

Lecture Four

NORMATIVE MIND SCIENCE?

Psychology, Neuroscience, and the Good Life

1. Can there be such a thing as Normative Mind Science?

In the first two lectures I've tried to make credible a certain way of conceiving of both the nature of persons and our aspirations to make sense of things and find meaning. I been trying, as promised, to keep my eye on various zones of conflict within and among the *spaces of meaning* that constitute the *Space of Meaning 21st century* -- the vast space we participate in and interact with as we, and our fellows, engage in the attempt to flourish, to achieve *eudaimonia*.

Tonight and in the next lecture, I ask this question: Can there be such a thing as normative mind-science? And if there can be is that a good thing? One might answer "obviously yes" to the possibility question since when mind-science tells us how a proper visual system works, or what the right balance of neurotransmitters is, or when work in "neuro-ethics" tells us that emotional centers are more active than reasoning areas when people solve ethical dilemmas, they are utilizing normative concepts such as 'properly functioning,' 'right balance.' They are also, in the case of neuro-ethics,' actually studying moral norms and moral values. True, both times. On the other hand, a scientist who is shy about doing anything normative could take cover in all three cases by claiming that such normative concepts are simply shorthand for what is statistically normal.

So let us ask a more revealing question: Can there be a mind-science that empirically studies what is statistically abnormal, but nonetheless, good, of great value: namely, the causes and constituents of *eudaimonia*? Yes. In fact, such work is ancient, and it is now on again. The great ancient attempts to do *eudaimonistic scientia* were, I think, for reasons having to do

with positivistic conceptions of science marginalized and treated as ‘merely philosophical.’

Later if there is time I’ll explain what I mean.

As for the second question: would it be a good thing if mind-science tells us how to achieve *eudaimonia*? -- one can imagine a host of worried responses. When (in lecture 1) I first talked about science and what gets some people edgy about it, I pointed to attempts by science to act as if it can do more work than it can in fact do, like explain everything. I also said that science describes and explains but does not, except in engineering, trade in “oughts.” This, I now admit, was not quite right. There are, and always have been sciences, such as medicine and more recently psychiatry, clinical psychology, and social work that incorporate norms in the relevant sense. Where do the norms of physical and mental health come from? From inside medicine and psychiatry and its mates or from outside these sciences? I think the answer is that the norms come from within these sciences and from outside them. The right answer is ‘both’. What I have in mind is this -- to keep things simple I’ll use only the example of norms of mental health. First there is observation of what is statistically normal, what the average person is like. Second, the image of the “average Joe and Joanne” is considered in light of norms that come from comprehensive views of mental health and, what is different of good persons. These are in the air in public social space and often do not abide giving the *imprimatur* of ‘healthy’ to what is statistically normal.

People commonly have problems in the areas of love and work. There are norms about being a good husband, wife, father, mother, and child. These are culturally and socially variable. It may be statistically normal for husbands to be domineering, even violent. Perhaps, hopefully, the best ethical, political, and religious thought in the culture says this is wrong. The norms that people need to abide are not the ‘normal’ ones. Mental health workers see the distinction and

work with patients to reconfigure themselves and their lives to be better persons. Greater tranquility and happiness might be offered as rewards. So first pass, we can say somewhat artificially (because it makes the process temporally two-step and sequential when it is in fact dynamic) that scientifically trained mental health workers develop norms of mental well being that use the mental health equivalents of homeostasis and equilibrium from ordinary medicine and adjust, modify, and enhance them by utilizing the best available cultural wisdom about what makes for a good person, for living a good life, and so on. If one says this is highly fallible, I agree. But what method isn't? One attractive feature of the picture meanwhile is that all the spaces in the Space of Meaning ^{21st century} have a say in the norms.

It is worth mentioning that psychiatrists and clinical psychologists, on the one hand, and social workers, on the other hand, typically have a different clientele. The first typically have clients from the high-end of the socio-economic ladder for whom the external conditions of life are good; social workers typically deal with people whose psyches are damaged or in jeopardy because their environments are dreadful. Different circumstances, different treatment issues.

One more comment before I simply proceed full-steam-ahead to talk about how I see the prospects for normative mind-science in our time, and to my argument for why further development of such a science or sciences would be a good thing. The objection I hear in my ear is this: "Owen, the sciences you are talking about are not really sciences, not really 'scientific' -- even practitioners will often refer to what they do as a 'craft' or even worse as 'an art.'" There is a lot to say, I'll just say this: Compare to structural engineering. The laws of physics, material science, and engineering are background for the design of a skyscraper. Siting is key, call it a particular initial condition; then there are weather issues, matters of which ways the winds blow, whether other structures create wind tunnels, and so on. I remember distinctly when

I was in graduate school in Boston just after the John Hancock Tower was finished; its windows kept blowing out. Someone hadn't thought through the particulars about the wind.

When it comes to mental health, or to what is more, to flourishing, each person is a unique spatio-temporal site, with a particular life history, a certain "normal weather within," situated inside a certain culture that draws for display of its own norms, and so on.¹ It is more complicated than siting and constructing a building. The 'artful' or 'craft-like part' comes from applying the scientific or empirical knowledge, wisdom, or generalizations in the nuanced way required by the complex particularity of any human life.

2. Two Ancient Examples of *Eudaimonistic Scientia*

I said that normative mind-science is ancient. I offer two examples that come from approximately 2500 years ago. The relevant texts are Aristotle's *Nicomachean Ethics* and the Buddhist *Abhidhamma*.

Both of the ancient texts provide psychologies of the statistical normal. But in both Aristotle's *Nicomachean Ethics* and the Buddhist *Abhidhamma* there is also provided a theory of human excellence. For Aristotle such a person is *eudaimon*, one who is truly happy, one who genuinely flourishes, one who is rational and virtuous. In the Buddhist case we are provided with divisions among *wholesome*, *unwholesome*, and *neutral* mental states, as well as various lists of afflictions and virtues. The most excellent person is the one who has released herself

¹ There are good generalizations in mind-science. Although the quality and precision of the generalization depends on what domain one looks at. The understanding of the visual system is in the best shape among the five sensory modalities. Because there are 20+ areas of visual cortex involved in "seeing" the laws about how normal processing occurs are functional and system-level. There are many excellent generalizations about, for example, how shape detection is done here and color detection is done there and how they bind unconsciously. There are no laws like $f=ma$ and the consensus now is that as in medicine there won't ever be such laws. Why? Not because visual scientists are not smart enough, but because the biological system in question is a complex functional system that behaves in a normal and regular manner, but that admits of irregular well-functioning designs; and this is so because that is the nature of biological things. If a person has a stroke that affects a particular aspect of visual processing, say motion detection, neurologists know where the damage will show up and do a "look-see" to confirm it. But because each brain is slightly different at each and every level, the damage may be sited slightly differently in two individuals who show *exactly* the same deficit.

from various poisons and mental afflictions and lives among the four divine illimitables of compassion, loving-kindness, sympathetic joy and equanimity. And like Aristotle's *eudaimon*, she – think of her as a *bodhisattva* -- is statistically abnormal, but embodies the most excellent norms. Both flourish in ways normal folk don't. One wonders how is it possible to do normative science of this sort, of these two sorts? I'll explain. It may surprise you, but I'll claim one does such normative mind-science, when one does it well, empirically.

First Aristotle: The Aristotelian picture is familiar; I've used aspects of it already in previous lectures, so I'll be brief:

Everyone agrees that the greatest good is *eudaimonia*, 'happiness'. But there is disagreement about what brings happiness or, what is different, constitutes happiness. Every idea is a contender -- sensual hedonism, solitude, great wealth, power, and so on. Indeed, they are all out-there embodied in how different people are and live. None are the answer. The answer is that our human end -- the way we actualize our full potential is by living a life of Reason and Virtue. Virtue consists of the set of dispositions to perceive, feel, think, judge, and act in the right way, at the right time as the particulars of the situation that calls for the virtue in question warrants. Aristotle's list of virtues includes four from Plato:

- Courage
- Justice
- Temperance
- Wisdom

To which these six are added:

- Generosity
- Wit
- Friendliness

- Truthfulness
- Magnificence (spending lavishly of worthy things, e.g., sacrifices, warships, public buildings –).
- Greatness of Soul (believing because it is true that one can accomplish great things and thus is worthy of honor)²

² In *The Therapy of Desire* (1994), Martha Nussbaum presents a compelling case for understanding the post-Aristotelian Greek and Roman philosophers as doing much more than simply advancing and refining Aristotle's ethics. Post-Aristotelian ethics advances a view of the good life that is open to everyone, not just the well bred. Despite the universal access to a life of virtue, the Epicureans and Stoics especially (See Becker's 1998 masterpiece that shows that, and how, a noble Stoicism is still a live option for us), paint a more demanding picture of virtue than Aristotle does. This more demanding ethical conception requires much deeper psychic change than Aristotle thought necessary in order to alleviate suffering and bring happiness in its place. In part, the need for greater direct attention to an individual's psychic economy is due to the fact that Aristotle was insufficiently attentive to the way certain destructive states of mind, for example, greed and avarice, cause suffering and bad actions, but are nonetheless subject to voluntary control. According to the therapists of desire more than good socialization, even as supplemented by attending Aristotle's lectures on ethics, is required for virtue. Direct therapy on the minds of adults to quell or eliminate negative desires is needed as well. In addition, the expansion of the list of virtues to include universal compassion requires work to expand and enhance whatever tendencies of fellow feeling are rooted in our nature, but that are enhanced insufficiently by local (Aristotelian) moral conventions.

Fortunately, the transformation of the psyche required for true virtue and happiness is possible so long as the philosopher equipped with a more expansive set of instruments than argument alone, plays the role of a trainer or physician for the soul.

