What can we know about ourselves and how do we know it?

Włodzisław Duch,
Department of Informatics, Nicolaus Copernicus University, Poland,
School of Computer Science, Nanyang Technological University, Singapore
Google: W. Duch

Abstract

Recent developments in cognitive neuroscience radically changed the perspective on understanding human nature. For the first time in history many philosophical questions can be placed on scientific, rather than on philosophical grounds. These questions include understanding of the mind, self, free will, religious and cultural beliefs, morality, politics and social organization. Scientific consensus based on these discoveries is slowly being developed and will have far reaching consequences. Evolutionary perspective explains how *homo sapiens* has evolved, why do we have specific structures of the body, brain, sensory abilities, and how the mind emerges from embodiment and social interactions. Social neuroscience shows that there is emergent causality: biology determines affective and cognitive abilities, preferences and beliefs, personality, but it is itself influenced by the environment that changes our brains and bodies. All these mechanisms are deeply hidden from ordinary introspection, creating a wrong perception of human nature. Traditional views on human nature are briefly summarized and radical reductionist interpretations of neurobiologists presented, comparing humans to a bag of chemicals. Scientific discoveries cannot be ignored, but their interpretation is not so obvious. Problems with describing our mental states and knowing ourselves are analyzed. Treating brains as a substrate that enables partially autonomic mental processes, and identifying oneself with the whole organisms rather than some abstract model of self, allows for more optimistic interpretation of .

1. Traditional views: mind over matter.

Cognitive sciences revealed many facts about human nature that do not fit to traditional picture based on ancient creation myths. They will clash with traditional views that have contributed to social stability for many centuries, perpetuated by cultural and religious beliefs and organizations. Traditional views on human nature were based on general ideas that formed in the ancient Greece:

- all that occupies space is matter;
- all that moves is alive;
- regular motions are a sign of intelligence.

Matter without something to push it was seen as inert. Living bodies had to differ from the dead ones, and the simplest explanation was to postulate existence of the soul, that could bring dead bodies to life, and in humans could have a rational part, or mind. Aristotle in his *De anima* treatise (Aristotle 350 BC) argues that the mind is immaterial and immortal because it acts with no bodily organ, seems to exists without the body, and thus cannot be corrupted. These ideas led him (and later also St. Thomas Aquinas) to the conclusion that there are at least three different kinds of souls: nutritive, plant souls responsible for growth, animal souls responsible for sensing, desires and reactions, and rational souls or minds that only humans

possess. In his view soul was essential essence of things, form rather than matter, responsible for functions of things, like seeing in case of eyes. Aristotle was convinced that touch was the primary sense, but evolutionary biology tells us now that it must have been something rather similar to taste and olfaction, the ability to distinguish different chemicals that even the most primitive animals need to maximize their chances for finding nutrition. Minds have their organs, nerves and glands, through which various bodily subsystems are activated and regulated. Although until recently mental activity could only be inferred through overt behavior functional neuroimaging techniques are now able to show some of its aspects. Indeed, mind seems to be more form that matter, resulting from specific organization of neural tissue and the way it functions by electrical activations that do not occupy space. In this sense mind is not a material object, although it requires brain and body as a material substrate through which it may act and in which it can develop its form.

Tendency to reify ideas led to the popular view of a ghost in the machine, a soul that animates the body, making free choices according to one's will, the idea most clearly formulated by Descartes. John Locke argued, and in the 20th century behaviorists believed, that mind is *tabula rasa*, a blank slate without innate traits on which environment may imprint any pattern. It was impossible to deny that bodies are not equally perfect, but minds were invisible, and the belief that they should have the same potential (they cannot be corrupted, as Aristotle wrote) was very appealing. The idea that human nature is good thanks to the (almost perfect) divine design complemented traditional beliefs, leading J.J. Rousseau to the image of a noble savage corrupted by the society (Pinker, 2002). All these ideas are still commonly believed, although they have all been proven to be wrong.

The interdependence of mind and body has been largely neglected until quite recently, when embodied intelligence and enactivism, the view that interactions with the environment are responsible for structuring body and mind, become very influential in philosophy, psychology, social science and even robotics. Genetic differences are all too clear, and they are also quite clear at the mental level. Some people are born with strong empathy becoming saints and social workers, and some people are born completely without it, becoming ruthless psychopaths. Like any other trait in biology, empathy resulting from brain activity is not expressed to the same degree in different people. Natural selection in stable environment may create species that are perfectly adopted to that environment but will be wiped out if conditions change. Humans have adopted to very diverse geographical conditions and ways of living and that means that all their traits must show great divergence. The existence of mentally retarded people, sadists, psychopaths, drug addicts is the other side of the same process that creates geniuses, talented people, great leaders, compassionate physicians and tireless social workers. Even such binary traits like masculinity and femininity are in fact multidimensional and almost continuous.

