

Hoa Lo Prison and the Stanford Prison Experiment – Thoughts on Morality, Institutional Life, the Hermetic Experience and Human Nature

The contrast could not be more striking, or suggest more contradictory theses about human nature. Consider two prisons. In each, ill defined directions are given to the prison guards. Little or no guidance is given with regard to how they are allowed to treat inmates. One common element of instruction exists. In both prisons, the guard are instructed to exercise their creativity in service of one essential goal; causing in the inmate population a feeling of helplessness, and lack of control or power over their individual fates. The further purposes served by this essential goal are quite different in the two prison settings, but success in this one task is deemed necessary to the further goals.

Equipped with the accouterments of prison authority (uniforms, batons and shades that cut off the possibility of eye contact) the guards in our first prison engage their creativity in efforts to carry out their directive. They are to run this prison for a set period of two weeks. They make use of time-worn techniques intended to strip identity, (drab prison uniforms, prisoner numbers used in place of names, for instance). They also use other techniques intended to disorient and confuse inmates (lighting at all hours, sleep disruption, transport while blindfolded). Other methods are improvised, including forced nudity, collective punishment, verbal insult and humiliation, solitary confinement. All conspire against the prisoner's sense of identity and security.

In this prison, the hapless inmates quickly dissolved under this stress, and any nascent unified front dissipated. After a brief period of rebellion, and standing up for themselves, prisoners largely looked to 'survive' 'get out in one piece' and most of them were pliant to the guards' demands and manipulations. Some small number of inmates rebelled, stood up against the abuse, but they are effectively isolated from the group, sometimes literally, but always at least in the figurative sense that the other prisoners would join the guards in condemning the rebels. Quite a few prisoners suffered emotional and psychological breakdowns while in captivity. More often than not prisoners adopted the moral evaluations of fellow prisoners that were on offer from the guards, evaluations clearly offered as part of the methodology of control. If a prisoner was considered a 'bad prisoner' by the guards, so it was in the eyes of his fellow prisoners. Prisoners readily followed guard directives to deride the recalcitrant.

Upon release from this first prison, prisoners had understandable resentment toward guards, and prison administration, but considerable guilt about their own behavior with regard to fellow prisoners, particularly the rebels, people that would otherwise have been viewed as courageous, and perhaps heroic.

The somewhat Hobbesian picture of human nature that emerges from our first prison is grim. Under pressure, powerless human beings will be reduced to largely self interested motivations, will be willing to overlook abuse of others and in fact sometimes participate, or treat the occasional brave rebel, not as a hero, or a person of moral and physical courage, but as a morally bad individual. An entire life of moral behavior is negated or at least thrown into question, altruism and social cohesion collapse, concern for the welfare of others is dispensed with. The veneer of civilization is thin, and morality appears to be a sham, something that will fall by the wayside under pressure.

Looking at prison number one, lessons can be learned about proper responsible 'architecture' of systems of incarceration, and similar institutions where there are stark power relations, and indeed, much has been made of these matters in the literature that has grown up around this particular prison. But that will not be the primary focus of this paper.¹ The focus here will be upon the prisoners, and lessons we can learn from them, lessons we can learn about life in pressure situations. In our first case, as was just said, the lessons look to be quite grim. This brings us to our second prison:

In the second prison, all of the methods used in the first were utilized, to degrees more intense, and for a much longer and indefinite period of time. There were also additional harsher environmental factors and methods used over most of that same time period. These included undernourishment, life-threatening and crippling torture, the stocks, isolation for months or years, extreme punishments for communications between inmates, uncooperative behavior or rebellion. Bathing was a luxury allowed only occasionally. Extremes of temperature were endured for years.

In our second case, upon release, there was, as with prison number one, an understandable sense of moral outrage toward the guards and administration of the prison, even as there was a recognition that they were something like 'professionals' doing their jobs. There was even a level of forgiveness by some prisoners. But, in stark contrast to the inmates of our first prison, inmates of our second prison did not have the burden of living with considerable guilt vis-à-vis their behavior with regard to fellow inmates. For, the prisoners in our second case exhibited extreme levels of unity, were incredibly cohesive, presented a frustrating and always united front to their jailers, regularly put the welfare of fellow inmates above their individual welfare, and quite a few came away with a paradoxical appreciation for the experience. For they believed that otherwise, they would probably not have been able to experience either the level of love for their fellow man, or develop the courage, moral and physical, that we can only describe as heroic, nor plumb the full depth of human potential. Yet, to a man, none of the former inmates of this second prison would describe themselves in these exemplary terms.

¹ Zimbardo, The Lucifer Effect

They consider themselves to be ordinary human beings. Most people, they say, would react in the same way to the circumstances. As we will see, there is reason to doubt this. It seems their background was crucial to the outcome. But, even if we grant that to be the case, there are certain more universal lessons to be gleaned, that can be of value.

The picture of human nature that emerges from our second prison is anything but grim. Using its story, we can infer that under extreme pressure, powerless human beings will pull together, work toward their collective survival, put service to others before self, will not stand idly by while others are abused, and will rally round the rebel, the hero, serve his cause, and indeed become heroic themselves. Human beings in such circumstances will not allow their moral sense to become perverted. They will not succumb to the coercion of propaganda paired with pain. They will not be convinced that morally good people are morally bad. Civilization and morality do not succumb as mere thin veneers, melting away under the heat, and indeed this experience seems to suggest that they find their origins, as potencies within humanity that flower in times of great stress. They are innate features of humanity that will be coaxed to life in such harsh circumstances. In short; there is a deep and abiding optimism in the view of human nature that seems to be a natural reading of the events surrounding prison number two.

