

---

# Can We Be Epigenetically Proactive?

Kathinka Evers

---

The human brain is an essentially evaluative organ endowed with reward systems engaged in learning and memory as well as in higher evaluative tendencies. Our innate species-specific, neuronally-based identity disposes us to develop universal evaluative tendencies, such as self-interest, control-orientation, dissociation, selective sympathy, empathy, and xenophobia. The combination of these tendencies may place us in a predicament. Our neuronal identity makes us social, but also individualistic and self-projective, with an emotional and intellectual engagement that is far more narrowly focused in space and time than the effects of our actions. However, synaptic epigenesis theories of cultural and social imprinting on our brain architecture suggest that there is a possibility of culturally influencing these predispositions. In an analysis of epigenesis by selective stabilisation of synapses, I discuss the relationships between genotype and brain phenotype: the paradox of non-linear evolution between genome and brain complexity; the selection of cultural circuits in the brain during development; and the genesis and epigenetic transmission of cultural imprints. I proceed to discuss the combinatorial explosion of brain representations, and the channelling of behaviour through “epigenetic rules” and top-down control of decision-making. In neurobiological terms, these “rules” are viewed as acquired patterns of connections (scaffoldings), hypothetically stored in frontal cortex long-term memory, which frame the genesis of novel representations and regulate decision-making in a top-down manner. Against that background I propose the possibility of being epigenetically proactive, and adapting our social structures, in both the short and the long term, to benefit, influence, and constructively interact with the ever-developing neuronal architecture of our brains.

## Keywords

Cultural circuits | Empathetic xenophobia | Epigenetic proaction | Epigenetic rules | Neuroethics | Precaution | Responsibility | Selective sympathy | Species-specific identity | Synaptic epigenesis

## 1 Introduction

Contemporary neuroscience no longer views the brain as an input-output processing device but as an autonomously active, self-referential, and selectional system operating in a projective style, which is in constant social interaction and in which values are incorporated as necessary constraints. The idea that evolution by natural selection has given rise to an essentially evaluative cerebral architecture raises the question whether, in the human species, such neurobiologically-based predispositions have further developed the means to generate novel specific values on higher cognitive levels. The concept of “value” would then play a central role as something that is taken into account in decision-

making and that influences a choice, selection, or decision, that can occur on many levels—non-conscious as well as conscious—as a basic biological function or as a feature of advanced moral reasoning. But, if we are born evaluators, to what extent can these predispositions with which we are all born be culturally controlled?

In this article, I suggest that our innate species-specific neurally based identity disposes us to develop universal evaluative tendencies, such as self-interest, control-orientation, dissociation, selective sympathy, empathy, and xenophobia. The combination of these tendencies may place us in a practical and moral predicament. Our neuronal identity as persons makes

## Author

[Kathinka Evers](#)  
kathinka.evers@crb.uu.se  
Uppsala Universitet  
Uppsala, Sweden

## Commentator

[Stephan Schleim](#)  
s.schleim@rug.nl  
Rijksuniversiteit Groningen  
Groningen, Netherlands

## Editors

[Thomas Metzinger](#)  
metzinger@uni-mainz.de  
Johannes Gutenberg-Universität  
Mainz, Germany

[Jennifer M. Windt](#)  
jennifer.windt@monash.edu  
Monash University  
Melbourne, Australia

us social, but also individualistic and self-projective, with an emotional and intellectual engagement that is far more narrowly focused in space and time than the effects of our actions.

However, the neuronal organisation of our adult brain develops in the course of a fifteen year-long period following birth, during which, and, to a lesser extent, after which it is subject to cultural influence, both on the individual level and, at the social group level, across generations (Lagercrantz 2005; Lagercrantz et al. 2010; Collin & van den Heuvel 2013). Synaptic epigenesis theories of cultural and social imprinting on our brain architecture (which differ from less discriminative epigenetic modifications of nuclear chromatin) (Changeux 1985; Kitayama & Uskul 2011) suggest that there is an interesting possibility, which, in my opinion, has hitherto been underestimated. That is, we could potentially be *epigenetically proactive* (Evers 2009) and adapt our social structures, in both the short and the long term, to benefit, influence, and constructively interact with the ever-developing neuronal architecture of our brains.

## 2 The social individualist

### 2.1 An egocentric evaluator

The human brain is intrinsically active: it produces electrical and chemical activity both in response to external stimuli and, spontaneously, independently of them. The brain is an autonomously-active motivated neuronal system, genetically equipped with a predisposition to explore the world and to classify what it finds there (Changeux 1985, 2004). On-going spontaneous activity is present throughout the nervous system. In the embryo, spontaneous movements (Narayanan & Hamburger 1971) and waves of endogenous retinal activity (Galli & Maffei 1988; Goodman & Shatz 1993) are thought to play an important role in the epigenesis of neural networks through synapse selection (see below). On-going spontaneous activity is also present in the adult brain, where it is responsible for the highly variable patterns of the electroencephalogram (EEG; Berger 1929; Raichle et al. 2001). Thalamocor-

tical networks generate a variety of oscillations, whose rhythms change across the sleep-wake cycle (Llinas & Paré 1991). Optical imaging methods in anesthetized animals also reveal fast spontaneous states of neuronal activity that, far from being random, exhibit patterns that resemble those evoked by external stimuli. In parallel, functional neuroimaging studies in humans have shown a globally-elevated brain metabolism at rest, with localized patterns suggesting that particular cortical regions are maintained in a high, although variable, state of activity referred to as “default mode” by Raichle et al. (2001).

Hypotheses of knowledge acquisition posit that patterns of spontaneous activity, referred to as “pre-representations”, arise in the brain and are selected by reward signals as “representations” confirmed by both external experience and internal processes of evaluation within a conscious neuronal workspace (Dehaene & Changeux 2011). Such “models of the world” are stabilised through “cognitive games” by analogy with Wittgenstein “language games”, as permanent features of the developing cognitive apparatus, according to a process referred to as “mental Darwinism” (Changeux 2004).

Anticipation of reward signals introduces a delay between the elaboration of tacit plans of action and actual interaction with the world performed by the organism, which presupposes a distinction of temporal states: awareness of the present, remembrance of the past, and anticipation of the future (Barto & Sutton 1982; Schultz et al. 1997; Dehaene & Changeux 2000; Schultz 2006). Without any capacity to evaluate stimuli, the brain could neither learn nor remember: it has to prefer some stimuli to others in order to learn. This classical idea in learning theory has been expressed in neuronal terms by Dehaene & Changeux (1991), and by Edelman in his accounts of primary consciousness (Edelman 1992). In these accounts, learning is a change in actual behaviour, or the storage of a trace subsequently unveiled (Dudai 1989, 2002) through brain categorizations of stimuli. These are given in terms of positive or negative values, understood as something that is taken into account in decision-making and that influences a choice, selection or decision, which can occur on

many levels. Through its intense and spontaneous activity, the brain has also been described as a narrative organ, spinning its own neuronal tale (Evers 2009). The narrations will vary greatly between individuals, but each will be self-projective.

The natural egocentricity or individualism of the human brain appears quite pronounced. In its projection of autonomously-produced images, the brain refers all experiences to itself, that is, to its own individual perspective. This self-projection is a biological predisposition that humans possess innately and that is closely connected to our predisposition for developing self-awareness, which Edelman suggests is a necessary condition for developing higher-order consciousness (Edelman 1992; Denton 2006; see also Tulving 1983). The existence of a self-projecting systems monitoring internal processes in the brain was suggested by an early Positron Emission Tomography (PET) study of self-generated actions showing hemodynamic activity in the posterior cingulate cortex (Blakemore et al. 1998). This observation was confirmed and extended by magneto-encephalography following synchronization in the gamma range (55–100 Hz), thus defining a major network of the brain: the paralimbic interaction between the medial prefrontal/anterior cingulate and medial parietal/posterior cingulate cortices and subcortical regions (Lou et al. 2004; rev. Changeux & Lou 2011). Damasio (1999) distinguished a “core consciousness” (core self) from an “extended consciousness” (extended self) that we consider as analogous to the “minimal self” and “extended self” of Gallagher (2000). Minimal self-awareness is prereflexive, immediate and normally reliable, while still involving a sense of ownership of experience (Gallagher 2000). The “extended self” is a coherent self that persists across time and requires a system that can retrieve long-term memories of personal experiences—namely, episodic memory (Gardiner 2001). Consequently, episodic memory retrieval becomes an indispensable component of the more complex forms of self-awareness and consciousness (Tulving 1983).

In the course of growing up, the infant develops the capacity to focus its attention; it

learns to distinguish between and recognise objects in its environment, such as faces, and becomes aware of itself as standing in various relations to these objects. Conscious processing develops into auto-distinction (when “this-here” is distinguished from “that-there”). When further developed, the individual becomes aware of itself as a subject of experience and ascribes mental states to itself: auto-distinction evolves into self-awareness (when “this-here” becomes “I”) usually at around one and a half years of age (Lagercrantz 2005), and possibly even earlier (Falck-Ytter et al. 2006; see also Rochat 2001). From the age of six to twelve months, the child typically sees a “sociable playmate” in the mirror’s reflection. Self-admiring and embarrassment usually begin at twelve months, and at fourteen to twenty months most children demonstrate avoidance behaviours. Finally, at eighteen months 50% of children recognize the reflection in the mirror as their own and by twenty to twenty-four months this rises to 65%—this is revealed, for instance, by them trying to evince marks on their own nose, taking advantage, in all these instances, of their episodic memory abilities (see Tulving 1983).

An evolved survival function that adds an evaluative element to our brain’s self-projective mode of operation is self-interest, expressed as a desire to survive, to be well-fed, safe, to reproduce, and so on. This is not a defining characteristic, for there are exceptions, for example subjects who have a very poorly developed self-interest (Damasio 1994; Damasio & Carvalho 2013). Nor is it necessarily rational, since biological evolution is circumstantial. There is an abundant literature on the phenomenologically rich concept of self-interest in philosophy and ethics, in terms e.g., of enlightenment, egoism, capacity for altruism, etc. Such issues are relevant and interesting but beyond the scope of this discussion. In the present context, self-interest is understood in a minimalistic sense, as an evolved survival function that adds an evaluative element to our brain’s self-projective mode of operation.

Self-interest is also a source of the urge to control the immediate environment, and of the need for familiarity, security, and preference for

the known. The subjective experience of some level of control and the security that this provides is in fact a necessary condition for the individual to develop in a healthy manner and to consolidate an integrated sense of self (Ledoux 1998). When the external circumstances become severely disturbing, we feel increasingly threatened and have a defence mechanism that is eventually activated: *dissociation*, here understood as a process whereby information—incoming, stored, or outgoing—is actively prevented from integration with its usual or expected associations.

The human being is, in this sense, a “dissociative animal”: we spend a considerable amount of intellectual and emotional energy on distancing ourselves from a wide range of things that we consciously or non-consciously fear or dislike (Evers 2009). When an experience is too painful to accept, we sometimes deliberately do not accept it; instead of integrating it into our ordinary system of associations, we push it away from us, and prevent it from being integrated into our consciousness. Pushed to an extreme, this tendency may become pathological, e.g., in the development of Dissociative Identity Disorder (cf. DSM-IV), but as a non-pathological process it is an important adaptive function, and a valuable evolutionary asset allowing us to survive events that we would otherwise be unable to endure (Putnam 1989; Evers 2001).

So far, I have described the brain as an autonomously active, self-projective, and selectional neural system with innate evaluative tendencies, e.g., self-interest, control-orientation, and dissociation. These cerebral features characterize the individual, but they are also reflected in the social relationships proper to the human species.

## 2.2 Selective sympathy & empathetic xenophobia

In social animals, self-interest is a source of interest in others. In the case of humans, this social interest focuses primarily on those to whom the self can relate and with whom it identifies, such as the next of kin, the clan, the community, etc. The human brain conjugates op-

posite tendencies: first, embodied in the human subject, it is engaged in highly individualistic and self-projective actions, such as the search for water or food. But it also mediates co-operative social relationships: the “I” is extended to endorse the group, as a “we”, and distinctions are drawn between “us” and “them” (Ricoeur 1992; Changeux & Ricoeur 2000). Sympathy and aid is typically extended to others in proportion to their closeness to us in terms of biology, e.g., face recognition (Michel et al. 2006; Hills & Lewis 2006), racial out-group versus in-group distinctions (Hart et al. 2000; Phelps et al. 2003), culture, ideology, etc.

Imagining an action or actually performing that action both have similar neural circuits (which include the premotor cortex, supplementary motor area, cerebellum, parietal cortex, and basal ganglia) to those activated when one observes, imitates, or imagines actions performed by other individuals (Jeannerod 2006; Decety 2012). The model mechanism suggested is that actions are coded in terms of perceivable effects (Hommel et al. 2001). Performing a movement leaves a memory of the association between the motor pattern by which it was generated and the sensory effects that it produces. Such stored associations can then be used to retrieve a movement by anticipating its effects. This perception-action coupling mechanism, which includes active sensing and motor-sensory loops (Gordon & Ahissar 2012) and to which may be added the motor theory of language (Lieberman & Mattingly 1985), offers a mechanism for intersubjective communication and social understanding by creating functional links between first-person and third-person information (Decety & Sommerville 2003; Jackson & Decety 2004).

Functional Magnetic Resonance Imaging and magneto-encephalography among other methods have led to the demonstration that when children or adults watch other subjects in pain, the neural circuits mobilized by the processing of first-hand experience of pain are activated in the observer (Singer et al. 2004; Cheng et al. 2008). This sharing allows mapping of the perceived affective cues of others onto the behaviours and experiences of the self-oriented

response. Decety (2012) argues that, depending on the extent of the overlap in the pain matrix, and complex interactions with personal dispositions, motivation, contextual information, and self-regulation, this can lead to personal distress (i.e., self-centred motivation) or to empathic concern (i.e., an other-oriented response). This basic somatic sensorimotor resonance plays a critical role in the recognition and sharing of others' affective states.

There is an important neural distinction between apprehending and caring that makes it possible to understand the affective state of another without feeling engaged in it. Studies in the neurobiology of empathy (here understood as the ability to apprehend the mental states of others), and sympathy (the ability to care about others) suggest that these abilities involve complex cognitive functions with large individual and contextual variations that depend on both biological and socio-cultural factors (Jackson & Decety 2004; Singer et al. 2004; Singer et al. 2006; Iacoboni et al. 2005; Jackson et al. 2006; Lawrence et al. 2006; Parr & Waller 2006; Engen & Singer 2013). Such results are important, because appreciating the brain's role in apprehending and responding to the affective states of others can help us understand people who exhibit social cognitive disorders and are deficient in experiencing socially relevant emotions such as sympathy, shame, or guilt.

However, even in supposedly healthy human brains the capacity for other-oriented responses, such as sympathy, is pronouncedly selective and limited by spontaneous aggressive tendencies (Panksepp 1998; Lorenz 1963). When sympathy and mutual aid is extended within a group, they are also (de facto) withheld from those that do not belong to this group. In other words, interest in others is ordinarily expressed positively or negatively towards specific groups—but very rarely are attitudes extended to universal coverage, for example as attitudes towards the entire human species, or towards all sentient beings.

Understanding does not entail compassion, but is frequently combined with emotional dissociation from “the other”. We can easily understand, say, that a child in a distant country

probably reacts to hunger or pain in a way that is similar to how children in our own country react to hunger or pain, but that does not mean that we care about those children in equal or even comparable measures. Indeed, if understanding entailed sympathy, the world would be a far more pleasant dwelling place for many of its inhabitants. By nature, we are “empathetic xenophobes” (Evers 2009): we are empathetic by virtue of our intelligence and capacity to apprehend the mental life of a relatively wide range of creatures, but far more sympathetic to the closer group into which are born or choose to join, remaining neutral or hostile to “out-group” individuals.<sup>1</sup>

Thus, in spite of our natural capacity for empathy, sympathy, and mutual assistance, the human being can also be described as a self-interested, control-oriented, dissociative xenophobe. In view of their historic prevalence, it is not unlikely that these features have evolved to become a part of our innate neurobiological identity and that any attempt to construe social structures (rules, conventions, contracts, etc.) opposing this identity must, in order to be realistically implemented, take this biological challenge into account in addition to the historically well-known political, social, and cultural challenges.

A major practical problem is that the effects of our actions are not limited, as are our capacities for engagement. The difficulty of wide involvement due to the brain's self-projective egocentricity is matched by a capacity to cause large-scale effects, which poses serious problems whenever large-scale or long-term solutions are needed—say, to improve the global environment, reduce global poverty, or safeguard future generations. Our societies are importantly construed around egocentric and short-term perspectives—political, economical, etc.—making it extremely difficult to put global or long-term thought and foresight into practice. This is of course only to be expected, since our brains'

<sup>1</sup> I am here discussing social attitudes in terms of subjective evaluators, but they can also be discussed in terms of non-conscious non-feeling units. Some current neuroscience literature may prefer to discuss the issue not from the point of view of subjective definitions but rather from the perspective of relevance detection and evaluation that is objectively observed.

neuronal architectures are engaged in social interactions and determine the social structures that we can and do develop.