Traditional views are still strongly maintained and perpetuated by most religions. Critique of scientific discoveries based on biblical texts in the past should be a humiliating lesson for theologians who try to base their understanding of human nature on such grounds. It has been documented already in 1894 by Andrew Dickson White, co-founder of Cornell University and the first president of the American Historical Association in his two-volume book "History of the Warfare of Science with Theology in Christendom" (White 1896). While the main thesis of this book – warfare between science and theology – is grossly oversimplified, analysis of naïve religious views clashing with discoveries in virtually every branch of science: ancient history, anthropology, archeology, astronomy, chemistry, ethnology, geology, geography, meteorology, mythology, medicine, physics and other fields has been documented quite well.

Many Christian denominations still present views that are deeply rooted in ancient thinking. Moderately progressive denominations, such as the Roman Catholic Church, declare that there is no contradiction between science and religion. However, credible claims may only be made by experts who know theology, neuroscience and other relevant areas. If there are no people who studied such interdisciplinary subjects one cannot claim that religious views of human nature can really be reconciled with science. The new Catechism of the Roman Catholic Church writes about the heart: "Out of the heart come evil thoughts, murder, adultery, fornication", in many places sounding not like a metaphor, but like an echo of ancient theory of Galen. Can one forget about essential differences between people and claim that: "Natural law, presented in the heart of each man and established by reason, is universal in its precepts and its authority extends to all men" on ideological basis of particular religion, ignoring cultural diversity?

Statements like: "Man occupies a unique place in creation: (I) he is 'in the image of God'; (II) in his own nature he unites the spiritual and material worlds" seemed to be self-evident or at least reasonable in the recent past, when animals were called "brutes". However, development of ethology, work of Frans de Waal (2005, 2009) and other primatologists showed the evolutionary basis of morality, cultural differences, personality and self-consciousness in animals. Each species and each culture is unique, and at the same time shows significant similarities due to the common biological basis and functional requirements. There is almost perfect similarity at the molecular level (same atoms, molecules, proteins), very similar DNA, cells and their structure, same basic construction of the brain. Human family included species now extinct, with *homo sapiens* genome containing significant admixture of genes from the Neanderthal and Denisova human species. Endosymbiosis, symbiotic incorporation of bacteria that added useful organelles to form eukaryotic cells is now well established theory. Traces of this early evolution are everywhere in human bodies. Our sensory receptors: rods and cones in the eye retina, the hair cells of the cochlea, the olfactory cells, use a modified bacterial flagella (little tails used for movements). Mixing genes with related species speeds up the evolution facilitating development of new functions at many levels.

Biology is very complicated and thus all statements based on naïve observations are wrong or just approximate. Catechism of the Roman Catholic Church states that men is created "male and female", but even in antiquity people knew about hermaphrodites. Many third sex variants have been known in all cultures around the globe: Akava'ine (Maori), Hijra (India, Pakistan), Kathoey (Thailand), Khanith (Oman), Mahu (Hawaii), Waria (Indonesia), Muxe (Mexico), Two-Spirit people of North America and many others. Some countries, including India and Pakistan, have now acknowledged the third gender by adding a separate box in passports and identity documents, great progress has been made in Malaysia and Thailand in accepting their existence. Gender identities, sexual identities have their roots in natural biological diversity and claiming that man is created 'male and female' has little basis in reality. Acknowledgment of the other genders has been conspicuously absent in the Western cultures, meeting even now with widespread hostility. From the religious perspective one may argue that the first commandment "You shell have no other Gods before me", given by God that introduces himself as "I Am That I Am", means that all preconceived ideas should be rejected in favor of the scientific search for understanding of the world and the human nature. This understanding will certainly not come from analysis of Scriptures, it has to come from empirical sciences. Yet virtually all religions insist on scripture-based understanding of human nature, trying to protect traditional cultural biases that are mistakenly identified with religious wisdom.

Traditional view of Young Earth Creationism is still accepted by half of the USA population although it clearly contradicts all scientific evidence. The view of human nature with souls controlling the body, fully responsible for all free human decisions, surviving death to be re-

born again or continue living forever in some form, has even more support. Dualism is very appealing, and in the 2009 surveys 65% students from the University of Edinburgh (UK) agreed with the statement "Each of us has a soul that is separate from the body" (Demertzi et al. 2009). Reconciling neuroscience with religious anthropology seems to be difficult but is not impossible (Brown et al., 1998). Popular beliefs about human nature may be very diverse and rather far from scientific views that leaves no function for the souls. Frazer (1922) described many folk beliefs from around the world illustrating diversity of views on human nature. In Chinese culture during the Hungry Ghost month (in the seventh lunar month) dead bodies had to be tied with a rope in fear that their souls may come back from hell.