Obviously, these readings of our two results are so different as to be apparently contradictory. One then is naturally led to assume that there must be some important differences in the two cases that are sufficient to account for the radically different results. The search for these differentiating and explanatory features is the intent of this essay. One cannot hope to give a detailed hypothesis, but a reasonably compelling sketch. It will rely on the reflections of two people that have thought long and hard about their experiences in these two prisons. The first person is Dr. Phil Zimbardo, research psychologist and Principle Investigator of the Stanford Prison experiment. We will see that his thoughts about proper exercise of authority and architectonic of institutional structure will have some analogs with the situation at our second prison, the Hoa Lo prison (Hanoi Hilton) and will have relevance to things we can say about the differences between the two prisoner populations. The second expert is Admiral James Stockdale, who was prisoner of war in Hoa Lo for much of the duration of the Vietnam War. He too has things to say about responsible leadership or authority within the prison, in particular the inmate population, and effective building or architectonic of institutions that will mediate moral survivability in situations of powerlessness and duress.

By reflecting on the contributions of these two men, I think we can glean something of significance to say about human nature under stress, and the significance of institutions, roles and rules, for a proper understanding of and coping with the temptation to fall into moral error when under stress. We can also glean answers to the more broad and philosophical question of the relationship that exists between institutions, morality and human nature.

But, before we talk about the views of these two men, I do want to reinforce the contrary seeming results of our two cases. For, if you were to be asked to bet upon which prison population would exhibit more rebellion, and present a more unified front to the constituted authority; you would probably place your chips on the space occupied by the Stanford prisoners. For, their circumstance was not only significantly less harsh than was the case at Hoa Lo, but they were always free to leave. It was a consensual imprisonment for which they were being paid, albeit a very modest sum. What is more, the guards were recruited from the same pool as were the prisoners. All were volunteer college students from 1970 Stanford. So, ideally, you would expect that they would have knowledge of this fact operating somewhere in their conscious mind, realize there were no serious repercussions for standing up to 'the man', and would therefore be more disposed to act 'heroically'. As events unfolded, this was most decidedly not the case.²

At Stanford, prisoners and guards lost sight of the artificiality of their situation. Why? Zimbardo's claim is that the guards lost themselves in their roles because at some level more or less conscious, they wanted to contribute to science, play assigned roles properly, and wanted to succeed in carrying out the instructions they were given, pursuant to their roles. One guard even claimed he had grown interested in conducting 'mini-experiments' of his own, seeing how far he could push with abuse before prisoners would push back. Whether or not this was a rationalization for a desire to behave sadistically, the fact remains that respect for the purposes of the experiment, and their assigned roles in that experiment did play a part in damping moral awareness.

Interestingly, Zimbardo himself reports he gradually lost sight of the fact that he was a scientist, becoming engrossed in his other more 'artificial' role, a role he now says he should not have taken on in the first place; the role of the superintendent of the jail. This loss of perspective in the Stanford case is remarkable and in need of explanation. If anyone would have been in a position to be meta-cognitively aware of the risks involved in taking on a role 'inside' the experimental institution, it would appear to be the PI. Remarkably, this loss of perspective, and an attendant loss of moral sensitivity, was something that was an issue not only for the guards and superintendent, but the prisoners as well.

It is clear that most of the prisoners involved lost sight of the fact that the jail was not real. This forgetfulness might help explain the lack of heroics or rebellion, through the sense of despair having formed due to the loss of perspective. But this just leads us to the other horn of our paradox:

² This is all the more remarkable when you consider that they were young college students during the height of the countercultural youth movement of the 1960s to 1970s, with its suspicion for authority, and especially law enforcement.

You would probably not have bet on the Prisoners at Hoa Lo presenting a unified front, or significantly rebelling for any period of time, if you compared their severe, genuine and long lived plight to the lesser and sham plight of the Stanford prisoners. Yet, the inmates of Hoa Lo far and away take the prize for exhibition of unified rebellion and heroics.

How can this be?

As intimated before, I think part of the answer can be found as a result of careful meditation on the nature of the several institutions involved in our two “laboratories”, and the powers of these institutions to mold cognition and behavior of individuals contained therein. Institutions have ubiquitous and sometimes overlooked power over human beings. Why? Essential to fully grasping this power, it is crucial that we come to terms with some rather obvious facts. Institutions have ends for which they exist, as well as rules, roles and leadership geared toward those ends. Leadership functions to perpetuate and interpret the institutional ends and frameworks. Institutions provide purpose and status thanks to their long lives. Individual human beings live within them for a time, contribute to their lives, and then leave them. Living within institutions confers status. All of this is trivially true, but also uniquely true of human beings. A meditation upon these facts may help to explain the pitfalls in our first case, and the excellences on display on our second case.

It is to be noted that Zimbardo’s analysis, (one carried out primarily in terms explicating how the artificial jail and its institutional system had damping effects on the moral awareness of guards and prisoners) is also of profit when we look at how the Hoa Lo experience, particularly the active role taken by POW leadership and the general prison population, in fact heightened moral awareness in that prison population. Admiral Stockdale has much to say in this regard. He makes clear to us that there are two sides to the situational “influence” a system or institution can have on moral cognition and behavior. His story, his account of institution building within the confines of Hoa Lo accentuates the positive side, while Zimbardo’s cautionary tale emphasizes the negative. We will see that competent and careful institutional creation and leadership tends toward the positive heightening aspect, while incompetent, hasty and hurried institutional creation and leadership tends toward the negative and damping. And, crucially, we see by comparing our two cases, that laudable ends do not guarantee that institutions created to serve those ends or actions taken by individuals in their roles of service to those ends will also be laudable.

All of this has a bearing on moral science, more broadly construed, because we all live firmly ensconced in a world filled with human institutions and multitudinous pressures exerted on those institutions and individuals within them. What is remarkable about our two cases is that they, like scientific experiments in other fields, in effect isolated the prisoners from factors other than the prisons’ institution and pressure (imperfectly of course), and thereby allowed us

to see the effects of such isolation and concentrated mono-sourced pressure. Once in this isolated state, not unlike physical objects in the artificial setting of chemical or physical experiments, prisoners were subject to intense moral, emotional and psychological coercion, carried out by the two institutions in question, institutions that were created with the express purpose of creating such pressures. We are allowed to watch the results, that is; how the group of individuals responds to this pressure. We observe, and can answer psychological or sociological questions: Do the prisoners respond as a group? Do they respond as individuals? Do they support each other? Do they look out primarily for self? Do they betray? Do they defy? Do they become compliant and docile? What effect is there on their moral awareness, or moral purpose? And for each of these questions, we can also answer the inevitable questions as to why we see some results in one case, but not the other. We can formulate answers in terms of the deliberate design elements of the institutions in question. This might be seen as a pair of morally compromised scientific experiment, but that is exactly what we have; scientific experiments concerning institutional design.