However, our brain identity incorporates social influence. Culture and nature stand in a relationship of mutual causal influence: whilst the organisation of our brains in part determines who we are and what types of societies we develop, our social structures also have a strong impact on the brain's organisation; notably, they impact upon cultural imprints epigenetically stored in our brains. The genetic control over the brain's development is subject to epigenetic evolutionary processes; that is to say, to a coordinated and organised neuronal development that is the result of learning and experience and that is intermixed with the action of genes. The door to being epigenetically proactive is, accordingly, opened. In the following analysis of epigenesis by selective stabilisation of synapses I shall discuss the relationship between genotype and brain phenotype; the paradox of non-linear evolution between genome and brain complexity; the selection of cultural circuits in the brain during development; and the genesis and epigenetic transmission of cultural imprints.

### 3 Neuronal epigenesis

#### 3.1 Genotype & brain phenotype: The paradox of non-linear evolution between genome & brain complexity

The comparison between what we presently know about human genomes and the brain phenotype raises the paradox of a non-linear evolution between the complexity of the genome and that of the brain (Changeux 1985, 2012b). From a molecular neurobiologist's perspective, the cognitive abilities and skills required for the highest functions of the human brain are built from a cascade of events driven by a "genetic envelope", which makes the difference between *Homo sapiens* and the human family's earliest ancestors, but which cannot be simply related to genome size, nor to the number of genes.

The total amount of DNA housed in the haploid genome is approximately 3.1 billion

base pairs, but no more than 20,000–25,000 gene sequences (1.2% of our genome code for exons—the DNA components of genes), and this number does not significantly differ from mouse to human. Moreover, the difference in full DNA sequences are very limited: between humans and chimpanzees they comprise no more than 4% of the genome. However, the total number of neurons in the human brain is in the order of 85 billion, compared to about 70 million in the brain of the mouse (Azevedo et al. 2009). Yet, notwithstanding the increase in cell numbers, with each neuron possessing its particular connectivity and its set of genes expressed, mammalian brain anatomy has evolved dramatically from a poorly corticalized lissencephalic brain with about 10–20 identified cortical areas to a brain with a very high relative cortical surface, multiple gyri and sulci, and possibly as many as 100 identified cortical areas (Mountcastle 1998). Thus, there exists a remarkable nonlinear relationship between the evolution of brain anatomy and the evolution of the genome organisation.

Molecular and cellular explanations have been suggested to account for this nonlinear relationship. One is the combinatorial expression of spatio-temporal patterns of genes that affect development (Changeux 1985; Edelman 1987; Tsigelny et al. 2013). Another, non-exclusive explanation, is the contribution of "epigenetic mechanisms" driven by interaction with the environment in the course of the long postnatal period of brain maturation—circa 15 years in humans—during which critical and reciprocal relationships take place between the brain and its physical, social, and cultural environment. It is on these epigenetic mechanisms that I shall focus here.

#### 3.2 The epigenesis of neuronal networks by selective stabilization of synapses

The word "epigenesis" can be traced back to William Harvey (1651), who stated in contrast to contemporary preformationist views that the embryo arises by "the addition of parts budding out from one another". It was subsequently used by Conrad Waddington (1942) to specify the re-

relationship between the genes and their environment to produce a phenotype. This is also the meaning adopted in the theory of the epigenesis of neuronal networks by selective stabilization of synapses, according to which the environment affects the organisation of connections in an evolving neuronal network through the stabilization or elimination (pruning) of labile synapses, under the control of the state of activity of the network (Changeux et al. 1973). This meaning, which I shall use henceforth, contrasts with the more recent and biochemically distinct meaning of the word *epigenetic*, **which** refers to the status of DNA methylation and histone modification in a particular genomic region. This concerns the neuronal nucleus, but not the diversity of individual synaptic contacts (Sassone-Corsi & Christen 2012). The modulatory role of chromatin modifications in long-term memory has already been described (see e.g., Levenson & Sweatt 2005), but the informational content involved—which relies upon cell bodies—is expected to be in orders of magnitude smaller than that of synaptic epigenesis, based upon the combinatorial power of individual synapses.

During embryonic and postnatal development, the million billion ( $10^{15}$ ) synapses that form the human brain network do not assemble like the parts of a computer, that is, according to a plan that precisely defines the disposition of all the individual components. If this were the case, the slightest error in the instructions for carrying out this program could have catastrophic consequences. On the contrary, the mechanism appears to rely on the progressive setting of robust interneuronal connections through trial-and-error mechanisms that formally resemble an evolutionary process by variation selection (Changeux et al. 1973; Changeux & Danchin 1976; Edelman 1987; Changeux 2012a). At sensitive periods of brain development, the phenotypic variability of nerve cell distribution and position, as well as the exuberant spreading and the multiple figures of transiently-formed connections originating from the erratic wandering of growth cone behaviour, introduce a maximal diversity of synaptic connections. This variability is then reduced by the selective stabilization of some of the labile con-

tacts and the elimination (or retraction) of others. The crucial hypothesis of the model is that the evolution of the connective state of each synaptic contact is governed globally, and within a given time window, by the overall “message” of signals experienced by the cell on which it terminates (Changeux et al. 1973).

One consequence of this is that particular electrical and chemical spatiotemporal patterns of activity in developing neuronal networks are liable to be inscribed under the form of defined and stable topologies of connections within the frame of the genetic envelope. In humans, about half of all adult connections are formed after birth at a very fast rate. The nesting of these multiple traces directly contributes to forming and shaping the micro- and macroscopic architecture of the wiring network of the adult human brain, thus bringing an additional explanation to the above-mentioned non-linearity paradox.

Another consequence of the synapse-selection model (originally presented as a “theorem of variability”) is that the selection of networks with different connective topologies can lead to the same input-output behavioural relationship (Changeux et al. 1973). This accounts for an important feature of the human brain: the constancy or “invariance” of defined states of behaviour, despite the epigenetic “variability” between individual brains’ connectivity.

Finally, both the spontaneous and the evoked activity may contribute to synapse selection. In this framework, a suggestion has been made that reward signals received from the environment may control the developmental evolution of connectivity (Gisiger et al. 2005; Gisiger & Kerszberg 2006). In other words, reinforcement learning would modulate the epigenesis of the network. The model has been implemented in a case of the learning of a visual delayed-matching-to-sample task (see below). This process of synaptic selection by reward signals may concern the evolution of brain connectivity in single individuals, but it also concerns the exchange of information and shared emotions or rewards between individuals in the social group (Changeux 1985, 2004; Gisiger et al. 2005). This is an important part of our argument; it may

thus play a critical role in social and cultural evolution.

### 3.3 The selection of cultural circuits in the brain during development & the epigenetic transmission of cultural imprints

There is an abundance of experimental studies that are consistent with, or directly support, the model of synapse selection. In humans the maximum synaptic density is reached within three years, then steadily declines until the total number stabilises around the time of puberty (Huttenlocher et al. 1997; Bourgeois 1997; Petanjek et al. 2011). Yet the process of synaptic refinement goes far beyond puberty: learning is life-long (Petanjek et al. 2011). The observed global decline in synaptic numbers during childhood plausibly reflects a rich cascade of elementary steps of learning by selection. Numerous studies have shown that when neuronal activity is experimentally modified, synaptic elimination is altered (Benoit & Changeux 1975, 1978; Stretavan et al. 1988; Purves & Lichtman 1980; Luo & O’Leary 2005; Innocenti & Price 2005; Collin & van den Heuvel 2013). At variance with the classical Lamarckist-constructivist scheme (Quartz & Sejnowski 1997), blocking the activity maintains a high number of connections: it is activity that enhances synaptic elimination (Benoit & Changeux 1975, 1978; Stretavan et al. 1988; Luo & O’Leary 2005). Thus “to learn is to eliminate” (Changeux 1985).

Among the cortical connections established in post-natal life are the long-range tracts between the frontal areas (Miller & Cohen 2001; Fuster 2008) and other brain cortical areas (including sensory ones) (Goldman-Rakic 1987; Goldman-Rakic 1999; Hagmann et al. 2008; Collin & van den Heuvel 2013). Some years ago, it was suggested, according to the “global neuronal workspace” hypothesis, that these long-range connections, by broadcasting signals to multiple brain areas, yield subjective “conscious” experience by allowing sensory inputs—seeing, hearing and so on—global access to many brain areas (Dehaene et al. 1998; Dehaene

& Changeux 2011). The long-range connections would provide a structural basis for the global experience known as conscious access.

These long-range connections are particularly important in the case of the prefrontal areas which contribute to planning, decision-making, thought, and socialisation. The ontogeny and postnatal development of long-range connectivity expectedly reveal phases of exuberance and phases of selection and axonal pruning (Collin & van den Heuvel 2013). In human newborns evolution is slow, and it has been suggested that the phase of exuberant long axon removal is largely completed at the age of two years, accompanied by increasing information processing and cognitive development (Collin & van den Heuvel 2013). Evolution continues during adolescence until adulthood with decreasing segregation and increasing integration, mainly but not exclusively driven by modulation of connections strength (local synaptic elimination persists in the adult; Petanjek et al. 2011). It is expected to have major consequences on the laying down of cultural imprints including the “epigenetic rules” associated with socialisation.

The acquisition of reading and writing may be viewed as a typical example of epigenetic development of “cultural circuits”. Writing and reading are recent cultural inventions (about 5000 years old) that evolved into distinct sub-systems and put considerable demands on our cognitive system. Historically, the first evidence for specialized writing and reading circuits in the brain was the discovery by the French neurologist Dejerine (1895) of pure alexia, also known as alexia without agraphia. Individuals with pure alexia suffer from severe reading problems while other language-related skills such as naming, oral repetition, auditory comprehension or writing are typically intact. Alexia results from cerebral lesions in circumscribed brain regions including the angular and supramarginal gyri. New specialized sets of connections are present exclusively in individuals that have learned written language and have been selected and consolidated in the course of development at sensitive periods (4–6 years) as a consequence of an intensive period of education.

The human brain did not evolve to learn to read, but possesses enough epigenetic variability in the course of its development (and also—though to a lesser extent—in the adult) to incorporate a cultural invention of this kind. During the acquisition of reading and writing by Western subjects, representations for visual forms of words progressively settle into the occipito-temporal cortex, recruiting a subset of functionally-appropriate object recognition regions in the temporo-parietal junction (Dehaene et al. 2010). The group of illiterate individuals is consistently more right-lateralized than their literate controls (Pettersson et al. 2007). Interestingly, alphabetic writing systems recruit circuits that differ in part from those mobilized by the Chinese ideographic systems. In French readers reading French, activations were enhanced in left-hemisphere visual area V1, with the strongest differences between French words and their controls found at the central and horizontal meridian representations. In contrast, Chinese readers reading Chinese showed enhanced activations in intermediate visual areas V3v/hV4, which was absent in French participants (Szwed et al. 2014). Also, the capacity to read sheet music is selectively altered in music-specific forms of alexia. Neuronal circuits specific to a given culture may thus become epigenetically established in the brains of social group members. Written language-learning is only one of the many cultural imprints acquired during the development of the human brain (Changeux 1985). For instance, cross-cultural differences between Asian and Western participants manifest themselves as differential increases of fMRI in the medial prefrontal cortex with reference to self-judgment (Zhu et al. 2007; Ray et al. 2010) and also to diverse brain recordings in mind reading (Kobayashi et al. 2007), holistic attention (Hedden et al. 2008), or facial photo recognition (Na & Kitayama 2011). The adult human brain thus builds up from a complex intertwining of cultural circuits progressively laid down during development within the framework of a human-specific genetic envelope.

There is no compelling evidence that culturally-acquired phenotypes will sooner or later

be genetically transmitted. What the evidence does show is that they have to be learned by each generation, by children from adults, and epigenetically transmitted from generation to generation, beginning in the mother's womb and up until the adulthood. Teaching reading and writing to circa five-year-old children requires elaborate pedagogic strategies, which in a general manner are absent in non-human primates (Premack 2007).

In short, cultural imprints have a physical reality in the human brain. Cultural imprints have also been demonstrated in non-human brains, e.g., by Peter Marler's work on birds' song-learning (Marler 1970). Yet the importance of cultural imprints on behaviour are comparatively much more important in humans compared to non-humans, in particular due to the long postnatal period of brain maturation. They play a critical role in shaping the brain phenotype in relation with the social group, through oral and written language but also through diverse culture-specific habits, traditions, and symbolic systems, including the ethical and social norms embodied in the adult brain.

I shall now proceed to discuss issues raised by the combinatorial explosion of brain representations and the channelling of behaviour through *epigenetic rules* and top-down control of decision-making.

epigenetic rules =<sub>Df</sub> In neurobiological terms, these “rules” shall be viewed as acquired patterns of connections (scaffoldings), hypothetically stored in frontal cortex long-term memory. They frame the genesis of novel representations and regulate decision-making in a top-down manner.

## 4 “Epigenetic rules” and top-down control of decision-making

### 4.1 The hierarchical architecture of the brain

It has been suggested that ethical and social norms are, from a perspective in which the brain is central, ultimately encoded as spatio-

otemporal patterns of neuronal activity that can be mobilized within the conscious neuronal workspace (Dehaene & Changeux 2011). Yet from a neurobiological standpoint, this view hinges upon the classical issue of the combinatorial explosion raised by the immense network of almost a million billion ( $10^{15}$ ) interconnected synapses of the human brain. The question that arises, then, is how the particular patterns of neuronal activity, which, for instance, encode defined actions or perceptual events and ultimately ethical rules, are selected within this gigantic neural network. In my view, the concept of a hierarchical organisation of the brain needs to be taken into consideration more closely.

Analysis of the neurological deficits caused by lesions discloses hierarchical and parallel neural architectures that help us understand higher brain functions (Shallice & Cooper 2011). Among these is the inhibition of automatic (or reflex) actions and the elaboration of goal-directed behaviours and their control. In the brain, an evolutionary-recent territory of cerebral cortex architecture, the lateral prefrontal cortex, has been shown to play a critical role in the temporal control of behaviour. It serves as a “temporal buffer” between past events and future actions, allowing behaviours that follow internal goals to occur (Fuster 2001; Goldman-Rakic 1987; Petrides 2005). Moreover, the lateral prefrontal cortex exerts top-down control of cognitive processes associated with hierarchically-lower regions distributed in more posterior territories on the basis of internal plans, goals, or what may be referred to as “rules” (Miller & Cohen 2001; Passingham 1993; Shallice 1988; Dehaene & Changeux 1991; Koehlin et al. 2003). It thus contributes to decision-making within the actual context of a given individual history and stored memories (Damasio 1994) and to “neurally encoded rules” that can associate a context with a specific behavioural response and the ability to generalize a rule in novel circumstances.

An early formal model of learning by selection according to a rule was devised in the Wisconsin Card Sorting Task, which is commonly used as a test of the integrity of frontal lobe functions (Dehaene & Changeux 1991). It

requires subjects to infer a “rule” according to which a deck of cards must be sorted, i.e., colour, shape, or number. Feedback from the experimenter takes the form of a simple positive or negative reward (correct or incorrect). The goal for the subject is to get as many “right” responses as possible. Initially, cards must be sorted according to, say, colour. When performance is successful, the “sorting rule” is changed, for example from colour to shape; the subject must notice the change and find the new rule. The global architecture of a network that passes the task comprises two distinct levels of organization: a low level (level 1) that governs the orientation of the organism toward an object with a defined feature and which would correspond to a visuo-motor loop, including visual areas and the premotor cortex; and a high level (level 2) that controls the behavioural task according to a memory rule, and which would be homologous to the prefrontal cortex or closely-related areas (Dehaene et al. 1987; Dehaene & Changeux 1989).

A key feature of the model is that the high level contains a particular category or cluster of neurons, referred to as “rule-coding clusters”, each of which codes a single dimension (e.g., number, colour, or shape). During the acquisition step, the layer of rule-coding neurons is assumed to play the role of a “generator of diversity”. The spontaneous activity then plays a critical role in the activation of a given rule-coding cluster; and because of lateral inhibition only one cluster is active at a time. A search by trial and error takes place, until a positive reward is received from the environment (here the experimenter). Then, the particular cluster active at this precise moment is selected (for discussion see Monchi et al. 2001; Asplund et al. 2010; Fuster 2008). The number of trials necessary to learn the current rule is small (1–2), and single trial learning may occur in normal subjects as it does with the model (Dehaene & Changeux 1991). This learning of short-term rules based upon the fast (millisecond to second) allosteric transitions of synaptic receptors may also be transferred to long-term stores as epigenetically-acquired patterns of connections (see above).