The concept of the soul was introduced at the time when the nature of movement was not understood. Things do not move by themselves, when they are not pushed they stop, so movement always requires a pusher, and body movement an internal pusher, or a soul (*anima* in Latin) to animate the body (Duch, 1999). In Egypt essentially everything that moved (including heart, body, name, and even shadow) had separate metaphysical essence and was considered to be alive (Bunson, 2002). Greek philosophers adopted this idea in a simplified form. Souls had the power of originating movement (understood as any change), as Aristotle writes in his treatise De Anima (Aristotle, 350 BCE): "what has soul in it that differs from what has not, in that the former displays life." Although in fact no mechanisms of life are explained in this way such apparent accounts are deceptively satisfactory: the body stopped moving because something has left it. Our intuitions agree with those of Blaise Pascal: "It is impossible that our rational part should be other than spiritual; and if any one maintains that we are simply corporeal, this would far more exclude us from the knowledge of things, there being nothing so inconceivable as to say that matter knows itself. It is impossible to imagine how it should know itself." (Pascal, 2006).

Our intuitions tell us that the Earth is flat, bit we know that this is not true, so we should not be strongly attached to intuitions. What do we really known about ourselves and how can we know it? How do we learn who we are?

2. How can we know about ourselves?

The word *psychology* has been derived in 17th century from the Greek words *psyche* (breath, spirit, soul) and *logia* (study of). The first book *Psychologia empirica* (Wolff, 1732) already described it as the study of the mind. Initially psychology focused on conscious experience and psychophysics. Trying to be objective the major schools – Paris (Binet), Cornell (Titchener) and Würzburg (Mayer, Orth, Külpe) – developed "systematic or experimental introspection". Titchener created a 1600 pages long manual for experimental psychology training people in systematic introspections, trying to integrate subjective descriptions with objective measurements. A few psychophysical laws have been discovered, but quantification of experience proved to be difficult. The perception of colors can be described using hue, saturation and brightness (although other 3 dimensions may be used), but can similar characterization be done to describe other types of experience? Early psychologists could not reach consensus on dimensions characterizing emotions, or thoughts (Schwitzgebel, 2011). A big debate ensued about the existence of "imageless thought" but results obtained in each laboratory strongly depended on assumptions of the researchers.

Despite all the efforts at the beginning of the 20th century psychology failed to find reliable methods to create a comprehensible theory of mind and did not become the science of human experience. Not all brains are created equal, but it was very hard to abandon the assumption that minds work in a similar way in all people, allowing for conscious access to

all mental processes. It has created a framework in which diversity of results based on introspection could not be accommodated. Mental processes depend on many hidden factors that cannot be easily controlled in experiments, resulting from complex non-linear interactions in the brain. Behaviorists had no ambitions to explain human experience, in fact they excluded the study of subjective experience from scientific for many decades. The development of cognitive psychology clearly showed that explanation of behavior is not possible without invoking complex internal processes. But how these processes create the self, which can only be another process in the brain, and how this self is informed about results of multiple parallel information processes in the brain? Does it have access to all information, or are there limitations on what can we know about ourselves? Can this knowledge be really presented using ordinary concepts used in everyday language (called "folk psychology" by philosophers), concepts that have been created mainly for communication between people, expressing their needs, rather than for precise discrimination of mental states? In short, how can we speak about the mind, subjective inner experience, and what is there to speak about? Is science of human experience, aimed at explaining phenomenology of mental events accessible through introspection, possible?

Psychology, philosophy and neural sciences have turned away from such questions and even now their importance has not be fully realized. The science of consciousness, actively developed now for almost two decades, tries to replace mental states with brain activity searching for neural correlates of conscious experience. Philosophers subscribing to eliminative materialism (Churchland, 1981) pointed out that common-sense understanding of mental processes, including all folk psychology concepts, are poorly defined and do not correspond to real, coherent neural phenomena. In this way all questions related to naive understanding of the mind are shoved away as non-scientific, reducing psychology to objective observation of brain states. There is a strong tendency in experimental psychology to focus on correlation of neuroimaging and behavioral data. It is however hard to imagine how such approach could lead to a satisfactory description of subjective experience. Folk psychology description of mental events certainly does not do justice to brain neurodynamics responsible for these states, but it plays an important role in communication. Understanding how to map brain states on folk psychology concepts, and how these concepts influence our decision processes is a great challenge for neuropsychology. To talk in a meaningful way about subjective mental processes a new level of description is needed, connecting neurodynamics with inner experience.