In one case the pressure, in a sort of ironic development, prompted deliberate design of an institution. This was carried out by prisoners. Under incredible moral, emotional and psychological pressure Admiral Stockdale and the other American officers at Hoa Lo quite deliberately and consciously set about creating a “civilization” and an institution which proved vital in setting the inmates’ moral bearings in relations to the North Vietnamese institution within which they were being held and tested. At the core of this small civilization was the Code of Conduct for American prisoners of war that had been formulated in the wake of the Korean conflict. Nothing like this occurred at Stanford. We need to know why this is. The answer is of some amount of practical significance.

As Zimbardo emphasizes, we must take care to pay attention to the fact that institutions shape behavior, not only of those in the “ranks”, and those at opposite ends of power relations, but leaders as well. But, leaders are expected to consciously guide institutions, serving as their ‘hearts and minds’, the rational guide and moral conscience of the organizations. They should have a meta-level grasp of things.

Being human, leaders must be aware that institutions can damp their moral cognition, just as much as they can similarly affect the rank and file. This fact, along with a consideration of the fairly complex interplay between the goals of the conflicting cultural/military institutions that we find in the Hoa Lo case, a conflict that did not exist at Stanford, can allow us to sketch an explanation of the seemingly paradoxical results of our two cases. Finally, reflection on the results of these explorations tell us something about the essence, limits and potentialities of human nature, and allow us to sketch a picture of man as (please pardon the pretentious neologism) ‘*homo moralis institutionalis*’, a moral creature and a creature of institutions.

A brief roadmap of the way forward: We will first sketch in some detail the nature of institutions, that is; the roles played by the ends for which they are created; the importance and attraction of tradition; the functional role played by...well...roles within institutions; and the centrality of leadership for healthy long lived institutions. We also will sketch how these factors can explain the behavior both of prisoners and of those in authority in our two prisons. In particular, we will see how carefully crafted rules and clear and engaged leadership allowed for success and a morally sound prisoner's institution in the Hoa Lo prison population, where no such rules and weak inept or detached leadership, if not a complete lack thereof rendered success with integrity less likely, and produced no prisoner institution in the Stanford case. Indeed similar institutional design and leadership shortcomings on Zimbardo's side of the power relation in the Stanford case risked not only the psychological and moral integrity of prisoners, but his own and that of the guards as well. Having looked at all this, we can then draw some general lessons concerning human nature, and its relationship to institutions. But first a backgrounder or primer of sorts on institutions, with apologies for the many statements of the obvious:

Institutional life- humanity's accomplishment.

First, a brief sketch of institutions, a ubiquitous feature of human life: They are our creations. They are all around us, and contain us. Some are large, others local. They are nested, some containing others. All institutions have ends for which they exist. What is more, in some cases, there are certain broad purposes or institutional goals for which individual institutions serve as intermediate means. Trivial examples: A psychology laboratory exists to generate psychological research. That serves the broader goal of furthering scientific knowledge, adding to the objective body of knowledge that is the institution of science. The institution of marriage serves, among other things, the goal of societal perpetuation, by way of allowing for procreation in a generally favorable environment for child development. Marriage practices serve their containing cultures in this way and others.

Institutions enjoy an ontological status that is an interesting mix of the objective and subjective.³ Institutions are a bit like cities, in that they predate many if not all of the human beings who live within them. Also, like cities, they contain parts that are older, others more recent, and adapt themselves to changing environmental factors or technologies. We are born into most of them. We inherit them from our parents and more distant ancestors.

Institutions leave behind vast amounts of objects that record and perpetuate their existence. Documents and other physical objects constitute their continued existence. The objectivity of

³ For an interesting extended discussion of institutions, see John Searle (The Construction of Social Reality, and Making the Social World: The Structure of Human Civilization]. I rely heavily on his work.

institutional life can be attributed to the objectivity of what Karl Popper terms “world 3” objects.⁴ Institutions survive because we create physical objects that record the bodies of laws, rules, regulations, standard operating procedures and other ‘facts’ of our various institutions. The ‘objective’ life of institutions is also perpetuated by use of non-linguistic symbols, and clothing. In the SEP, for instance, uniforms and glasses were recognized symbols of the authority of law enforcement officials.

Yet, this objective status enjoyed by institutions, and the objects that help constitute institutional objectivity and history are dependent, for their status as institutional object, upon ‘subjective’ factors and behaviors such as beliefs, language use, communications, general acceptance, acquiescence and something like mutual ongoing ‘promising,’ expectation, or collective presupposition of certain reciprocal forms of behavior. Without all this, institutions would fail to exist. This all reflects the fact that we are beings of fairly complex psychology with an amazing ability to compass vast stretches of time and community into our day-to-day business and self-image. It is also safe to say that, of all the creatures on planet Earth, we are unique in being an institutional animal.