In the course of the modelling of the Wisconsin card-sorting task, an additional architecture was introduced in the form of an auto-evaluation loop, which can short-circuit the reward input from the exterior. It allows for an internal evaluation of covert motor intentions without actualizing them as behaviours, but instead by testing them by comparison with memorized former experiences (Dehaene & Changeux 1991).

In these early formulations, the “rule-coding clusters” were pre-wired in the neuronal network. Subsequent models, however, opened the range of possible epigenetic rules to a brain-wide space of combinations made available within the global neuronal workspace (Baars 1988). This is of importance when we consider the ability to coordinate thoughts or actions in relation to internal goals, which is referred to as “cognitive control” and is a rather infrequent phenomenon. This discussion thus illustrates how rules encoding ethical norms may originate from the brain. Against this background—which shows how ethical rules might be epigenetically built from brain organization—I propose the possibility of being epigenetically proactive, and adapting our social structures, in both the short- and long-term, to benefit, influence, and constructively interact with the ever-developing neuronal architecture of our brains.

## 4.2 A cascade model of top-down cognitive control

Cognitive control has been further investigated by Koechlin et al. (2003) using a set of more complex tasks than the Wisconsin Card Sorting Task, and which span (at least three) nested levels of complexity. They consist in the presentation of series’ of coloured visual stimuli (squares or letters) organized into blocks, with an increasing importance of contextual signals: from “sensory control” with little if any contextual signal, to “contextual control” and, at the higher level, to “episodic control”. Brain imaging fMRI recordings with healthy human subjects revealed that the lateral prefrontal cortex contributes to a hierarchical cascade of executive processes that involve at least three nested

levels of processing. These are neurally implemented in distinct regions, from posterior premotor to rostral lateral prefrontal cortex regions (typically Brodman’s area 46; Koechlin et al. 2003; Badre & D’Esposito 2007; Badre et al. 2009). Patients with focal lateral prefrontal cortex lesions performed cognitive tasks with sensory, contextual, and episodic deficits associated with focal damage to Brodman’s areas 6, 45, and 46, respectively—as is expected from the cascade model (Azuar et al. 2014; Kayser & D’Esposito 2013).

By analogy with the Wisconsin Card Sorting Task (WCST) model mentioned above, behavioural rules are also sorted, but at different nested levels of information processing, the highest level rules “controlling” in a top-down manner the underlying rules closer to the senses. Hypothetically, ethical norms may be viewed as some particular kind of “control rules” developed within a social context, though this possibility still deserves to be explored by Koechlin, D’Esposito and colleagues.

Recently Collins & Koechlin (2012) have further suggested a computational model of human executive functioning associated with the prefrontal cortex, which integrates multiple processes during decision-making, such as expectedness of uncertainty, task switching, and reinforcement learning. The model reveals that the human frontal function may monitor up to three or four concurrent behavioural strategies and infers online their ability to predict action outcomes: whenever one appears more reliable than unreliable, this strategy is chosen to guide the selection and learning of actions that maximize rewards (see also Miller & Cohen 2001; Passingham 1993; Shallice 1988; Fuster 2008; Dehaene & Changeux 2011).

In their original paper, Collins and Koechlin do not explicitly mention social interaction. Yet we may consider an extension of their model to the social context by assuming that ethical or social norms are part of the “concurrent behavioural strategies” that they postulate exist in decision-making. The selection and learning of actions would then be more elaborate than the simple maximization of immediate rewards.

The developing baby is exposed very early on to a defined social and cultural environment, possibly even pre-natally (Lagercrantz & Changeux 2009; Lagercrantz et al. 2010). At this stage of development an intense synaptogenesis steadily occurs in the cerebral cortex, and epigenetic selection of neuronal networks accompanies the acquisition of the “maternal” language as well as of the common rules of the social community to which the child’s family belongs. The developing baby/child is “impregnated” with the current ethical rules of the social community, and this is often linked with the symbolic (philosophical/religious) system of representation character of the community to which it belongs. These early traces may last for the lifetime of the individual and sooner or later create conflicting relationships with a fast-evolving environment aggravated by the increased longevity of the individual (Changeux 1985). On the basis of the neurobiological data mentioned above, one may define these rules as epigenetically-acquired patterns of connections (scaffoldings) stored in frontal cortex long-term memory, which frame the genesis of novel representation and “cognitively controlled” decision-making in a top-down manner.

Against this background I propose the possibility of being epigenetically proactive and adapting our social structures, in both the short- and the long-term, to benefit, influence, and constructively interact with the ever-developing neuronal architecture of our brains.

## 5 A naturalistic responsibility

### 5.1 Proactive epigenesis

The first sentence in the 1948 Universal Declaration of Human Rights states: “All human beings are born free and equal in dignity and rights.”

Read as a description of the actual situation of human beings, this is blatantly and tragically false. Read as a normative ideal that we should strive for, it is noble but tragically unrealistic: considering our present cerebral structure, we are not likely to acknowledge in actual social practice the equal dignity and rights of all

individuals independently of race, gender, creed, etc. Life conditions may have improved for many humans over time, yet the present global situation remains appalling, notably, with respect to poverty, unequal distribution of health care, and the predominantly non-egalitarian or bellicose relations between individuals or groups. The vast majority of human beings appear reluctant, unable to identify with, or show compassion towards those who are beyond (and sometimes even towards those who are within) *their* sphere. While some societies or individuals may be more prone than others to developing a strong ethnic identity, violence, racism, sexism, social hierarchies, or exclusion, all exhibit some form and measure of xenophobia.

What I have here suggested, however, is that we might make presently unrealistic ideals, such as equality in dignity and rights, somewhat more realistic by selecting them for epigenetic proactivity.

Synaptic epigenetic theories of cultural and social imprinting on our brain architecture open the door to being epigenetically proactive, which means that we may culturally influence our brain organisation with the aim of self-improvement, individually as well as socially, and change our biological predispositions through a better fit of our brain to cultures and social structures.

I suggest that certain areas of research are especially important to pursue with the goal of “epigenetic proaction” in mind. They aim at integrating recent advances in neuroscientific research into normative debates at the level of society. This does not necessarily mean that my level of explanation is “neurocentric” or “neuroreductionist”. My aim is more “encyclopedic” in the sense that I wish to illustrate the benefits that neuroscience can bring to the humanities and social sciences and conversely. I do not see myself as either neuro-“centric” or “reductionist”—which would mean an exclusion of other categories of determinants at the social or historical levels—but I am more modestly willing to unify knowledge between the humanities and the neurosciences, which are too often deliberately omitted from the debate. This can be illustrated by two examples: violence in adoles-

cents in relation to their social environments, and violence in adults associated with interconfessional conflicts.

Violence in adolescents is a common phenomenon in our societies and it is frequently repressed through police and judiciary means, often resulting in incarceration. But this approach to juvenile violence simply omits the scientifically-established fact that adolescence is also a time of “neurodevelopmental crisis”. Evidence from anatomical and functional-imaging studies has highlighted major modifications of cortical circuits during adolescence. These include reductions of gyrification and grey matter, increases in the myelination of cortico-cortical connections, and changes in the architecture of large-scale cortical networks—including precentral, temporal, and frontal areas. (Klein et al. 2014). Uhlhaas et al. (2009) have used MEG synchrony as an indicator of conscious access and cognitive performance (rev. Dehaene & Changeux 2011). Until early adolescence, developmental improvements in cognitive performance are accompanied by increases in neural MEG synchrony. This developmental phase is followed by an unexpected decrease in neural synchrony that occurs during late adolescence and is associated with reduced performance. After this period of destabilization follows a reorganization of synchronization patterns that is accompanied by pronounced increases in gamma-band power and in theta and beta phase synchrony (Uhlhaas et al. 2009). These remarkable changes in neural connectivity and performance in the adolescent are only just being explored and may lead to special unexpected proactive care from society. In turn, this requires active research, including a social educative environment adequate to adolescents’ special needs. This may include adequate physical exercise, cultural games, educational training, and new kinds of therapies yet to be invented.

Violent interconfessional conflicts have raged throughout human history. They continue to plague our modern societies and are presently an important cause of wars and other forms of violence throughout the world. One should remember that every newborn and child brain incorporates critical features of its biolo-

gical, social, and cultural environment including, in addition to spoken and written language, symbolic systems and religious rituals (which include dietary and vestimentary practices as markers of the faith). These epigenetic traces are almost irreversibly laid down and may persist throughout the whole life of the individual. Yet they might be renewed through epigenetic transmission from adults to newborns. In this context, early proactive epigenetic imprinting through education is of critical importance. The aim of that education should not be to abolish faith or emotional convictions (e.g., moral, political, or religious) but only to control the fervour, intolerance, and fanaticism in their expression. The problem, as I see it, is not a belief itself, but the emotional intensity to which it gives rise and the manner in which it is expressed. Influencing a child brain to reduce its propensity to ideological violence or fanaticism and enhance its tolerance to others’ differences also requires special proactive care from society that per force involves active research—including a social educative environment adequate to this particular goal.

These are only two illustrations of the many that are possible, chosen because they have been problematic throughout the history of humankind and show no signs of disappearing.

At the individual level, the social conditions of an infant, or an adolescent, are of crucial importance in their cerebral development, and adequate conditions can in principle be provided. The factual realism of this application is largely a matter of political will and social agreement. The scientific challenge will be to further develop the knowledge of these conditions and their effects on the developing infant and adolescent brain. Also, the challenge will be to develop our knowledge of how social conditions affect the adult brain, e.g., to prevent neurodegeneration.

On a more general level, when applied on a larger scale to a society, a population, or to the entire human species, the argument follows the same logic and is no less important—but it becomes considerably more complicated to apply, theoretically as well as practically.

If new cultural imprints were epigenetically stored in our brains (say, less violent or less sectarian features), future generations would presumably develop societies that reflect them (i.e., become more peaceful and inclusive). A weakness of this optimistic reasoning is its circularity, since we would already need to be peaceful in order for a peaceful society to be maintained. A crucial question then becomes: how long does it take for a cultural characteristic to leave a cerebral trace? In some measure stable and enduring cultural structures are needed in order to effect stable neurobiological changes and store cultural imprints in the brain that might give evolution a push in the right direction, but the chances of maintaining societies that conflict with the present nature of its inhabitants—say, maintaining a peaceful egalitarian rule in a society of violent xenophobes—are arguably slim.

The challenges involved in trying to be epigenetically proactive by culturally influencing the future actions of human genes and neuronal structures, with the aim of altering higher cognitive functions and their resulting behaviour seem formidable, at least if enlarged sympathy is on the agenda. Still, within the epigenetic neuroscientific framework, at least the theoretical possibility exists, and it is worthy of consideration by many other disciplines beyond neuroscience. Depending on how we choose to develop our culture, one day epigenetic rules that enlarge the presently-narrow realm of human sympathy might perhaps emerge.

## 5.2 Conclusion: A naturalistic responsibility

The origins of norms and the relationship between facts and values have been much debated in philosophy. Reasoning that weds scientific theory with normative considerations has been accused of committing the logical error of confusing facts and values, which is known as “the naturalistic fallacy”.

The expression “the naturalistic fallacy” was coined by the British moral philosopher G. E. Moore and refers in his work to the identification (or reduction) of goodness with (or to)

another property such as utility, pleasure, or happiness (Moore 1903). That issue is not relevant in the present context. In the interpretation of the naturalistic fallacy that is relevant here, the fallacy consists in deriving an “ought” from an “is”, or a value from a fact, and letting descriptive properties entail normative properties, which confuses the distinction between facts and values in a fallacious manner. This argument is reminiscent of David Hume’s claim that what *is* is entirely different from what *ought to be*, for “the distinction of vice and virtue is not founded on the relations of objects, nor is perceiv’d by reason” but is fundamentally a matter of feelings and as such is neither true nor false (Hume 1739, III, I). I agree that it is fallacious to derive “ought to be” from “is”, and consider this a conceptual mistake that our theory of epigenetic proaction must and indeed does avoid. I do not assert that factual descriptions of the brain’s architecture are tantamount to yielding recommendations or assertions of norms, do not confuse “is” with “ought”, and consequently do not commit the naturalistic fallacy in this formulation.

We should observe that a *value* may be represented on many levels: non-conscious as well as conscious, as a basic biological function or as a feature of advanced moral reasoning. When discussing the naturalistic fallacy, value as a feature of advanced normative reasoning is the relevant sense of the term. The logical distinction between fact and value could collapse if the term is defined differently—say, if it features as a non-normative biological function. The logical error in the naturalistic fallacy concerns the fact/value distinction as it is drawn between normative and descriptive statements, namely between *ought* and *is*; not between facts that are/are not biological values, where that concern would presumably not arise.

However, eagerness to avoid the naturalistic fallacy must not prevent our normative reasoning from being informed by scientific theories. Normative judgments should be informed by facts, even though they cannot be entailed by them. If certain evaluative tendencies are innate in the normal human brain’s architecture, such as self-interest and selective sympathy, this

fact (if it is one) about the human being's neuronal structure would admittedly entail that every healthy, sufficiently mature individual will to some degree feel both self-interest and sympathy towards some other creature. However, this is not the entailment of a norm, but an empirical entailment of another fact. It does not entail that it is good (or bad), or that we ought to conceive it as good (or bad) that we are thus construed. Similarly, if it is true that we are, for example, and as we have argued, self-projective xenophobes, knowledge of this (presumed) fact is not in itself a justification of it. Understanding is not the same as justification: to know, or to understand, is not to approve. On the contrary, knowledge about our neural structures' predispositions should increase our awareness of the need for stable and realistic social structures and agreements to keep us in check.

We should also observe that a belief in the approximate universality of certain values, or preferential tendencies as innate features of the human neurobiological make-up, is logically compatible with a belief in maintaining the description/norm distinction.

My primary focus has been on the important empirical connections between biological facts and norms. Norms are brain constructs elaborated by human societies, biologically as well as culturally embedded in and constrained by the contingent evolution of socio-cultural structures—in particular, by the multiple symbolic philosophical and religious systems that have developed. This fact, and the realisation that normative judgments should be informed by facts even though they cannot be entailed by them suggests that science, philosophy and—not least—neuroethics—have a major responsibility: namely to decipher the network of causal connections between the neurobiological, socio-cultural, and contingent historical perspectives that allow a moral norm to be enunciated at a given moment in human history; and to evaluate their “universal” character as pre-specified in our genome and shared by the human species in distinction from those relative to a given culture or symbolic system. The “fallacy” of the naturalistic approach is thus inverted into a naturalistic responsibility (Evers 2009): the re-

sponsibility to connect facts and values, biology, and socio-cultural structures, and to use that enriched understanding for the benefit of ourselves and our societies.

We may hope that through the rational exchange of arguments between partners with different cultures and moral traditions debating together, a species-specific “human core” could become dominant beyond individual differences and converge on a common structure (Changeux & Ricoeur 2000). At the same time, we must note that the diversity of human individuals and societies is enormous and must be respected while we strive to find this common ground that might allow coexistence.

The idea of proactively selecting those specific dispositions or capacities (such as sympathy) that we all share as human beings which that, if properly developed, may benefit our global co-existence while respecting individual and ideological diversities, is well in line with Darwin. Darwin wrote in *The Descent of Man*:

As man advances in civilization, and small tribes are united into larger communities, the simplest reason would tell each individual that he ought to extend his social instincts and sympathies to all members of the same nation, though personally unknown to him. This point being once reached, there is only an artificial barrier to prevent his sympathies extending to the men of all nations and races.

Lewontin (1993) argues that while traditional Darwinism has portrayed the organism as a passive recipient of environmental influences, a correct understanding should emphasize that humans are active constructors of their own environment—in particular the social and cultural environment. I agree and argue further that, in line with Darwin, we can be active constructors of our own brains through using our environment and culture, in a relationship that is reciprocal.

In this article, my main focus has been on feasibility—that is, on whether we *can* be epigenetically proactive. If we assume an affirmative answer to that question, an important fol-

low-up question arises: whether we *should* be so. My basic position, that I have here tried to express, is that epigenetic proaction could be a very promising, powerful, and long-term way of influencing human nature and of improving our societies. However, in order to pursue this in a responsible and adequate manner, caution is required, along with careful analyses of the relevant social and ethical issues. Science can be, and has throughout history repeatedly been, ideologically hijacked, and the resulting dangers increase with the strength of the science in question. If, say, humans learn to design their own brain more potently than we already do by selecting what we believe to be brain-nourishing food and pursuing neuronally-healthy life-styles, we *could* use that knowledge well—that is, there is certainly room for improvement. On the other hand, the dream of the perfect human being has a sordid past, providing ample cause for concern about such projects. Historic awareness is of the utmost importance for neuroethics when assessing suggested applications in a responsible and adequate manner. Moreover, what we mean by “responsible and adequate” is open to interpretation. The traits we choose to favour epigenetically, and the social structures we choose to develop, depend on who “we” are, and in what society we wish to live.