Creation of such language will be difficult for many reasons. Cognitive functionalism and neuroimaging techniques are the dominant style of research today, but mere localization of brain functions is not yet an explanation of mental function. Each brain is unique, processing information in a flexible way depending on the context. The same task may be solved using different strategies, and this creates a big problem for all statistically-oriented experimental approaches (including neuroimaging techniques) that average over many uncontrolled factors and search for average activity of the brain. All major brain areas and structures have been implicated in a large number of functions, depending on the context (Anderson, 2010). Understanding covert processes has proved to be much more difficult than originally expected, and seems rather unlikely without deeper understanding of the information flow in the brain that requires large scale computer simulations.

The first-person experience demands explanation. Empirically inclined philosophers tried to provide it using systematic reflection to study phenomena, or acts of consciousness. The phenomenology school split into the transcendental approach by Edmund Husserl and more existentially grounded analysis of the ways of being by Martin Heidegger, Maurice Merleau-Ponty and others. Some of these ideas led Francisco Varela and his co-workers to the

development of neurophenomenology, trying to connect neural events with mental events. Their book on the pragmatics of experiencing (Depraz & Varela, 2003) contains a lot of methodological remarks how to "modernize introspection", using the ancient wisdom traditions, Eastern and Western concentration techniques and awareness increasing practices. Still it does not give any insights or new concepts that could be used for description of mental processes. It is not yet clear that mindfulness training is a way to discover more reliable results than systematic introspection techniques gave in the past, although it may be important in affective sciences (Davidson, 2010) and may have great value in psychotherapy.

Another trend that emerged from ecological approach to perception (Gibson, 1986) stresses the development of cognition based on specific embodiment and the ability to explore the world by acting instead of passively receiving information. Enactivists rejected classical representational views of the mind that was dominating in cognitive science and artificial intelligence for a long time in favor of dynamic perception-action cycle, tuning the organism or a robot to the affordances of the environment. The importance of embodiment and enactive ideas (cognition is acted out) is clear not only in cognitive science, but also in robotics, where robot brains are trained by active exploration of the environment. Understanding of how we learn to perceive the world, and what we learn by exploring it, should help to create a verbal account of the phenomenology of this process, linking it to subjective experience.

Since a long time no-one has tried seriously to describe inner experiences in some details, naively assuming that conscious experience is in some sense obvious, that we can search for mental correlates of consciousness. This may be possible only in a few simple situations, like binocular rivalry. Experimentally oriented philosophers discovered formidable obstacles in attempts to answer even simple questions about the nature of conscious experience. In the book "Describing Inner Experience? Proponent Meets Skeptic" (Hurlburt and Schwitzgebel, 2007) the authors made another attempt to answer the question "Can inner experience be accurately apprehended and faithfully described?" in an empirical way. They asked a person to describe details of her mental states prompted by a beeper at random moments. Despite numerous suggestions and additional clarifying questions her descriptions have been vague, frequently confusing and at times contradictory. Folk psychology concepts are useful to communicate important information, but people are surprisingly poor judges of their own phenomenology, easily confused or mistaken about their own experience. Schwitzgebel showed it in more recent book, the "Perplexities of Consciousness" (Schwitzgebel, 2011). Analyzing a wide range of mental phenomena he argues that we really do not know much of what is going on in ourselves, and are unable to describe and understand our stream of conscious experiences. Frequently we are not able to describe what and how we feel, give unambiguous description of experiences related to dreams, imagery, peripheral vision, perspective and visual illusions, untypical use of sensory data (like echolocation), or describe emotional experiences. More experiments with beepers with larger groups of people showed that even a simple distinction between sparse and abundant experience, between marginal and dominating unconscious perception, is impossible to make.

John Watson, famous behaviorist, initially claimed that he had vivid visual imagery, but after he rejected the existence of mental images on theoretical grounds he claimed that he had none. Prior beliefs may change character of introspective reports, although in Schwitzgebel experiments initial belief that experience should be abundant or sparse did not seem to matter. Blind people may avoid obstacles analyzing faint sound echoes, false belief may direct their attention to facial pressure or some other sensations, making it hard to learn correct phenomenology. Believes about ourselves may not only influence, but completely

determine interpretation of our mental states. As with the scientific observations that are theory-laden and cannot be interpreted without theoretical framework, our own internal phenomenology seems also to depend on our theory of mind (Carruthers and Smith 1996). This is indeed surprising and goes contrary to the common belief that we cannot be mistaken about our own experience, especially in direct perception. But how do we actually known what is visual, auditory or tactile signal? In the brain all such information is changed into electrical activity. The brain has to learn to interpret where this activity comes from and what structure does it have to be able to correctly categorize it into different modalities.