It is trivially true that institutions would cease to exist if human beings ceased to exist. What would be left over would be physical objects, such as books buildings and articles of clothing, but the institutions would have perished. Institutions can also cease to exist if human beings cease to take them seriously, or if they can no longer ‘read’ or understand the objective instantiations of the institutional life as having the status intended. Mayan civilization is no more, even though its artifacts survive. The institution of phrenology no longer exists, because no one takes it seriously.⁵

By taking part in institutions, human beings participate in, contribute to or become recorded in the objective aspects of institutional life. When we marry, this is recorded publically by documents. When researchers conduct experiments, these not only follow established guidelines of research methodology, but the results become part of the objective ‘world’ of science. Indeed, language, itself an institution, makes much if not all of this ‘living for posterity’ possible. Indeed, language makes institutions possible. This living beyond the biological horizon gives participation in institutional life allure because we are aware of our own mortality and would naturally like to overcome it.⁶ By way of the undeniable fact that they are ubiquitous,

⁴ Karl Popper, *Objective Knowledge: An Evolutionary Approach*, 1972

⁵ At least I think so. This claim will probably be falsified by someone pointing out that somewhere there is a small but determined band of phrenologists doggedly preserving their ‘discipline.’ Such, so I here, is the case with the “Flat Earth Society.”

⁶ For an interesting poetic, if somewhat obscure discussion of the effects of our awareness of mortality on the course of individual human lives; consider Heidegger’s discussion of “being-toward-death.” *Being and Time*.

institutions can exert tremendous influence on our behavior. In many ways we live in and through institutions. They provide meaning and purpose for much of what we do.

Institutions also make possible states of affairs, powers or accomplishments that would not otherwise be likely to occur. For instance, governments make possible roads, which make possible trans-continental travel in relatively short periods of time. The institutionalized system of incentivizing that is a monetary system makes possible, among other things, long term complex and unplanned cooperation in technological innovation. Institutions provide the automobiles as well as the roads.

A last obvious point and reminder; Institutions come into conflict with one another when their aims are incompatible, or when the values they embody are incompatible. It is also trivially true that individuals and institutions can come into conflict.

Institutional ends and the power of “ideology”.

In his writings concerning the Stanford Prison Experiment, Dr. Zimbardo, with enviable objectivity concerning his own behavior, argues that an overarching or driving ‘ideology’ purpose or end toward which an institution is oriented can exert a strong damping influence on the moral awareness of people within the institution.⁷ Often people believe that through their participation in an institution, they can bring about significant benefits for their local society, or perhaps for mankind, benefits the value of which it may be arguable, allow for behaviors not normally tolerated. This drive to create benefits for mankind is among the motivations for scientific research. We see, in the SPE, evidence that this motivation existed, not only for the scientist, Zimbardo, but for the guards, indeed for prisoners as well. There are several times when interviewed participants reported ‘pressing on’ in the interests of science.⁸

In fact, this is one telling area of disanalogy between our two cases. All of the people involved in the SEP; investigator, guards, and prisoners knew they were taking part in a scientific experiment, something the results of which we can surmise they believed may prove to be important for science as an institution, and society more generally. It appears to be the case that recognition of this fact, on the one hand, inhibited prisoners from rebelling as soon as they would have if they were not involved in a scientific experiment, and on the other hand, encouraged guards to persist in objectively abusive behavior longer than they would have in a non-experimental setting. What is more, Zimbardo admits that recognition of the scientific intent of his work shaped his own behavior, leading him to persist longer than he normally

⁷ It’s important to note here that Zimbardo uses the word “ideology” quite loosely. Where we would be inclined to reserve use of that word for very general political or moral values or principles, he uses it indiscriminately to refer to these as well as very specific purposes for which institutions may be created, such as generating scientific discovery, or preventing a specific sort of harm.

⁸ Lucifer Effect

would have in allowing the abuse that grew up in the Stanford basement. He was made aware of this by a reminder from a colleague (later to become his wife), who had not been involved in the day to day running of the experiment, but who had visited during a 'bathroom run.' He tells us that he slid into his role of superintendent, unawares, and, in hindsight, surmises that he did so primarily in the interests of pursuing the science, and lost his moral bearings in the process. His case is an example of the 'damping' influence of institutional life on moral awareness.

Contrast this with the situation at Hoa Lo: There was no overarching single ideology or end that the North Vietnamese and the American POWs were both involved in bringing to fruition. The North Vietnamese, steeped in Communist ideology, felt that the American POWs were an arm of a reactionary force of international capitalism, and consequently either in the grips of false consciousness, or deliberately working against the inevitable coming of the proletarian revolution and the eventual liberation of mankind from class conflicts, an event the North Vietnamese believed would usher in a new golden age. In either case, they felt justified in using torture and other methods to break false consciousness or punish willing criminals, and extract propaganda cooperation. Utilizing what is in essence a utilitarian calculus, they rationalized their actions as 'breaking a few eggs to make omelets.'⁹ The hastening of the world revolution morally justified their actions.

The American POWs did not share this world view, and most saw in Communism a threat to human well-being, a totalitarian system that lent itself all too easily to barbarism, a system and ideology that needed to be contained. In addition, they saw themselves as preserving an island of their own American culture, and its ideology, to put it somewhat paradoxically, an island of freedom in a setting of coercion and tyranny.

Zimbardo sees the "ideological" influence in his case, and has spent the better part of his career studying the morally dissociative effects of systems or institutions built to serve 'ideologies.' His focus is on the damping effects of what he terms 'the System' upon the 'cogs' of such institutions, that is, mid and lower level individuals wielding power over others in situations like his prison. Yet, his primary focus has been on those 'cogs', and on the architecture of institutions in relative power. Less attention has been focused on the prisoners, those with no power.

From within this theoretical stance he goes to great pains to draw parallels of varying levels of plausibility between his prison, 20th Century totalitarian regimes, events during the Vietnam

⁹ Attributed to New York Times correspondent Walter Duranty during the Stalin era.
<http://www.nytc.com/company/awards/statement.html>

era, and more recently events in “the so called war on terror”¹⁰. This is done in service of fashioning cautionary lessons for those in power; leaders and creators of criminal justice and national security related institutions. Indeed, he crafts ‘indictments’ of national leadership during the Bush administration, once again, of varying levels of plausibility, based upon his findings.¹¹

In short the indictments run that under the pressure of the threat of further terror attacks, post 9-11, the Bush administration created institutions or interrogation practices that in some way were ultimately responsible for the abuses at Abu Ghraib prison. Without arguing the merits of that indictment, we can nevertheless glean that Zimbardo counts prevention of significant harm as an overarching “ideology” or purpose of some institutions. The lesson, indeed a valuable lesson we can take away from Zimbardo’s lifetime of ruminations on the Stanford experiment, is that such overarching purposes can, once adopted, and absent effective oversight and engaged leadership, exert a damping effect on our moral awareness. This is surely worth keeping in mind.

