HUL 253: Moral Literacy and Moral Choices
Term Paper

Morality of Unsettling Interventions

Abhishek Kumar Sharma, Sahil Loomba and Sajal Gupta

Author Contributions:

- Abhishek Kumar Sharma has contributed to Sections 1, 4.2 and 5.
- Sahil Loomba has contributed to Sections 2 and 3, and overall structuring of the paper.
- Sajal Gupta has contributed to Section 4.1.

Contents

1 What is an Unsettling Intervention?

Unsettlement is quite commonplace, and can be found at any street corner. All it needs is one individual to gain some information that conflicts with his *fundamental beliefs*. A *belief*[1] is a claim which is mentally accepted without a proof. Since it is not supported with a proof, any belief can potentially be eroded with a clearly conflicting proof against it. Production of such a proof against a belief is an *intervention*, for it asks of change in the present state of beliefs of an individual. An intervention becomes *unsettling* when the beliefs in concern are *fundamental*. A belief is fundamental when it forms the basis of a great degree of *synthetic* knowledge a person possesses. The fundamentality is greater when a given belief forms the basis of a greater quantity of proof-supported acquired and/or inferred knowledge.

A popular example of unsettlement is that of Neo from the cyberpunk cult classic *The Matrix* (The Matrix, 1999). His entire synthetic knowledge about his life is challenged by Morpheus who brings him out of the simulated reality, eroding his conception of the erstwhile believed reality. This is also a characteristic of situations where one may cause unsettlement among people who unreasonably (and unknowingly) associate a meaning to otherwise absurd objects and activities by arguing them out of their ignorance. It is quite equivalent to letting a happy Sisyphus realise the absurdity (Camus, 1955) in his regime of working the boulder up the hill again and again.

Unsettlement is associated with a repulsion from it, often taking a form of fear. Transition of a minor to a major (in Kantian terms) is one such kind, where, as discussed by Kant himself (Kant, 1784), there is a fear of taking even the first few steps as if they were uncertain leaps.

---

[1] It is interesting to look into the etymology of the word *belief*, which is related to *lief*, which has a connotation of love, of coherence in thought, means the thing with which we think along. Hence beliefs can be seen as our first companions in thought.

2 Abstraction of the Mind

From time immemorial, human cognition has been an active area of study in different fields of the natural and social sciences. In its essence, cognition is a set of mental faculties responsible for cultivation and creation of knowledge. Starting with perception, to attention, to computation, to reasoning, to memory and going right up to action: linguists, neuroscientists, psychologists, computer scientists and philosophers have all tried to understand and model human cognition.

Celebrated German philosopher Immanuel Kant (1724-1804) has had a lot of influence on modern cognitive sciences, with his views on mind and consciousness, presented in his *Critique of Pure Reason* (1781) (Christman, 2015). The central thought of his model was that the mind is a complex set of functions, requiring both concepts and intuition, in the form of *synthesis*. Synthesis is defined as "the act of putting different representations together". Kant believed in the existence of *a priori* knowledge, which preceded, rather than follow, experience. (Since a certain amount of knowledge is required to comprehend the experience itself). We can call this knowledge as *genetic knowledge*: that which has been passed down to an individual from his ancestors through the process of evolution. Owing to our faculty of sensibility, we possess concepts and intuitions, such as that of *space* and *time*, which allow us to gain and make sense of this *a priori* knowledge. For example, the statement "2 = 2" is an analytic judgement. On the other hand, when we perceive input through our sense organs, we cultivate or *synthesise* more knowledge, also called *a posteriori* knowledge. Examples of such synthetic judgements are "all human beings have two legs", since the attribute of having two legs is not within the concept of being human. Such synthesis by each individual human can be termed as *epigenetic knowledge*: the one gained by a human from personal experiences, over space and time. Besides personal experience, there is also knowledge which one borrows from the extracorporeal milieu of human culture, history and heritage. Certainly, one could place religion, cultural systems and even human scientific work under this third type of knowledge: the *epiphylogenetic*.