The Hellenistic philosophical schools in Greece and Rome -- Epicureans, Sceptics, and Stoics -- all conceived philosophy as a way of addressing the most painful problems of human life. They saw the philosopher as a compassionate physician whose arts could heal many pervasive types of human suffering. (1994, p. 3)

These Greek and Roman compassionate philosophers -- Epicurus, Lucretius, Zeno of Sidion, Chrysippus, Pyrrho, Seneca, Cicero, Epictetus, Sextus Empiricus, Marcus Aurelius, among them -- were, founders, luminaries, and practitioners of schools that arose after Aristotle's death (322 B.C.E.) and that remained highly influential into the 2nd or 3rd centuries C.E. In various ways these philosophers retained admirers until the 19th century. Descartes, Spinoza, Adam Smith, Kant, Nietzsche, and Marx all engaged the work of one or more of these philosophers. And the American Founding Fathers who had read Plutarch's *Moralia*, which includes his famous "Lives" of the major Hellenistic philosophers, as well as in all likelihood Cicero, and Seneca, reveal that influence in well-known ways in the "Declaration" and the "Constitution."

The therapists of desire provided, indeed insisted upon providing strong -- ideally, valid and sound -- arguments to support their diagnoses, prognoses, and therapeutic practices. In part, like their Indian counterparts, especially Buddhists, this was because they believed that mistaken views (*moha*, Pali) are often at the root of human suffering, for example, money is widely thought to bring happiness but. But they also recognized that argument alone does not always produce the necessary change. Even if false belief -- what Buddhists call delusion or wrong view (*moha*) -- is lifted at some conscious level ('Ok, money doesn't bring happiness. Now I get it'), there are typically long-standing emotional and conative tendencies and attitudes associated with the false belief (possibly antecedent to it), that in virtue of being deep-seated and partly unconscious may still control the motivational circuits. Even if the false belief is exposed as false, acquisitive desires and behavior may not abate ('I know that money doesn't bring happiness, but I keep trying to accumulate wealth, and I feel vacant, empty, dissatisfied'). Here the therapists of desire rightly saw the need to bring to bear techniques, in addition to arguments, to adjust or change the economy of desire, often working to outright eliminate certain destructive emotions by antidotes that were psychologically incompatible with them. Wishing someone ill and feeling deep love and compassion for him at the same time are psychologically inconsistent -- at least they comprise a highly unstable tandem.

Michel Foucault refers to this style of doing philosophy, which involves working to form or restructure the self as utilizing techniques of self-work, *techniques de soi*. Nussbaum agrees, but warns that, then as now, there

One might ask how did Aristotle, for that matter how does anyone, generate such a list, a view about what -- as rare as it might be -- *eudaimonia* consists in. The process is one of what Rawls calls “reflective equilibrium.” One observes lives that work and those that don’t work to produce what a reflective person thinks is *eudaimonia*. One examines the conditions of the world and of character that seem to do the job, critically evaluates these with that initial reflective conception of *eudaimonia* in hand, and adjusts one’s conception of both the ‘conditions’ productive of and/or constitutive of *eudaimonia*, and one’s conception of ‘true happiness/flourishing’ along the way. Close inspection begins with an initial hypothesis about what *eudaimonia* is, what leads to it, constitutes it, and so on. Call the starting conception *eudaimonia*^{initial reflective conception}. One deploys *eudaimonia*^{initial reflective conception} in examining the roots, sources, and character of *eudaimonia*^{initial reflective conception}, and this process yields a more sophisticated view, call it, *eudaimonia*^{mature reflective conception}. The “more mature reflective conception” is better; but there is no guarantee it is the best picture we can come up with. I’ll return to this point shortly when with the Buddhist picture as a contender, at which point we can ask which is better.

I said that *eudaimonistic scientia* is empirical. Here, first pass, is why: One starts with a hypothesis about what constitutes a healthy or good person (compare to botany on what a healthy plant of a particular variety is), and one asks questions about what causes and constituents contribute to the well-functioning form (how much sun, water, what soil type(s), etc.). One tests these ideas and reaches a well-confirmed hypothesis, possibly adjusting one’s initial conception (these plants look very beautiful but they are short-lived being prone to such and such fungus).

were *techniques de soi* that relied on mesmeric force and hocus pocus without the requirement that sound arguments also be offered warranting soul change of a particular sort by way of a suitable technique.

The Moral Psychology of the Buddhist *Abhidhamma* (Pali: Sanskrit: *Abidharma*):³

It is in moral conformity to the “Four Noble Truths” that guide all Buddhist practices, that the current Dalai Lama writes: “The principal aim of Buddhist psychology is not to catalog the mind’s makeup or even to describe how the mind functions; rather its fundamental concern is to overcome suffering, especially psychological and emotional afflictions, and to clear those afflictions” This is exactly right. To the puzzlement of many Westerners, the *Abhidhamma* taxonomizes mental states into *wholesome* and *unwholesome* (and, to a lesser extent *neutral*) kinds. I mischievously try to get a rise out of scientific friends by telling them that this is so and that I approve of this. It always works. “Owen, have you lost all sense of standards!” Many colleagues have a knee-jerk response, thinking it shockingly irresponsible to mix scientific psychology with ethics. Very naughty! I ask: “Have you read any books in abnormal psychology or psychiatry?” The point as I see it is this: Just as we might criticize a psychiatry text on the grounds that it assumes an unwarranted conception of mental health, any concerns with the ascriptions of ‘*wholesomeness*’ or ‘*unwholesomeness*’ require showing what is wrong in the Buddha’s conception of the good life. “Put up, or shut up.” The fact that the *Abhidhamma* combines descriptive with normative insights gathered from the Buddha’s teaching in the first two baskets is not itself an objection of any sort.

In any case, the *eudaimonistic* impulse of Buddhist psychology cannot be emphasized enough. The *Abhidhamma* is a masterpiece of phenomenology. And despite what His Holiness

³ Normally, I use Sanskrit terms rather than Pali because they are more familiar to Western eyes. Here I reverse the procedure and use Pali rather than Sanskrit throughout since the *Abhidhamma* is composed in Pali. Furthermore, it is a sign of respect for the tradition, since the historical Buddha who lived around 500 B.C.E. is thought to have spoken in Pali, not Sanskrit. The first texts, the three baskets of the Pali canon are composed/written down around 100 C.E. The *Abhidhamma* consists of cosmology, metaphysics, and psychology. The prefix *abhi* means, or suggests, the drawing of distinctions. It is attached to *dhamma* (Pali; *dharma*, Sanskrit), which in this context refers to the teachings about the bare ontology of things. Books 1 & 7 are the main ones devoted to psychology, the other books are devoted mostly to Buddhist views on time, causation, cosmology, and so on). The other two baskets of Buddhist wisdom contain parables (*suttas*; Sk: *sutras*) and moral rules for monks, nuns, and the laity.

says about not being concerned with taxonomy, the *Abhidhamma* remains arguably the best taxonomy of conscious mental state types ever produced. Most of my remarks pertain to books 1 and 7, which deal with psychology. All Buddhists treat books 1 and 7 of the *Abhidhamma* as a psychological masterpiece combining at once deep phenomenology, analytic acuity, and classification of mental states in terms of the ‘wholesome’ and the ‘unwholesome’ in accordance with how they fit into the Buddhist view of *eudaimonia*.⁴

What first catches the eye of the Western reader is the extraordinary number of distinctions drawn among states of consciousness. The book begins with a taxonomy of consciousness (*Citta*) into conscious mental state types (*cittas*). These number 89 initially, and reach 121 after some adjustments. Each type is characterized in terms of the sort of object it takes in (so visual and auditory consciousness differ in an obvious way); its phenomenal feel (e.g., sad and happy); its proximate cause or root, e.g., there is greed-rooted and hatred-rooted consciousness; and its function or purpose (avaricious consciousness is thirsty and aims to suck in and swallow what it desires).⁵

As one studies the *Abhidhamma* one gets into the spirit of drawing distinctions upon distinctions, and, indeed one could really start to believe the Tibetan joke that a master phenomenologist might be able to discern 84,000 (the number is akin to us saying ‘a gazillion’)

⁴ This deep respect for the *Abhidhamma* obtains despite the fact that many Mahayana Buddhist see the original manual, penned by Theravadin monks, as too glowing in its treatment of the monastic life, as well as still embracing remnants of the doctrine of *atman*, not having really absorbed the idea that at bottom all things are empty (*sunyata*). The complaint is that the decomposition of mental states bottoms out in, what seem to be, indestructible psychological atoms which themselves should be considered conditioned (because of the doctrine of dependent origination = everything is in flux, dependent for its ephemeral being of prior conditions), and thus that are further decomposable *ad infinitum*.

⁵ *Citta* and the *cittas* are analytically distinguished from the mental factors (*citasekas*) that they, as it were, can contain. So, roughly, Joy-Consciousness might contain joy-at-an-infant’s-birth-in-my-family or joy-at-a-friend’s-success.⁵ Joy-Consciousness is a type of consciousness, thus a *citta* of *Citta*, whereas joy-consciousness-about-family and joy-consciousness-about-friends would be two sub-types (factor, *citaseka*). Even the *citasekas* admit of lower-level distinctions that happily are not made! But to give a feel: I might be happy [that sister Nancy has a baby] and [that sister Kathleen has one]. The feeling is of the same type, but the intentional content, marked off by brackets, differs in the two cases.

types of anger or craving – anything for that matter! The second thing (or perhaps it will happen first) that will strike the Western reader is that the words ‘*wholesome*’ and ‘*unwholesome*’ (as well as, but less frequently, ‘*neutral*’) are used in the process of classification itself. Again one can easily imagine the objection that ‘real,’ that is, scientific psychology describes and explains and predicts, but does not judge the various kinds of sensations, perception, emotions, and learning it analyses in normative terms.