To sum up this section on ideology, we see, in the case of his Stanford experiment, a situation where the overarching ideology or purpose of the consensually created institution seems to have led all parties to lose sight of the reality of the situation, and the moral status of their actions. Where Zimbardo may go too far is in his sometime seeming to claim that people embedded in such “systems” lose significant levels of moral responsibility concomitant to losing significant levels of freedom, being buffeted about by third person forces, in the guise of ‘the system.’ In fact, his own book militates against this semi-deterministic view, in his account of heroism which appears in the later chapters. Each case he sites is a person embedded in the same offending system as the supposed lost ones. These heroic individuals, apparently quite ‘ordinary’ people like Hugh Thompson¹², were able to exercise freedom, and take responsibility to change the system or circumvent its actions. The later chapters of Zimbardo’s book are devoted to such folks, and in particular, Army reservist Joe Darby, who blew the whistle on abuses at Abu Ghraib. It would seem cases like this show that ‘systems’ need not be as smothering as Zimbardo sometimes tends to paint them.¹³ Additionally, he seems very willing to absolve cogs of moral responsibility, due to the damping effects of the pressures of being

¹⁰ Zimbardo’s use of scare quotes. See <http://www.zimbardo.com/lucifer.html> He considers the war on terror to be largely overblown “fear mongering”, thus the scare quotes. This writer does not concur. The threat of Islamism is real, and needs to be soberly dealt with.

¹¹ *Lucifer*, chapter 15

¹² The man who stopped the Mai Lai massacre.

¹³ An interesting example of this tension in Zimbardo’s thought can be seen in his response to the first question posed by a Midshipman in the question and answer period following Dr. Zimbardo’s presentation at the Naval Academy in 2008. The entire video can be accessed here. That Q&A begins in part 4: <http://www.usna.edu/Ethics/publications.htm>

contained in 'systems,' yet seems to accord to those higher up, those that create or head systems, greater levels of autonomy and moral responsibility, despite the fact that they too are subject to pressures, pressures that theoretically at least, could also be used to absolve them of significant blame. While it may be said in response to this that we cannot afford to allow those with such power too much laxity, due to the likely consequence that such laxity would be taken as precedent and taken advantage of, this does not directly address this *theoretical* objection. Lastly, even if we are to accept this deterrence based 'bad precedent' argument against laxity when it comes to holding people morally and legally responsible, it seems then, that we should apply it also to those cogs he seems more willing to excuse, for that exact same reason; deterrence. For if future 'cogs' are aware of earlier 'cogs' having been let off due to the extenuating circumstance of having been perceived or described as helpless cogs in a 'system' that damped moral responsibility, then, that perception too would seem to be encouraging of future bad behavior.

The power of tradition

As stated before, it is a truism that institutions outlast the lives of those individuals that inhabit and partially constitute them at any given time. Institutions, like human beings, have lives and histories. Institutions have records of accomplishment, and failure. Institutions are more or less rationally constructed and maintained. Institutions are long-lived things temporarily constituted, administered and guided by those now residing within them. We normally have a concern to pass them to posterity undamaged and improved.

All of this does work toward giving institutions not only a certain value for human societies, but a certain allure, a certain ability to convey status. They have value because they convey benefits. They convey status because they persist, and allow one to become a part of that persisting structure, a part of institutional history, an agent of their benefits. They allow one to outlive biological necessity.

In the case of academic and scientific institutions, there is a status conferred by way of earning a Ph.D. This status allows one to enter a select society, and contribute to its growth of objective knowledge. One can benefit society, and be remembered for it. Similarly, to become an officer in the Navy is to have met some rigorous standards of training and education. It allows one to enter into a select group, and wield very important powers in the service of the greater institution of one's nation, and humanity. Once again, posterity will record and remember.

When one enters an institution, one is given access to elements of its structure; one is allowed to add to its historical accretion. One can 'give back' by taking part. One is also allowed to benefit from membership. This engenders a sense of obligation or loyalty to that institution. This is a double edged sword. It leads to many positive things. But, Zimbardo argues, on the

basis of a reflection on his own case, that this sense of obligation and loyalty toward an institution can lead to the sort of lapses we see in the SPE. Subscription to the ends or ideology that buttress an institution, giving it its *raison d'etre*, can induce lapses into what would otherwise be considered unethical behavior. Subscription to the communist ideology played such a role in the North Vietnamese POW prison systems. Zimbardo self reports similar tendencies, as a committed member of the scientific community.

Yet, as we focus on the inmate populations, we see that a lack of tradition on the one hand and a full and healthy tradition on the other hand, may help account for the quite different results in our two inmate populations, and, most encouragingly, may also account for the heightened moral awareness evidenced by the prisoners at Hoa Lo. The prisoners in Stanford did not come into the situation with anything like the personal status, institutional framework, loyalty, sense of obligation or impetus to contribute that even Dr. Zimbardo came to his experiment with. They were not scientist, nor were they intent on trying to add to the edifice of world 3 knowledge maintained by the scientific community.

They were all simply twenty-something college boys, leading their separate lives, pursuing differing degrees, no doubt with varying degrees of seriousness. There was no strong drive to contribute bricks to the edifice of objective knowledge. There was an awareness that results might add to the edifice, but no strong loyalty to the institution, to science, to psychology.