This description leads us to the question of *how* synthetic knowledge is absorbed by an individual, that is, how does synthesis take place? One popular learning model is that of *Connectionism*. This model studies human cognition mathematically, as a unit of densely connected neurons, with *synapses* forming the connections themselves. Similar to the arrangement of neurons in a human brain, there are weights corresponding to every synapse, which when modified by firing neurons hold the ability to learn a variety of tasks common for human agents, like language recognition. This idea of weight-setting by repeated firing between correlated neurons is called *Hebbian Learning*. When mapped to the brain, this behaviour manifests in the form of modules expressing specialised functions: say the occipital lobe, which has developed the ability to process vision perception.

Akin to this behaviour of formation of strong and robust connections on consistent training of neurons, we present the following model of human learning. In infancy, a human baby is born with some ground genetic knowledge, concepts and intuition. This analytic *a priori* is then followed by the process of "putting representations together": synthesis. Over time, as packets of information are received by this human, they add weights to certain *connections*, thus leading to strengthening of such experiences and knowledge. Over a longer stretch of time, certain ideas like "The sun will rise tomorrow", or "This woman is my mother" have been trained so repeatedly over the mind's network, that they have formed an almost *inert block* of epigenetic knowledge. Whereas certain packets remain loosely bound to one another, such as "I do not like dancing", possessing connections that are subject to change, reorganisation and up/downgradation by future experiences-to-come. Moreover, there also exist certain accepted tenets, such as those of religion and science, which form a strong

foundation of our epiphylogenetic knowledge: over years of education, religious upbringing, and so on.

2.1   Variant and Invariant Knowledge

We term beliefs which hold an invariance to new experiences, that is, those which remain identical through space and time, under *invariant knowledge*. These beliefs are inert: resistive to change, having been strengthened by repeated connection-weighting. On the flipside, we term beliefs which are lose identity through space and time under *variant knowledge*. These beliefs are moving: subject to change under the duress of new information (experiences). What separates the two is the *line of inertia*: a boundary on the left of which inertia is very high, and on the right of which it is very low. What governs the behaviour of this line, within the abstraction of the mind?
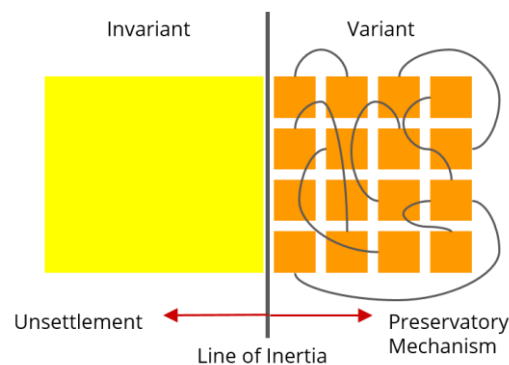

Figure 1: A schematic of abstraction of the mind

2.2   Unsettling and Preservative Forces

There are two forces which work in opposition to one another, over the line of inertia: unsettling and preservative forces. The action of the former is evident in when a new block of information is so perverse that it uproots certain invariant beliefs, causing the shift of the line of inertia to the left. The latter force is a product of biological evolution: a natural drive which supports preservation of knowledge. The force's motivation is evident in not only that it keeps us sane, in not having to restart from ground zero every now and again, but also in the way it minimises computational cost of decision making.

Usually, these opposing forces balance each other and keep the entire system in equilibrium. But when an intervention occurs, the balance can be disturbed, and if the preservatory mechanism of the one intervened is weaker than the power of unsettlement of the intervention, then the line of inertia shifts to the left, causing a *de-inertialisation* of the uprooted belief.
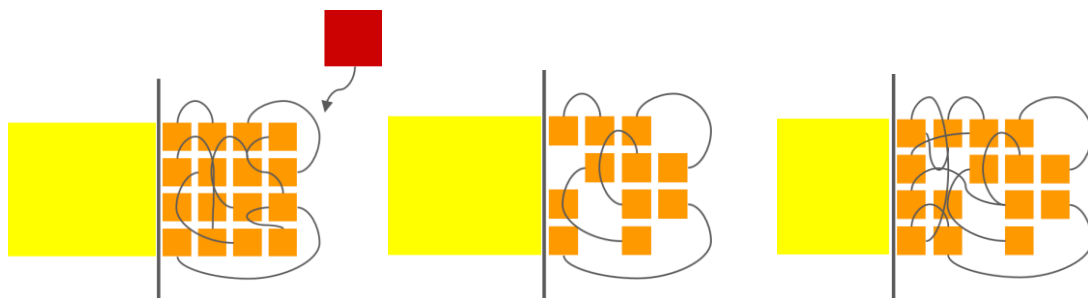

