to critical scrutiny, no argument can be suppressed except by the rational force of a better argument, and everyone involved is genuinely motivated to find the best argument, whomever it comes from and regardless of their desire to "win."

---

[22] Armies, police, intelligence networks, *etc.* are so complicated that they require centralization, and so expensive that they must be funded by the populace through taxation or state coercion such as a draft; furthermore, in *democratic* states, such institutions are deliberately set up to serve the public interest, but the degree to which they are actually *responsive to* the demands of citizens – at least in some democracies – is a matter for debate.

Under these conditions, Habermas maintains, the best arguments will naturally come to be accepted as valid by all parties. When the discourse is specifically *ethical*, i.e., concerning how we ought to live and to treat each other, the norms that prevail under such an ideal discourse form the basis for a *universally valid code of ethics*. *Moral authority*, therefore, is simply the ability to persuade others in an ongoing free exchange of information, to the extent that it approximates this ideal discursive process. To influence people using moral authority is to *convince* them by appeal to rational argumentation, crucially, within a discursive context of mutual respect, free of coercion, and equally open to all.

In today's world, governments have a monopoly on violence (and are often unafraid to deploy it against their own citizens), and both wealthy private citizens and multinational corporations are able to use the force of money to influence governments and to dominate critical discourse (to suppress dissent and amplify consent). NGOs obviously cannot avail themselves of the first two forces, given that NGOs are *independent* of governmental control and are *not* profit-making private enterprises; therefore it follows that NGOs (especially those with distinctively ethical aims) can only (and *ought to*) use moral authority to achieve their ends, even in cases where these ends conflict with the wishes of governments and global capital.

A consequence of the NGOs reliance on moral authority to project its influence is that to maintain moral authority, the NGO's organizational culture must respect the norms that govern truly free critical discourse: equal toleration, authenticity and sincerity, and non-coerciveness. So it then follows that an *individual executive* engaged in critical discourse on behalf of an NGO must respect these norms.

Given this special obligation, we may finally return to the *specific question* of well-intentioned lying. In virtue of being committed to the norms of an open, free, and critical discursive process, does the NGO executive therefore acquire any *special obligations* that restrict further the conditions under which she may lie for a good cause? We argue that it is quite clear that she does. For consider whether the NGO executive ought to simply adopt the common-sense **MRP** as a norm while engaged in the discursive process: she may lie only when there are no alternative actions she could take to prevent some great harm or injustice (recall, such as death, deprivation of fundamental needs, or deprivation of basic human rights) from befalling other(s). Owing to the extraordinary moral dilemmas faced by the leaders of NGOs discussed above, such

leaders could easily face situations wherein the extreme conditions of the **MRP**, so rarely met in "ordinary" life, are actually frequently satisfied.

The Malaria Epidemic Scenario:

Picture an NGO focused on humanitarian relief in the poorest areas of sub-Saharan Africa. The malaria epidemic is predicted to be especially bad this year, and the organization does not have the resources it would need (antimalarial drugs, mosquito nets) to protect many of the villagers they work with. Suppose that the organization has *exhausted* its efforts in trying to raise awareness, to secure additional funding from big donors and government grants, *etc*. – and due to the *time sensitive* nature of the problem, thousands of people may die.

Sarah, the executive, realizes that there is *one* last-ditch effort they could take to secure the needed funds before people start dying. The organization's annual fundraising drive is fast approaching, wherein the NGO will inform member contributors and other possible small donors of the good causes their donations will serve. Sarah knows that its donor base is much more likely to respond much more generously to the problems of famine and starvation than an issue like deaths from malaria. In fact, Sarah has learned from experience that images of emaciated children tend to provoke a much more empathetic response than facts and figures about the human cost of malaria. The organization could not raise anywhere close to the funds it would need by focusing on the looming malaria epidemic.

Therefore, Sarah decides to convince her employees working on the drive to exaggerate the organization's efforts on famine relief this year, to lie about what the *real* humanitarian crisis is, and to mislead donors as to *how*, exactly, their contributions would go to save lives.

Arguably, Sarah, the imaginary executive, has lived up to her moral obligations as defined by the **MRP**. By stipulation, she has exhausted all alternative means available to her (this is a plausible scenario, given the extreme time-pressure). Therefore, the "necessary means" condition is satisfied: there is (we assume) no available means of preventing the harm except by lying to members and other small donors. Second, the requirement that such a lie is intended to avert a "great" harm or miscarriage of justice is satisfied, given that preventing death and serious bodily harm are paradigmatic types of the sorts of harms for which we are willing to excuse a lie, in

accordance with the **MRP**. Sarah ought to lie, from the perspective of consequentialism, deontological ethics, and likely virtue theory.

We will not engage in a process of reflective equilibrium here, but we will at least pause to note how *unintuitive* this result seems. On the one hand, the executive accomplishes a very good thing by sparing many people from illness and death. But it may strike the reader (and we would suggest reflecting on this point) that her strategy of *massively deceiving* the organization's support base of concerned citizens, who volunteer their resources in support of the NGO's cause, is to swindle and betray the trust of exactly those people who have but the most trust in the executive and her organization's mission to help make the world a better place. True, the executive is not guilty of *stealing* the money, in the sense that she is personally profiting, but it still seems intuitively wrong to take money from people with the promise that it will be used towards some noble end, with the full intention of diverting the money towards *some other* noble end, to which the donors did not originally consent.

We can explain what has gone wrong here by appealing to Habermas' notion of *moral authority*, which the NGO must project in order to effectively promote its causes, through shaping public opinion by means of rational persuasion. An NGO can lay claim to moral authority *only* if it adheres to the norms of critical public discourse. The crucial norm that is violated in the above example is that of honesty. *The executive acting on behalf of the NGO is not being forthright with supporters and with the broader civil society.* In fact, she actively distorts the exercise of public reason by implanting false or misleading information into the discourse. Whenever someone in a leadership role at an NGO does this, they risk the NGO's credibility within public discourse. For within the sphere of public reason, an individual or an organization must depend upon rational persuasion alone to convince others of the worthiness of the recommendations they propose. And lying is ultimately manipulation through linguistic communication, a subtler form of coercion – which has no place in a discourse that is supposed to be free of such power imbalances. In losing credibility, the NGO loses the influence it has within the ongoing discussion, and ultimately its ability to promote the good through rational persuasion – i.e., its moral authority. Truth and the power of the better-reasoned argument are an NGO's most powerful tools for achieving social change. In deceiving others, the NGO ultimately sabotages its best possible strategy for achieving moral aims. The stakes for breaching the public's trust are very high within the sphere of discursive public reason.