Contrast that with the inmates at Hoa Lo. They were, by and large U.S. military officers, and aviators. As such, they had high degrees of loyalty not only to the United States, its constitution, and the U.S. Navy and Air Force, but to the culture that gave rise to them. An essential core of the ethos of the U.S. military is to put the welfare of the seminal founding document of the United States as an institution before concern for self. Indeed, for Stockdale and his men, concern for welfare of fellow inmates became the local embodiment of the sort of concern and service for the welfare of the nation that the U.S. military oath expects. These men fell quite naturally into the habit of putting the welfare of their comrades above their own individual well being. This is assuredly at least partially attributable the force of this noble tradition. But, as we will see, it was also, ironically, a result of intensive efforts, on the part of the North Vietnamese, toward utter isolation of the prisoners from one another. That very pressure elicited the outreach, the desire to communicate, and the altruism, even though it was designed to cause prisoners to fall back into narrow self-interested behavior patterns.¹⁴

¹⁴ Evidently, such pressure had worked, in at least some cases during the Korean War. In response, the six article Code of Conduct was created. The necessity for such a code became apparent in light of disturbing reports of American POWs cooperating with the enemy, turning on each other, fighting for food, and leaving comrades exposed to the elements at North Korean prison camps. <http://www.pacaf.af.mil/news/story.asp?id=123241828>

So, while Admiral Stockdale describes his efforts, and those of the other senior leadership in Hoa Lo as building a civilization from the base, it is more accurate to say, as he also does, that it was an attempt to transplant a previously existing 'civilization' or institution into hostile territory, create and maintain a colony, a lonely outpost of American civilization in the midst of an avowedly hostile and powerful ideologically contrary civilization.¹⁵ Tellingly, he describes the code of conduct as codifying for POWs an interpretation of their situation according to which they had not been removed and isolated from the war, but indeed were still firmly in the teeth of the war effort, fighting behind enemy lines. This code allowed them to see the machinations of the North Vietnamese for what they were, attempts to use the men as propaganda weapons aimed squarely at the home front, and U.S. media. The officers at Hoa Lo saw themselves as an integral part of the ongoing efforts of their nation and their military service institutions. This gave them much needed psychological purchase in the face of barbarity, and prevented the sort of moral fall that occurred in Korean prisons.

Yet, we see too, in the common tradition of 'service to country' shared by the North Vietnamese and American officers, grounds for a sort of mutual respect or understanding. Both sets of officers were cognizant of what the others consider to be their duties. They recognized the obvious conflicts of ends, and that duty required that attempts must be made to thwart their conflicting ends. This was something both sides respected.

In short, the power of tradition, in the case of the prisoners at Hoa Lo was a determinative factor in their success in resisting the wiles of the North Vietnamese over the course of 7 to 8 years. What is more, we see an uneasy tension in our two cases. On the one hand a reverence and loyalty to scientific tradition led to moral blindness. On the other hand, reverence for tradition and loyalty led to moral excellence. (Just as obviously, a reverence for tradition led the North Vietnamese to atrocity.) We have to ask why there is a moral difference.

In short, we will see that it is the architecture of the institutions and quality of leadership that count. The scope and reach of the rules that dictate roles is vital. This in turn, is dependent upon those that create that architecture, and interpret it, the leaders of the institution. In the one case, there was little architecture, in the pursuit of the institutional or traditional end, while in the other, a very carefully crafted blueprint, served as a moral bearing for those within the institution as they pursued its end. For lack of this sort of guidance, literally lived 'from the front' by those that lead, and thus impose it, institutions run the risk of leading those within them into moral failure. Engaged leadership, and relatively simple, clear and specific direction via rules governing roles within institutions can preserve moral and psychological integrity. This is far too brief. Let me expand:

¹⁵ Stockdale, Master of my Fate II

Rules, Roles, Leadership and lawgiving

Once one buys into an institution, decides to play a role in an institution, or finds oneself in a previously existing institution, he/she expects or hopes that there is information as to the roles in the institution and attendant expectations. Generally speaking, the more specific information there is about the roles, the more confident one can be in making judgments as to how successfully one is carrying out the roles. Too little information will fail to clearly delineate roles, nor allow for assessment of success and failure in playing roles. It will create an epistemological unease, which will, if left untended, undo the institution. For, moral self-evaluation is dependent upon such clarity. Lack of clarity give purchase to uncertainty, guilt, fear and ultimately, self loathing.

Now, the sort of information we are concerned with here are rules, laws, and standard operating procedures. These create roles or 'places' if you will, within institutions. Human beings can be cognizant of these places within institutions because the rules that create them are cultural artifacts, what we have called "world 3 objects," objects constructed via language, a basic institution. In virtue of these, we can read or communicate about roles, and determine our place in an institution relative to them. We can read or communicate about the roles in an institution, and decide whether we want to join that institution. We can read or communicate about roles, and form judgments as to our effectiveness in any roles we have adopted, or have involuntarily received.

There is a threshold of richness in such guiding information above which we can say that a viable institutional framework exists. Below that threshold, however, we would say that the preconditions of a viable institution are insufficiently met. Below that threshold, the institution is set up to fail.¹⁶ We see instances of this relationship between informative content and institutional health, both within the story of the SPE and within Admiral Stockdale's narrative of life at Hoa Lo. We look at each in turn:

Zimbardo discovered that a mere assigning of roles did little toward constituting a healthy institution. He gave minimal instructions to the students who would play the guards. This gave them wide latitude to improvise. He also did not actively oversee, and, in particular, when the guards felt they were not being monitored, during night shifts, the egregious behavior tended to increase.

¹⁶ There also seems to an upper limit, of informational content beyond which, paradoxically, institutions stifle the life of those within, inhibiting creativity initiative and action, thus killing the institution's ability to adapt to its environment. The Soviet system was a case in point. The hyper-regulatory modern bureaucratic state creates epistemological difficulties that belie the old Aristotelian saw that 'ignorance of the law is no excuse,' as well as making for an overcautious citizenry when it comes to risk taking and initiative.