The three poisons that come with our natures are hatred (*tanha*), craving (*lobha*), and delusion (*moha*). These three, uniquely perhaps, are always bad or unwholesome. Furthermore they ramify and interact with other mental states, indeed with one’s overall sense of well being, in ways that produce un-ease.⁶

At the other end of the spectrum are the “Four Divine Abodes” (*brahmaviharas*) -- also called the “illimitables” or “immeasurables” (*appamanna*).⁷ The four are:

Loving-kindness (*metta*)

Compassion (*karuna*)

Appreciative Joy (*mudita*)

Equanimity (*upekkha*).

⁶ The three poisons are first elaborated as giving rise to “*The Six Main Mental Afflictions*” Attachment of craving, Anger (including hostility and hatred), Pridefulness, Ignorance and delusion, Afflictive doubt, and Afflictive views. These in turn are roots for the “*The Twenty Derivative Mental Afflictions*” Anger which comes in five types: Wrath, Resentment, Spite, Envy/Jealousy, Cruelty; Attachment which also comes in five types: Avarice, Inflated self-esteem, Excitation, Concealment of one’s own vices, Dullness; Four kinds of Ignorance: Blind faith, Spiritual sloth, Forgetfulness, and Lack of introspective attentiveness. Finally, there are six types caused by Ignorance + attachment: Pretension, Deception, Shamelessness, Inconsideration of others, Unconscientiousness, and Distraction. The point is that there are a lot of ways one can go wrong. The tools required for the therapy in virtue will not surprisingly need to be abundant and multifarious in kind.

⁷ Of the four sublime states two, compassion (*karuna*) and appreciative joy (*mudita*) are considered as *sui generis* mental types. The other two (loving-kindness (*metta*)) is a mode of non-hatred which as such does not always show itself as -- or give rise to -- loving-kindness. Although we can rightly say that it one way of expressing non-hatred or that a non-hating state of mind is a necessary condition of loving-kindness. The same analytic situation applies to equanimity (*upekkha*), which is a sub-species of mental neutrality.

One might think that just as the “three poisons” are categorically bad, the “divine abodes” are categorically good. Maybe. But an important caveat is in order. Wisdom (*panna*) and virtue (*sila*)—require as a necessary condition avoiding the three poisons. It seems constitutive of loving-kindness, compassion, and appreciative joy at the success of others that these states rule out (are psychologically incompatible) with hatred and envy (a form of greed). But delusion (*moha*) can mitigate the sublimity of even the “divine abodes.” Suppose one achieves equanimity because she fails to notice certain particulars about her own character, or the character or plight of others, that she ought to be noticing or paying attention to. Here delusion surfaces and might make us question whether the equanimity is warranted. It feels sublime, but it is supported and sustained by failing to see what one ought to see. There is an unwholesome aspect to such equanimity.⁸

For reasons such as these, Buddhist psychology pays a considerable amount of attention to the *causes* of mental states, especially before moral assessment is made, as well as to causes of epistemic disruption (internal and external), and to matters of intention and epistemic as well as moral responsibility. All states rooted in hatred or greed are *unwholesome*, as are states caused by wrong view. The case of the magic pill that makes one happy is not discussed, but similar cases are. If a seizure causes me to experience euphoria, it is deemed “rootless,” and rootless states are unwholesome.⁹ Notice, this means that one can be in a state that feels good, has positive valence, but which is, nonetheless, unwholesome. One reason for judging the state

⁸ Similarly one might feel happy about one’s friend’s successes, but have failed to notice that the successes were not achieved in an honest way. Whether one would be judged culpable for this sort of ignorance depends on what was in view and what wasn’t. The point is that certain epistemic deficiencies can undermine the warrant, and thus the sublimity, full stop, of being in an (otherwise) divine state of mind.

⁹ David Wong wonders if this isn’t excessively moralistic: why not just say a state of euphoria caused by a seizure is neutral? Good question.

unwholesome is that it is normative – both psychologically and ethically – that happiness be produced by goodness or self-work I engage in, not by aberrant neural firings.

One other point: all four abodes are said to involve states of mind towards others. One might agree, while emphasizing that, at the same time, all four are in fact states of mind of the individual who has them, and they have unique first-personal phenomenological feel for that person. Their object is, of course, the good of some other. But this analysis seems to run into trouble with equanimity (*upekkha*), which might seem to be a pure state-of-my-soul, and thus not directed at, for, or towards anything outside me. To be sure, my being calm and serene might make me more pleasant to be around, but it is not constitutive of equanimity, as we understand the state, that it has this aim. But this is not how Buddhists understand equanimity. Equanimity (*upekkha*) means more than serenity. It is constitutive of equanimity that I feel impartially towards the well being of others. If I am in the state of equanimity, interpreted as *upekkha*, I am in a state that involves, as an essential component, equal care and concern for all sentient beings.¹⁰ We might translate *upekkha* as equanimity-in-community if it helps avoid confusion with our understanding of equanimity as a purely self-regarding state of mind.

¹⁰ In classical Buddhist sources, there are three separate disciplines for the study of consciousness. The *Abhidhamma* focuses on examination of the causal processes of the hundreds of mental and emotional states, our subjective experience of these states, and their effects on our thoughts and behavior. It is related to what could be called psychology including cognitive therapy and phenomenology. Second, Buddhist epistemology analyzes the nature and characteristics of perception, knowledge, and the relationship between language and thought in order to develop a conceptual framework for understanding the various aspects of consciousness – thoughts, emotions, and so on. Finally, Vajrayana uses visualization, thoughts, emotions, and various physical techniques such as yogic exercises in an intense meditative effort to accentuate wholesome ways of being and to transmute the afflictions of the mind. It is concerned not with discovering an independent permanent entity called “the mind” but rather with understanding the nature of the ordinary mind and effecting its transformation into a non-afflicted, clearer state.

3. The “Internalist Objection” to *Eudaimonistic Scientia*.

Before proceeding I want to consider an objection one might make (I would) about the two *eudaimonistic* views I’ve sketched and that are thus on the table. The objection is this: Suppose one accepts the idea that the work discussed has empirical grounding: there is observation of how the “many” and the “few” are, and how they fare, governed by the method of “reflective equilibrium.” Once done, no answer is given as to which of the two contenders is the right one. The reason: the only measure of what constitutes *eudaimonia*, flourishing, excellence, enlightenment is, in fact, what norms are avowed, practiced, considered best from *inside* the culture. Who gives a better picture of *eudaimonistic* excellence, Aristotle or Buddha? Normative mind science can give no answer. The Aristotelian answer is right for Aristotelians; the Buddhist view for Buddhists. Science resolves such problems. Who had it more right Newton or Einstein? Easy. What can the *eudaimonistic* empiricist do?

Let me make the worry worse before making it better. Consider these two relatively recent quotes from the Dalai Lama:

“Now, we are made to seek happiness. And it is clear that feelings of love, affection, closeness, and compassion bring happiness.” (*The Dalai Lama and Howard Cutler, 1998, p. 52*)

“For our life to be of value, I think we must develop basic good human qualities --- warmth, kindness, compassion. Then our life becomes meaningful and more peaceful --- happier.” (*The Dalai Lama and Howard Cutler, 1998, p. 64*)

There is a reading of both these statements as empirical, that is, as statements about what sort of things produce *eudaimonia* (read closely they can be read as making different claims, one about what produces ‘happiness’; the other about makes for a ‘meaningful life’ – but no bother now). This way of reading the statements is to be contrasted with reading them as endorsements or recommendations for living a certain way. Now, if the empirical claims¹¹ were true, a sensible person would want to adopt certain norms. Consider this possibility: once any culture or social group or moral or spiritual tradition has settled on a certain conception of *eudaimonia*^{mature reflective conception}, then it is almost guaranteed to create a social environment that pulls for endorsement of the virtues, values, and norms said to enable, cause, or constitute *eudaimonia*^{mature reflective conception}. This will make it true that to find genuine happiness you need to be-a-certain-way and do thus and so, but not because being and doing thus and so produce “true happiness” for everyone, but because the contingencies of social reinforcement around here are such that only in this way is it granted. This is what I call “the internalist objection.”

At this point one usually quotes this passage from Aristotle. Thus I dutifully do so by way of my own free translation:

We should consider our discussion adequate insofar as we make things perspicuous enough as regards our subject matter. We do not seek or expect the same degree of exactness in all sorts of arguments (compare: mathematics, physics, history), just as we do not expect exact sameness in the products of different crafts (compare pressing coins, to knitting clothes)...[I]n ethics and political science each of our generalizations ought to

¹¹ I use the plural because, as I say parenthetically, these statements make somewhat different claims at least on the surface. The first is a claim about what brings happiness. The second about what brings a sense of meaning and purpose with happiness possibly added as “gravy.” Anyone interested in testing these two hypotheses would need to unpack them further.

be understood as holding true usually. And because this is the nature of our premises [that such and such holds generally, but not universally}, we must be satisfied with probabilistic conclusions of the same sort (*NE*, I, 2, 1094b-1095a).

Aristotle's aim is to distinguish the inferences of mathematics and physics from *eudaimonistic scientia* (which we now see includes both ethics and political science). Mathematics involves deductive arguments and yields necessary conclusions. Physics starts with universal generalizations: Always and everywhere and for all things, $f=ma$.¹² Arguments with the latter as a major premise can yield a certain conclusion if and only if two of the three variables can be specified exactly. If not, we deal with approximations, and our conclusion yields the same.