Combining the facts that (**a**) due to the nature of their work, NGO executives are often in a position where they ought to tell noble lies, even given the strict conditions of the **MRP**, that (**b**) NGOs depend upon their *moral authority* within the sphere of public reason to compete with the forces of *violence* and *money* and to perform good works, and finally (**c**) that lies, no matter how well intentioned, have *no* place within the sphere of discursive public reason, and an executive who tells a lie to advance the cause of her NGO imperils her organization's moral authority, we derive the following result: *In order to effectively promote good causes*, *NGO executives must be held to an even more stringent standard restricting well-intentioned lying than the **MRP** (when acting on behalf of their organization in their professional role, not as private citizens).*

What restrictions ought we to add, in order to give a satisfying special analysis of the ethics of lying within the NGO sector? Given that lies are inadmissible within the sphere of discursive public reason, should we recommend that NGO executives follow the **ERP** – i.e., *never lie, under any circumstance*, or is there some tenable middle ground? We believe that the key to understanding the additional moral obligations that the exercise of moral authority imposes is to recognize that these new limitations must necessarily be circumscribed to *non-coercive* situations, where at least some form of rational persuasion is possible – that is, where the NGO has some ability to exercise its moral authority. If the executive categorically lacks the ability to exercise its moral authority, then by definition it has no recourse to counteract through critical discursive processes *at all*, and hence one's moral obligations are relaxed back to the "default" **MRP**.

Now, these non-coercive situations extend beyond policy debates, outreach efforts, petitioning democratic representatives, and all such forms of participation in civil society, which are all normal activities of NGOs and that form part of the process of influencing critical public discourse. Here, on our argument, deception is strictly forbidden. But there can also be some situations where the scales of power are unfairly tipped against the NGO's interest in promoting good causes – typically by the forces of money and violence. In many or most of these situations, however, the NGO still has recourse to the court of public opinion, where it can draw attention to (or expose) human suffering and injustice, and by courting public reason, it can attempt to marshal support against such forces. This is so even if the NGO is "outgunned" by these forces – e.g., when a corporation spends vast sums of money on a media campaign to

discredit NGOs that oppose corporate malfeasance, or when a government clamps down on public expressions of dissent. Even if these two forces are very effective at countering the NGO's exercise of moral authority within the sphere of public reason, it remains true that an NGO's moral authority is the most powerful weapon in its arsenal (indeed, it is the *only* weapon). Practically, then, even if the fight is unfair, NGOs need to fight their battles in the court of public opinion, subject to the norms governing discursive public reason, win or lose. And so long as the leaders of NGOs are bound by the norms of honest and non-manipulative communication, lying is impermissible and counterproductive.

However, by way of analogy to the way in which we relaxed the **ERP**, by taking into consideration extraordinary circumstances, *there could be* coercive situations wherein the NGO has absolutely no recourse to the sphere of public reason. The simplest example would be if a government uses violence to the point of effectively "shutting down" civil society, say, by suspending freedom of the press and of assembly, by instituting martial law and suspending basic protections for civil and human rights. Suspend the very institutions and norms that make the exercise of discursive public reason possible, and trivially, the NGO is instantly dispossessed of moral authority in Habermas' sense of the term. Consider, however, a more interesting case, quite literally analogous to the "murder at the door" counterexample that falsified the **ERP**. Imagine an NGO working for the liberation of homosexual and transgender people in a country where they regularly fear persecution, both from the law and from mob violence. The NGO works intimately with LGBT activists to help advance their mutual cause. Suppose that an executive abruptly learns from an inside source that military officers are coming to the NGO's headquarters within the next hour, to demand that the NGO turn over the addresses of the activists and any other records they may have, so that the activists can be swiftly rounded up and sent to prison camps before the activists have a chance to flee. The executive in charge has such information, but could potentially falsify it to send the government on a "wild goose chase," while the NGO helps the activists go underground and seek amnesty. Does the executive's commitment to stringent norms of honest communication force her to betray her compatriots?

Intuitively, we think that the executive would be doing a very brave thing to lie to the government in this particular scenario. Here, there is no time to (say) write a letter of protest about the government's impending action for publication around the world, or to lobby the

United Nations to impose economic sanctions, and so on – and if the government is determined enough to use violence, no matter the state of public opinion, such actions would have virtually *no* chance of preventing the calamity, perhaps even if there were there more time. *The relevant moral difference* is that the NGO executive in this example is *deprived of all possible recourse* to exercise its moral authority through rational persuasion within the public sphere. *Nearly* all of the time, the NGO can take *some* action to marshal public opinion in support of its moral pursuits. Even where success is unlikely and the situation is very dire, by its very nature, the NGO is compelled to effect change in a specific, non-manipulative and non-coercive way.

Accordingly, deception is almost *never* a legitimate means for an NGO executive to achieve some moral end. The *only* exception to this constraint is when the NGO is forced into a situation where any possible exercise of moral authority is completely thwarted or rendered impossible (as in the "martial law" case). These extraordinary, exceptional circumstances place the NGO in a context where the possibility of rational persuasion of public opinion is foreclosed. If and when such circumstances are truly realized, NGO executives are no longer operating within the sphere of public reason, and hence not bound by its norms. In such extraordinary cases, the **MRP** of course still serves as the best moral guideline for NGO executives:

The Special Principle (SP)*: An NGO executive may never to lie in order to promote a moral end, so long as they have recourse to promote that end by exerting their organization's moral authority. In exceptional circumstances wherein a party uses coercive means to foreclose the executive's ability to promote its cause(s) using its moral authority, the executive may use deceptive means in accordance with the* ***MRP****.*

To clarify, and to avoid a potential misunderstanding: the justification for this exception is the NGO's complete *lack of any alternative action* that would allow it to exercise its moral authority (analogous to the **MRP**'s requirement that there be no *alternative non-deceptive* means of preventing harm or injustice). Crucially, the nature of this exception is *not* based in some "eye for an eye," retaliatory ethic, as if the use of overwhelming force (in the form of violence or money) could somehow nullify the NGO's responsibility to use rational persuasion. On this misunderstanding, if some party uses coercion to frustrate the NGOs pursuit of moral aims, then, because that party has stepped outside the sphere of discursive public reason, the NGO may take up the fight there as well by using *deception*, a subtle but still coercive means of

communication. *This is not our argument.* As in so many areas of ethics, the old adage "two wrongs do not make a right" applies here.