Figure 2: (a) Incoming information, termed as an Intervention (b) A non-unsettling intervention, which disturbs only the variant (moving) beliefs (c) An unsettling intervention, which disturbs even the invariant (inert) beliefs, causing de-inertialisation

3   On Autonomy and Paternalism

Autonomy is considered to be a vital attribute of an individual: the capacity of a person to make decisions for himself, in accordance with his own beliefs, through self-employed reason, uncorrupted by external forces. Etymologically, autonomy refers to self-law (*autos →* self, *nomos →* law). In Kantian philosophy, as explained in his essay (Kant, 1784), being autonomous would mean the ability to put the moral law upon oneself, without external manipulation. Does an intervention, then, convert individuals from being autonomous to being heteronomous?

Before answering this question, it is also important to discuss the social practice of *Paternalism*. It refers to the interference into an individual's condition, usually contrary to their own will, under the belief that the interfered would be better off in this situation. Instances of paternalism are evident in interpersonal relations, as well as in a political setting, say when the Government decides to provide in-kind benefits rather than in-cash benefits to the poor. Traditional Kantian moral philosophy seems to put up a straight objection to the idea of paternalism: a deep intervention for the benefit of the individual. Since it seems to infringe on the autonomy of the individual. However, a look at the underlying secondary layer of this argument would lead to a different conclusion: such paternalism is necessary to induce autonomy from heteronomy. Kant's description of *minority*, "[the] inability to make use of one's own understanding without direction from another, [...] is self-incurred" (Kant, 1784), paints a picture of immense inertia on the part of the minor. The person enjoys certain *nominal* autonomy, with the self-law being governed by the comforts of being a minor, however he lacks *real* autonomy, which is the ability to put law upon oneself.

An unsettling intervention therefore, much like paternalism and the public use of reason, might temporarily infringe upon an individual's *nominal* autonomy, but it permanently pushes him towards a purer form of *real* autonomy.

It would be interesting to draw over here an analogy to the idea of *social contract*, as put forward by the French Philosopher Jean-Jacques Rousseau (1712-1778). A compression of Rousseau's argument would go a bit like: under the assumption of social interdependence, and citizens possessing common knowledge that they share a conception of the common good, we can have a society where everybody is equally free if everybody puts their power in the direction of the *general will*. On a cursory read, this statement seems paradoxical, since "the standard to which will must conform is itself non-voluntaristic." (Cohen, 1986). However, as Cohen points out using self-effacing Hobbesianism, "individuals are initially motivated solely by their own long-term interests in protecting their person and goods. There comes a point at which they recognise that mutually beneficial coordination is possible, and that it is in their own long-term self-interest to cooperate on mutually beneficial terms if they can rely on the cooperation of others." Thus, the nominally-autonomous (akin to a *minor*) may feel that their self-interest of self-incurred complacency is of paramount importance, but when the really-autonomous (akin to a *major*) unsettles the former, the overall good has the potential to be realised[2].

---

[2] Note that although there is no direct correspondence of an *unsettler* to a major and an *unsettled* to a minor, there exists a strong implicit correlation between the two concepts.

4 Agents in the Moral Problem

4.1 The *Second Person* Moral Problem

The S*econd Person* moral problem is essentially the following question: "Is it right to unsettle someone else's invariant beliefs?" Certainly, there cannot be a simple yes/no answer to this question, else it wouldn't be a dilemma. However, let's begin by shedding some light on both sides of the argument using some realistic case settings.

Suppose there is a small community which is primarily comprised of uneducated people. Some superstitions like Sati and child marriage are prevalent in this community. Now, if a social activist were to try and convince these people to let go of these beliefs by giving some arguments which are logically sound, we are tempted to say that this unsettlement is morally justified. We are causing people to give up their traditions (an epiphylogenetic belief), which will be very hard for them to do as they have been following them for centuries: these practices hold sentimental and cultural value. However, this change would probably lead to the improvement in lives of the people: an increase in value of life.