The first point about eudaimonistic science is that it has no universal generalizations. The problem could be practical, epistemic, or ontological. Aristotle is a naturalist so I think it best to understand him as thinking that the lack of universal generalizations is a practical and/or epistemic matter. It would be false to say "Money never brings subjective happiness." Why? Given our practical methods for divining states of subjective happiness, it does appear to make some people happy.¹³ Now it may be, when neuroscience matures, and we can more precisely state what neural state is a state of subjective happiness that these outliers will show a small difference in neural activity from others whom we deem subjectively happy. Then we will be able to state that "Money never brings subjective happiness," now defined to exclude these outliers.

¹² Aristotle, of course, did not know the law in this form.

¹³ My friend Jonathan Haidt (personal communication) explains that happy people make more money. In all likelihood this is because they are happy, appealing, extraverted, and so on. This is what is called a "reverse correlation." Two other points from Haidt: (1) comparative economic advantage has some definite effect on judgments of subjective well-being; (2) some data show that money might be (for more than a few) about as important a causal contributor to subjective well-being as the number and quality of close relationships. A final observation: wealth and health go together. Health, especially when it comes to longevity, has a strong relation to stress levels (cortisol is the nasty culprit), so it may well be that wealth allows people to escape certain statistically normal stressors, such as not having enough to pay the bills.

What I am calling epistemic problems might arise in minor premises. I'm testing a statistical generalization or a universal one and I am just not sure, given my observations, whether for some person or group of persons they meet my criterion for being subjectively happy. So I say this: Most of the Forbes 500 richest American seem pretty happy, subjectively speaking. Or again back to those who seem like outliers: I can't possibly measure all the causal contributors to these individuals' states of mind. So although money does for some look as if it is the cause of their subjective happiness, it may in fact be caused by genes, unusual maternal care, or by the fact that they always get enough sleep, or by a whole bunch of such things (including the money).

Whatever the cause, we should just accept that most "laws" in *Geisteswissenschaften* are statistical generalizations, not universal laws. This doesn't make any findings less empirical, it just makes them less precise.

With this much in hand, let's return to the "internalist objection." The objection can now be stated this way: Ok, we'll give you imprecision: your science is immature¹⁴ and your subject domain is very complicated. What we won't give you is this: you can't give any external, objective reasons for your judgments about what states-of-affairs, norms and ways of life, are best or most worthy that doesn't privilege a class of norms or ways of life that are part of your data set. You mix norms with facts.

This is true. The question is, should it worry us? Note we do not commit Hume's fallacy. There is nothing in either Aristotle or the Buddhist approach that involves attempting to *derive* 'oughts' from 'is's', or values from facts. This is because deductive logic is not in play.

¹⁴ In general, I recommend not accepting the "immaturity" charge. Those who say such things as that "psychology-as-a-science began in Leipzig in 1849" commonly make it. I am not saying there isn't some truth to the charge. But overall I think it best to see the problems with mind-science, as well as with all *Geisteswissenschaften*, as having to do with the epistemic difficulties of studying very complex dynamic systems. Also, sometimes what we seek is "local knowledge." Even when we get the latter right, it doesn't (necessarily) generalize.

In fact, it is my opinion that in the history of ethics (almost) no one has been stupid enough to actually commit Hume's fallacy, although some people smuggle 'oughts' in a bit too easily without adequate argument.¹⁵

The reasoning involved is inductive, abductive, statistical, and probabilistic. And most of science except theoretical physics is like this. Recall the method of "reflective equilibrium" as described above. Rawls, as I said, names his method as such, but it was Aristotle who invented it and who first described how it works: we survey the views and aspirations of the 'many' and the 'wise' (1098b27, 1153b31, 1173a14, 1179a16). We treat all these beliefs and ways of life as worth taking seriously, but as revisable (1095a22, b16 1113a33, 1153b35, 1159a19, 1163b26, 1167b27, 1172b3). We use techniques such as these to uncover problems with the beliefs of both the many and the few: do people who think money brings happiness, usually gain it. No. Instead they get stuck on what we nowadays call the hedonic treadmill. Is honor satisfying? Yes, but normally only if deserved, and so on.

We begin responding to "the internalist objection" by admitting we are looking for norms, values, and practices that are the best. The analogies are from engineering and the crafts. Given that we need/want bridges that work and shoes that don't leak, we cull from all the possible ways of accomplishing these things the ones that work best. Were Athenian bridge-builders or cobblers the best ever? Of course not. But the best were the best at that time.

The situation is almost Kuhnian. Just as a scientific theory stands as best until there is a competitor that does better, so with bridge-building, shoe-making, and ethics.¹⁶ At this point the

¹⁵ Hume was keenly aware of creepy clerics who tried to pull the wool over peoples eyes by seeming to derive categorical moral prohibitions from alleged facts such as in the Bible God says 'such and such', therefore you ought not ever do 'this and so'. That God said 'such and such' is, if he said it, a fact, a description of something that happened. Nothing deductively follows from it. Nada. Zero. Well you can keep repeating it. If p, then p.

¹⁶ I don't want to claim that there is always progress. Once there were many shoemakers who would put an extra 1/4" heel on a shoe. My Dad needed this. So do I. But now such cobblers are hard to find. So too certain virtues, such

comparative ethicist can usefully chime in. One worry I gave voice to (although truth be told it doesn't normally worry me all that much) is this: Aristotle offers one set of virtues and the Buddha another. Perhaps they are not inconsistent, but they are not the same. And so far no answer has been given as to which is best. The critic of *eudaimonistic scientia* will say that this problem is to be expected, indeed it is inevitable, given the objectionable internalist methodology being deployed.

Here's a sketch of a reply: First, we are not required to simply accept Aristotle's picture because he offers it as his best theory -- even as the correct theory for Athenians. Take his last two virtues, magnanimity and great-souledness. The latter might be criticized, or at least warrant watchfulness, because it seems as if the great-souled person might more than others be prone to self-puffery. Aristotle is careful to say that the expectation of honor is to be based on correct perception of merit. But given that *hubris* is bad, we are entitled on grounds internal to the Greek situation to worry about the virtue's stability. Regarding magnificence, we have resources for criticism that are both *internal* and *external*. What I mean is this: In order to be magnificent in Aristotle's sense, requires that one be rich. One might think wealth comes from merit. But given the attention paid by the Greeks to luck-in-life's-circumstances, one might see that wealth is often a matter of luck (*eutuches*). One response is to say that this is why magnificence is a virtue, the excellent person gives away a lot of his stuff for the common good, both because this is good, and because it involves recognition that his wealth is not a measure of his merit -- it comes from a lot of good luck. These critical resources are internal to (but perhaps not fully visible in) the tradition. But again suppose that the matter of luck is just not connected in the

as honor and loyalty within and to one's community might get watered down, even lost, as we move from small agrarian communities to industrial urban life.

minds of the Athenians to wealth distribution. The internal dialectic doesn't, for whatever reason, let those who regularly think about 'luck' see its role in the case of wealth.

Here we can go *external*: First, there is the small matter of the economy built on slaves (actually, at this time, usually prisoners of war). Aristotle is silent on the legitimacy of slave ownership; it is part of the-taken-for-granted background. So we raise some questions from outside about that practice. The most promising external tactic, however (since the ones about slaves = prisoners of war might not yield a grip; consider Mr. Bush), would be to point to economic arguments about how wealth accrues. Wealth acquisition does not reliably track merit, and it can, if we are not watchful, lead to sufficiently uneven distributions such that there are prudential reasons (safety and security given that the poor and oppressed might rise up) to do some evening out. If we think there is also ethical knowledge, we can form a tandem, think of Amartya Sen's work: we now know enough about ethics and economics to know that the distribution of wealth can and should be reconfigured if the wealth of the few is based on mistreatment or oppression of the many. Aristotle didn't see or know this. We do. And thus when his list of virtues is placed into "wider reflective equilibrium," we can make a normative recommendation that does not challenge the virtue of magnificence. It challenges the social conditions that make the virtue one only a few can display.

The strategy of seeking "wider reflective equilibrium" can now be used in comparing Aristotle's conception of *eudaimonia* with the Buddhist conception. I simply won't discuss the possibility that the views might be incomparable or incommensurable. I don't think they are, although they definitely have features that make most sense internal to the traditions themselves. And this is predicted by conceiving of ethics as a form of human ecology, a line I have been pushing for over a decade. The idea is that there are almost certainly certain universal

necessities across all human environments that pull for and thus make rational certain prohibitions such as ones against murdering innocents, stealing rightfully owned property, and so on.¹⁷ Beyond these “big ticket” items, local ecological conditions will create their own pressures on normative construction. I’ll elaborate this point later when I discuss Jonathan Haidt’s important work on cross-cultural positive psychology.

In the case of the Aristotelian-Buddhism comparison we can note first that being asked to compare them puts us in Kuhnian space. Neither theory, taken on its own, we’ll suppose, had in its time and place a serious, worthy competitor. So we put them into competition across space and time. Here -- without fully engaging in the exercise -- is what you should expect to find in such a competition (this is based on work my colleagues and I do regularly at Duke’s Center for Comparative Philosophy). You’ll find that both theories have strengths and weaknesses that one only sees when they are discussed together. I can’t and thus won’t elaborate here, except to say this much. If we take everything on Aristotle’s list, we see most of his virtues in the *Abhidhamma* as *wholesome* states. What we see in Buddhism, but not in Aristotle, are the four divine abodes. If we think as I do that Jesus of Nazareth and the Buddhist tradition (possibly Mencius and, even Mozi, in China) saw the merit of expansive love and compassion in ways that Aristotle didn’t (because of how he was positioned internally to the Athenian ecology of value), then we can recommend expanding Aristotle’s list. Examined from a perspective of wider reflective equilibrium, these virtues can be seen as excellent.¹⁸

This then is how I respond to “the internalist objection.” I concede that we are engaged in an empirical inquiry that involves methods not normal or required in what we think of as

¹⁷ Even when one sees such “universals,” there are always questions about who the “innocents” are, what kind of property ‘rights’ are acknowledged, food produced by my hands, my slaves, my wife?