Why it could be hard to reach a judgment about what is in our mind, to be clear about how things appear to us? What can the brain learn by observing its own states? Animals distinguish relatively small number of their brain states, needed for survival and useful in communication of their emotional states using voice and body gestures (de Wall 2009). With much bigger human brains and more complex behavior the need for communication and explanation led to creation of a growing number of concepts. They serve as labels distinguishing classes of brain states resulting from perception of objects, their properties, sensory stimulations, emotional states, and more abstract ideas such as color, self and mind. To understand why people make specific comments about their experience we need to:

1) categorize quasi-stable brain states,
2) understand the dynamics of their succession, and
3) understand what kind of comments may be generated about this process.

Transitory states of the brain are too short to create associations and enact mental states (Dobosz and Duch 2011), continuous experience is analyzed as a sequence of meaningful scenes, segmented in cinema-like fashion (Zacks et al. 2010). Categories of quasi-stable brain states that tend to repeat themselves often are labeled by concepts that facilitate communication. A good part of this process is innate, with hearing phonemes or seeing shapes, colors and movement learned in an unsupervised manner, if sufficient stimulation is provided during the development. Another part, responsible for the phenomenology of internal brain states has to be learned. Categorization of colors in different cultures significantly differs, despite the same visual perception (Berlin, 1969), and this may influence perceptual learning. Learning new phenomenology, for example learning to distinguish amaranth from alizarin or amber colors on an artist palette, requires more precise categorization of states of visual system. If we learn to pay attention to the subtle differences between these colors, and if we learn their names, we will be able to describe our inner experience with greater precision. Research on experiencing color in dreams is very confusing, it seems that people frequently are not sure if they experienced color or sound in their dreams.

In some cultures spatial orientation is always given in absolute terms (east-west, like in sea navigation) rather than egocentric coordinates (left-right). Some languages require mandatory specification of gender and this leads to association of many feminine or masculine properties with neutral objects. There are languages that require specification where exactly the information comes from: direct observation, inference, conjecture, hearsay etc. (Deutscher, 2010). Such language habits are reflected in brain neurodynamics (Liu et al. 2010), influencing the way mental states are associated, experienced and described. Linking experience with language requires supervision, someone to provide the names, and it must influence how brain states can be categorized, how neural activation may spread, changing subjective perception. Any new brain activations, and thus learning any new concept or fact, changes brain dynamics and thus changes experience. The interplay between language helping to categorize better perception and make appropriate actions has been studied in neurorobotics (Tani, 2008). Brains must extract salient information making it explicit for

fast processing. How do I know that pain should quickly be stopped, before permanent mutilation will be done? Arousal of many brain areas by secondary somatosensory cortex prepares organism for action, interrupting all ongoing activity by a strong signal that is interpreted as unpleasant. Damage to this cortical structure leads to pain asymbolia, a condition dangerous for health.

Experience depends on activation of complex brain network, but experimental psychologist still try to describe it in folk psychology terms constructing verbal models instead of mathematical dynamical systems. Observation of even relatively simple continuous non-linear dynamical systems are not sufficient to create verbal models that will allow for prediction of their evolution. Therefore psychology in the longer run has no choice but to turn to mathematical modeling. Paying attention to weak signals one learns to analyze them better and changes the brain dynamics. Some information comes from internal flow of activations, but essential part of information comes from observations of our own actions in the world. This has been well established in any motor learning situations, where and prioprioceptive visual information flows to the premotor cortex and modifies the activity of motor cortex. We learn about ourselves and thus change ourselves partially by observing the result of our activity. However, this feedback loop is so fast compared to the time necessary to establish synchronization of large groups of neurons forming quasi-stable brain state that linear causality does not make here much sense. Strictly speaking separability between organism and environment is an illusion that lies at the root of problems of traditional epistemology (Edelmann, 2006).