One last example will make this point, to be contrasted with the above case.

The Sexual Education Scenario:

Suppose an NGO constructs and operates schools in an underdeveloped country, which just so happens to be ravaged by the AIDS epidemic. One of the NGO's proudest accomplishments is that it provides comprehensive sexual education in the area of STD transmission, and to distribute condoms for free in the classroom. However, the NGO is extremely dependent upon government grants for its continued funding, and a new incoming administration has taken executive action to deny funds to any organization that dispenses condoms and does not teach "abstinence only."

Here, an organization (the government) is again using coercive means (the force of money) to frustrate the NGO's moral goals, imposing an agenda that will effectively result in more tragic deaths. The NGO cannot afford to forgo the money without completely shutting down operations, leaving children without an education. In this case, the NGO's executive would *not* be morally justified in agreeing to the terms of the grant money but to secretly continue the sexual education and condom distribution at the school. The executive is unfortunately bound by special moral obligations to resist the changes within the sphere of discursive public reason – e.g., by attempting to change the political landscape, or by seeking alternative means of funding – relying on the force of rational argument the power of truth to shape public opinion. That the new administration is effectively strong-arming the NGO and wields undue influence is morally irrelevant, precisely *because the NGOs strongest weapon is the truth*.

V. Conclusions

In summation, we find that NGO executives are actually much more ethically constrained in their ability to lie to promote good causes than are ordinary citizens. Persons in general have strong moral obligations to be honest with each other, owing to the corrupting influence of lies on an individual's respect for the truth and the need for all people to respect a robust norm of reciprocal honesty, which is a precondition for any stable social order. We have found that

several different perspectives on ethics lead us to these same conclusions as to why lying is typically a serious wrongdoing.

However, we also found (in agreement with common-sense morality) that in some rare circumstances lying may be morally permissible or morally obligatory – whenever the *only* way to protect someone from befalling some very serious harm or injustice is to deceive. This exception exists because while having a strong norm of reciprocal honesty is essential for leading a meaningful human life, it can be of no value to us unless we can lead lives *worth living* in the first place; this fact requires us to value human life, liberty, and well-being to such an extent that we should opt to value these moral goods, and these alone, above the truth – when it is practically impossible to simultaneously protect both. This was the "Moderate Restrictive Principle," (**MRP**) which we believe provides the ideal conditions for judging when lies with the intent of effecting moral good is permissible.

Our investigation into the distinctive features of NGOs turned up several key insights, which further narrowed these restrictions on well-intentioned lying, when NGO executives would lie to advance their organization's cause, acting as a representative of the organization (not a private individual). The fact that NGOs are specifically oriented towards the furtherance of some moral goal makes it highly likely that executives will be in situations where common sense morality (per the **MRP**) dictates that they *ought to* lie. On the other hand, using Jürgen Habermas' theory of discourse ethics, we showed how NGOs can only effectively compete with other powerful institutions, such as governments, corporations, and the mass media, through exerting their *moral authority* – their ability to influence popular opinion through arguments advanced within an ongoing discourse of public reason.

Precisely because moral authority derives from the ability to *rationally persuade* without *coercion*, engagement in deceptive practices causes an organization to lose its moral authority, for deception is a form of manipulation and thus an insidious form of coercion. It follows that there must be much more stringent restrictions on lying in the pursuit of moral ends for professionals acting on behalf of NGOs. In general, NGO executives have a moral obligation to never lie, precisely because of their reliance on moral authority – that they must *use* the truth to convince the public of the importance of their cause, not *distort* it. The single exception to this **Special Principle** is (as one might expect) when an NGO executive is coercively denied the ability to exercise its moral authority – that is, when a party takes action to completely prevent the

NGO from promoting their cause in the public sphere through rational argument. In these and only these quite exceptional circumstances, the executive of an NGO may use deceptive means if it is the only way to prevent extreme threats to human life, liberty, or wellbeing.

Recommended Reading

Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy.* Cambridge, MA: MIT Press.

Kant, I. (1996). On a supposed right to lie from philanthropy. In *Practical Philosophy*, Ed. Mary J. Gregor, pp. 605-616. Cambridge University Press.

Parfit, D. (2011). *On What Matters*, vol. 1, Oxford University Press.

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

Part Two: The Role of Truth in the Pursuit of Moral Ends: an Empirical Investigation

Introduction

The ethical analysis concluded that NGO executives are much more ethically constrained in their ability to lie to promote good causes than is the general population. However, do NGO professionals actually behave according to these ethical guidelines? Put differently, would NGO professionals consider lying in NGO related situations less permissible than in other scenarios?

The empirical study investigated NGO professionals' tendency to lie for moral causes. Previous research has shown that people's lying aversion tends to decrease when other people are expected to benefit from the lie. This effect was found both in lab (i.e., Erat & Gneezy, 2009[23]) and field research (e.g., Sherman & Hickner, 2008[24]). Moreover, it appears that people's lying aversion further decreases if they benefit from the lie in addition to other people's benefit (e.g., Erat & Gneezy, 2009). Prima facie, there is no reason to believe that NGO professionals don't share the same fundamental inclination to lie for the benefit of others.

As stated above, the prevalence of lying to benefit others has been investigated in the past. Erat and Gneezy (2009) explored this type of lies (termed: "white lies") in the general population through the use of psychological-economic experiments (controlled experiments conducted in the laboratory). White lies were divided into two sub-types: lies that help others at the expense of the person telling the lie ("Altruistic White Lies"), and lies that benefit other people as well as the liar herself ("Pareto White Lies"). The researchers discovered that 33% of their participants used altruistic white lies, whereas 65% used Pareto white lies, meaning that people were more prone to lying white lies when they were also the beneficiaries of the lies.

---

[23] Erat, S., & U. Gneezy (2009). White Lies, Rady School of Management Working paper. In (2012): "White Lies," *Management Science, 58(4)*, 723–733.