Even less guidance was given to the inmates. They were told to behave as they assumed inmates would behave. Zimbardo postulates a reliance on the fallback of cultural representations of prisoners, things encountered in books, movies or television shows. In both cases, individuals were left to improvise more often than not. Neither the guards nor the inmates were given a body of rules to reference, that delineated bounds of acceptable behavior. As a result, the guards began to improvise these. The prisoners didn't create such bodies of rules for themselves. It seems we can say the greater freedom or latitude given the guards actually worked against the goal of creating a functional prison that would last the full two weeks. The prisoners, it would seem, exhibited the degree of emotional and psychological stress they did, because they failed to form a governing institutional entity and body of rules from within which they could resist, provide individual guidance or bearings for acceptable behavior, and that would have provided them some psychological or emotional armor. This created fertile soil for guilt as well.

Similarly, we hear from Admiral Stockdale: When he first arrived at the Hanoi Hilton, in "new guy village", he was told in no uncertain terms by the POWs already in Hoa Lo, that it was simply unacceptable to give vague and general guidance, leaving it up to each individual to do the best he could to follow the existing code of conduct, one created after the Korean War.¹⁷ Instead, he had to carefully consider the end for which that code was created, the realities and limitations of human psychology and physiology, and crafted a code of conduct that was at once more specific than the official, and sufficiently detailed to give the sort of guidance craved by the inmates of Hoa Lo. His code had, built in, a clear recognition of the limits of human endurance, and directions as to what could and could not be revealed under the pain of torture. Yet, the directions were not overbearing, or weighed down with detail, leaving some room for reasonable interpretation. There was an easily memorized acronym, created for ease of transmission via the tap code. It governed all interaction with the North Vietnamese.¹⁸ The system was by and large successful. Without the system, Stockdale argues that guilt would have played upon the conscience of individual prisoners, to the detriment of the POWs and the goals of the code of conduct. For, each POW was given an initial round of torture using "the ropes" maintained until he volunteered information, and was then placed in solitary.¹⁹ Each man felt

¹⁷ Stockdale, *Master of My Fate II*

¹⁸ The acronym: "BACKUS" don't **B**ow in public; stay off the **A**ir; admit no **C**rimes; never **K**iss them goodbye, and **U**nity over **S**elf, that is; never negotiate for self but for all.

¹⁹ The barbarism of 'the ropes' cannot be overstated. The standard procedure was to place the prisoner sitting on the floor of the interrogation room, pull his arms behind. They were then tightly bound together above the elbow so as to cut off circulation. Then, the arms were pulled up and rotated away from the back, toward the head against the normal direction of rotation while the torturer would place his foot on the upper part of the prisoner's back, and force his upper body down between his legs, which were extended together in front. The prisoner would experience incredible pain, arms would often dislocate, tendons snap, and prisoners would be unable to breath. Many would vomit. Panic was the inevitable result.

guilt for having “broken”. With the system that Stockdale put in place, these vulnerable prisoners could delineate acceptable from unacceptable behaviors consequent to having been ‘broken’²⁰ by torture. They not only needed this structure, but moral reassurance.

This is the essence of effective and engaged institutional leadership. A familiarity with the overall goal of the institution in question, and a knack for creating simple, easy to comprehend, yet informative and relatively rich bodies of law, rules or guidelines the following of which is by and large possible for those within, and which will also allow for attainment of the goals of the institution without stifling the life of the institution or the individuals that inhabit it.

It is very interesting, and telling that the men at Hoa Lo, when purposively isolated from one another, and put under the direst of threat for attempted communications, nevertheless were compelled to seek out their neighbors, almost as tap roots seek water, while also insisting on the sort of order and body of rules here described. It is interesting that more than one described the thrill and deep sense of satisfaction in creating a civilization in the midst of barbarity.

One wants to say that this was a spontaneous expression of something deep within the human psyche, an expression of man as *Homo moralis institutionalis*, as intimated before. Stockdale himself uses such imagery, comparing the circumstance to the hermetic tradition. He provides this striking analogy in a commencement address.²¹ He likens the situation to that of the process of growing laboratory created diamond. One can place an ordinary lump of coal in a confined space, expose it to extreme pressure and heat, and then rapidly cool it. The end result is a diamond, an object that has instantiated a potential for beauty or excellence that was contained in the coal. Under those conditions the diamond takes up an organization or structure that lays dormant within the carbon itself.

Similarly, one can argue, pressures bring forth latent civilizing and ennobling aspects of human nature. Hoa Lo is a case in point, even if SPE was not. Extending the hermetic analogy, Stockdale argues that the set of conditions that produces the morally exemplary result is not something that naturally obtains of some necessity, but takes the careful and judicious adjustments of variables. For him, the crucial variables are recognition of the limits of human

²⁰ The term here is disputed by some POWs. I believe this is because it carries with it an implication that the breakage, once accomplished was permanent. What is more accurate to say is that torture sessions eventually succeeded in creating through inflicting immense pain and physical duress, a sense of imminent death, that, not surprisingly, would yield a panicked reaction, complying with demands for information or sometimes cooperation in propaganda efforts. Key to the rules promulgated by the POW command structure was recognition of the inevitability of this panic reaction, and rules governing the nature of allowable information or cooperation. It was also made clear, the amount of torture that must be endured before one “broke.” Superhuman perfection was not expected, nor codified. Yet, one had to endure torture, for to too readily acquiesce would be to the detriment of the group and the war effort.

²¹ Stockdale, “The Melting Experience”

nature, and generation of appropriately calibrated rules and institutions thereby constituted. Institutions that are better adjusted to human nature and limitations will survive and flourish, those that are less well calibrated will fail, and occupy the dustbin of history.

But, can we safely make such broad inferences? Countering the general nature of this claim, one could cite, (as we did in the previous section on tradition), the fact that the men at Hoa Lo were not truly representative of raw human nature. Before they were placed in the hellish hermetic space of Hoa Lo, they were U.S. military officers. Their behavior was more a manifestation of that culture than a manifestation of some innate impulse toward morality and institutional life as a basic aspect of human nature.