Learning gradually makes fuzzy and hard to categorize experiences more precise and easier to distinguish. Initially I may not be sure of my own experience, if it is not a result of well established, typical brain activation. Stereotyped perception may easily miss even large changes, paying attention to typical elements, as experiments with change blindness clearly show (O'Regan and Noe, 2001). Interpretation of weak internal signals that has been masked by stronger signals, as in the case of blindsight (Weiskrantz, 1986), does not seem to be like anything experienced earlier. Brain activations that support it are weak and have a large variance, so blind people are not able to categorize them in a reliable way. Learning slowly creates new qualia, defining their experience in a clear way, with presumably stronger brain activations and less variance. To learn a new phenomenology I need to know what to look for in my experience, where attention should be focused. In the echolocation case discussed by Schwitzgebel (2011) a blind men was convinced that information about obstacles he is able to avoid comes from pressure he feels on his face. As this assumption has been false attempts to link his tactile impressions with obstacle avoidance could not have been successful, but if he had directed his attention toward auditory stimuli he should improve his orientation skills. In dreams only a subset of visual areas may be activated, leading to kind of visual experience that is quite hard to describe. For example, activation of the inferior temporal gyrus and MT areas without stronger activation of V1-V4 areas may give the feeling of object motion without shape or color, or of knowing that some people are acting without visualizing them. This is frequently my own feeling when I think about a movie, recalling the action and the people involved without real visual experience.

What do we really do when we try to describe our experience? We categorize or discretize brain states, assigning a phonological label to a dynamic distribution of brain activations. Experience is always much richer than language can express, requiring an infinite number of symbols to describe all possible brain activations. There is a tradeoff between the accuracy of description and mental states, and brain resources needed to pass this description to other people and to synchronize it with their own understanding of their mental states. Common concepts are useful for communication of important distinctions between mental

states, but they are precise only in matters of real behavioral importance. Specialized vocabulary to describe detailed phenomenology of brain states does not exist and would perhaps be too idiosyncratic to be useful. Verbal statements are comments on the successive brain states, or categories of experience that transmit socially relevant information. Communication is possible if there is an approximate correspondence between concepts and mental experience (and brain states) that they point to. Brains of babies do not support all the mental states that adults have, so many concepts must elude them. Conceptualization of mental states in different cultures and across various languages significantly differs for concepts of little behavioral consequences. As a whole modern languages cover similar space of possible cognitive and affective states. However, matching semantic maps between different languages for emotional words is a non-trivial task (Ploux and Ji 2003).

Brain activations in states of consciousness that are normally not involved in communication, like dream states, meditative states, hyponotic, somnabulic, or hallucinatory experiences, are rather different than normal conscious mental states that we all share and have learned to communicate. In such unusual states it may not be possible to establish correspondence with the symbols used to characterize brain states during normal perception. Visual experience engages coordinated activity of ventral and dorsal pathways in the visual stream; in macaque monkey there are over 30 visually responsive cortical areas. If something goes wrong with the top-down or bottom-up information flow all kinds of visual agnosia may arise. While classical apperceptive, color or associative visual agnosias due to the bottom-up stream impairment are well recognised, top-down impairments may lead to what I have called *imagery agnosia* (Duch 2009, 2011), that have not yet been investigated by neuropsychology. They have quite specific phenomenology that is hard to describe using standard language: the words that we use do not label these experiences with sufficient accuracy, have little meaning for people who cannot experience them, as they do not map well categories of brain states from impaired to normal brains. As a result people who do not have vivid imagery may experience untypical brain states that are hard to express in words. Large part of inner experience may be different for different people, but it will not create noticeable problems in communication.

Top-down flow of information in the brain helps to establish conscious percepts (Dehaene et al. 2006). Imagery of visual, auditory or other sensory information needs sensory cortex to re-create conscious qualia in proper modalities. Usually top-down processes are sufficiently strong to activate primary sensory cortices, establishing resonant states between associative memory and primary sensory brain areas (Janata 2010). However, if feedback connections are too weak to provide such activations without the help of bottom-up information a person may show all kinds of symptoms typical for agnosia when required to perform some tasks based on imagery. Problems with visual imagination describes as dim, pale and non-existing may probably be linked to the strength of the up-and-down stream resonant states (Ullman, 1996). Noisy, fluctuating activations that do not form quasi-stable states cannot be uniquely categorized, so in this realm experience remains ambiguous, reports will be confusing, as discussed by Schwitzgebel (2011). This may include one or many sensory modalities. Auditory imagery agnosia will be the inability to consciously imagine pitch and timbre of sounds. Evoked Response Potentials (ERPs) in the auditory cortex for sounds that are missing from well-known melody, similar to such potentials when sounds were not missing, showed that people are filling in the missing sounds using their imagination. I have conjectured that people who cannot imagine sounds will show no auditory ERP response, and that such responses will be correlated with the vividness of auditory imagery questionnaire, as they are in case of vision. So the only way that someone without auditory imagery may know that there is a melody in their head is by humming or

playing it on an instrument. We learn about ourselves not only from the internal information flow, but also from observing our own behavior.