[24] Sherman R., & Hickner J. (2008). Academic physicians use placebos in clinical practice and believe in the mind-body connection. *Journal of General Internal Medicine, 23(1),* 7-10.

Nevertheless, Erat and Gneezy's economic game experiments were very simplistic. In these experiments lying always resulted in improving other people's condition, but never worsen it. This is unlike the standard outcome of lying outside the laboratory. In the real world, lying to benefit others and/or promote good causes often results in worsening the condition of others indirectly or directly (e.g., receiving money by false pretense means that others, who did not lie, are deprived of it). Moreover, and as reflected by the ethical report, the lying behavior ***is rarely an isolated act***, thus ramifications are to be expected (to the liar herself or to her peers). This is especially true in the NGO world in which lying may result in local consequences (e.g., a donor ceasing to donate money to the lying NGO), and global ones (e.g., ramifications on public opinion).

Main objectives

The goal of the empirical component of the research was to investigate three main questions:

1) What is the prevalence of the belief that lying for the sake of good causes is permissible, among Human Rights and Global Development senior NGO professionals from the U.S. and the U.K.? Is it greater than in a control group (senior professionals from the private sector)?
2) Among NGO professionals, does the prevalence of this belief depend on the type of the good cause promoted by the act of lying (e.g., promoting human rights)?
3) To what extent are NGO professionals likely to use good causes as justifications (i.e., use good causes as a kind of rationalization or excuse) to promote their own personal benefit?

Basic research assumptions

To execute the research, three fundamental assumptions were made. First, we assumed that NGO professionals perceive themselves as pursuing moral ends in their professional capacity. Moreover, we assumed that they have more experience in confronting moral dilemmas involving the potential of lying to promote good causes in their professional life compared to the control group. Therefore, we assumed that it was possible that the belief that lying to promote good causes is permissible was more common among this distinct population.

Second, we assumed that in the case of NGO professionals (especially senior staff) who have broad decision-making power and are likely to face moral dilemmas that raise the possibility of lying, it is highly plausible to assume a high correlation between their moral beliefs about the

practice of lying and their actual practices of lying. Accordingly, the research investigated the NGO professional's personal beliefs, but not their ultimate behavior.

Third, in order to make valid comparisons between NGO professionals and the control group from the private sector, it is crucial to exploit moral dilemmas similarly accessible to both populations. Thus, one cannot compare the two populations by merely investigating reactions to scenarios rooted in the NGO-world. Consequently, any comparison between these two groups should include general context scenarios. Therefore, for both groups we examined beliefs that pertain to both types of scenarios: NGO specific and general context ones.

Method

Participants

295 participants took part in the online survey. 226 of the participants were NGO senior staff professionals from Human Rights and Global Development NGOs from the U.S. and the U.K., and another 67 participants were sampled from the private sector senior professional population. All the participants were fluent speakers of English and resided either in the U.S. or the U.K.

Research design

The research was based on online questionnaires. The questionnaires included four sections. The order of the section presentation was kept constant across all participants.

Section 1

General description:

The first part of the questionnaire consisted of two questions which were phrased as scenarios. The scenarios consisted of two descriptions of events in which an executive / senior decided to lie to promote good causes. The two scenarios depicted two entirely different stories about two characters (a male and a female).

The participant had to determine to what degree she agreed with the decision to lie (on a seven point scale: 1 = strongly disagrees to 7 = strongly agrees).

The participants were presented, based on random selection, with one of two parallel scenario descriptions. The first involved a senior (either Dan or Linda) who lied to promote good without any personal gain (termed ***Altruistic lie***). The second scenario description involved a senior who

lied to promote a *moderate* good (less than in the Altruistic lie case), but there was also a personal gain resulting from the act (termed ***Egotistic-Pareto lie***). The scenario descriptions referred either to events in the context of an NGO or to events in a more general context, also based on random selection. Below is an example of a possible scenario.

An example scenario:

This is an altruistic scenario, rooted in the context of the NGO world.

iLearn is a well-established NGO that helps poor people in developing countries secure decent employment through online training courses in technological skills. In order to make these services widely available, they build many computer centers in rural areas in the global south.

Dan, the highly regarded and experienced director of iLearn's Central Africa department, is about to retire very soon. He arrives in the customs offices in one of the cities in Bengala, a small country in Central Africa, in order to receive a shipment of computers that are meant to be installed in a remote village in the country. In the office, he is asked to report total value of the computers. **He is well aware that if he tells the truth, regulations require a long bureaucratic process that could last up to three months.** Such a delay would mean **losing one full semester of online teaching in Bengala**, and thus significantly delay the technological training program for that year. He also knows that if he lies about the cost of the computers and reports that they cost much less than their real cost, he will receive the computers immediately.

Dan decides to report that the computers cost much less than they actually do and receives the computers.

Experimental design of section 1 (methodological details)

Half of the scenarios were rooted in the NGO world, and half included a general context. In addition, the scenarios addressed dilemmas involving the possibility of lying two types of lies. The first type was pure Altruistic lies (lies that promote good causes, and are not expected to benefit the liar, at least not directly), and the second type was Egotistic-Pareto lies (lies which were expected to benefit the liar substantially, but were also expected to promote good causes though to a lesser extent). Thus, this part of the questionnaire included 2 factors (type of scenario: NGO-world & General; type of lie: Altruistic & Egotistic-Pareto), creating 4 conditions: Altruistic-NGO; Altruistic-General; Egotistic-Pareto-NGO; Egotistic-Pareto-General

Two item sets (i.e., versions of the same basic scenarios) were created. Each item set included 4 variant scenarios, one for each condition. The two items sets can be viewed in the Appendix.

To avoid complexity, in this section each participant saw only 2 scenarios, one from each item set. Thus, no participant saw more than one version of each item. However, the two scenario versions (one within the NGO context and one within the general context) were either from the two Altruistic conditions, or from the two Egotistic-Pareto conditions.

The two item sets (with the 4 scenario versions within each set) were divided into 4 different lists, each with one scenario variant from each item set. Within each list the two scenarios were shuffled, so that half of the times the first scenario was in the NGO context, and in the other half – in the general context.

Both items and 4 conditions were rotated across both groups of participants. Each participant (either from the NGO professional group or the control group) was randomly assigned to each list.