¹⁸ Note it is useless to think the request is being made of Athenian ethics that it be more expansive. That Athens is over. But in so far as we are heirs to that tradition the request can be made of us.

observer independent sciences where normative questions don't arise (as visibly). All the human sciences are reflexive; we do the inquiring, and we are the objects of inquiry. This is just a fact about what has to happen if you want to study humans. One important task is to describe and explain what goes on in perception, emotion, and cognition. Another is overtly normative, but it has very precise analogies, as I've indicated, in sciences such as engineering and botany. Given certain ends, how best can we achieve them? Engineering exists by combining knowledge from physics, material science, and so on, with the intention to satisfy non-frivolous practical desires. Botanical and agriculture science work the same way (botany, of course, incorporates aesthetic values). Same with *eudaimonistic scientia*. In all three cases, the method of reflective equilibrium is the only method ever proposed to do the job.

One might think that the use of the method cross-culturally is a way of testing what in science would be called "rival theories." And it is. But there is still a difference that makes a difference. Unless there are socio-moral or political problems or situations that afflict life in both cultures, where one theory does better than the other at solving it, then and only then is the situation one in which both sides can/might agree that the rival is more resourceful (see Lakatos, MacIntyre). In science the cases of "crucial experiments" are almost always ones in which two rival theories predict a different objective event (light curves or doesn't when the eclipse next month occurs). But ethics doesn't engage in prediction in this same way (although it does make predictions about *eudaimonia*) Furthermore, in ethics there are two familiar sorts of cases that don't seem to arise in normal science, at least of the sort that allow sometimes for "crucial experiments." Sometimes, the "problem" that would reveal the rival's strength is not a problem from the point of view of one culture's perspective. Genetic engineering is not an issue on the radar of most people in the Himalayas; they need cataract surgery, not gene tampering. But take

a question such as “how unfettered ought capitalism to be? -- asked by modern Chinese thinkers and Western libertarians; or whether active euthanasia, now that we have very merciful techniques, is ok? -- asked to Dutch secularists and Roman Catholics. In both cases, we know that unshared background norms color judgments about how to deal with the technology, economic practice, or whatever.

I can make this version of the internalist’s objection, perhaps now deflated some by appeal to the method of wide reflective equilibrium to “the internalist’s worry,” more visible by concluding tonight’s lecture by with a discussion of some recent work in “positive psychology” by Jonathan Haidt.

4. The Psycho-social Ecology of Virtue and Comprehensive Pictures of ‘The Good’

Jonathan Haidt has done important work on cultural differences in conceptions of *eudaimonia*. The internalist objection, and the analyses I’ve offered to make it less worrisome, arise again in considering Haidt’s model. According to Haidt and his colleagues, in the ethnographic and psychological literature, four domains of life appear universally as socio-moral domains that human build normative space around. These domains are suffering/compassion, reciprocity/fairness, hierarchy/respect, and purity/impurity.¹⁹

¹⁹ In new work (in press) Haidt and his colleagues have added an in-group/out-group module. I say nothing directly here about the wisdom of adding that module. But Haidt wisely never claimed completeness for his foursome; nor do I assume it. Some things I do say below might make one wonder whether in-group/out-group might not be a adaptive problem domain rather than a module. I reserve judgment.

Table 1: Four moral modules and the emotions and virtues associated with them

	Suffering	Hierarchy	Reciprocity	Purity
Proper domain (original triggers)	Suffering and vulnerability of one's children	Physical size and strength, domination and protection	Cheating vs. cooperation in joint ventures, food sharing	People with diseases or parasites; waste products
Actual domain, (modern examples)	Baby seals, cartoon characters	Bosses, gods	Marital fidelity, broken vending machines	Taboo ideas (communism, racism)
Characteristic emotions	Compassion	Resentment vs. respect, awe	anger, guilt vs. gratitude	Disgust
Relevant virtues	Kindness, compassion	Obedience, deference, loyalty	Fairness, justice, trustworthiness	Cleanliness, purity, chastity

The well-confirmed working assumption is this: The social life of *Homo sapiens* and our ancestors requires attention to these domains. It is essential to understand that these tendencies, conceived as “reactive attitudes” (Strawson, 1965; Flanagan 2002, 2002a, in press) originated in distant ancestors, not *de novo* in the lineage *Homo sapiens sapiens*. Indeed, Darwin (1871/1872) saw both facial and behavioral homologues of some of these attitudes in canines. Presumably the dispositional modules with their associated emotions evolved as adaptations in ancient environments to solve common, possibly universal, adaptive challenges. The modules are universal, and that is the level at which universality in a strict sense ends. The reason is this: different natural and social ecologies pull for different solution strategies to local adaptive challenges. That said, it would be a mistake to think all normative solutions are designed with ethical ends in mind. If I am well-off I may endorse norms to keep my self and my loved ones that way and keep those worse off in their place. If I have the power to endorse those norms and make the common folk think they are divinely inspired, I and my people are ‘winners’ and the

‘losers’ think things are as they are supposed to be. Thrasymachus in Plato’s *Republic*, Marx, and Foucault think this mode of normative design is quite common. It is something to be watchful for.

In any case the overall picture from Haidt is this: we have evolved with a set of dispositions to respond intuitively by moralizing these four (+) domains. This is done differently across different social landscapes and ecologies and this helps explain the plurality of moral conceptions. Haidt thinks at least some of the moralities that develop on top of the four dispositional modules are “incommensurable.” I am not sure. Even something less than “incommensurability” would help explain moral conflict. And either way there will be problems of moral communication. I’ll come back to these issues shortly.

One might, adopting language from John Rawls, think of our expansive moral conceptions as involving “comprehensive conceptions of the good.” Ethics, politics, and spiritual and aesthetic traditions interact to provide their basic structure, and this basic structure will determine for most people, in large measure, how they draw their norms from the spaces of meaning that constitute the Space of Meaning^{21st century}. Aesthetic attitudes, and attitudes towards science and technology, will be visibly affected by how persons locate themselves within the spaces they see as providing sense and meaning, and these attitudes will result in an interplay whose result is normally gradual transformation of our ways of making sense of things and finding meaning.

Inside America, Haidt sees evidence for differences in moral habituation between politically liberal and politically conservative adults, which explains differences in moral attitudes among generations. Haidt and Joseph write that one way in which cultures vary is by differential activation of the four modules:

In our own research we have found that American Muslims and American political conservatives value virtues of kindness, respect for authority, fairness, and spiritual purity. American liberals, however, rely more heavily on virtues rooted in the suffering module (liberals have a much keener ability to detect victimization) and the reciprocity module (virtues of equality, rights, and fairness). For liberals, the conservative virtues of hierarchy and order seem too closely related to oppression and the conservative virtues of purity seem to have too often been used to exclude or morally taint whole groups (e.g., blacks, homosexuals, sexually active women).²⁰

The larger “positive psychology” movement, within which Haidt is a key player – by my lights, the philosophically most sophisticated participant -- typically identifies key virtues and then lists what are called “strengths of character” which are ways the virtue is realized, or better perhaps, tools differentially required for embodying the virtue. Recall (from lecture two) that the Positive Psychology Movement, one major strand at any rate, claims that these virtues are universal:²¹

²⁰ Haidt and Joseph rightly cite George Lakoff’s *Moral Politics: What Conservatives Know That Liberals Don’t* (Chicago, 1996) as making closely related points. I recall reading in the mid-70’s (with shock) Lord Patrick Devlin’s *The Enforcement of Morals* and being sent for a loop. Devlin, a very respected liberal British jurist, was asked to respond to the Wolfenden Report that had argued for decriminalizing consensual male homosexuality on Millian grounds, it didn’t harm anyone. Perhaps to his own surprise, Devlin recommended caution: If a practice -- he thought homosexuality fit the bill in the UK in the late 50’s -- provokes strong feelings of “intolerance, indignation, and disgust,” don’t legally permit it. *Rationale*: distinguish between the “reasonable person” and the “rational person,” where the first is the unreflective ordinary person and the latter is the reflective person who might be able to produce the argument to the effect that a practice does no real harm. The moral fabric of society is held together by what “reasonable people” think, not what the handful of “rational people” think. Haidt’s model doesn’t require endorsing Devlin’s conclusion. But it helps make sense of it. One observation on Devlin: he seems to think a practice like homosexuality would have aroused the disgust response in virtually everyone at the time he wrote. But for this to be true, the 5% homosexual population should have the disgust response too. This is possible, I guess, if even gay people self-ascribe disgust to what they are. But the rational person will, I think, be able to explain why this is sad, the result of lack of compassion, as well as due to false views (possibly self-ascribed) about the choice involved in sexually orientation.

²¹ Haidt does not endorse the idea that the universal virtues cited are, in fact, universal. Even Peterson and Seligman sometimes use the language of ‘ubiquity’ rather than universality. Haidt’s reasoning, which I endorse, is this: the intuitive modules are universal; some very few adaptive problems (needing to stay alive) might also be universal. Beyond that what norms, values, and virtues evolve will be variable, based on mixes of who has power,

- Wisdom
- Courage
- Humanity
- Justice
- Temperance

For each major virtue, there is a set of “strengths of character,” which typically enable the virtue.

Consider the “strengths of character” lists for these three virtues: *Humanity*, *Justice*, and

Temperance (Peterson & Seligman, 2004, pp. 29-30):

Humanity.