Arthur Koestler compares evolution to “a labyrinth of blind alleys” and suggests that “there is nothing very strange or improbable in the assumption that man’s native equipment, though superior to that of any other living species, nevertheless contains some built-in error or deficiency which predisposes him to self-destruction” (Koestler 1967, xi). In that light, steering evolution by influencing the cultural imprints to be stored in our brains appears to be an attractive option.

## Acknowledgements

I wish to thank Jean-Pierre Changeux for his important scientific contributions to this paper, and for his detailed scrutiny of the arguments expressed. I also wish to thank Yadin Dudai, Sten Grillner, Hugo Lagercrantz, and Arleen Salles for their valuable comments on earlier

versions of this manuscript. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 604102 (HBP).

## References

- Asplund, C. L., Todd, J. J., Snyder, A. P., Gilbert, C. M. & Marois, R. (2010). Surprise-induced blindness: A stimulus-driven attentional limit to conscious perception. *Journal of Experimental Psychology Human Perception and Performance*, *36* (6), 1372-1381. [10.1037/a0020551](https://doi.org/10.1037/a0020551)
- Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P., Filho, J. W., Lent, R. & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, *513* (5), 532-541. [10.1002/cne.21974](https://doi.org/10.1002/cne.21974)
- Azuar, C., Reyes, P., Slachevsky, A., Volle, E., Kinkingnehun, S., Kouneiher, F., Bravo, E., Dubois, B., Koechlin, E. & Levy, R. (2014). Testing the model of caudorostral organization of cognitive control in the human with frontal lesions. *NeuroImage*, *1* (84), 1053-60-1060. [10.1016/j.neuroimage.2013.09.031](https://doi.org/10.1016/j.neuroimage.2013.09.031)
- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge, UK: Cambridge University Press.
- Badre, D. & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, *19* (12), 2082-2099. [10.1162/jocn.2007.19.12.2082](https://doi.org/10.1162/jocn.2007.19.12.2082)
- Badre, D., Hoffman, J., Cooney, J. W. & D'Esposito, M. (2009). Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience*, *12* (4), 515-522. [10.1038/nn.2277](https://doi.org/10.1038/nn.2277)
- Barto, A. G. & Sutton, R. S. (1982). Simulation of anticipatory responses in classical conditioning by a neuron-like adaptive element. *Behavioral Brain Research*, *4* (3), 221-235. [10.1016/0166-4328\(82\)90001-8](https://doi.org/10.1016/0166-4328(82)90001-8)
- Benoit, P. & Changeux, J.-P. (1975). Consequences of tenotomy on the evolution of multiinnervation on developing rat soleus muscle. *Brain Research*, *99* (2), 354-358.
- (1978). Consequences of blocking the nerve with a local anaesthetic on the evolution of multiinnervation at the regenerating neuromuscular-junction of the rat. *Brain Research*, *149* (1), 89-86.
- Berger, H. (1929). "Über das Enkephalogramm beim Menschen" [On the use of the encephalogram in humans]. *Archiv für Psychiatrie und Nerven-krankheiten*
- Blakemore, S. J., Wolpert, D. M. & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1* (7), 10196573-10196573. [10.1038/2870](https://doi.org/10.1038/2870)
- Bourgeois, J.-P. (1997). Synaptogenesis, heterochrony and epigenesis in the mammalian neocortex. *Acta Paediatrica Supplement* 42227-33
- Changeux, J.-P. (1985). *Neuronal man*. New York, NY: Pantheon Books.
- (2004). *The physiology of truth: Neuroscience & human knowledge*. Boston, MA: Harvard University Press.
- (2012a). Synaptic epigenesis and the evolution of higher brain functions. In P. Sassone-Corsi & Y. Christen (Eds.) *Epigenetics, brain and behavior* (pp. 11-22). Dordrecht, NL: Springer.
- (2012b). *The good, the true, the beautiful: A neuronal approach*. New Haven, CT: Yale/Odile Jacob.
- Changeux, J.-P. & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature*, *264* (5588), 705-712. [10.1038/264705a0](https://doi.org/10.1038/264705a0)
- Changeux, J.-P., Courge, P. & Danchin, A. (1973). A theory of the epigenesis of neuronal networks by selective stabilization of synapses. *Proceedings of the National Academy of Sciences USA*, *70* (10), 2974-2978.
- Changeux, J.-P. & Lou, H. C. (2011). Emergent pharmacology of conscious experience: New perspectives in substance addiction. *Federation of American Societies of Experimental Biology Journal*, *25* (7), 2098-2108. [10.1096/fj.11-0702ufm](https://doi.org/10.1096/fj.11-0702ufm)
- Changeux, J.-P. & Ricoeur, P. (2000). *What makes us think? A neuroscientist and a philosopher argue about ethics, human nature and the brain*. Princeton, NJ: Princeton University Press.
- Cheng, Y., Lee, P., Yang, C. Y., Lin, C. P. & Decety, J. (2008). Gender differences in the mu rhythm of the human mirror-neuron system. *PLoS ONE*, *3* (5), e2113-e2113. [10.1371/journal.pone.0002113](https://doi.org/10.1371/journal.pone.0002113)
- Collins, A. & Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology*, *10* (3), e1001293-e1001293. [10.1371/journal.pbio.1001293](https://doi.org/10.1371/journal.pbio.1001293)
- Collin, G. & van den Heuvel, M. P. (2013). The ontogeny of the human connectome: Development and dynamic changes of brain connectivity across the life span. *Neuroscientist*, *19* (6), 616-628. [10.1177/1073858413503712](https://doi.org/10.1177/1073858413503712)
- Damasio, A. (1994). *Descartes error: Emotion, reason, and the human brain*. New York, NY: Putnam.
- (1999). *The feeling of what happens: Body and*

- emotion in the making of consciousness*. San Diego, CA: Harcourt.
- Damasio, A. & Carvalho, G. B. (2013). The nature of feelings: Evolutionary and neurobiological origins. *Nature Reviews Neuroscience*, *14* (2), 143-152. [10.1038/nrn3403](https://doi.org/10.1038/nrn3403)
- Darwin, C. (1871). *The descent of man*. London, UK: John Murray.
- Decety, J. (Ed.) (2012). *Empathy: From bench to bedside*. Cambridge, MA: MIT Press.
- Decety, J. & Sommerville, J. A. (2003). Shared representations between self and others: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7* (12), 527-533. [10.1016/j.tics.2003.10.004](https://doi.org/10.1016/j.tics.2003.10.004)
- Dehaene, S. & Changeux, J.-P. (1989). A simple model of prefrontal cortex function in delayed-response tasks. *Journal of Cognitive Neuroscience*, *1* (3), 244-261. [10.1162/jocn.1989.1.3.244](https://doi.org/10.1162/jocn.1989.1.3.244)
- (1991). The Wisconsin card sorting test: Theoretical analysis and simulation of a reasoning task in a model neuronal network. *Cerebral Cortex*, *1* (1), 62-79.
- (2000). Reward-dependent learning in neuronal networks for planning and decision making. *Progress in Brain Research*, *126*. [10.1016/S0079-6123\(00\)26016-0](https://doi.org/10.1016/S0079-6123(00)26016-0)
- (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70* (2), 200-227. [10.1016/j.neuron.2011.03.018](https://doi.org/10.1016/j.neuron.2011.03.018)
- Dehaene, S., Changeux, J.-P. & Nadal, J.-P. (1987). Neural networks that learn temporal sequences by selection. *Proceedings of the National Academy of Sciences, USA*, *84* (9), 2727-2731.
- Dehaene, S., Kerszberg, M. & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences USA*, *95* (24), 14529-14534. [10.1073/pnas.95.24.14529](https://doi.org/10.1073/pnas.95.24.14529)
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J. & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330* (6009), 1359-1364. [10.1126/science.1194140](https://doi.org/10.1126/science.1194140)
- Dejerine, J. (1895). *Anatomie des centres nerveux vol. 1*. Paris, FR: Rueffet Cie.
- Denton, D. (2006). *The primordial emotions: The dawning of consciousness*. Paris, FR: Flammarion.
- Dudai, Y. (1989). *The neurobiology of memory: Concepts, findings, trends*. Oxford, UK: Oxford University Press.
- (2002). *Memory from A to Z: Keywords, concepts and beyond*. Oxford, UK: Oxford University Press.
- Edelman, G. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York, NY: Basic Books.
- (1992). *Bright air, brilliant fire: On the matter of the mind*. New York, NY: Basic Books.
- Engen, H. G. & Singer, T. (2013). Empathy circuits. *Current Opinion in Neurobiology*, *23* (2), 275-282. [10.1016/j.conb.2012.11.003](https://doi.org/10.1016/j.conb.2012.11.003)
- Evers, K. (2001). The Importance of being a self. *International Journal of Applied Philosophy*, *15* (1), 65-83. [10.5840/ijap20011512](https://doi.org/10.5840/ijap20011512)
- (2009). *Neuroéthique. Quand la matière s'éveille*. Paris, FR: Éditions Odile Jacob.
- Falck-Ytter, T., Gredebäck, G. & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9* (7), 878-879. [10.1038/nn1729](https://doi.org/10.1038/nn1729)
- Fuster, J. M. (2001). The prefrontal cortex—An update: Time is of the essence. *Neuron*, *30* (2), 319-333. [10.1016/S0896-6273\(01\)00285-9](https://doi.org/10.1016/S0896-6273(01)00285-9)
- (2008). *The prefrontal cortex*. London, UK: Academic Press.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4* (1), 14-21. [10.1016/S1364-6613\(99\)01417-5](https://doi.org/10.1016/S1364-6613(99)01417-5)
- Galli, L. & Maffei, L. (1988). Spontaneous impulse activity of rat retinal ganglion cells in prenatal life. *Science*, *242* (4875), 90-91. [10.1126/science.3175637](https://doi.org/10.1126/science.3175637)
- Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: A first-person approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356* (1413), 1351-1361. [10.1098/rstb.2001.0955](https://doi.org/10.1098/rstb.2001.0955)
- Gisiger, T., Kerszberg, M. & Changeux, J.-P. (2005). Acquisition and performance of delayed-response tasks: A neural network model. *Neuroscience Letters*, *383* (1-2), 489-506. [10.1093/cercor/bhh149](https://doi.org/10.1093/cercor/bhh149)
- Gisiger, T. & Kerszberg, M. (2006). A model for integrating elementary neural functions into delayed-response behavior. *PLoS Computational Biology*, *2* (4), e25-e25. [10.1371/journal.pcbi.0020025](https://doi.org/10.1371/journal.pcbi.0020025)
- Goldman-Rakic, P. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.) *Handbook of physiology* (pp. 373-417). Washington DC: The American Physiological Society.
- (1999). The physiological approach: Functional architecture of working memory and disordered cognition in schizophrenia. *Biological Psychiatry*, *46* (5), 650-661. [10.1016/S0006-3223\(99\)00130-4](https://doi.org/10.1016/S0006-3223(99)00130-4)
- Goodman, C. S. & Shatz, C. J. (1993). Developmental

- mechanisms that generate precise patterns of neuronal connectivity. *Cell*, *72*, 77-98.  
[10.1016/S0092-8674\(05\)80030-3](https://doi.org/10.1016/S0092-8674(05)80030-3)
- Gordon, G. & Ahissar, E. (2012). Hierarchical curiosity loops and active sensing. *Neural Networks*, *32*, 119-129.  
[10.1016/j.neunet.2012.02.024](https://doi.org/10.1016/j.neunet.2012.02.024)
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J. & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, *16* (7), e159-e159.  
[10.1371/journal.pbio.0060159](https://doi.org/10.1371/journal.pbio.0060159)
- Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H. & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport*, *11* (11), 2351-2355.
- Harvey, W. (1651). *Exercitationes de generatione*. London, UK: Animalium.
- Hedden, T., Ketay, S., Aron, A., Markus, H. R. & Gabrieli, J. D. (2008). Cultural influences on neural substrates of attentional control. *Psychological Science*, *19* (1), 12-17. [10.1111/j.1467-9280.2008.02038.x](https://doi.org/10.1111/j.1467-9280.2008.02038.x)
- Hills, P. J. & Lewis, M. B. (2006). Reducing the own-race bias in face recognition by shifting attention. *Quarterly journal of experimental psychology*, *59* (6), 996-1002.  
[10.1080/17470210600654750](https://doi.org/10.1080/17470210600654750)
- Hommel, B., Müssele, J., Aschersleben, G. & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Science*, 849-937.
- Hume, D. (1739). *Treatise of Human Nature*.
- Huttenlocher, P. & Dabholkar, A. (1997). Regional difference in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387* (2), 167-178.  
[10.1002/\(SICI\)1096-9861\(19971020\)387](https://doi.org/10.1002/(SICI)1096-9861(19971020)387)
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C. & Rizzolatti, G. (2005). Grasping the intention with one's own mirror neuron system. *PLoS Biology*, *3* (3), e79-e79.  
[10.1371/journal.pbio.0030079](https://doi.org/10.1371/journal.pbio.0030079)
- Innocenti, G. M. & Price, D. J. (2005). Exuberance in the development of cortical networks. *Nature Reviews Neuroscience*, *6* (12), 955-965. [10.1038/nrn1790](https://doi.org/10.1038/nrn1790)
- Jackson, P. L. & Decety, J. (2004). Motor cognition: A new paradigm to study self-other interactions. *Current Opinion in Neurobiology*, *14* (2), 1-5.  
[10.1016/j.conb.2004.01.020](https://doi.org/10.1016/j.conb.2004.01.020)
- Jackson, P. L., Brunet, E., Meltzoff, A. N. & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain: An event-related fMRI study. *Neuropsychologia*, *44* (5), 752-761.  
[10.1016/j.neuropsychologia.2005.07.015](https://doi.org/10.1016/j.neuropsychologia.2005.07.015)
- Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford, UK: Oxford University Press.
- Kayser, A. S. & D'Esposito, M. (2013). Abstract rule learning: The differential effects of lesions in frontal cortex. *Cerebral Cortex*, *23* (1), 230-240.  
[10.1093/cercor/bhs013](https://doi.org/10.1093/cercor/bhs013)
- Kitayama, S. & Uskul, A. K. (2011). Culture, mind, and the brain: Current evidence and future directions. *Annual Reviews of Psychology*, *62*, 419-449.  
[10.1146/annurev-psych-120709-145357](https://doi.org/10.1146/annurev-psych-120709-145357)
- Klein, D., Rotarska-Jagiela, A., Genc, E., Sritharan, S., Mohr, H., Roux, F., Han, C. E., Kaiser, M., Singer, W. & Uhlhaas, P. J. (2014). Adolescent Brain maturation and cortical folding: Evidence for reductions in gyrification. *PLoS ONE*, *9* (1), e84914-e84914.  
[10.1371/journal.pone.0084914](https://doi.org/10.1371/journal.pone.0084914)
- Kobayashi, C., Glover, G. H. & Temple, E. (2007). Cultural and linguistic effects on neural bases of "theory of mind" in American and Japanese children. *Brain Research*, *1164*, 95-107. [10.1016/j.brainres.2007.06.022](https://doi.org/10.1016/j.brainres.2007.06.022)
- Koechlin, E., Ody, C. & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302* (5648), 1181-1185.  
[10.1126/science.1088545](https://doi.org/10.1126/science.1088545)
- Koestler, A. (1967). *The Ghost in the Machine*. UK: Arkana Books.
- Lagercrantz, H. (2005). *I barnets hjärna, Le cerveau de l'enfant*. Paris, FR: Éditions Odile Jacob.
- Lagercrantz, H. & Changeux, J.-P. (2009). The emergence of human consciousness: From fetal to neonatal life. *Pediatric Research*, *65* (3), 255-260.  
[10.1203/PDR.0b013e3181973b0d](https://doi.org/10.1203/PDR.0b013e3181973b0d)
- Lagercrantz, H., Hanson, M., Ment, L. & Peebles, D. (Eds.) (2010). *The newborn brain neuroscience and clinical applications*. Cambridge, UK: Cambridge University Press, 2nd Edition.
- Lawrence, E. J., Shaw, P., Giampietro, V. P., Surguladze, S., Brammer, M. J. & David, A. S. (2006). The role of 'shared representations' in social perception and empathy: An fMRI study. *NeuroImage*, *29* (4), 1173-1184.  
[10.1016/j.neuroimage.2005.09.001](https://doi.org/10.1016/j.neuroimage.2005.09.001)
- Ledoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Cambridge, UK: Cambridge University Press.
- Levenson, J. M. & Sweatt, J. D. (2005). Epigenetic mechanisms in memory formation. *Nature Reviews Neuro-*