Due to the length and complexity of the questionnaire (and especially Section 1 which was also the focus of the questionnaire), and to avoid fatigue effects, data from Section 1 were analyzed only from the first presented scenario (out of the two scenarios each participant saw) in each questionnaire. Nonetheless, it is important to note that during the data analysis, the same general pattern of results also emerged from analyzing data from both scenarios.

A cross-platform online survey was created to collect the data from the sampled participants.

Section 2

A second group of questions used the same 1-7 scaled format, but investigated a group of statements concerning general beliefs about the practice of lying to promote good causes. Participants had to determine to what degree they agreed with each of the statements. The same list of questions in the same order of presentation appeared in all lists.

The statements that appeared in Section 2 of the questionnaire:

Almost every person has lied at least a few times during their lifetime.
When a person is caught in a lie, even for a good cause, they run the risk of paying a heavy personal price.
Lies are harmful even when they are well intentioned, since they reduce interpersonal trust.
Many of those that condemn lies that are told for the sake of noble causes are hypocrites that

wouldn't hesitate to lie to promote their own self-interest.
It is never permissible to tell a lie.
Sometimes lying is better than telling the truth.
A lie can only be justified when told for the sake of a truly noble cause.
In the end of the day, a person will be judged by her achievements and not by whether or not she lied in order to achieve them.
A lie can only be justified when the liar herself does not significantly benefit from it.
It is impossible to promote important and significant causes of any kind while always telling the truth.

Section 3

This section included another scenario in the exact format described in Section 1. The scenario was set within ***the NGO context*** only and in a single condition we termed "***Super Egotistic***". In the Super Egotistic condition the immediate benefit to others was minimal (though existing), while the immediate benefit to the liar was considerable. We believe that any willingness to lie in this extreme NGO context scenario would be ascribed to lying in bad faith, with or without the involvement of self-deception. Exactly the same question appeared in all questionnaires.

The Super Egotistic scenario:

This super egotistic scenario is rooted in the context of the NGO world. All participants saw exactly the same scenario.

World Justice (WJ) is an NGO that focuses on advocacy efforts that are intended to influence affluent countries to increase their official development aid to developing countries. The NGO is run both by employees and volunteers.

Sam, the young and ambitious director of WJ's London office, is about to transfer to the organization's headquarters in New York for a short period, after which his next assignment will be decided. **He knows that the decision about his next position will be greatly influenced by the amount of funds he managed to raise in his current one**.

Sam is trying to initiate an interactive online seminar for all of the NGO's English speaking volunteers, focusing on effective ways for the volunteers to engage in advocacy activities around WJ's mission. If this initiative succeeds, it would garner much attention in the New York headquarters.

Sam arrives at a meeting with an important donor who is interested in donating to online activities focused on global development advocacy, and is considering a donation to WJ.
The donor asks Sam whether WJ was responsible for promoting the "food for all" bill in the U.S. congress, **a policy reform the donor strongly supports**.
In reality, while WJ supported the bill, it was **not significantly** involved in promoting it. Sam knows that if he tells this to the donor, **it would significantly reduce his chances of receiving the donation**.

Sam decides to tell the donor that WJ was significantly involved in promoting the bill.

Section 4

A final group of questions was used to gather demographic and occupational information about the participants.

Results

It took approximately 10 minutes for participants to complete the questionnaire.

Overall, 158 participants saw one of the Altruistic-NGO scenarios and one of the Altruistic-General scenarios: 118 were NGO professionals, and 40 participants were from the private sector (i.e., the control group). 137 participants saw one of the Egotistic-Pareto-NGO scenarios and one of the Egotistic-Pareto-General scenarios: 110 were NGO professionals, and the other 27 participants came from the private sector.

Basic demographic information

The mean age of the NGO professionals group was 56.54 (SD=9.00), and 54.54 (SD=11.07) for the control group. Consequently, no age difference was found between the NGO professionals and the private sector professionals (i.e., the control group). The NGO group consisted of significantly more women than in the control group (61% vs. 25%). Nevertheless, except for a gender related difference that was found in one of the statements in Section 2, no gender

related effects were discovered. Thus, with the exception of Section 2, gender will not be reported.

In addition, a higher percentage of NGO professionals held an academic degree (90% vs 77%).

The following is a brief summary of the results. Statistical details are available upon request.

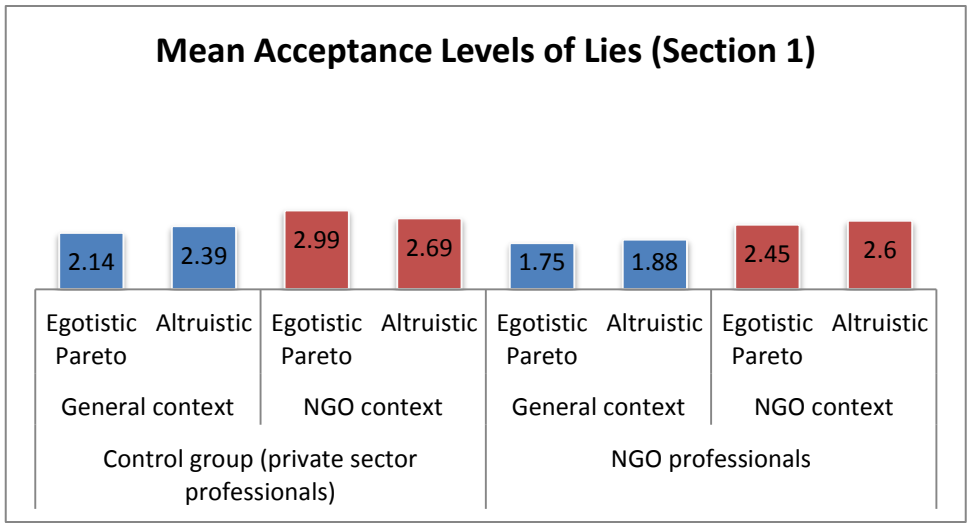
Section 1

Due to the relative complexity of the questionnaire (and specifically Section 1) and to avoid possible fatigue effects, the data reported in this section were collected only from the first scenario participants saw in the survey. (Note that different participants saw different scenarios, so that overall, scenarios from the two item sets and from all possible conditions were equally presented.) Nonetheless, a similar pattern was found when the two scenarios were analyzed together.