Consider the case of a young child leading a happy life with his parents. He is not aware of the fact that he is not their biological son, and was adopted (contrary to his strong epigenetic belief). There is a moral dilemma which the parents face. Should they tell him that he is not their naturally born son or withhold this information forever? Honesty is considered as one of the most vital moral maxims, but is the unsettlement that would ensue within the child morally acceptable? This is a very difficult question to answer. Discovering this fact might cause immense emotional distress to the young child. However, concealment of the truth also appears an immoral outcome. In this debate of truth versus good, which side should be weighed higher?

Another interesting and pertinent clash which could be generated is that between a religious and an atheistic individual. Plenty of moral questions can be put up by the atheist: should I go and question the religious beliefs of every other person? Is it morally justified of me trying to logically prove that God doesn't exist? By doing this, what *good* can I achieve? A lot of people derive motivation, hope and strength from the idea of God. Is this unsettlement morally appropriate?

Eventually, questions on our existence as a form of *life* (genetic beliefs) can also be raised, something as unsettling as "why do we exist?", or "what is the meaning of life?" Clearly, some of these interventions need not even come from an external source: they can be entirely self-synthesised. Does that render the *Second Person* moral problem useless? Perhaps it makes more sense then, to ask the alternate question of "should I *let* myself get settled?"

4.2 The *First Person* Moral Problem

The intervened person has a range of alternative behaviours he can have once intervened in a potentially unsettling fashion – he has to orient himself towards the new information that is conflicting with his existing beliefs.

One way is to have an ignorant attitude of self-deception (Hamlyn, 1985), where the person allows his original belief to be mentally accepted as true while knowing that the intervention has produced a proof against it. Essentially, a statement and its negation both coexist within the same mental content. This might be a voluntary (conscious) decision, or an involuntary (subconscious) one. This leads to an inconsistent system of concepts, hinders epistemic goals, if any.

One can thematise the act of unsettlement in a red pill vs. blue pill scenario as showcased in The Matrix. The pills represent whether or not the one who is intervened accepts

the intervention and updates his beliefs by embracing the newfound knowledge, indifferent to the fact that these new beliefs might unsettle his present world-view drastically, and whether it will be better or not is not certain. If he does so, then he is said to choose the red pill. On the other hand, the intervened is said to choose the blue pill if he refuses to receive the intervention, or receives it but commits self-deception (conscious or unconscious) and does not end up getting any change in mental content, effectively not getting any unsettlement. Hence, in the act of intervention the intervener (like Morpheus) presents the two pills to the prospective intervened (compared to Neo). The second person moral problem can be thus seen as decision of Morpheus's choice to present the pills or not. Similarly, the first person moral problem comes down to asking that upon being presented with the pills, which pill Neo takes.

From the perspective of the intervened, the problem is reducing to a choice between a state which is certain (which can potentially be an illusory one) versus an uncertain state (which is more real and truer). Further, this evidently reduces to whether the intervened would choose acquirement of truth over continuation of familiar conditions.

Why may one prefer continuation of familiar conditions? There can be several reasons for doing so. The simplest explanation would consist of saying that preservation allows for maintenance of integrity of knowledge that confers certain survival value, which makes its preservation an evolutionarily favourable trait which would amplify and preserve itself over time while enhancing the longevity of the host. We may have positive driving factors such as pleasure and happiness and negative ones such as fear and anxiety driving this preservatory instinct. The same can be said for such drives at an epiphylogenetic level such as cultural acceptance of a trait or behaviour as good.

In contrast, what could be the reasons for pursuing a truer world picture? One of the reasons could be that one considers epistemic goals concerning existence as the only ones which are worth pursuing. Taking up such a position is sourced from the idea that in the face of absurdity in existence (Camus, 1955), the only meaningful goal is the epistemic one to understand existence, and to do so efficiently, one must remain in pursuit of truth in a consistent fashion, asking one to be open for unsettlement.