The very existence of imagery agnosia has not yet been fully acknowledge in neuropsychology and may be directly linked to the "imageless thought" controversy of the early introspectionist debate (Schwitzgebel, 2011) and the whole philosophical and psychological tradition linking thought with images. Faw (2009) has pointed out that conflicting intuitions on imagery may be based on different individual abilities. He estimates that only 2-5% of people are not able to form visual images, but little statistical data on how many people are visual non-imagers have been collected since Francis Galton's 1880 paper on visual imagery. He also has made some simple surveys asking people about their ability to hear imagined sounds: 2% claimed no such image. 12% could not recalling and internally hear a song but could hum it. This has little to do with their musical abilities, although makes it harder to learn playing an instrument, especially playing melodies by heart. I have described phenomenology of auditory imagery agnosia based on my first-person experience (Duch, 2011), focusing on the difficulties in learning to play an instrument. Repeating the melody is always a matter of trial and error; it is easy to hear that the sound is wrong, but without auditory imagery impossible to remember where is the right sound. Following written musical score I could learn to play, but always feel like a listener, surprising myself, frequently improvising, with seemingly no conscious influence on what I am playing. Evidently at a higher cognitive level my brain was able to internalize a lot of musical knowledge that I am consciously not aware of. Each action is a step in the dark, without knowing the consequences. Some qualia are generated in this process, perhaps comparable to a blindsight situation (Weiskrantz, 1986) where a feeling where to make the next step is developed without visual experience. With auditory agnosia trying to play a piece of music is similar to maneuvering blindly in the auditory space, without the ability to imagine what will happen when next action is taken.

Faw himself has no visual imagery, a claim that has been met with incredulity by his colleagues in psychology and philosophy. I have personally met quite a few people who claim not to have visual imagery and I am certain that I also do not have it. There is nothing uncertain to it: shapes and colors do not appear in my imagination, I have no idea how people are dressed or what is the color of my wife's coat unless I will explicitly encode it in my memory by verbal naming, I cannot draw from memory but have no problem to draw from nature. Moreover, it is not just the lack of visual imagery, but also of imagery in other sensory modalities, including auditory, tactile and gustatory sensations.

These observations have wider implications for understanding how do we know ourselves. There is clearly some imagery in my head that I cannot see or hear. Like most other people I am able to dream and to daydream, although without sensory aspects associated with imagined situations. I seem to weakly feel the actions, movements, to know what is it about, know who is involved and what is happening, yet all this is on some abstract level. Sensory deprivation in isolation tanks did not induce in me any auditory or visual hallucination. A good part of what I know about myself comes from observation of my actions in the world and predicting these actions to create a model of my behavior. This situation may happen more often than we are willing to admit, with a whole range of problems related to the inability of consciously interpreting our own brain states, and the need to express and recognize them through bodily actions. Perhaps imagery agnosia may actually help in abstract thinking, as neural activity is not dispersed into sensory cortices. Better understanding of these issues may have far reaching implications for education, assessment of talent, and correlation of brain activity with conscious experiences. The distinction between "verbalizers vs. visualizers" is sometimes used in education and consumer research, with psychomet-

ric measures of imagery, processing style and daydreaming/fantasy content and frequency, but there is no research to correlate talent (or the lack of it) with vividness of imagery.

Although science has discredited the value and reliability of information gained via introspection we all suffer from the illusion that introspective information is sufficient for self-assessment, leading to overrating its value in judgment, free will and decision making, personal relationships, stereotyping, disregard information conveyed by our own behavior, contributing to prejudice and false understanding of our motives, leading to conflicts, discrimination, and lapses in ethics (Pronin 2009). Flooded with thoughts, feelings and intentions we naturally treat them as a direct, authentic source of self-knowledge. At the same time we are suspicious of other people introspections, understanding that what they think about themselves is not necessarily true. The "introspection illusion" is a barrier to self-knowledge, presenting perceived information in biased way, and suppressing important information that does not fit to the model of the self. Although such internal models are compelling to our self, they are far from being realistic.

"I think therefore I am", claimed Descartes, but am I what I think I am? To fully know oneself would then mean to internally build a perfect predictive model of one's own behavior, based on all internal and external information accessible to the brain. This is unlikely because predicting how such as complex system as the brain will behave in various contexts, especially in extreme situations, may not be possible. Another goal seems more feasible, although it may be very hard to achieve: acting with agreement of internalized values without consciously accessible model of oneself. This is one of the goals of spiritual training in Zen tradition (Austin, 1988), where a student is accepted as a new master after she/he embodies the teaching without any conscious effort.