Generally, the data showed a relatively low level of lie acceptance. On a scale of 1 to 7, the general level of acceptance was 2.36 (SE = 0.08), meaning that as a rule, participants did not approve of lying for good causes. No significant difference between men and women was found, and there was no difference between age groups.

The acceptance of lies was slightly higher among participants from the private sector group compared to NGO professionals (2.55 vs 2.17, respectively). Overall and across the two participant groups (NGO professionals and the control group), the acceptance of lying to promote good causes in the NGO context was significantly higher than in the general context (2.68 vs 2.04). Interestingly, no significant difference was found in the acceptance of Altruistic versus Egotistic-Pareto lies (2.39 vs 2.33). A breakdown of the results by groups of professionals, context and type of lie is presented in Figure 1.

Figure 1: Mean acceptance levels for different type of lies, contexts, and professionals groups. There is a significant difference (across the two participant groups) in acceptance levels of lies in the NGO context (red) and the general context (blue).



Section 2

Overall, participants tended to accept the statements regarding the costs of lying, as well as the statement that almost everyone had to tell a lie at least once. Results are presented in Figure 2.

A significant difference was found between the two participant groups in two statements (alpha<.05):

"Lies are harmful even when they are well intentioned, since they reduce interpersonal trust": 5.47 for the control group, and 5.9 for the NGO-professionals group, meaning that the latter tended to accept the statement more than the private sector population.

"Many of those that condemn lies that are told for the sake of noble causes are hypocrites that wouldn't hesitate to lie to promote their own self-interest": 4.63 for the control group compared to 3.89 for the NGO professionals, meaning that the private sector group tended to

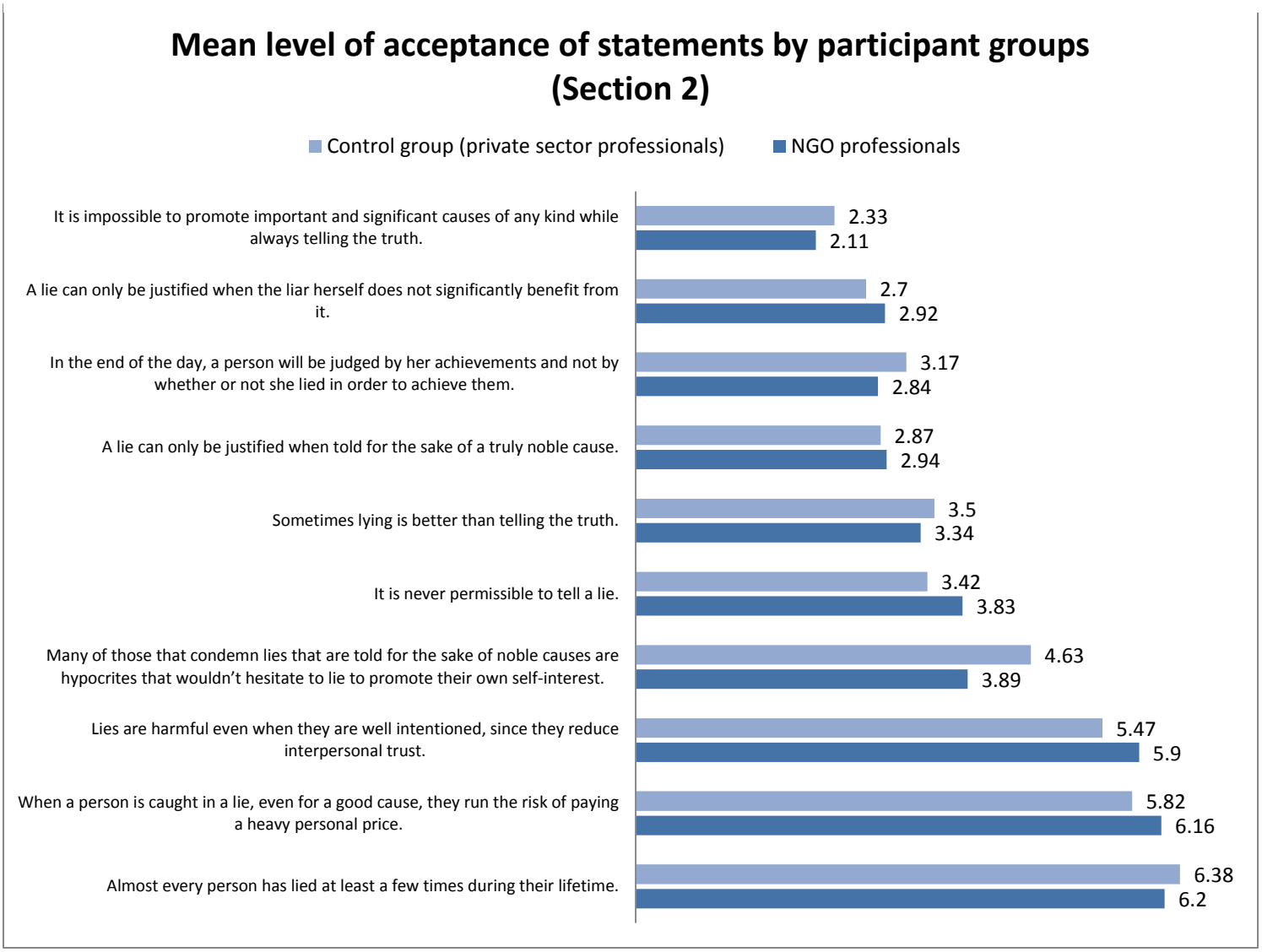
accept the statement more than the NGO professionals group, leaving the latter in a more neutral position. This statement also elicited a reliable effect of gender (4.44 for men, and 3.79 for women), but crucially, it did not interact with the two participant groups (the NGO professionals and the control group).

Section 3

As recalled, in Section 3 participants were presented a description of an event in which an NGO executive decided to lie in a situation where the lie was expected to lead to a substantial personal gain, yet with negligible positive effect on the goals of the organization (***Super Egotistic)***.