5   Prescription

The second person problem and the first person problem as described above are interlinked in the way that if the intervened is open to unsettlement then the intervener should face no dilemma in intervention. On the other hand, if the first person is not open to unsettlement, then it is prescriptive for the intervener that his intervention is not going to be welcomed anyway, so an attempt at intervention will be futile.

But can we really prescribe a behaviour towards unsettlement? We saw in the previous section that it boils down to a choice between pursuit of epistemic goals in relation to truth, and staying in accord with the preservational drives.

Radical positions on both ends of the spectrum are problematic. Consider for a moment that for epistemic goals, one is leaving the entire belief system open for variance, with no regard of any genetic or epiphylogenetic knowledge, where all information is to be acquired and inferred on a first person basis. Such a knowledge scheme would be inherently limited by the finitude of human condition. In the limited time and capability to explore space, only a finitely sized sphere of truth about the world can be explored, and if one takes up a completely sceptic stance towards any information that is received, this finite knowledge acquired by the given person is lost with his death. Different agents follow such a learning curve without any regard to predisposed knowledge and end up knowing similar things in an incomplete fashion again and again over generations.

On the other hand, if no variance is allowed, then it is very much evident that there would not be any room for growth and improvement in the belief system and this stagnancy would continue to stay. It would be a condition where, if all goes right, agents will get born and die, learning or retaining more or less the same content, without any regard of whether it is constituted by correct features or not. However, given the contingent nature of life systems, especially for genetic and epiphylogenetic routes that preserve driving factors (happiness and goodness respectively), it is not very probable that such a state of perennial stagnation happens to exist.
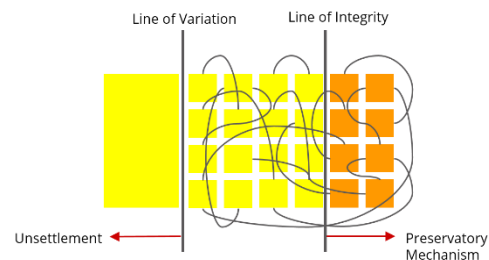

Figure 3: Decoupling Variation and Integrity

Hence, the desiderata for a behaviour must be to allow for improvement in the belief system to acquire new ideas and validate them, while using different avenues to enable continuity of the truths acquired beyond the limits of individual human finitude[3]. A prescription to fit for such desiderata would be to keep oneself open to intervention (and not fear unsettlement) but maintain integrity in acquired data by rigorous proofs, and often leave room for abductive reasoning supplanting a background scepticism when there is no conclusive proof for or against a claim. Such a system roughly overlays upon modern pure sciences[4] where scientists collectively move towards epistemic goals of understanding nature by collaboration, relaying and reviewing of scientific work, which always remains in a state of progress. Such a collective scheme may just remove any moral considerations pertinent to unsettlement, for if there is no unsettlement incurred by the intervened, there is no moral question for the intervener either. The bottom line being, an unsettlement happens only when one entertains it.

---

[3] This quite echoes with the traditional clash of political positions of conservativism and liberalism.

[4] Pure sciences are mentioned specifically because the applied sciences are often justified by invoking their practical benefits that are conducive to survival and reproduction (Woodward, 2003). These ends might just be vacuous as they just let the very practice propagate over generations while incurring epistemic stagnancy.

6 References

Camus, A., 1955. *The Myth of Sisyphus and Other Essays*. New York: Knopf.

Christman, J., 2015. *Autonomy in Moral and Political Philosophy*. [Online] Available at: http://plato.stanford.edu/archives/spr2015/entries/autonomy-moral/

Cohen, J., 1986. Reflections on Rousseau: Autonomy and Democracy. *Philosophy & Public Affairs,* 15(3), pp. 275-297.

Hamlyn, D., 1985. Self-Deception. *Journal of Medical Ethics,* 11(4), pp. 210-211.

Kant, I., 1784. An Answer to the Question: What is Enlightenment?. *Berlin Monthly*, December.

The Matrix. 1999. [Film] Directed by Wachowskis. America-Australia: Village Roadshow Pictures.

Woodward, J., 2003. Causation and Manupulation. In: *Making Things Happen*. New York: Oxford.