Love: Valuing close relations with others, in particular those in which sharing and caring are reciprocated; being close to people

Kindness [generosity, nurturance, care, compassion, altruistic love, “niceness”]: Doing favors and good deeds for others; helping them; taking care of them

Social intelligence [emotional intelligence, personal intelligence]: Being aware of the motives and feelings of other people and oneself; knowing what to do to fit into different social situations; knowing what makes other people tick

Justice

Citizenship [social responsibility, loyalty, teamwork]: Working well as a member of a group or team; being loyal to the group; doing one’s share

Fairness: Treating all people the same according to notions of fairness and justice; not letting personal feelings bias decisions about others; giving everyone a fair chance

Leadership: Encouraging a group of which one is a member to get things done and at the same maintain time [sic] good relations within the group; organizing group activities and seeing that they happen

what local conditions demand, what resources for problem solving exist, past historical solutions. Everything interesting will be in detailed philosophical-anthropological analysis.

Temperance

Forgiveness and mercy: Forgiving those who have done wrong; accepting the shortcomings of others; giving people a second chance; *not* being vengeful

Humility/Modesty: Letting one's accomplishments speak for themselves; *not* seeking the spotlight; *not* regarding oneself as more special than one is

Prudence: Being careful about one's choices; *not* taking undue risks; *not* saying or doing things that might later be regretted

Self-regulation [self-control]: Regulating what one feels and does; being disciplined; controlling one's appetites and emotions

I don't know if Peterson and Seligman would approve of this interpretation, but one way to think of their lists of 'virtues' and 'strengths of character' is this way: 'Virtues' name virtues at a superordinate level, and strengths of character names 'virtues' at a lower level. Courage is a superordinate virtue; in one place and time it will be embodied primarily as a military virtue, at another place and time as a form of "integrity," say, standing up for one's convictions. With regard to the three virtues picked out above -- Humanity, Justice, Temperance -- the strengths of character read my way explain such things as these:

- Humanity can be wide or narrow, local or global, in-group or out-group (see footnote above on Haidt's expansion of his model to include an in-group/out-group module). The virtues of empathy, sympathy, fellow-feeling and compassion are not specified as the virtues they are until a culture gives them form, scope, and range.
- Justice is compatible with hierarchy or egalitarianism.

- Temperance will depend on how in particular times and place “disgust” reactions are conceived of – whether they have narrow or wide scope, or really simply what their scope is.²²
- The instantiation of specific virtue forms under the master virtues will interact and mutually inform each other. If the conception of justice is hierarchical, then it may be particularly contemptuous not to show the elders proper respect. If humanness is conceived (under a particular conception of justice) to have a very wide scope, involving love for all sentient beings, then one might feel poorly (disgusted or contemptuous) towards family or nation-state chauvinists.

Haidt’s model explains all this, although I am not aware that he himself has mapped it all out.²³ I now use this much to revisit a debate I had several years ago with Paul Churchland concerning “moral network theory” (Flanagan, “The Moral Network,” and P.M. Churchland, “Flanagan on Moral Knowledge,” in McCauley (ed), *The Churchlands and Their Critics*, 1996).

Churchland used connectionist mental modeling, San Diego-style, to provide a neural network model of how a “moral system,” such as a person might acquire moral knowledge.

Churchland and I, as well as Haidt, think John Dewey was right when he insisted that morality is

²² On the first day of ethics class I tell me students about a marriage practice that has been common for 2000 years among nomads in Nepal -- at times as many as 25% of Nepalese engage in the practice -- where a daughter marries all the brothers in a family (my evolutionary psychology friends claim it is way too unstable to last more than a week! They are mistaken). American students find this form of polyandry disgusting. I probe. It makes the women feel yucky because of what seems like infidelity as well as because of prospects for sexual exhaustion; the young men simply hate the thought of sharing their wife, especially with their brothers. A different example: Suppose American college students started coming topless to class in the spring. I would find this distracting, but not disgusting. It would be unconventional, and if it just started with no warning – possibly a sign of bad manners, but I, at least, wouldn’t moralize it. At a nude beach it is not even distracting.

²³ I won’t take up the question of the status of Haidt’s modules as modules except to say this much: He acknowledges that his modules are not *special-purpose* in the way Fodorian perceptual modules are. Mistreatment of my sister (injustice) will be perceived as such and arouse disgust and contempt. If, at the same time, I see you as unworthy slime relative to my clan, hierarchy intuitions/judgments will be in play. The first point is that he who mistreats my sister is dead. The second point is that a single situation can activate (result in the interpenetration of) all four Haidt modules at once.

a natural phenomenon to be studied naturalistically. Two passages from Dewey are always worth quoting in tandem:

“For practical purposes morals means customs, folkways, established collective habits.

This is a commonplace of the anthropologist, though the moral theorist generally suffers from an illusion that his own place and day is, or ought to be, an exception” (1922, 55).

“Moral science is not something with a separate province. It is physical, biological, and historical knowledge placed in a humane context where it will illuminate and guide the activities of men” (1922, 204-205)

The basic idea is that morals are comprised in large measure of habits of heart, mind, and behavior. Moral habits in virtue of being complex social constructions on top of our basic psychobiological nature are not guaranteed to be ‘moral’ in the sense of being ‘good.’ If humans are correctly modeled by connectionist systems, we will learn what we are encouraged to learn and we will admire and avow norms that are socially endorsed. We will need therefore to understand how critical leverage is gained, if that is it can be gained.

A connectionist model can learn, and how it does so depends on (a) its initial settings, (b) experience, and (c) feedback on how well it is catching on to, in this case, moral norms that are socially endorsed *and* that the system is not strongly disposed to resist acquiring. The last point gets us to the vicinity of the difference of opinion between Churchland and me: The initial settings of a human learning device, conceived along connectionist lines, should have certain reactive attitudes set if they are to model us. Thus ‘disgust’ responses to certain stimuli will need to be set. How far can we learn to deviate from the initial settings? There is no general answer since deviation from each reactive attitude depends on a host of factors including the strength of

the particular initial settings. It just may be that a plausible set of initial settings as inspired by evolutionary psychology will make it highly improbable that human females can acquire norms like those of black widow spiders or praying mantises and learn to decapitate their mates after impregnation (forget about the fact that if the guys catch on they'll resist). Even if there would be no adaptive loss in terms of sex and child-rearing, the practice might just be too disgusting to catch on.

Churchland has no strong opinion (last time I checked) on what the initial settings should be for learning morality – he certainly does not subscribe to a “moral *tabula rasa*” picture, so I think it would be a friendly suggestion that the initial settings be the one’s Haidt proposes, priming in the four intuitive modules for suffering/compassion, reciprocity/fairness, hierarchy/respect, and purity/impurity, and maybe the recently proposed in-group/out-group module.

My initial disagreement with Churchland revolved solely around two issues – I had no problem with the possibility that connectionism was a good model for moral learning, and we are both well-bred naturalists. First, Churchland is overly optimistic about the moral community’s ability to arrive at high-quality moral knowledge by the ordinary processes of moral socialization; second, Churchland fails to emphasize sufficiently the local nature of much moral knowledge.

I still think these things. I’ll see if I can draw Churchland back into the discussion since, now armed with Haidt’s model, I can explain better why I have these two concerns about a connectionist moral-learning system. Churchland optimistically assumes that such a system in virtue of being a learning system with self-correcting capacities will naturally escape local chauvinism, and that the mechanisms of learning will lead naturally to moral improvement and

refinement. Set up the right way, with the right original architecture, such a system can have these positive results, but will need to supply the system with a few fancy tools, found to be necessary from other parts of psychology, social psychology in particular. I'll explain.

Haidt's model predicts and explains that different social ecologies will seek to solve their adaptive problems by tuning up or down the four modules. This explains the plural shapes of socio-moral-political systems. Assume, at the time a moral code forms, that it is, in fact, a good, satisfactory, solution to the adaptive problems facing the group or society.

But this is a much too generous assumption, since the adaptive problems a society with a caste system already in place faces will be very different from those faced by one with a prior commitment to egalitarianism in place. Without getting into my reasons for saying this, if it isn't obvious, I say this much: the culture with the caste system already in place is already off-on-the-wrong foot in solving its problems, unless the new adaptive problems happen to be ones that are revealing "problems" with the caste system.

But "make-believe" for a moment that societies respond to new adaptive problems in ways that are decent design solutions to the problem. Once a moral system is in place it is transmitted to the youth. Suppose, as seems likely, that their adaptive problems are different than those that initiated "the original good solution." How will the youth see that this is so and adjust their socio-moral-political perspective?

A non-naturalist might seem to have an easier time with this problem than the naturalist. She can say that there is really only one adaptive problem: being good in the 'eyes' of God and attaining eternal reward. When there are questions about the adequacy of a moral code, evaluate then in terms of divine standards and one will be able to set one's self straight. This is not a tactic available to the naturalist. We are always, because it is the nature of the human condition,

in “the internalist predicament.” Normative assessment, moral improvement, and so on, must come from inside the dialectical space of our own norms, the norms of those who live differently, and in the space of meta-norms for resolving disagreements or deciding to live tolerantly without agreement. Put less chillingly: the “internalist predicament” can be described this way. *If* one lives within a relatively homogeneous form of life, a relatively homogeneous *Space of Meaning*, then there are usually only two ways normative adjustment can/will occur: (1) There might be competitions in small open spaces within an arena where a norm and its rationale are settled, e.g., marriage is a good and it is forever. The small space might be one such as this: Everyone agrees about the latter, but everyone is also aware that there are some awful marital situations, so a discussion is begun about whether divorce should remain totally impermissible or might be permitted in rare “extreme” situations; (2) A new, unexpected adaptive challenge is faced, for example, new commercial practices intrude the formerly isolated space and new issues of “economic fairness” towards members of an out-group arise. As soon as there is a sufficient level of commercial interaction as envisaged between the two formerly isolated “relatively homogeneous” groups, they face (we’ll suppose) coordination problems regarding how to do business with each other. Each group could set up independent norms governing business transactions with the other group that have no effect of their in-group dealings,²⁴ or, as likely, these initially independent norms might affect and re-configure their own normative practices in different ways depending on their respective normative starting point. The dialectic is still internal to norms, but it is no longer internal to one set of normative practice. There are *external pressures* to make internal normative adjustments (in both directions).