- science*, 6 (2), 108-118. [10.1038/nrn1604](https://doi.org/10.1038/nrn1604)
- Lewontin, R. (1993). *The doctrine of DNA: Biology as ideology*. London, UK: Penguin Books.
- Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21 (1)
- Llinas, R. R. & Paré, D. (1991). Of dreaming and wakefulness. *Neuroscience*, 44 (3), 521-535. [10.1016/0306-4522\(91\)90075-Y](https://doi.org/10.1016/0306-4522(91)90075-Y)
- Lorenz, K. (1963). *On aggression*. London, UK: Methuen.
- Lou, H. C., Luber, B., Crupain, M., Keenan, J. P., Nowak, M., Kjaer, T. W., Sackeim, H. A. & Lisanby, S. H. (2004). Parietal cortex and representation of the mental self. *Proceedings of the National Academy of Sciences USA*, 101 (17), 6827-6832. [10.1073/pnas.0400049101](https://doi.org/10.1073/pnas.0400049101)
- Luo, L. & O'Leary, D. M. (2005). Axon retraction and degeneration in development and disease. *Annual Reviews of Neuroscience*, 28, 127-156. [10.1146/annurev.neuro.28.061604.135632](https://doi.org/10.1146/annurev.neuro.28.061604.135632)
- Marler, P. (1970). A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative & Physiological Psychology*, 71 (2), 1-25. [10.1037/h0029144](https://doi.org/10.1037/h0029144)
- Michel, C., Rossion, B., Han, J., Chung, C. S. & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, 17 (7), 608-615. [10.1111/j.1467-9280.2006.01752.x](https://doi.org/10.1111/j.1467-9280.2006.01752.x)
- Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Reviews of Neuroscience*, 24, 167-202. [10.1146/annurev.neuro.24.1.167](https://doi.org/10.1146/annurev.neuro.24.1.167)
- Monchi, O., Petrides, M., Petre, V., Worsley, K. & Dagher, A. (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 21 (19), 7733-7741.
- Moore, G. E. (1903). *Principia ethica*. Cambridge, UK: Cambridge University Press.
- Mountcastle, V. (1998). *Perceptual neuroscience: The cerebral cortex*. Cambridge, MA: Harvard University Press.
- Na, J. & Kitayama, S. (2011). Spontaneous trait inference is culture-specific : Behavioral and neural evidence. *Psychological Science*, 22 (8), 1025-1032. [10.1177/0956797611414727](https://doi.org/10.1177/0956797611414727)
- Narayanan, C. H. & Hamburger, V. (1971). Motility in chick embryos with substitution of lumbosacral by brachial by lumbosacral spinal cord segments. *Journal of Experimental Zoology*, 178 (4), 415-413. [10.1002/jez.1401780402](https://doi.org/10.1002/jez.1401780402)
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York, NY: Oxford University Press.
- Parr, L. A. & Waller, B. M. (2006). Understanding chimpanzee facial expression: Insights into the evolution of communication. *Social Cognitive and Affective Neuroscience*, 1 (3), 221-228. [10.1093/scan/nsl031](https://doi.org/10.1093/scan/nsl031)
- Passingham, R. (1993). *The frontal lobes and voluntary action*. Oxford, UK: Oxford University Press.
- Petanjek, Z., Judas, M., Simic, G., Rasin, M. R., Uylings, H. B., Rakic, P. & Kostovic, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences, USA*, 108 (32), 13281-13286. [10.1073/pnas.1105108108](https://doi.org/10.1073/pnas.1105108108)
- Petersson, K. M., Silva, C., Castro-Caldas, A., Ingvar, M. & Reis, A. (2007). Literacy: A cultural influence on functional left-right differences in the inferior parietal cortex. *European Journal of Neuroscience*, 26 (3), 791-799. [10.1111/j.1460-9568.2007.05701.x](https://doi.org/10.1111/j.1460-9568.2007.05701.x)
- Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360 (1456), 781-795. [10.1098/rstb.2005.1631](https://doi.org/10.1098/rstb.2005.1631)
- Phelps, E. A., Cannistraci, C. J. & Cunningham, W. A. (2003). Intact performance on an indirect measure of race bias following amygdala damage. *Neuropsychologia*, 41 (2), 203-208. [10.1016/S0028-3932\(02\)00150-1](https://doi.org/10.1016/S0028-3932(02)00150-1)
- Premack, D. (2007). Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences USA*, 104 (35), 13861-13867. [10.1073/pnas.0706147104](https://doi.org/10.1073/pnas.0706147104)
- Purves, D. & Lichtman, J. (1980). Elimination of synapses in the developing nervous system. *Science*, 210 (4466), 153-157. [10.1126/science.7414326](https://doi.org/10.1126/science.7414326)
- Putnam, F. (1989). *Diagnosis & treatment of multiple personality disorder*. London, UK: The Guilford Press.
- Quartz, S. R. & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral Brain Sciences*, 20 (4), 537-596.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences USA*, 98 (2), 676-682. [10.1073/pnas.98.2.676](https://doi.org/10.1073/pnas.98.2.676)
- Ray, R. D., Shelton, A. L., Hollon, N. G., Matsumoto, D., Frankel, C. B., Gross, J. J. & Gabrieli, J. D. E. (2010).

- Interdependent self-construal and neural representations of the self and mother. *Social Cognitive and Affective Neuroscience*, 5 (2-3), 318-323. [10.1093/scan/nsp039](https://doi.org/10.1093/scan/nsp039)
- Ricoeur, P. (1992). *Oneself as Another* (K. Blamey trans). Chicago, IL: University of Chicago Press.
- Rochat, P. (2001). *The infant's world*. Cambridge, MA: Harvard University Press.
- Sassone-Corsi, P. & Christen, Y. (Eds.) (2012). *Epigenetics, Brain and Behavior*. Springer.
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Reviews in Psychology*, 57, 87-115. [10.1146/annurev.psych.56.091103.070229](https://doi.org/10.1146/annurev.psych.56.091103.070229)
- Schultz, W., Dayan, P. & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275 (5306), 1593-1599. [10.1126/science.275.5306.1593](https://doi.org/10.1126/science.275.5306.1593)
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge, UK: Cambridge University Press.
- Shallice, T. & Cooper, R. (2011). *The organisation of mind*. Oxford, UK: Oxford University Press.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J. & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303 (5661), 1157-1162. [10.1126/science.1093535](https://doi.org/10.1126/science.1093535)
- Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J. & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439 (7075), 466-469. [10.1038/nature04271](https://doi.org/10.1038/nature04271)
- Stretavan, D. W., Shatz, C. J. & Stryker, M. P. (1988). Modification of retinal ganglion cell axon morphology by prenatal infusion of tetrodotoxin. *Nature*, 336 (6198), 468-471. [10.1038/336468a0](https://doi.org/10.1038/336468a0)
- Szwed, M., Qiao, E., Jobert, A., Dehaene, S. & Cohen, L. (2014). Effects of literacy in early visual and occipitotemporal areas of Chinese and French readers. *Journal of Cognitive Neuroscience*, 26 (3), 459-475. [10.1162/jocn\\_a\\_00499](https://doi.org/10.1162/jocn_a_00499)
- Tsigelny, I. F., Kouznetsova, V. L., Baitaluk, M. & Changeux, J.-P. (2013). A hierarchical coherent-gene-group model for brain development. *Genes, Brain and Behavior*, 12 (2), 147-165. [10.1111/gbb.12005](https://doi.org/10.1111/gbb.12005)
- Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.
- Uhlhaas, P. J., Roux, F., Singer, W., Haenschel, C., Sireteanu, R. & Rodriguez, E. (2009). The development of neural synchrony reflects late maturation and restructuring of functional networks in humans. *Proceedings of the National Academy of Sciences USA*, 106(24), 9866-9871. [10.1073/pnas.0900390106](https://doi.org/10.1073/pnas.0900390106)
- Waddington, C. H. (1942). The epigenotype. *Endeavour*, 18-20.
- Zhu, Y., Zhang, L., Fan, J. & Han, S. (2007). Neural basis of cultural influence on self representation. *NeuroImage*, 34 (3), 1310-1317. [10.1016/j.neuroimage.2006.08.047](https://doi.org/10.1016/j.neuroimage.2006.08.047)

---

# Should we be Epigenetically Proactive?

A Commentary on Kathinka Evers

Stephan Schleim

---

“Can we be epigenetically proactive?”, is the question asked by Evers in her paper in this collection. After describing an original approach to using insights from the epigenesis of neural networks to develop new training and treatment programs, in particular to educate children and adolescents to become less violent and more sympathetic, the author suggests that there is a naturalistic responsibility for using science in this manner. In this commentary, I relate her proposal to the human enhancement debate at large, with a focus on the prevalent concept of human wellbeing. After a discussion of the factors that account for people’s quality of life and the role of research that allows them to decide the priorities for a good life themselves, three caveats against Evers’s approach are presented: (1) that epigenetic intervention carries the risk of psychological side-effects; (2) that people’s autonomy must be respected; and (3) that the world’s situation may not be as bad as suggested by the author when describing the benefits of her proposal. It is therefore concluded that, at least for the time being and until these challenges are met, we should not be epigenetically proactive.

## Keywords

Adaptation | Autonomy | Neuroenhancement | Social engineering | Wellbeing

## Commentator

[Stephan Schleim](#)  
s.schleim@rug.nl  
Rijksuniversiteit Groningen  
Groningen, Netherlands

## Target Author

[Kathinka Evers](#)  
kathinka.evers@crb.uu.se  
Uppsala Universitet  
Uppsala, Sweden

## Editors

[Thomas Metzinger](#)  
metzinger@uni-mainz.de  
Johannes Gutenberg-Universität  
Mainz, Germany

[Jennifer M. Windt](#)  
jennifer.windt@monash.edu  
Monash University  
Melbourne, Australia

## 1 Introduction

[Kathinka Evers](#) [this collection](#) discusses the possibility of changing people epigenetically. In particular, she discusses the option of increasing sympathy and decreasing xenophobia and violence. The term “*epigenetics*” is often used to describe processes affecting the activity of genes such as DNA methylation, which might enable the inheritance of acquired properties ([Bird 2007](#)). In contrast to this meaning, Evers uses the term more narrowly, with reference to the epigenesis of neural networks by selective stabilisation of synapses as an essential mechanism of

brain development ([Changeux & Danchin 1976](#)). The idea of affecting people’s development—or *ontogenesis*—through this mechanism, in order to achieve a desired state (e.g., an increase in sympathy) and/or to avoid an undesired state (e.g., a decrease in xenophobia or violence) can then be called *epigenetic proactivism*.

After describing human beings as social individualists and egocentric evaluators predisposed for selective sympathy and xenophobia, Evers explains neuronal epigenesis in detail. By influencing synaptic selection, this process may

critically affect social and cultural evolution. The central brain area for this is, according to the author, the prefrontal cortex, which is involved in planning, decision-making, thought, and socialisation; in particular, lateral prefrontal areas are associated with behaviour control. With respect to a task developed to test prefrontal cortex functioning, namely the Wisconsin Card Sorting Task (Dehaene & Changeux 1991), Evers discusses how neuronal epigenesis could explain rule-learning and top-down control. Finally, she devises two examples—adolescent violence in relation to their social environments and violence in adults associated with interconfessional conflicts—to illustrate what epigenetic proactivism may mean in practice. She eventually invokes a *naturalistic responsibility* to use the respective scientific and philosophical knowledge for the benefit of ourselves and our societies.

In this commentary, I will start out by relating Evers’s proposal to the *human enhancement* debate, which has received much attention recently—in particular within neuroethics. After summarising the general assumptions and caveats of this debate, I will elaborate on the definition of people’s wellbeing prevalent in the discourse on human enhancement and present an alternative based on social science research.

Finally, I will discuss epigenetic proactivism, Evers’s original proposal for changing people, in more detail. Arguing that the actual means—whether neurobiological, psychological, or social—do not matter very much, while issues related to adaptation, autonomy, and instrumentalisation are of essential ethical and philosophical relevance, I will emphasise the role of an individual’s *informed decision*. I will discuss in particular the three theses that (1) their proposed epigenetic intervention carries the risk of psychological side-effects; (2) that people’s autonomy must be respected; and (3) that the world’s situation may not be as bad as suggested by the authors when describing the benefits of their proposal. My conclusion will therefore be that the ethical justifiability of epigenetic proactivism critically depends on whether people can freely choose themselves whether or not to become epigenetically proactive, in a

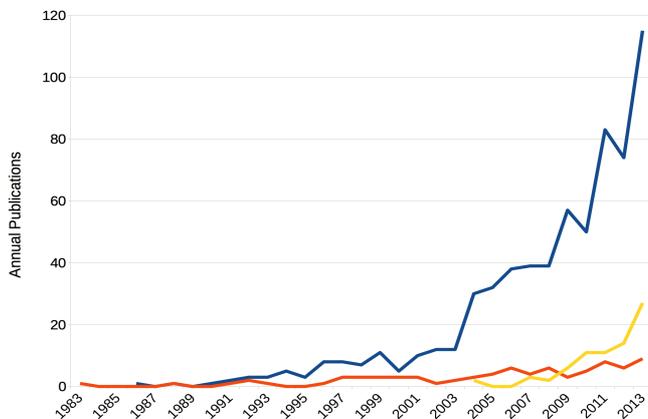
situation sufficiently free from social coercion and in sufficient awareness of the likely outcomes—effects as well as side-effects—of that intervention.

## 2 The human enhancement debate

In a paper on the “biopolitics” of cognitive enhancement, Peter Reiner recently referred to Plato’s *Phaedros*, where Socrates discusses what we nowadays might call the psychological side-effects of writing, namely the risk that our memory skills will deteriorate when we rely more on written texts (2013). Interestingly, Socrates’s concerns—voiced some 2400 years ago—seem to be confirmed by recent experiments indicating that people are less likely to remember information when they expect it to be easily accessible with the aid of computers (Sparrow et al. 2011). It goes without saying that everything we do has some psychological or neural impact, whether transient or permanent. However, writing—and, more recently, digital information processing—can be seen as an enhancement technology, as it enables asynchronous and distant communication with contemporaries as well as saving thoughts and ideas for the future.

We should keep in mind, though, that the very notion of *cognitive enhancement* was introduced only recently into the scholarly debate and its increasing prevalence coincided with the institutionalisation of neuroethics in the early 2000s (Figure 1). In the meantime, some authors criticised the exaggerated promises of the debate, pointing out misperceptions in the assessment of pharmacological enhancement behaviour, the complexity of the brain’s neurotransmitter systems, and the insufficient success of the much larger bio-psychiatric paradigm of improving psychological functioning in those looking for treatment (Lucke et al. 2011; Quednow 2010; Schleim 2014a). The latter means that even when the aims of the intervention are clearly circumscribed—e.g., decreasing the severity of the symptoms characteristic of a disorder—and research funds are abundant, bio-psychiatric research has unfortunately not been as successful as expected. This may relativise the hopes for effective bio-

psychological enhancement in the healthy in the near future.



**Figure 1:** Publications on enhancement. Publication data from the ISI Web of Science show a steep increase in publications covering “cognitive enhancement” (blue) that coincides with the institutionalisation of neuroethics (Farah 2012). “Mood” or “affective enhancement” (orange) and “neuroenhancement” (yellow) are addressed much less frequently, although these topics also are increasingly discussed. (ISI Web of Science Topic Search)

While describing writing as a means of cognitive enhancement may seem plausible at first glance, it also carries the risk of neglecting several distinctions that may be ethically and socially important. Such distinctions are, for example, those between learning the use of an instrument to achieve a certain aim and oneself becoming an instrument for the aims of others; between using an external device and directly interfering in the body; and between defining ends autonomously and being adapted to another’s ends heteronomously. Distinctions in actual cases will not always be clear and often fall into a grey zone, but this does not mean that possible interventions cannot be discussed against these concepts. These may be understood as marking the ends of a spectrum: for example, from full autonomy to full heteronomy. Indeed, while some scholars frame the consumption of stimulus drugs such as amphetamine, methylphenidate, or modafinil by students as individual choices for better cognitive functioning (Greely et al. 2008), that is, in an autonomous fashion, several results suggest that stu-

dents might rather respond to the demands of a competitive academic environment, and thus heteronomously. I will argue later that this opposition between freedom and coercion is the crucible of ethically assessing epigenetic proactiveness.