Indeed, other examples exist, of uprising organized by military men in similar hellish circumstances. During the Gulag uprisings of 1952 to 1954, we see repeated instances of such events. Of particular interest is the Kengir uprising of 1954, which saw quite literally the flowering of a small temporary culture, complete with various governmental institutions, indigenous religious ceremonies, marriages, and propaganda efforts.²²

Yet, we should not be so quick to dismiss the postulation that these examples evidence something basic to human nature, something not attributable merely to the presence of individuals with a background in military culture. At Kengir, prisoners of many backgrounds took part, and in fact initiated the rebellion before placing a military man at the head of the temporary government. There are similar cases of spontaneous relatively successful governing institutions in threatened populations that were not by and large military. One is reminded of the relatively complex and healthy governing structure created by the prison population at Sobibor concentration camp.²³ One person was of military background, and indeed did help lead a successful uprising. However, the majority of the governing inmates were not military. Indeed, women played vital roles in that body. The governing structure actually predated the arrival of the one leader of military background. Additionally, we must not forget that the Warsaw uprising involved civilians of both genders, and all ages. Indeed, even in the Kengir uprising, many of the leading prisoners were civilians with no military background. Also, in times of disaster, institutions or organizations are spontaneously created by civilian populations that are cut off from aid.

Indeed, more broadly construed, pressure or stress in the form of environmental or climatic challenge was probably an impetus toward the formation of larger civilizations as man's technical prowess, and increased concentrations of populations necessitated cooperation in the business of survival. Military organizations grew up with civilization.

²² <http://english.turkcebilgi.com/Kengir+Uprising>

²³ <http://www.ushmm.org/wlc/en/article.php?ModuleId=10005192>

To round out this section, we see that Hoa Lo had what SPE lacked. At SPE there was a lack of effective and engaged institutional leadership and specific architecture. While the “ideology” or goals of the institution were laudable (contribution to scientific knowledge), the lack of these two features increased the likelihood of the morally problematic results. Neither the guards nor the prisoners had a sufficiently robust institutional framework against which to measure their behavior. There was an ill defined immediate set of directions to ‘act the part’ and little else. On the other hand, the Hoa Lo prisoners carefully crafted realistic expectations in the form of a small set of easily transmitted rules all prisoners were expected to follow. These rules took as their basic anchor the spirit and intentions of the Code of Conduct. They were tempered with a realistic assessment of the physical and psychological limitations of human beings. There was great effort put toward dissemination of these rules and continual communications between isolated prisoners. There was a clear chain of command and equally clear explanation of the role of the prisoners as fighters ‘behind the lines.’ This savvy improvisation upon the established norm of the Code of Conduct was contributory to the prisoners’ success in resisting the coercion of the North Vietnamese. Lack of this careful sort of institutional architecture and leadership was contributory to the psychological and moral breakdown of the prisoners (and others) in the SPE.

What does this say about human nature and morality?

Aristotle noted that man is the rational social animal, the animal of the polis. Kant argues that man is the animal that is capable of grasping and formulating universal laws (both scientific and moral). Popper describes us as inhabiting three worlds, one of which we create. It is populated with among other things, laws, rules and institutions. What is essential to all three views is that we are creatures that produce and live within institutions, more or less complex structures of mutual promise, and action. Just as much as coral produce reefs or bees produce hives and colonies, we produce and live within institutions. What sets us apart, though, is that our ‘reefs’ our institutions exist only because we mutually acquiesce or agree to limit and adjust our behaviors and implicitly or explicitly promise to act within, maintain or build upon previously existing institutions. We live in a very thick temporal dimension. Furthermore, institutions can only exist because we humans can formulate goals or ends, in light of which the institutions find their reason for existence. Institutions can only survive if the individuals that populate them survive. Institutions that stifle that life will suffer decline and eventually die. Institutions that strike an appropriate balance between the lives of the individuals contained therein and the institutional life will tend to flourish. We can also argue, based upon relative levels of survivability, that some institutional frameworks are better than others, that there is a ‘way of life’ for which humans are suited, whether by design or chance. It is the job of leaders to craft and helm institutions that serve that end.

It was claimed earlier, that language is a basic institution, without which no other institution could exist. Institutions are essentially fabrics of mutual promising. Promises cannot be executed except within language. Elemental to the possibility of a language, is a very basic mutual implicit promise (or collective intention to use a phrase of John Searle's)²⁴ between language users. In order to effectively communicate I must agree to use words consistently with others' usages. I cannot just decide to assign novel meanings to words without destroying the very possibility of communicating, and indeed language itself. So, implicit in the basic human institution is something very like what Kant would call a universal moral imperative. One can postulate that there is a system of such basic moral imperatives. We can call this 'the basic moral system' or 'basic moral institution.' Most, if not all elements of this basic system can be seen as reciprocal promises (implicit or explicit) to behave toward other human beings in certain ways. One can take the measure of an institution, and indeed individual human beings by assessing the extent to which they honor these collective intentions or mutual implicit promises.

Now, if, as seems entirely plausible, it is natural for human beings to develop and use language, then, it is natural for them to develop the aspects of the basic moral system or institution that it requires. Now, if language is the basic institution, and is dependent on the basic moral institution, then it should then not be at all surprising that it is natural for language users to develop the sorts of second order institutions that depend on language and the basic moral institution for their existence, if conditions are right, that is; if conditions dictate creativity by way of bringing about or instantiating a 'hermetic' setting.

So, yes, indeed there are diamonds innate in our human lumps of coal. This is not to say that we will always end up with diamonds. No, that should no more be expected than that most parameter setting used in a lab experiment will prod coal to produce gem stones. Be that as it may, judicious and rational exploration of the parameters of human being, and institutional design, will allow us to approach that gem, balancing the interests, rights and responsibilities of the individual against the interests, rights and responsibilities of our institutions, keeping both 'healthy, wealthy and wise'. That is the task of ethical leadership.

²⁴ John Searle, *The Construction of Social Reality*. See also "[What is an institution?](#)" *Journal of Institutional Economics* (2005), 1: 1, 1–22