²⁴ By “independent” I mean something like this the new commercial norms govern only trade with the new group, but don’t pertain to the original groups commercial practices with each other. There are many current examples of trade agreements of this form.

Then there are prospects both in the ‘relatively homogeneous’ singleton cases, and in cases where there is socio-normative interaction among groups that abide different norms, and especially in very heterogeneous cultures, or when globalization is in play, to develop meta-norms. These take many forms: they include norms that are thought to yield peace rather than war (even without normative agreement), as well as norms for ‘rational’ discussion of what is merely conventional, makes sense, passes certain agreed upon epistemic tests, and so on. Meta-norms in so far as they are divined and abided constitute some form of external perspective. They can be political, ethical, epistemic, or most often, mixed.

I can draw out the important implications of these points for *eudaimonistic scientia* by using Haidt’s model to deepen the earlier discussion of “the internalist objection,” and at the same time to explain why norms typically have more local coloration than Churchland thinks, and why I don’t think we should be as optimistic as Churchland is that the process of normative development, conceived along connectionist lines, will yield reliably moral progress (some of the reasons involve the role of self-interested power maneuvers in the construction and enforcement of norms).

Background: In 1971 when Rawls first published his *Theory of Justice* he explicitly acknowledged making use of the method of “reflective equilibrium” as I have described it, and he claimed that *all* rational people -- if placed behind the “veil of ignorance” into the “original position,” where they have no idea how they will fare in the lottery of life – will choose the same two principles of justice to structure socio-moral-political life. The liberal theory of justice will be chosen because it best allows each person to develop and live in accordance with his or her own “comprehensive theory of the good.”

By the time I began to have regular discussions with Rawls in the early 80's, he would say this: the two principles of justice advanced in *Theory of Justice* would be endorsed by citizens of liberal democratic states such as America and Sweden *if* they were to go behind the veil of ignorance into the original position, and were to use the method of reflective equilibrium. What happened? How did what were initially to be the principles of justice all rational people would accept become ones that persons already living in liberal democratic states would choose?

Haidt's model offers an explanation, although I don't know if he himself has given it or anticipated it. The straightforward way to make the point is this way: There simply may be no such thing as universal ethical intuitions at the level Rawls was initially looking to locate them. What I mean is this: at some basic level we may get universal moral reactions: Consider, Mencius's example of the child falling into a well, one may gain universal assent that 'this is bad' and 'I (even Hitler) feel like doing something to stop it'. Disgust at watching innocents murdered or revulsion at gratuitous thievery, may get similar reactions. But Rawls was looking for something much more sophisticated: a universal theory of justice (which it seemed might be stated in terms of two principles, a principle of justice – equal liberty, and “the difference principle” -- no social or economic inequality unless it helps the worse-off). His concession that the theory is one likely to be seen, espoused, or intuited by Americans and Swedes is best understood as born of the recognition that there is no method that accords with the mandatory practice of seeking “wide reflective equilibrium,” which is not informed by prior settled intuitions.

Indeed, once this concession is made, one sees the rationale behind this sort of resistance to Rawls's procedure of asking folks to go behind the “veil of ignorance into the original position:” Either I might object from the start to the procedure because you are in fact, asking me

to “put aside” some of my most cherished convictions, such as my belief in the moral neutrality of capitalism, my opposition to certain kinds of sexual freedom, and so on. Or, I exclaim after I engage in the game according to your rules – “Conjurer! Trickster!” Why? Because your procedure required that I change, give up, etc., some of my most cherished moral intuitions.²⁵ And I see that you tricked me by making me set aside for purposes of the exercise my most cherished moral intuitions.

Ok, the question is, what implications does all of this have for the possibility of *eudaimonistic scientia*, especially for such matters as the prospects for making legitimate moral criticism and moral progress, and what implications, if any, does it have for the legitimacy of the method of wide reflective equilibrium? Even if there is no strict moral incommensurability, there is abundant evidence that there is moral conflict. How do we deal with it?

Now Haidt makes much of the fact that much ordinary “moral judgment” is not judgment at all. Faced with moral problems, human react from the intuitive moral system. Because we are socialized to offer reasons for our judgments when called upon to do so, we do so. But our reason giving is normally not best understood as citing what *caused* our reaction (intuitions did that). Rather, reason giving is often just the giving of post-hoc rationalizations.

Haidt doesn't so much approve of this, he simply thinks this is what the data show.²⁶ On the one hand, this social-intuition theory reveals why one might worry, as I do, that models like Churchland's explain moral agreement and moral consensus, but leave us poorly equipped to deal with moral conflict. A different, but related point is this: many of the adaptive pressures

²⁵ I don't recall any discussion in Rawls about getting the latter sort of reaction from Americans. It is curious that Rawls thought that Americans will have the “right intuitions,” since few Americans seem attracted to “the difference principle.” Northern European welfare states in fact embody some form of the “Difference Principle,” so I could see the “Difference Principle” having intuitive appeal there. Many Americans read the principle, correctly, as “socialist, and American intuitions about desert (from Locke et al.) hardly favor it.

²⁶ See also Joshua Greene, (200x). I am not convinced yet by the data. Most of it comes from fMRI on people reacting to moral dilemmas of the trolley-problem variety. Most moral problems are not dilemmas and trolley type problems are rare. The problem with the data set at this point is time pertains to ecological realism and validity.

and much moral socialization has decidedly local causes and thus local or regional coloration. In America much political and ethical haggling is best understood by paying close attention to local and regional history and current adaptive challenges.

Here is a thoroughly naturalistic idea that might help us understand and deal with both moral conflict and failures of moral insight. There are two findings in cognitive science that can be blended: (1) There are well-known attribution biases that lead people to believe false things. One bias, the perseverance bias will lead people to continue thinking that they are good or bad at a task, a social intelligence task, for example, even if they are told that the data that led them to make the initial judgment were a complete fabrication! The only reliable way to get the false perseverative belief to yield is to teach the victims of the effect about the perseverance effect itself; (2) “Team Damasio” has shown that bad social decisions are produced in people with disturbances in communication between emotional and cognitive systems.

Putting (1) and (2) together gives some support to the philosopher’s old adage that moral insight requires hard work. If nature has gifted us with a moral system that operates mostly intuitively and if intuitions are not always reliable, then we are positioned to propose a meta-norm: Pay close attention to your intuitive moral responses and to the confidence you experience about the validity of your norms and values.²⁷ Evolutionary dispositions in concert with hard work at socialization on the part of the moral community are guaranteed to make you feel strong moral convictions, reinforced by your fellows who share your “comprehensive view

²⁷ Hagop Sarkissian points this out to me: “You seem to use ‘intuitions’ in the sense of innate ‘reactive-attitudes’, but Damasio seems to use ‘intuitions’ interchangeably with ‘somatic markers’, which are forged through learning. If so, then what we have is not, strictly speaking, a continual exercise of gaining rational foothold over these base intuitions, but rather of continually reforming them.” This is a really important point, which I need to think about. First pass: I may be unwise to adopt without comment Haidt’s use of ‘intuition’ and intuitive.’ I normally read him as using these terms in a way that could cover the initial settings of the reactive attitudes as the way they, still considered as reactive attitudes, are modified by learning. The fact is that Haidt usually sees the relevant action in moral response as occurring in the emotional regions of the brain. Damasio’s somatic marker hypothesis allows that learning moral habits is more systemic involving PFC as well as, say, the amygdala, etc. More needs to be thought, then said.

of the good.” This situation works for compassionate Buddhists and Nazis in pretty much the same way.²⁸ It will be wise to recognize this much, and thus be willing to bring, as necessary, your moral intuitions under rational critical scrutiny.

How? engage in “wide reflective equilibrium” whenever the opportunity presents itself. The method is not a guarantee of success at getting to “the right view” – remember it gives a large role to intuitions itself, but it asks that the intuitions pass certain reflective tests. If there is any method that can lead to moral justification, moral improvement, and the like, it is the one.²⁹

Sometimes the right meta-norms, perhaps gathered from cross-cultural reflection, can help engender sensitivity to false beliefs, e.g., that some humans are less than human. Slavery and caste systems can in theory yield to such scrutiny. There are two other ways to gain rational leverage in intuitively well-oiled machines: Teach the youth both the norms abided by the elders while at the same time giving them some instruction in being watchful for internal consistencies and adaptive inadequacy.³⁰

Internal inconsistencies: Here one works with the fact that our psychology about socio-moral norms is ordinarily intuitive. So we instruct the youth to watch out for intuitive discomfort, cognitive and affective dissonance, cramps in the heart-mind. Internal inconsistencies in the relevant sense might be narrow, relating pretty much to a single norm or practice, or wide and systemic. Take the example mentioned earlier of strict prohibitions against

²⁸ The insight here is ancient: One sees it throughout Confucius’s *Analects* when he worries about the effects of “village worthies,” charismatic leaders, teachers and so on who lack *ren, li, and de* – humaneness, respect for the rites, and virtue. The “village worthies” may seem to possess virtue but they don’t, and as a result, the citizens and young people are misguided in their socio-moral-political development. Mozi, to my mind, the most underestimated of the classical Chinese philosophers sees this problem clearly and accuses Confucius of not seeing how serious this problem is (see Flanagan, 2004 mss. (rightly unpublished)).