There is already empirical evidence from representative surveys or interviews with students that emphasises the relevance of this distinction. For example, M. Elizabeth Smith & Martha Farah describe in their extensive review on “smart pills” that the largest nationwide study identified admissions criteria (competitiveness) as well as two other social factors as the strongest predictors of stimulant drug consumption (2011). Interviews with non-medical consumers of stimulant drugs at an “elite” college carried out by Scott Vrecko suggest that people use stimulants for emotional and motivational ends rather than for cognitive enhancement, in particular to increase motivation to begin with or to complete boring tasks (2013). Finally, reviewing forty studies on public attitudes toward pharmacological cognitive enhancement, Kimberly J. Schelle and colleagues found that coercion to use drugs is a consistently mentioned concern (Schelle et al. 2014). This evidence associates the availability of enhancements like stimulant drugs with the pressure to adapt people to given standards of performance. Yet in the scientific literature the notion of cognitive enhancement is much more prevalent than the emotional and motivational aspects frequently mentioned in practical contexts (Figure 1).

Scientists and policy-makers in the *UK Foresight Project on Mental Capital and Well-being* note that globalisation increases demands for competitiveness as well as the pressures in our working lives (Beddington et al. 2008; Foresight Project 2008). They conclude that in a rapidly changing world like ours, we must make the most of all our resources in order to keep up with competitors; whole countries have to capitalise on their citizens’ cognitive resources. To achieve this aim, John Beddington and colleagues see vast possibilities in improving a country’s “mental capital” for all members of the population. They identify the possibility to do so at each stage in life, such as the early

identification and treatment of people with learning difficulties or the governmental support of those who want to work longer—though, notably, not shorter. A failure to react in a timely way to the challenges would come at a high cost for society, while early intervention in education could improve productivity at work and avoid costs related to a loss of mental capital (Beddington et al. 2008).

This view on performance enhancement for individual and social welfare reflects the focus of influential papers in neuroethics, emphasising the potential improvement of attention, memory, or wakefulness through the consumption of stimulant drugs or other pharmacological substances and neuroscientific technologies affecting the nervous system (Farah et al. 2004; Greely et al. 2008). Assumptions regarding the possible benefits of such substances are frequently based on trials employing test designs from *clinical psychology*, developed to identify and trace impairment in psycho-behavioural functioning, whether the investigated sample consists of patient populations, healthy people, or both (Bagot & Kaminer 2014; Repantis et al. 2010; Smith & Farah 2011).

Even if such test designs are of high clinical value, it is much less clear what statistically significant, yet often subtle, improvements in such experimental tasks, for example, in planning or memory games, mean for the *living environment* of the healthy. Whether such improvements indeed translate into an increase in individual wellbeing or the mental capital of a nation has yet to be shown. Indeed it is not even clear what a reliable and ecologically valid way of answering this question would look like. While this is still quite challenging after much debate on pharmacological enhancement, it is presently even less clear what such a standard could look like for epigenetic proactivism. In addition to measuring the benefits, neuroscientists frequently address the possibility of a psycho-behavioural trade-off—that is, the risk that an improvement in one domain would come at a loss in others (Brem et al. 2014; Hills & Hertwig 2011; Quednow 2010; Wood 2014). Given these complexities in the empirical research on

enhancement, it will be helpful to introduce an explicit definition for further discussion.

Human Enhancement =<sub>DF</sub> A change in the biology or psychology of a person which increases the chances of leading a good life in the relevant set of circumstances.

Notice how this definition, proposed by Julian Savulescu and colleagues in the introduction to a recent edited volume on human enhancement (Savulescu et al. 2011), relates the good life of an individual—its biology or psychology—to the context in which that individual lives: human enhancement is something done to or with a particular person in a fixed set of circumstances, namely, a change in her or his biology or psychology. This choice already predisposes the debate and research on enhancement with respect to adapting an individual to her or his environment.

To provide an illustrative and provocative counterexample: under this definition the “treatment” of a homosexual suffering from social exclusion by instigating heterosexual acts and relations, as was routinely performed by clinical psychologists and psychiatrists until the 1970s (Barlow 1973; Hinrichsen & Katahn 1975), would qualify as a form of human enhancement—inasmuch as it succeeds in “helping” the subject to avoid the undesired sexual behaviour that instigates social exclusion and the suffering probably caused by it. With respect to this historical example we already know that leading psychiatrists later acknowledged that there was nothing inherently wrong with homosexuals, but that their suffering indeed originated from social exclusion; this reasoning eventually led to the decision not to consider homosexuality a mental disorder any longer (Friedman et al. 1976). It is instructive to contrast the definition proposed by Savulescu and colleagues with the following inverted alternative.

Human Enhancement-Inverted =<sub>DF</sub> A change in the relevant set of circumstances that increases the chances of a person to lead a good life according to her or his preferences.

This alternative is not meant to be a logical inversion, but instead switches the levels of intervention, of that which is malleable and that which is considered as given. In an experimental fashion, one could also say that it is about a switch of dependent and independent variables, from the individual to its life context. Yet the aim of the intervention remains unchanged: increasing the chances of leading a good life. It goes without saying that both definitions, when put into practice, are constrained by available means and ethical principles, for example also requiring that we take the likelihood of other people's chances of leading a good life into account. It is not necessary here to argue that the inverted definition is better than the original; my intention is merely to show that we need not focus on bio-psychological changes alone. Instead, we can target the *social context* as well, decreasing the risk of adapting people to a social standard. Please note that this in itself does not imply a normative judgment, but rather widens the perspective for further analysis by taking alternative levels of intervention into consideration. As mentioned before, the balance between freedom and coercion, and autonomy and heteronomy will be essential with respect to epigenetic proactivism.

Here I have described some basic assumptions and criticism of the neuroethics debate on human enhancement, including the association of wellbeing with standards developed in clinical contexts that focus on individuals rather than on their social contexts. In the next section I will introduce research aimed at describing and understanding what people themselves consider to be quality of life, which poses an alternative to the standard adapted from clinical psychology.

### 3 Who defines wellbeing?

The position paper on cognitive enhancement by Henry Greely and colleagues starts out with the claim that “[s]ociety must respond to the growing demand for cognitive enhancement” (Greely et al. 2008, p. 702). The article by Beddington and colleagues on the mental wealth of nations begins with the conclusion that “[t]o

prosper and flourish in a rapidly changing world, we must make the most of all our resources—both mental and material” (Beddington et al. 2008, p. 1057). Both statements are similar in that they frame recent developments in such a way that they necessitate a reaction: we “must” respond in a particular manner. Greely and colleagues call for a “responsible use of cognitive-enhancing drugs by the healthy” (Greely et al. 2008, p. 702), though the majority of readers responding to their paper understood them as exaggerating the benefits of drug use generally or as being financially influenced by drug companies (Greely 2010). Beddington and colleagues call for the maximisation of our resources. All these authors want to increase benefits and decrease harms. However, who defines what counts as a benefit, as wellbeing, or as a good life? This is an essential and fundamental question that will influence every benefit-risk-analysis on human enhancement (Nagel 2014; Schleim 2014b).

As mentioned in the previous section, several scholars discuss the potential of means for enhancement, particularly psychopharmacological drugs, with respect to studies employing clinical test designs—whether investigating healthy people, those with a mental disorder, or even animals. Such tests measure reaction times or error rates in tasks requiring, for example, attention, memory, or planning. That is, the experimental setting frequently originates from a pragmatic context guided by identifying, treating, and/or predicting the development of a certain mental disorder. The underlying *mental disorder concept*, which is in itself controversial and subject to recurrent modifications, essentially hinges on a subject's clinically significant distress or functional impairment in the domain of cognition, emotion, and behaviour (American Psychiatric Association 2013; Stein et al. 2010). However, benefit, wellbeing, or a good life as discussed in the debate on human enhancement at large are not merely the opposites of clinically significant impairment; a five percent increase, say, in a task where a subject has to memorize as many digits as possible, and that may identify memory problems, does not reflect an increased performance in a real test, not

even a maths exam at school or university. Much less is it a suitable indicator of a benefit for the quality of life, although such a finding may be sufficient for publication in a peer-reviewed pharmacological journal.

However, there are advanced, direct, and representative measures of the quality of life. One example is the United Nations *World Happiness Report*, which compares the situation in 156 countries. The variables GDP per capita, social support, healthy life expectancy at birth, freedom to make life choices, generosity, and perceptions of corruption together explain 75.5% of the international variance of world happiness in 2012 (Helliwell et al. 2013). A more recent development is based on the OECD *Guidelines on Measuring Subjective Well-being* (OECD 2013). These allow people to create their own *Better Life Index*, prioritising eleven pre-defined domains such as education, jobs, housing, or safety.

More than 60,000 citizens from OECD countries have so far submitted their preferences, yielding important regional differences.<sup>1</sup> For example, people from the USA valued housing (on average 7.8 on a scale up to 10 points) and income (10.0) the highest, but work–life balance comparatively low (5.3). By contrast, people from Denmark, which is number one in the World Happiness Report, prioritised work–life balance higher than all others (9.8), and also valued life satisfaction (9.4) and community (10.0) very highly, while considering income less important (4.0). One may raise the question, of course, whether such statements are biased by social stereotypes or social desirability, but what could be a better measure of what people find important for leading a happy life than asking them directly? This is particularly so when they participate in the survey entirely on their own account.

These results emphasise two essential points for the human enhancement debate: first, people differ individually as well as regionally on what they find important for their wellbeing. Second, many of these aspects are not directly based on bio-psychological factors, but on social

factors. Indeed, the OECD construct of subjective wellbeing focuses on income, health status, social contact, employment status, personality type, and culture as determinants of life satisfaction, affect, and eudaimonic wellbeing. Unlike clinical measures of psycho-behavioural performance, they do not primarily rely on functional impairment.

Most importantly, the Better Life Index allows people to indicate themselves what they find important for their subjective wellbeing; and it turns out that many of these aspects, like housing or safety, are actual social factors that can only very indirectly be targeted by bio-psychological intervention. Therefore it becomes clear that a biased or narrowed concept of human enhancement carries the risk of missing the point of what determines or enables a better life. Further systematic analysis beyond the scope of this paper is required to show whether the factors identified are more amenable to individual psychobiological intervention, such as targeted by Savulescu and colleagues (Savulescu et al. 2011), or socio-political initiatives. Yet, while Greely and colleagues or Beddington and colleagues merely assume that increased cognitive performance will increase people's quality of life (Beddington et al. 2008; Greely et al. 2008), an initiative like the OECD Better Life Index allows people to autonomously express their own views on the issue and thus provides robust empirical evidence. This strategy helps to avoid two normative fallacies: first, that a parentalistic decision is possible when it comes to what should be good for others and, second, the idea that just because some intervention leads to a higher test score it is therefore good.

This section has highlighted, again, the tension between individual freedom and social adaptation, between autonomy and heteronomy. While most scholars would emphasise that people should be free to choose for themselves, fundamental definitions as well as the framing of human enhancement can implicitly narrow freedom, for example by introducing a limited standard for quality of life or by constraining the target for intervention. That is, when people apparently have free choice, because they are asked to choose from a number of alternatives

<sup>1</sup> <http://www.oecdbetterlifeindex.org>  
accessed July 18, 2014

that choice may actually be quite limited, because the offered options neglect important alternatives.

As described in the previous section, people are well aware of the threat of coercion when discussing the prospects of enhancement. Coercion does not only exist at gunpoint, when acting under duress in a strong legal sense, but it can also come in a much less direct manner: For example, by telling people that they *must* choose from a limited set of options, because otherwise something bad is going to happen. Referring to what, putatively, many people are already doing or what globalisation requires increases the pressure on individuals. There are meaningful and evidence-based alternative views on human enhancement, beyond those focusing on functional impairment, as shown in this section. In the next section, I will focus on the epigenetic proactivism proposed by Kathinka Evers in more detail.

#### 4 Epigenetic proactivism

Evers starts out their description of the naturalistic responsibility to become epigenetically proactive with a reference to the *Universal Declaration of Human Rights*. She criticises that, understood as a description of the present world, it is false to assume that all humans are born free and equal in dignity and rights; and if we understood this as a normative ideal, it would be unrealistic to guarantee these rights for every human being, given our present cerebral structure. In contrast to the human rights ideal, many people suffer from poverty and insufficient health care, and live through serious conflicts. Most people lack the sympathy necessary to respect the rights of others and all humans exhibit some kind of xenophobia. In the end, Evers even refers to the idea that humans might be subject to some built-in error or deficiency, predisposing us to self-destruction. Against this background, she proposes her epigenetic proactivism as follows:

Synaptic epigenetic theories of cultural and social imprinting on our brain architecture open the door to being epigenetically proactive, which means that we may culturally influence our brain organisation

in the aim of self-improvement, individually as well as socially and change our biological predispositions by a better fit of our brain to cultures and social structures. (Evers [this collection](#), p. 12)

She discusses two examples in more detail, namely violence in adolescents and violent inter-confessional conflicts. Referring to neurodevelopmental research on children and teenagers' brains, she suggests that different educational measures such as physical exercises, cultural games, and new therapies amount to a kind of proactive epigenetic imprinting that increases control of aggression, emotion regulation, sympathy, and tolerance. It would be largely a matter of political will and social agreement, Evers claims, to develop the research enabling such educational programs and to apply them in practice. If successful, epigenetic proactivism would make societies more peaceful and inclusive, but the author also points to a problematic circularity, namely that we perhaps first need to live in an already peaceful society in order to enact such educational programs to maintain peace.

If we had to choose between epigenetic proactivism and the destruction of humankind, the decision would probably be easy; and the humbler prospect of avoiding adolescent violence and interconfessional conflicts also has some seductive allure. However, for three reasons I hesitate to agree with the conclusion that we have a naturalistic responsibility to improve ourselves epigenetically, assuming that science will develop enough at some point and offer the novel educational measures suggested by Evers: first, decreasing the disposition towards aggressive behaviour and increasing sympathy might have unexpected psychological side-effects; second, the value of human autonomy has to be considered by epigenetic proactivists, too; and third, the human condition might not be as bad as the author describes. I will discuss these three caveats in the following sections.

##### 4.1 Side-effects of epigenetic proactivism

At first glance, who would disagree that a world with less aggression and more sympathy would

be a better world? If we could indeed decrease adolescent and interconfessional violence, why shouldn't we put such an educational program into action? Evers refers to Darwin and evolution several times in her paper. Consequently, this biological framing also raises the question of the possible evolutionary value of aggression and violence (Eibleibesfeldt 1977; Smith & Harper 1988). Darwin's original idea of the survival of the fittest emphasises the very notion of securing access to scarce resources—often at the cost of other living beings, which may even lead to the extinction of a whole species. It may well be that aggression is an essential driver of evolutionary development.

It goes without saying that from the fact that something leads to an increased survival value it does not follow that it is morally good. But it is clear that, even from a social perspective, aggression might have a function, or might be necessary for achieving some desirable ends. In the famous novel *A Clockwork Orange* by Anthony Burgess, we learn about a fictional case where a cruel and ruthless juvenile delinquent—Alex—is successfully treated bio-psychologically to stop being violent. This is carried out in a pharmacologically enhanced operant conditioning program that associates scenes of violence with aversive stimuli, such that the former delinquent feels severe nausea whenever he is confronted with aggression, including assaults against himself. This has the side effect that after the treatment Alex cannot defend himself anymore and he therefore becomes a victim of severe humiliation.

While this example is different from the case of interconfessional violence discussed by Evers, it is directly related to her other example of violence in adolescents. It is a complex biopsychological question whether negative facets of aggression can be extinguished without also affecting people's capacity for self-defence. The author is aware of the problem of circularity, that a world may first have to become peaceful for epigenetic proactivism to be successful—and the present caveat emphasises this dilemma: if only some people were educated to avoid violence and conflicts, this could easily be abused by others.

How about increasing sympathy, then? Evers is critical about the fact that people are xenophobic and restrict their sympathy to small groups, while they should ideally extend it to human society at large. As disappointing as it may be from an ethical point of view, it could well be that a distinction between one's own or one's group's welfare from that of others is essential for psychological wellbeing. A dysfunctional self–other distinction, drawing a clear line between oneself and others, may play a role in schizophrenia (Decety & Sommerville 2003; Jardri et al. 2011). Furthermore, several investigations reported an association between emotional empathy and depression or decreased life-satisfaction (Gawronski & Privette 1997; Lee et al. 2001; O'Connor et al. 2002).

These links with mental health may be speculative to some extent, yet they illustrate that even a *prima facie* positive capacity may become negative when increased too much. Accordingly, it has become common wisdom within psychopharmacology that there is an optimal level of neurotransmitter concentration in the brain and that both a decrease and an increase may be dysfunctional and/or lead to unexpected side-effects (Wood et al. 2014). Even if ethicists, in line with Evers, presented strong arguments in favour of considering the welfare of those far away from oneself or one's group (Greene 2003; Sidgwick 1907; Singer 2002; Unger 1996), it should be born in mind that an increase of sympathy might lead to a decrease in subjective wellbeing.