The mean level of acceptance of the lie was relatively low, with no difference between the NGO professionals and the control group: 2.09 and 2.07 respectively. These acceptance levels are indeed significantly lower than the acceptance levels of lying in the other two conditions (namely, the Altruistic and the Egotistic-Pareto conditions) in the context of the NGO world. Thus it seems that both groups made the distinction between lying in good and in bad faith.

Figure 2: Mean level of acceptance of statements by participant groups.



Discussion

The empirical study aimed at deepening our understanding of the belief systems of senior NGO professionals with regards to lying to promote good causes.

It seems that the prevalence of the belief that lying for the sake of good causes is permissible is relatively low. Nonetheless, the acceptance level of lies was lower in the NGO group compared to the control group. This outcome is in accordance with the differences found between NGO professionals and the private sector group in their acceptance of certain statements presented

in Section 2, and specifically those regarding the problematic consequences of lying at the personal level as well as the social one, even when it is done to promote good causes.

However and most interestingly, we found a unanimous consensus among the NGO professionals and the control group that lying is more permissible in the NGO world than in other contexts, even when the benefit to others in both cases (as exemplified in the scenarios that can be viewed in the Appendix) was similar. This belief seems to reflect a general social convention, as we found no distinction between the NGO professionals and the control group. Note, however, that this finding clashes with the conclusion reached by the ethical analysis, namely that NGO professionals ought to be more restricted in lying to benefit others than the average person.

Another notable outcome is our inability to find differences between participants' willingness to lie in Egotistic-Pareto and Altruistic situations. Admittedly, it is possible that this is due to false reporting, resulting in a relatively low reported acceptance level of lying in the Egotistic-Pareto scenario. However, we don't think that this is the case since the altruistic scenario also invoked a relatively low level of acceptance of lying, which was consistent with the acceptance levels found in the statements presented in Section 2. Thus, it seems that peoples' lying aversion is genuine and reflects the realistic and complex properties of the scenarios described in Section 1.

The results in the Super Egotistic scenario further corroborate this hypothesis. In the only situation in which lying for good causes in (obviously) bad faith was tested, a significantly higher degree of lying aversion was observed, compared to the Altruistic and Egotistic-Pareto scenarios.

The main goal of the empirical study was to assess the levels of acceptance of lying for the sake of good causes in good faith. The assessment of participants' inclination to lie in bad faith is limited. Methodologically, this is because the measurement was only introduced in the last sections of the questionnaire, making participants highly susceptible to fatigue effects. Moreover, the assessment was restricted to a single scenario in a single condition (i.e., the NGO context for both participant groups). A future study may develop this point further.

To conclude, the empirical study sought to explore NGO professionals' prevalence to lie for moral causes. We found that NGO professionals presented a higher lying aversion than the control group. Nonetheless and similar to the control group, NGO professionals perceived lying

for moral causes in the NGO world as more permissible than in non-NGO contexts. This is in contrast with the recommendation of the ethical analysis.

Appendix

Below is a list of the two items sets in all 4 conditions in Section 1.

Item set 1

Altruistic, NGO context

iLearn is a well-established NGO that helps poor people in developing countries secure decent employment through online training courses in technological skills. In order to make these services widely available, they build many computer centers in rural areas in the global south.

Dan, the highly regarded and experienced director of iLearn's Central Africa department, is about to retire very soon. He arrives in the customs offices in one of the cities in Bengala, a small country in Central Africa, in order to receive a shipment of computers that are meant to be installed in a remote village in the country. In the office, he is asked to report total value of the computers. **He is well aware that if he tells the truth, regulations require a long bureaucratic process that could last up to three months.** Such a delay would mean **losing one full semester of online teaching in Bengala**, and thus significantly delay the technological training program for that year. He also knows that if he lies about the cost of the computers and reports that they cost much less than their real cost, he will receive the computers immediately.

Dan decides to report that the computers cost much less than they actually do and receives the computers.

Egotistic-Pareto, NGO context

iLearn is a well-established NGO that helps poor people in developing countries secure decent employment through online training courses in technological skills. In order to make these services widely available, they build many computer centers in rural areas in the global south.

Dan, the young and ambitious director of iLearn's Central Africa department is about to transfer to the NGOs New York headquarters for a short period, during which his next promotion will be decided. He arrives in the customs offices in one of the cities in Bengala, a small country in Central Africa, in order to receive a shipment of computers that are meant to be installed in a remote village in the country. In the office, he is asked to report total value of the computers. **He is well aware that if he tells the truth, regulations require a long bureaucratic process that could last up to three months**. Such a delay would mean **losing one full semester of online teaching in Bengala**, and thus significantly delay the technological training program for that year. In addition, **if the delay occurs, the person who will get the credit for opening the new center will be his successor**. He also knows that if he lies about the cost of the computers and reports that they cost much less than their real cost, he will receive the computers immediately.

Dan decides to report that the computers cost much less than they actually do and receives the computers.

Altruistic, general context

Dan, the highly regarded and experienced principal of a school in the state of Florinia, is about to retire very soon. After many efforts, he managed to secure a large procurement of technological equipment from the ministry of education that would enable him to open a computer lab in his school. This lab will help his students acquire technological skills that will be of great help in finding a job after graduation.

He receives forms from the ministry that ask him to report the average grade his students scored in a standardized test. **He is well aware that the actual average is lower than the standard set by the**

ministry, and that if he reports this score he would be required to go through a long bureaucratic process that would delay the arrival of the new equipment by 3 months. Such a delay would mean **losing one full semester of using the new computer lab**, and thus significantly delay the technological training program for that year. He also knows that if he lies about grade average and reports a significantly higher grade, he will receive the equipment immediately.

Dan decides to report that the grade average is significantly higher than it actually is, and receives the equipment.

Egotistic-Pareto, general context

Dan, the young and ambitious principal of a school in the state of Florinia, is about to finish his tenure at his current job and, after going through an evaluation by the ministry of education, receive his next assignment. He knows that his performance in his current position will greatly affect the level and prestige of his next one.

After many efforts, he managed to secure a large procurement of technological equipment from the ministry of education that would enable him to open a computer lab in his school. This lab will help his students acquire technological skills that will be of great help in finding a job after graduation.