²⁹ Compare to Reichenbach’s pragmatic justification of induction: It is impossible to prove that induction will work; what we can show is that if any method can produce knowledge it can/will.

³⁰ Introducing such principles will of course work best if one is already a fallibilist and a naturalist. Some normative systems make the mistake of thinking they have everything right and, because they are based on divine instruction are unrevisable. There is some historical evidence that external pressure can sometimes get such beliefs about, say, literal interpretation to yield. But it *is* a serious problem.

divorce. This is bound to create a certain amount of misery. One solution is to use the “misery” to begin discussion of allowing certain exceptions consistent with the rationale behind the prohibition. A wide, systemic inconsistency might reveal itself this way: Nussbaum’s work on women and human development reveals that many objectively ill, malnourished Indian widows think their health is ok; widowers, on the other hand, get that they are in poor health. One might think this is a straightforward epistemic mistake (and at a certain level it is). Hinduism, especially as practiced in villages but not in cities, has all sorts of odd rules regarding the behavior and status of widows. But even without bringing in certain norms of equality from Delhi to the villages, this interesting phenomena occurs: give both the men and women good water, decent nutrition, and the most modest health care, and now the widows and widowers both judge their health the same, sadly but accurately as not very good. The ethical point is that a small change in the objective conditions in this case, despite sexist norms in the villages, cause the intuitive judgments of the widows to change and, in this case, to become more accurate. Finally, I could use the same example, but this time bring the villagers into normative conversation with their fellow Hindu city-dwellers. If the city dwellers can legitimately say that there is nothing internal to Hinduism (even of the form the villagers accept) that warrants the village widows to self-conceive in the way they do, then there are reasons of consistent commitment to Hindu practice for the village widows to change their low estimate of their status. This is not a problem for Churchland’s connectionist moral network theory so long as the system is built to model the acquisition of critical meta-norms as well as norms. I don’t think it was designed to model this important feature in its first iteration

Adaptive inadequacies: Examples of adaptive inadequacies are a dime a dozen and affect all human practices. Before Roman aqueducts, everyone wanted and needed to transport water

over difficult terrain. In retrospect, most even workable ideas were not very good. The rise of the modern industrial state along with capitalistic economies brought new problems of poverty and poor care. In the 17th century in France and Holland there was immediate attention to the problems of the poor. These were motivated by both prudential concerns (the poor are dangerous and disgusting) and moral concerns (Jesus says care for the poor – trust me, it is in the relevant French and Dutch transcripts (or if you don't trust me, see Gouda, 19xx). Both countries solved the adaptive problem by building into their respective infrastructure high tax rates, which enables to the present day strong welfare states. America, when faced with similar problems and similar possible solutions has continued to fail to solve, some would say, even address, the problem. Why don't we have universal health care? One reason is that most Americans have false views about desert and merit. Our socialization bears too much weight from John Locke and Horatio Algier children's stories. Our system is broken and it will have to change. I am not saying that we will do so for moral reasons, although those are in play in some of the discussion. Even if the "public reasons" for changing our practices are stated as purely prudential – pertaining to public safety and good economic practices, we will eventually think that universal health care is a 'good thing,' something we as a society ought to provide. That's just the way things work.

Ok, I am going on a diatribe, and in any case that is enough for tonight. What I have claimed to have done is this: I asked whether "normative mind-science," in particular *eudaimonistic scientia* is possible. My answer is 'yes'. In fact, it is actual and has been practiced – empirically if not in the usual scientific ways – in the East and in the West for 2500 years. I have tried to explain why the methods of *eudaimonistic scientia* are different, indeed why they must be different, from the methods of non-normative, observer-independent sciences.

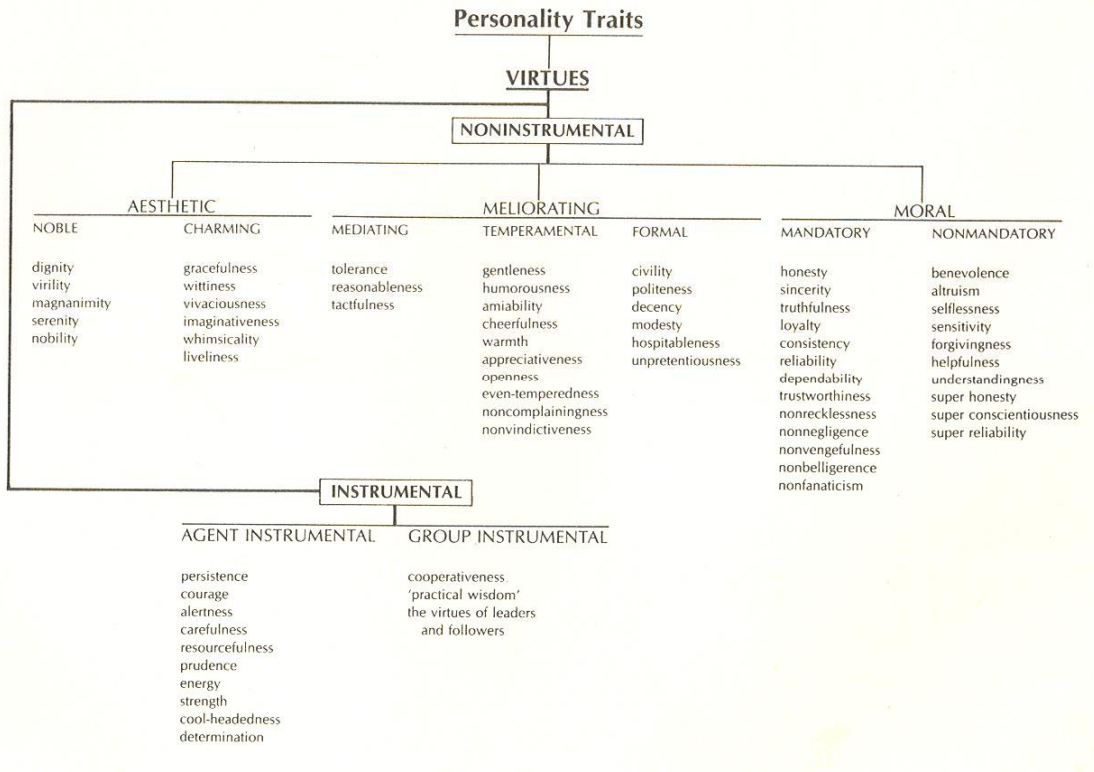
The method of *wide reflective equilibrium* plays a key, if not the main role in advancing moral knowledge. I tried to tame the concern that this method shows how pathetically methodologically impaired *eudaimonistic scientia* is. Finally, by using Haidt's social intuitionist model of morality, I showed how we can understand better how morality operates and how best, given our resources, we can deal with such matters as normative conflict and normative improvement.

In the next lecture I do two things: First, I examine new work on the neuroscience of happiness. *Eudaimonistic scientia* should be able to speak about what underpins well-being. And it does. But, there are predictable problems about measuring happiness due in part to disagreement about what it even is. Second, I examine the literature on "positive illusions." Apparently many healthy-minded people have false beliefs. These may even be partly constitutive of what makes them happy and healthy minded. Illusions, unlike delusions, are subject to modification, albeit with resistance. But even they will make epistemically conservative types nervous.

This discussion will nicely set us up for the topic of the final lecture: "Spirituality Naturalized?" since many worry that spirituality and religion involve objectionable illusions and/or delusions. We'll see.

Thank you very much for your attention this evening.

Appendix 1: Pincoff's List of Virtues Canonized (in various ways) in the West



Appendix II: Buddhist *Abhidhamma* COMPENDIUM OF MENTAL FACTORS

**TABLE 2.1:
THE 52 MENTAL FACTORS AT A GLANCE**

ETHICALLY VARIABLES—13

Universals—7

- (1) Contact
- (2) Feeling
- (3) Perception
- (4) Volition
- (5) One-pointedness
- (6) Life faculty
- (7) Attention

Occasionals—6

- (8) Initial application
- (9) Sustained application
- (10) Decision
- (11) Energy
- (12) Zest
- (13) Desire

UNWHOLESOME FACTORS—14

Unwholesome Universals—4

- (14) Delusion
- (15) Shamelessness
- (16) Fearlessness of wrong
- (17) Restlessness

Unwholesome Occasionals—10

- (18) Greed
- (19) Wrong view
- (20) Conceit
- (21) Hatred
- (22) Envy
- (23) Avarice
- (24) Worry
- (25) Sloth
- (26) Torpor
- (27) Doubt

BEAUTIFUL FACTORS—25

Beautiful Universals—19

- (28) Faith
- (29) Mindfulness
- (30) Shame
- (31) Fear of wrong
- (32) Non-greed
- (33) Non-hatred
- (34) Neutrality of mind
- (35) Tranquility of mental body
- (36) Tranquility of consciousness
- (37) Lightness of mental body
- (38) Lightness of consciousness
- (39) Malleability of mental body
- (40) Malleability of consciousness
- (41) Wieldiness of mental body
- (42) Wieldiness of consciousness
- (43) Proficiency of mental body
- (44) Proficiency of consciousness
- (45) Rectitude of mental body
- (46) Rectitude of consciousness

Abstinences—3

- (47) Right speech
- (48) Right action
- (49) Right livelihood

Illimitables—2

- (50) Compassion
- (51) Appreciative joy

Non-Delusion—1

- (52) Wisdom faculty