## 4.2 Human autonomy

The vision of a scientifically enhanced world, where people are better at controlling their emotions, particularly aggression and other impulses that might lead to violent behaviour, is a recurrent topic in the history of science. For example, in the 1960s and 1970s, neuroscientists, psychologists, and sociologists all discussed the problem of delinquency and aggression, also with respect to adolescents, and proposed different solutions for coping with it. The pioneer of *brain stimulation*, José Delgado, tested the effects of electrical inhibition or excitation of dif-

ferent brain areas associated with emotion processing, such as the amygdalae, in several animal species as well as in humans (Delgado 1965, 1971; Delgado et al. 1968). His discussion of the social implications of such technology is surprisingly reminiscent of epigenetic proactivism:

Understanding of biology, physics, and other sciences facilitated the process of ecological liberation and domination. Man rebelled from natural determination and used his intelligence and skills to impose a human purpose on the development of the earth. We are now on the verge of a process of mental liberation and self-domination which is a continuation of our evolution. Its experimental approach is based on the investigation of the depth of the brain in behaving subjects. Its practical applications do not rely on direct cerebral manipulations but on the integration of neurophysiological and psychological principles leading to a more intelligent education, starting from the moment of birth and continuing throughout life, with the pre-conceived plan of escaping from the blind forces of chance and of influencing cerebral mechanisms and mental structure in order to create a future man with greater personal freedom and originality, a member of a psychocivilized society, happier, less destructive, and better balanced than present man. (Delgado 1971, p. 223; reference omitted)

He and others (e.g., Mark & Ervin 1970; Valenstein 1973) were convinced that therapeutic need would drive the development of such neurotechnology. The envisioned “psychocivilized” world would be so beneficial for individuals and society at large, Delgado believed, that the advantages overruled any social and ethical caveats (Delgado 1971). At the same time, the psychologist Burrhus Skinner wrote a best-selling book on his vision of a peaceful society realised through *social engineering* and inspired by behaviourism rather than neurotechnology (Skinner 1971). Through rewarding the right

kind of actions, Skinner suggested, the socially desired behaviour would become more likely, and the undesired behaviour more unlikely. To avoid a totalitarian regime, the people subject to this social engineering should in turn control the reward structures, the so-called contingencies of a society. Yet, in spite of the book’s popularity, it was strongly criticised by Noam Chomsky for confusing science and politics and for a misapplication of central notions such as freedom and dignity (1971).

The two utopian proposals by Delgado and Skinner, the part of the human enhancement debate discussed above that describes a need for adaptation as without alternative, and epigenetic proactivism have in common that people should be changed in such a way that they contribute to a (putatively) desired social aim: a macroscopic state with better performance, competitiveness, peacefulness, and/or caring for others. This is in obvious conflict with the notion of autonomy that is so fundamental to Immanuel Kant’s moral philosophy: no human being must be treated only as a means to another end; all humans must also be treated as an ends in themselves (1785/1994). Given the description of epigenetic proactivism by Evers, stating that our brains shall fit better to our cultures and social structures, one may well ask whether those enhanced in this manner would not become mere instruments for the present system, with its social norms and values. Also with respect to John Stuart Mill’s utilitarian liberalism, interventions to improve people seem problematic, as Mill formulated the principle:

[...] that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will

make him happier, because, in the opinions of others, to do so would be wise, or even right. These are good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him, but not for compelling him, or visiting him with any evil in case he do otherwise. [...] Over himself, over his own body and mind, the individual is sovereign. (1859/1989, pp. 17–18)

Interestingly, Mill explicitly formulated the exception of self-protection and harm to others, to which Evers refers in her paper as well. However, I doubt that epigenetic proactivists can base their ethical justification on this case, as the harm they want to avoid is very indirectly related to intervention—which will most likely be applied to many people who would not have posed a threat to others without it. Furthermore, it can be doubted how imminent the danger is at all; this last point will be elaborated in the next subsection. Although other and more recent versions of “utilitarianism”, such as preference utilitarianism, place less emphasis on autonomy than Kant or Mill, they also lend the inner core of a person, for example, her or his preferences and values, a status of special protection (Singer 2011). This core is likely to be affected by changing people’s predisposition to aggression and sympathy, as the brief description of psychological side-effects in the previous subsection suggests.

Therefore, the essential question for epigenetic proactivism seems to be whether people can autonomously consent to the intervention. Evers’s title asks whether we *can* be epigenetically proactive; I have reformulated this to ask whether we *should* be epigenetically proactive. Here it is particularly relevant that her two examples, adolescent and interconfessional violence, explicitly address the development of children and teenager’s brains—that is, people whom we do not usually consider to be (fully) autonomous. The question of whether parents can take this decision, aimed at rewiring the nervous system of their children for a social aim, is too complex to be discussed here, but it calls for a solution before we can really think

about putting epigenetic proactivism into practice.

For our present purposes it shall suffice to suggest that it is unlikely that all parents would consent to such a measure. What would then happen to those who declined to participate in epigenetically proactive educational programs? Even today, some families resist education because they see a conflict between their values and teaching on, for example, sex education or evolutionary theory. In particular, those who benefit from the present social order would be unlikely to consent to a measure that might lead to a loss of power for them. As mentioned earlier, this may make those who are made less aggressive and more empathic more likely to be exploited by those who are not. Therefore, it is an essential challenge for epigenetic proactivism to take autonomy, informed consent, and the further complexities of intervening in the core of a person’s personality into account—and to consider that people’s views on these issues will be diverse!

Until these challenges of autonomy and informed consent in particular are met, I draw the tentative conclusion that we should not be epigenetically proactive. It should be noted, though, that while I am discussing the proposal by Evers here, the argument from autonomy is independent of the means actually used to enhance people—whether biological, psychological, or social. Rather, it is essential that people are free from coercion and can decide for themselves whether or not they want to become the kind of human being envisioned by proponents in the human enhancement debate, and that they have sufficient knowledge on the implications of that choice. Evers particularly focusses on children and adolescents when discussing examples of epigenetic proactivism, but it appears to be most difficult to describe what autonomous and informed choice means in precisely this group of human beings.

### 4.3 The human condition

Evers emphasises that many people live in precarious circumstances, even more than sixty years after the Universal Declaration of Human

Rights; in the end, she even refers to Arthur Koestler's idea that humans might have some built-in deficiency, predisposing us to self-destruction. Obviously, against that prospect, the promises of epigenetic proactivism look seductive. Indeed, we must concede that even some twenty-five years after the Cold War international conflicts have not abated altogether—in some areas they have even multiplied, and terrorism or economic instability are a concern for many. However, from the perspective of cultural evolution, universal human rights are a rather novel development and it may be too early to take a pessimistic stance on their success and effect. Returning to the UN World Happiness Report (Helliwell et al. 2013), one may ask whether the difference between the leading countries—Denmark, Norway, Switzerland, the Netherlands, and Sweden (ranked 1<sup>st</sup> to 5<sup>th</sup>)—, those in the middle—Libya, Bahrain, Montenegro, Pakistan, and Nigeria (ranked 78<sup>th</sup> to 82<sup>nd</sup>)—, and those at the bottom—Rwanda, Burundi, the Central African Republic, Benin, and Togo (ranked 152<sup>nd</sup> to 156<sup>th</sup>)—can be explained or even overcome by means of human enhancement like epigenetic proactivism rather than internationally-aided institutional development.

One shared rhetorical feature of those visions of a better humankind is a claim that all has somehow gone wrong, and even to predict an imminent catastrophe. For example, the various *Humanist Manifestos* of the 20<sup>th</sup> and early 21<sup>st</sup> century described serious threats to human survival.<sup>2</sup> Delgado emphasized an imbalance between our material and mental evolution, putting humanity at risk (1971), and Skinner started out by referring to problems related to population growth, pollution of the environment, and nuclear armament (1971). It probably lies in the eye of the beholder to speculate whether humankind has not yet destroyed itself because or in spite of unprecedented technological powers.

It is a matter of fact that we have not yet done so, and although many things have gone

wrong, others have gone right. Steven Pinker recently gathered evidence that, particularly when viewed in relation to the vast population growth of humanity, our present times are much more peaceful than the past (2011). He describes processes of pacification and civilization as well as a humanitarian and rights revolution that can provide hope that things will change for the better, not only for the worse. Therefore, even if human enhancement in general or epigenetic proactivism in particular may offer genuine improvement of the human condition in several ways, they are probably not necessary for human survival.

## 5 Conclusion

Kathinka Evers summarises research on the epigenesis of neural networks to describe a vision of epigenetical proactivism, a development of new training and therapeutic programs to improve humans. She asks whether we *can* be epigenetically proactive, pointing out the benefits of decreasing the prevalence of adolescent and interconfessional violence, and in so doing develops her answer: yes, in principle, we can be epigenetically proactive. However, she also describes a naturalistic responsibility to do this, which is the point at which my discussion of her proposal diverged from her view. Particularly with respect to autonomy and free choice I think that, for the time being, we should not be epigenetically proactive; and we should be even more cautious when interventions in children's and teenagers' brains are at issue. Minor caveats are related to the possible psychological side-effects of decreasing our disposition towards aggression and increasing that of sympathy, as well as a more optimistic view of how humankind is developing.

In this paper, I also related epigenetic proactivism to the human enhancement debate more generally, which has become much more comprehensive than can be addressed in such a brief commentary. It was important to examine the definition of wellbeing and the framing of urgency, as well as the primary level of intervention—bio-psychological or social—, issues that are also related to autonomy. This does not

<sup>2</sup> See the three Manifestos of 1933, 1973, and 2003 of the American Humanist Association on <http://americanhumanist.org/Humanism/> (accessed July 21 2014).

mean that knowledge on epigenetics could not be used in another manner for the purposes of enhancement, in situations where people can make an informed decision for themselves whether and how to engage in a certain kind of training. In this sense, it would be interesting to compare epigenetic proactivism to other non-pharmaceutical means of enhancement, such as nutrition, exercise, sleep, or meditation (Dresler et al. 2013). Generally speaking, the knowledge described by Evers could also be related to debates on improving school education neuroscientifically (Hook & Farah 2013; Posner & Rothbart 2005). Furthermore, when targeting human capacities that are also salient for moral cognition, the debate on *moral enhancement* may be an important reference point with overlapping prospects and concerns (Douglas 2008, 2013; Harris 2011).

Evers warned that science has been hijacked repeatedly throughout history and that in particular the dream of creating perfect human beings has a sordid past. Here I wholeheartedly agree with her and her related call for historic awareness. I hope that I have succeeded in showing why, beyond this awareness, it is also essential to take people's own views and autonomy into account. It may not only be the case that too much focus on enhancing people makes them sad by focusing too much on their deficiencies (Schleim 2014b; Schopenhauer 1874), but in the attempt to create superhuman beings a human catastrophe might also be provoked.

## Acknowledgments

I would like to thank the two editors as well as two anonymous reviewers for their extraordinarily helpful and constructive comments on a previous version of this paper.

## References

- American Psychiatric Association, (Ed.) (2013). *Diagnostic and statistical manual of mental disorders (5<sup>th</sup> ed.)*. Arlington, VA: American Psychiatric Publishing.
- Bagot, K. S. & Kaminer, Y. (2014). Efficacy of stimulants for cognitive enhancement in non-attention deficit hyperactivity disorder youth: A systematic review. *Addiction*, 109 (4), 547-557. [10.1111/add.12460](https://doi.org/10.1111/add.12460)
- Barlow, D. H. (1973). Increasing heterosexual responsiveness in the treatment of sexual deviation: A review of the clinical and experimental evidence. *Behavior Therapy*, 4 (5), 655-671. [10.1016/s0005-7894\(73\)80158-3](https://doi.org/10.1016/s0005-7894(73)80158-3)
- Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R. & Thomas, S. M. (2008). The mental wealth of nations. *Nature*, 455 (7216), 1057-1060. [10.1038/4551057a](https://doi.org/10.1038/4551057a)
- Bird, A. (2007). Perceptions of epigenetics. *Nature*, 447 (7143), 396-398. [10.1038/nature05913](https://doi.org/10.1038/nature05913)
- Brem, A. K., Fried, P. J., Horvath, J. C., Robertson, E. M. & Pascual-Leone, A. (2014). Is neuroenhancement by noninvasive brain stimulation a net zero-sum proposition? *NeuroImage*, 85 (3), 1058-1068. [10.1016/j.neuroimage.2013.07.038](https://doi.org/10.1016/j.neuroimage.2013.07.038)
- Changeux, J. P. & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for specification of neuronal networks. *Nature*, 264 (5588), 705-712. [10.1038/264705a0](https://doi.org/10.1038/264705a0)
- Chomsky, N. (1971). The case against B.F. Skinner. *The New York Review of Books*, 17 (11), 18-24.
- Decety, J. & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7 (12), 527-533. [10.1016/j.tics.2003.10.004](https://doi.org/10.1016/j.tics.2003.10.004)
- Dehaene, S. & Changeux, J. P. (1991). The wisconsin card sorting test: Theoretical analysis and modeling in a neuronal network. *Cerebral Cortex*, 1 (1), 62-79. [10.1093/cercor/1.1.62](https://doi.org/10.1093/cercor/1.1.62)
- Delgado, J. M. (1965). Sequential behavior induced repeatedly by stimulation of the red nucleus in free monkeys. *Science*, 148 (3675), 1361-1363. [10.1126/science.148.3675.1361](https://doi.org/10.1126/science.148.3675.1361)
- (1971). *Physical control of the mind; Toward a psychocivilized society*. New York, NY: Harper & Row.
- Delgado, J. M., Mark, V., Sweet, W., Ervin, F., Weiss, G., Bach, Y. R. G. & Hagiwara, R. (1968). Intracerebral radio stimulation and recording in completely free patients. *Journal of Nervous and Mental Disease*, 147 (4), 329-340.