He receives forms from the ministry that ask him to report the average grade his students scored in a standardized test. **He is well aware that the actual average is lower than the standard set by the ministry, and that if he reports this score he would be required to go through a long bureaucratic process that would delay the arrival of the new equipment by 3 months**. Such a delay would mean losing one full semester of using the new computer lab, and thus significantly delay the technological training program for that year. In addition, **if the delay occurs, the person who will get the credit for opening the new lab will be his successor**. He also knows that if he lies about grade average and reports a significantly higher grade, he will receive the equipment immediately.

Dan decides to report that the grade average is significantly higher than it actually is, and receives the equipment.

Item set 2

Altruistic, NGO context

Linda is the Chief Financial Officer of the NGO “Rights for Prisoners of war” (RPW). The NGO's staff has been engaged in a long negotiation with the government of Cubistan, an internationally isolated country that is in conflict with the country of Squaristan regarding the right to visit in a detention facility in Cubistan in which prisoners of war from Squaristan are held. RPW’s headquarters are located in Roundistan, which is neutral regarding this conflict. Due to this neutrality, the government of Cubistan decided to allow a psychologist from RPW to visit the detention facility to observe and report on the conditions of one group of prisoners. **Due to several restrictions imposed on the visit, it is uncertain to what degree the psychologist would be able to reliably assess the condition of the prisoners he will be allowed to see, which represent only one group from the total number of prisoners detained.**

Before the visit takes place, the ambassador of Cubistan calls RPW and requests that they send him a letter **signed by all the executives in the NGO (including Linda)**, stating that the organization and the observing psychologist do not have any ties to Squaristan.

Even though the organization doesn’t have official ties to Squaristan, **Linda knows that the psychologist does have family members residing in Squaristan**. However, no one else in RPW is aware of this fact.

Linda decides to sign the letter and not share this information with anybody.

Egotistic-Pareto, NGO context

Linda is the Chief Executive Officer of the NGO "Rights for Prisoners of war" (RPW). The NGO's staff has been engaged in a long negotiation with the government of Cubistan, an internationally isolated country that is in conflict with the country of Squaristan regarding the right to visit in a detention facility in Cubistan in which prisoners of war from Squaristan are held. RPW's headquarters are located in Roundistan, which is neutral regarding this conflict. Due to this neutrality, the government of Cubistan decided to allow a psychologist from RPW to visit the detention facility to observe and report on the conditions of one group of prisoners. **Due to several restrictions imposed on the visit, it is uncertain to what degree the psychologist would be able to reliably assess the condition of the prisoners he will be allowed to see, which represent only one group from the total number of prisoners detained**.

This visit is **very important to Linda**. This will be the first time external representatives are allowed to visit prisoners in Cubistan, and she has just received an invitation to give a series of lectures on prisoners of war in U.S. universities, and even started negotiations with a large publishing company to publish a book on the topic.

Before the visit takes place, the ambassador of Cubistan calls RPW and requests that they send him a letter **signed by all the executives in the NGO (including Linda)**, stating that the organization and the observing psychologist do not have any ties to Squaristan.

Even though the organization doesn't have official ties to Squaristan, **Linda knows that the psychologist does have family members residing in Squaristan**. However, no one else in RPW is aware of this fact.

Linda decides to sign the letter and not share this information with anybody.

Altruistic, general context

Linda is the VP of sales for "Anti-Viral", a pharmaceutical company that, among other things, is working on finding a cure for HIV-AIDS. Due to cash flow problems that do not fall under her area of responsibility, the company will soon be forced to reduce the salaries of several employees (**not including Linda herself**).

It is well known that in Cubistan, an internationally isolated country that is in conflict with the country of Squaristan, there is an ethnic minority which **has an especially low rate of HIV infection**.

Anti-Viral has been engaged in a long negotiation with officials from Cubistan in order to allow a genetics expert from the company to visit Cubistan and conduct a genetics study on the ethnic minority in order to gain valuable data that would support Anti-Viral's efforts to advance their research into finding the cure. **Any data gathered would significantly enhance Anti-Viral's ability to raise additional funds and thus avoid the salary reductions**. However, **the chances that the data would actually lead to a cure for HIV-AIDS are low**, and in any case a long development process would be required.

Anti-Viral's headquarters are located in Roundistan, which is neutral regarding the conflict between Cubistan and Squaristan. Due to this neutrality, the government of Cubistan has agreed to allow the expert to conduct the research in its territory.

Before the visit takes place, the ambassador of Cubistan calls Anti-Viral and requests that they send him a letter **signed by all the executives in the company (including Linda)**, stating that the organization and the expert geneticist do not have any ties to Squaristan (due to the risk of espionage).

Even though the company doesn't have official ties to Squaristan, **Linda knows that the expert does have family members residing in Squaristan**. However, no one else in Anti-Viral is aware of this fact.

Linda decides to sign the letter and not share this information with anybody.

Egotistic-Pareto, general context

Linda is the chief executive officer for "Anti-Viral", a pharmaceutical company that, among other things, is working on finding a cure for HIV-AIDS. Due to cash flow problems, the company will soon be forced to reduce the salaries of several employees. Linda knows that she will be held accountable for her role in this problem, and **is not sure whether she will be able to keep her job.**

It is well known that in Cubistan, an internationally isolated country that is in conflict with the country of Squaristan, there is an ethnic minority, which has an especially low rate of HIV infection.

Anti-Viral has been engaged in a long negotiation with officials from Cubistan in order to allow a genetics expert from the company to visit Cubistan and conduct a genetics study on the ethnic minority in order to gain valuable data that would support Anti-Viral's efforts to advance their research into finding the cure. **Any data gathered would significantly enhance Anti-Viral's ability to raise additional funds and thus avoid the salary reductions**. However, **the chances that the data would actually lead to a cure for HIV-AIDS are low, and in any case a long development process would be required.**

Anti-Viral's headquarters are located in Roundistan, which is neutral regarding the conflict between Cubistan and Squaristan. Due to this neutrality, the government of Cubistan has agreed to allow the expert to conduct the research in its territory.

Before the visit takes place, the ambassador of Cubistan calls Anti-Viral and requests that they send him a letter **signed by all the executives in the company (including Linda)**, stating that the organization and the expert geneticist do not have any ties to Squaristan (due to the risk of espionage).

Even though the company doesn't have official ties to Squaristan, **Linda knows that the expert does have family members residing in Squaristan**. However, no one else in Anti-Viral is aware of this fact.

Linda decides to sign the letter and not share this information with anybody.