- Douglas, T. (2008). Moral enhancement. *Journal of Applied Philosophy*, 25 (3), 228-245. [10.1111/j.1468-5930.2008.00412.x](https://doi.org/10.1111/j.1468-5930.2008.00412.x)
- (2013). Moral enhancement via direct emotion modulation: A reply to John Harris. *Bioethics*, 27 (3), 160-168. [10.1111/j.1467-8519.2011.01919.x](https://doi.org/10.1111/j.1467-8519.2011.01919.x)
- Dresler, M., Sandberg, A., Ohla, K., Bublitz, C., Trenado, C., Mroczko-Wasowicz, A. & Repantis, D. (2013). Non-pharmacological cognitive enhancement. *Neuropharmacology*, 64, 529-543. [10.1016/j.neuropharm.2012.07.002](https://doi.org/10.1016/j.neuropharm.2012.07.002)
- Eibleibesfeldt, I. (1977). Evolution of destructive aggression. *Aggressive Behavior*, 3 (2), 127-144. [10.1002/1098-2337\(1977\)3:2<127::AID-AB2480030204>3.0.CO;2-Y](https://doi.org/10.1002/1098-2337(1977)3:2<127::AID-AB2480030204>3.0.CO;2-Y)
- Evers, K. (2015). Can we be epigenetically proactive? In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- Farah, M. J. (2012). Neuroethics: The ethical, legal, and societal impact of neuroscience. *Annual Review of Psychology*, 63, 571-591. [10.1146/annurev.psych.093008.100438](https://doi.org/10.1146/annurev.psych.093008.100438)
- Farah, M. J., Illes, J., Cook-Deegan, R., Gardner, H., Kandel, E., King, P. & Wolpe, P. R. (2004). Neurocognitive enhancement: What can we do and what should we do? *Nature Reviews Neuroscience*, 5 (5), 421-425. [10.1038/nrn1390](https://doi.org/10.1038/nrn1390)
- Foresight Project, (2008). *Final project report*. London, UK: The Government Office for Science.
- Friedman, R. C., Green, R. & Spitzer, R. L. (1976). Reassessment of homosexuality and transsexualism. *Annual Review of Medicine*, 27, 57-62. [10.1146/annurev.me.27.020176.000421](https://doi.org/10.1146/annurev.me.27.020176.000421)
- Gawronski, I. & Privette, G. (1997). Empathy and reactive depression. *Psychological Reports*, 80 (3), 1043-1049. [10.2466/pr0.1997.80.3.1043](https://doi.org/10.2466/pr0.1997.80.3.1043)
- Greely, H. (2010). Enhancing brains: What are we afraid of? *Cerebrum*, 14, 1-10.
- Greely, H., Sahakian, B., Harris, J., Kessler, R. C., Gazzaniga, M., Campbell, P. & Farah, M. J. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature*, 456 (7223), 702-705. [10.1038/456702a](https://doi.org/10.1038/456702a)
- Greene, J. (2003). From neural “is” to moral “ought”: What are the moral implications of neuroscientific moral psychology? *Nature Reviews Neuroscience*, 4 (10), 846-849. [10.1038/nrn1224](https://doi.org/10.1038/nrn1224)
- Harris, J. (2011). Moral enhancement and freedom. *Bioethics*, 25 (2), 102-111. [10.1111/j.1467-8519.2010.01854.x](https://doi.org/10.1111/j.1467-8519.2010.01854.x)
- Helliwell, J., Layard, R. & Sachs, J. (Eds.) (2013). *World happiness report 2013*. New York, NY: Sustainable Development Solutions Network, a Global Initiative for the United Nations.
- Hills, T. & Hertwig, R. (2011). Why aren't we smarter already: Evolutionary trade-offs and cognitive enhancements. *Current Directions in Psychological Science*, 20 (6), 373-377. [10.1177/0963721411418300](https://doi.org/10.1177/0963721411418300)
- Hinrichsen, J. J. & Katahn, M. (1975). Recent trends and new developments in the treatment of homosexuality. *Psychotherapy-Theory Research and Practice*, 12 (1), 83-92. [10.1037/h0086413](https://doi.org/10.1037/h0086413)
- Hook, C. J. & Farah, M. J. (2013). Neuroscience for educators: What are they seeking, and what are they finding? *Neuroethics*, 6 (2), 331-341. [10.1007/s12152-012-9159-3](https://doi.org/10.1007/s12152-012-9159-3)
- Jardri, R., Pins, D., Lafargue, G., Very, E., Ameller, A., Delmaire, C. & Thomas, P. (2011). Increased overlap between the brain areas involved in self-other distinction in schizophrenia. *PLoS One*, 6 (3), e17500. [10.1371/journal.pone.0017500](https://doi.org/10.1371/journal.pone.0017500)
- Kant, I. (1994). *Grundlegung zur Metaphysik der Sitten*. Hamburg, GER: Meiner.
- Lee, H. S., Brennan, P. F. & Daly, B. J. (2001). Relationship of empathy to appraisal, depression, life satisfaction, and physical health in informal caregivers of older adults. *Research in Nursing & Health*, 24 (1), 44-56. [10.1002/1098-240x\(200102\)24](https://doi.org/10.1002/1098-240x(200102)24)
- Lucke, J. C., Bell, S., Partridge, B. & Hall, W. D. (2011). Deflating the neuroenhancement bubble. *American Journal of Bioethics Neuroscience*, 2 (4), 38-43. [10.1080/21507740.2011.611122](https://doi.org/10.1080/21507740.2011.611122)
- Mark, V. H. & Ervin, F. R. (1970). *Violence and the brain*. New York, NY: Harper & Row.
- Mill, J. S. (1989). *On liberty*. London, UK: The Walter Scott Publishing Co.
- Nagel, S. K. (2014). Enhancement for well-being is still ethically challenging. *Frontiers in Systems Neuroscience*, 8 (72). [10.3389/fnsys.2014.00072](https://doi.org/10.3389/fnsys.2014.00072)
- O'Connor, L. E., Berry, J. W., Weiss, J. & Gilbert, P. (2002). Guilt, fear, submission, and empathy in depression. *Journal of Affective Disorders*, 71 (1-3), 19-27. [10.1016/s0165-0327\(01\)00408-6](https://doi.org/10.1016/s0165-0327(01)00408-6)
- OECD, (2013). *OECD guidelines on measuring subjective well-being*. Paris, FR: OECD Publishing.
- Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. New York, NY: Viking.
- Posner, M. I. & Rothbart, M. K. (2005). Influencing brain networks: Implications for education. *Trends in Cognitive Sciences*, 9 (3), 99-103. [10.1016/j.tics.2005.01.007](https://doi.org/10.1016/j.tics.2005.01.007)

- Quednow, B. B. (2010). Ethics of neuroenhancement: A phantom debate. *BioSocieties*, 5 (1), 153-156. [10.1057/biosoc.2009.13](https://doi.org/10.1057/biosoc.2009.13)
- Reiner, P. B. (2013). Biopolitics of cognitive enhancement. In E. Hildt & A. Franke (Eds.) *Cognitive enhancement: An interdisciplinary perspective* (pp. 189-200). Dordrecht, NL: Springer.
- Repantis, D., Schlattmann, P., Laisney, O. & Heuser, I. (2010). Modafinil and methylphenidate for neuroenhancement in healthy individuals: A systematic review. *Pharmacological Research*, 62 (3), 187-206. [10.1016/j.phrs.2010.04.002](https://doi.org/10.1016/j.phrs.2010.04.002)
- Savulescu, J., Sandberg, A. & Kahane, G. (2011). Well-being and enhancement. In J. Savulescu, R. H. J. ter Meulen & Kahane (Eds.) *Enhancing human capacities* (pp. 3-18). Oxford, UK: Wiley-Blackwell.
- Schelle, K. J., Faulmüller, N., Caviola, L. & Hewstone, M. (2014). Attitudes towards pharmacological cognitive enhancement – A review. *Frontiers in Systems Neuroscience*, 8 (53). [10.3389/fnsys.2014.00053](https://doi.org/10.3389/fnsys.2014.00053)
- Schleim, S. (2014a). Critical neuroscience – or critical science? A perspective on the perceived normative significance of neuroscience. *Frontiers in Human Neuroscience*, 8 (336). [10.3389/fnhum.2014.00336](https://doi.org/10.3389/fnhum.2014.00336)
- (2014b). Whose well-being? Common conceptions and misconceptions in the enhancement debate. *Frontiers in Systems Neuroscience*, 8 (148). [10.3389/fnsys.2014.00148](https://doi.org/10.3389/fnsys.2014.00148)
- Schopenhauer, A. (1874). *Parerga und Paralipomena, Band I*. Zurich, SUI: Haffmans.
- Sidgwick, H. (1907). *The methods of ethics*. New York, NY: Macmillan and Co.
- Singer, P. (2002). *One world : The ethics of globalization*. New Haven, CT: Yale University Press.
- (2011). *Practical ethics*. Cambridge, UK: Cambridge University Press.
- Skinner, B. F. (1971). *Beyond freedom and dignity*. Toronto, Canada: Bantam.
- Smith, M. E. & Farah, M. J. (2011). Are prescription stimulants “smart pills”? The epidemiology and cognitive neuroscience of prescription stimulant use by normal healthy individuals. *Psychological Bulletin*, 137 (5), 717-741. [10.1037/a0023825](https://doi.org/10.1037/a0023825)
- Smith, J. M. & Harper, D. G. C. (1988). The evolution of aggression: Can selection generate variability? *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 319 (1196), 557-570. [10.1098/rstb.1988.0065](https://doi.org/10.1098/rstb.1988.0065)
- Sparrow, B., Liu, J. & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333 (6043), 776-778. [10.1126/science.1207745](https://doi.org/10.1126/science.1207745)
- Stein, D. J., Phillips, K. A., Bolton, D., Fulford, K. W. M., Sadler, J. Z. & Kendler, K. S. (2010). What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychological Medicine*, 1759 (1765), 1759-1765. [10.1017/s0033291709992261](https://doi.org/10.1017/s0033291709992261)
- Unger, P. K. (1996). *Living high and letting die: Our illusion of innocence*. New York, UK: Oxford University Press.
- Valenstein, E. S. (1973). *Brain control*. New York, NY: Wiley.
- Vrecko, S. (2013). Just how cognitive is “Cognitive Enhancement”? On the significance of emotions in university students’ experiences with study drugs. *American Journal of Bioethics Neuroscience*, 4 (1), 4-12. [10.1080/21507740.2012.740141](https://doi.org/10.1080/21507740.2012.740141)
- Wood, S., Sage, J. R., Shuman, T. & Anagnostaras, S. G. (2014). Psychostimulants and cognition: A continuum of behavioral and cognitive activation. *Pharmacological Reviews*, 66 (1), 193-221. [10.1124/pr.112.007054](https://doi.org/10.1124/pr.112.007054)

---

# Understanding Epigenetic Proaction

A Reply to Stephan Schleim

[Kathinka Evers](#)

---

Epigenetic proaction can be described as a way of steering evolution by influencing the cultural imprints stored in our brains. It is not to be confused with “human enhancement”. It is a process on the societal level that need not conflict with the notion of autonomy, nor suggest any “superhuman” ideal. Risks of misuse justify precaution, not abandonment of constructive scientific pursuits. Scientific knowledge can help us improve our life conditions in the long-term. A naturalistic responsibility is born out of science’s strong social relevance.

## Keywords

Autonomy | Enhancement | Epigenetic proaction | Precaution | Responsibility

## Author

[Kathinka Evers](#)

kathinka.evers@crb.uu.se

Uppsala Universitet  
Uppsala, Sweden

## Commentator

[Stephan Schleim](#)

s.schleim@rug.nl

Rijksuniversiteit Groningen  
Groningen, Netherlands

## Editors

[Thomas Metzinger](#)

metzinger@uni-mainz.de

Johannes Gutenberg-Universität  
Mainz, Germany

[Jennifer M. Windt](#)

jennifer.windt@monash.edu

Monash University  
Melbourne, Australia

## 1 Introduction

Epigenetic proaction can be described as a way of steering evolution by influencing the cultural imprints that are stored in our brains. The question analysed in my target article is what exactly this means and whether it is possible. Can we adapt our societies to constructively interact with the ever-developing neuronal architecture of our brains? The issue of whether such interaction is desirable is also raised but not discussed in depth.

In order to decide whether an action should be pursued it would be wise to first attempt to understand its nature and implications. Regrettably, in his commentary to my article, Stephan Schleim fails to acknowledge the main concern of my paper, namely the sci-

entific issue, moving instead to the normative question via some less relevant detours. The commentary therefore becomes misleading. Rather than engaging with the scientific points I make, Schleim takes as a starting point a flawed understanding of epigenetic proaction and tries to show how undesirable it would be. The arguments have little to do with the article on which he purports to comment.

## 2 Confusing epigenetic proaction with human enhancement

After making the assertion that “the actual means—whether neurobiological, psychological,

or social—do not matter very much” in his philosophical analysis of epigenetic proaction, [Schleim](#) proceeds to relate my position to the general debate on “human enhancement” ([this collection](#), p. 2). A long discussion follows about this debate that, although quite popular amongst some contemporary philosophers, is here out of context. In the target article, there is no mention of individual cognitive, moral, or performance enhancement, nor any mention of pharmaceutical “smart pills” and so on. The target article does not speak of epigenetic proaction as an individual opt-in/opt-out thing at all, nor does it speak of enhancement. And it certainly does not recommend, as [Schleim](#) suggests at the end of his commentary “the attempt to create superhuman beings” ([this collection](#), p. 15). The statement that my theory proposes methods for parents “aimed at rewiring the nervous system of their children for a social aim” ([Schleim this collection](#), p. 10) is a caricature. Perhaps the author has not read the target paper quite thoroughly enough. This would explain why the author does not specifically address any of the scientific issues raised in the paper.

### 3 Well-being and exaggerated virtues

In the commentary, the subsequent discussion is about who defines well-being and how. Whilst this in itself is an interesting question that deserves careful consideration from many perspectives, it is not directly relevant to the target article. The article raises the question of whether epigenetic proaction is possible, and presents scientific data and theories to explain what this means. On that basis, I suggest that they can be taken to support the view that it may indeed be possible. The questions of defining well-being or of specifying who should be in charge of defining well-being, whilst interesting, fall out of this scope.

In contrast, the question of “side-effects” can with some effort be considered at least somewhat relevant to the article under debate. Here, [Schleim](#) wonders: is it possible, e.g., to reduce aggression without making a person weak or meek? Can a less aggressive person defend

him- or herself against a more aggressive person? He seems to be doubtful, but my short reply is: obviously, yes. Much education, of children in particular and in human societies in general, includes attempts to check aggression—it does not thereby create either wimps or zombies. Even martial arts focus explicitly on checking aggression, whilst by definition aiming to make students excellent in combat. [Schleim](#) also wonders about the risky side-effects of increasing sympathy. He warns that increasing sympathy too much could perhaps lead to a “dysfunctional self–other distinction” that “may play a role in schizophrenia”. However, even if this were the case, this is not a necessary—or even very common—side-effect of increasing sympathy. Certainly, when we bring our children up to sympathise with others, we may increase their distress at the sight of suffering in others, but I do not believe that we thereby increase their risk of developing schizophrenia. Moreover, as a general principle, that an initially positive value can become negative if exaggerated does not entail that we should stop seeking it altogether. If that were the case, we would have little to strive for.

### 4 Epigenetic proaction: A process on the societal level

[Schleim](#) compares my theory to the famously misconceived social engineering projects of Skinner and Delgado, for whom, [Schleim](#) says, the goals blessed the means. He argues ([Schleim this collection](#), p. 9) that these “utopian proposals” stand “in obvious conflict with the notion of autonomy”, as understood by Immanuel Kant: no being must be treated only as a means to an end, but as an end in itself. I agree with Kant’s principle and see no conflict between it and the notion of epigenetic proaction. There is nothing in the idea of epigenetic proaction as I develop it in my article that suggests treating people as mere means to a social end, or of allowing them to “become mere instruments for the present system” ([Schleim this collection](#), p. 9). The idea in itself is neutral in this regard: of course the idea can be misused—all science can be misused—but it is no part of the theory to

have this negative consequence. In other words, there is no essential conflict between human autonomy and human epigenetic proaction properly understood.

As for the issue of informed consent that Schleim raises in that context, it does not directly arise through the topics I address in my article, but it would arise in the research that I recommend be pursued. Epigenetic proaction is a process on the societal level. When, for example, educational structures and methods are adopted in a functioning democratic society, people are invited to express their views through political elections, public debates, consensus conferences, etc.; but we do not ask each citizen for an individual informed consent. Nor do we ask for it when laws are passed. For example, in 1979, corporeal punishment of children became illegal in Sweden. The decision was preceded (and followed) by public debate and, as with most rules and regulations, some agreed with the ruling, while others did not—but the question of informed consent does not here arise. In contrast, if research in the natural and social sciences collaborate, e.g., to develop educational structures to assist and protect adolescents during that difficult phase of cerebral development, insofar as such research requires the use of human subjects individual informed consent will be needed. That this is the case is not a specific problem of the theory, but an ethical regulation (amongst many others) that all research must respect.

## 5 Opposing world-views

Concerning the human condition, surprisingly, Schleim criticises me for being overly concerned about the present states of poverty, war, and the many current violations of human rights around the world. He dismisses these worries as “rhetorical” (again comparing my arguments to those of Skinner and Delgado). Schleim seems to be at relative ease with the present state and future of humanity and, referencing Steven Pinker, draws the conclusion that there is hope that things will change for the better, so there is no need to be epigenetically proactive. Different world-views here confront one another.

Schleim concludes in what seems to me again a spirit of denial that people might be saddened by “focusing too much on their deficiencies” and ends his commentary by saying that “in the attempt to create superhuman beings a human catastrophe might also be provoked” (Schleim this collection, p. 12). True, no doubt—as, notably, Germany’s recent past illustrates. But this is not particularly relevant to my article: there is nothing in the theory of epigenetic proaction to suggest that we either should or could create “superhumans”.

## 6 Conclusion

Trying to understand and influence human norms in the light of what we today know about the brain is not an easy task. The scientific challenge is increased by the remarkable emotionality with which this whole area of research is permeated and which can apparently make it hard to see clearly what is actually being said. This emotionality is in part understandable: the notion of improving the human condition, including our biology, comes in some very sordid versions, as ideas of “racial purity” or “ethnic supremacy” serve to illustrate, and which remain present in various societies around the world. Historic awareness is indeed essential to safeguard constructive and hope-inspiring scientific ideas from being hijacked by nefarious ideologies (or, indeed, interpretations) and abused for unscientific purposes. However, the risk of misuse justifies precaution, not abandonment of constructive scientific pursuits.

Research collaborations between neuroscience, genetics and social science, notably, today provide rich and multifaceted knowledge about the human being and an increasingly integrated view of us as biological organisms interacting in complex natural and cultural environments in constant evolution. The resulting knowledge could further help us improve our life conditions, e.g., by assisting us in finding remedies for the developmental crises of adolescents, or excessive societal violence. What I call our “naturalistic responsibility” is born out of science’s strong social relevance. Whether or not in the future we shall use this knowledge

soundly remains to be seen. Which traits we decide to favour epigenetically, or what social structures we choose to develop, depends on who “we” are, and on the society in which we wish to live. We may hope that young scientists and philosophers shall rise well to that challenge, and develop the idea of epigenetic proactivity into a dynamic and socially responsible area of research.

## References

- Schleim, S. (2015). Should we be epigenetically proactive? A commentary on Kathinka Evers. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.