3. Scientific view: matter over mind.

Physics and chemistry has develop sophisticated experimental methods for rather simple systems, but in psychology nothing is simple. Naïve observations lead to wrong intuitions creating insurmountable obstacles. Some questions are like "a pole to which a donkey is tied for eons" (according to old Zen saying). This may include popular questions about consciousness, qualia, free will and the question: can the matter know itself?

Francis Crick in his popular book "The Astonishing Hypothesis: The Scientific Search for the Soul" (1995) writes: "You, your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cell and their associated molecules". Joseph LeDoux, an expert in the neuroscience of emotions, states: "You are your synapses" (LeDoux 2003). Anthony Cashmore, an eminent biologist, claims that belief in free will is like belief in magic or vitalism, discarded 100 years ago. He wrote: "Biological systems obey the laws of chemistry and physics; as living systems we are nothing more than a bag of chemicals" (Cashmore, 2010). Consciousness may give us the illusion of responsibility, beneficial for society. There is evolutionary advantage in thinking that we are free to do what we will, but all our decisions are determined by genetic and environmental history. Apparently already Darwin realized that the concept of free will cannot be defended (quoted after Wright, 1994), commenting: "This view should teach one profound humility, one deserves no credit for anything. Nor ought one to blame others". He also realized that from social point of view this is a useful illusion. Fortunately only a few gentlemen could understand this knowledge, "man who has thought very much, and he will know his happiness lays in doing good and

being perfect, and therefore will not be tempted, from knowing everything he does is independent of himself to do harm."

Is everything we do independent of ourselves? First we need to define what do we mean by self. Is the feeling of our special status as human beings just a useful illusion? Our behavior results from motor activity that is explained by neurophysiological processes, ultimately reducible to neurochemistry and physics. Are we automata (Duch, 2010), biological robots and mental life only epiphenomenal? Such reductionist materialism, characteristic of the philosophical view of many scientists is repugnant to Christians and other religious dominations, but do they have any rational arguments against it, or just fears that their traditional world view will collapse? These fears may actually be justified, feeling free facilitates positive social activity while disbelief in free will works on opposite way (Baumeister et al. 2009). However, there is no hope that neurobiological facts will simply go away. For many reasons traditional dualistic views cannot be maintained any longer, as it has never explained anything. The embodied and embedded view of mind explains a lot of psychology and behavioral data, leaving no space for detached mind that simply observes the world and makes decisions for us (Edelman 2006, Damasio 2010). The residue of unexplained higher human functions without neurocognitive correlates is now close to zero. Science has a coherent picture of how brains make up their minds, and in the coming decades will verify it by building artificial brains. Special status of the mind as non-physical entity cannot be defended on scientific grounds (Duch 2010).

Edelman (2006) has pointed out that minds are not caused by brains, they are logically entailed in the same way as absorption spectrum of a molecule is not caused, but entailed by its structure. Zombi water molecule without absorption spectrum is logically impossible. I have made similar arguments about the mind (Duch 2005): every system with brain-like structure must comment its own internal physical states using qualia-like descriptions and thus will claim to be conscious. Zombi-like brain is just logically not possible, but philosophers of mind frequently relay on such wrong intuitions (Dennett 1995). Understanding of free will or volitional processes is also relatively simple from the neuroscience perspective. It follows from understanding of spontaneous processes: how they arise, what causes them, why they win competition with other, alternative plans for action, how are they related to "self" that claims to originate them. These questions may be studied experimentally and theoretically through computer simulations. Brains constantly make plans for actions; actions are expressed externally by movements resulting from the activation of motor areas in the brain (including verbal responses), by thoughts that do not activate brain areas responsible for speech sufficiently strongly to generate motor activity of the vocal apparatus, or by changes in internals states experiences as feelings. Some stimuli require fast reflexive responses, there is no time for planning, and they do not generate the feeling of "I want to do this" or even "I'm acting".

An understanding of our own nature is the basis of ethics, politics, and education. It is ultimately the most important factor that guides human activity. Science has coherent picture of how brains work, but is this picture also explaining human nature? Understanding human nature requires interpretation of scientific facts, not suppressing or denying it. Minds are entailed by the brains, but was Darwin's interpretation that everything we do independent of ourselves correct? What is meant in this case by "we" and to "know"? I know how to walk, run and swim, and it is a non-trivial knowledge that cannot be verbalized. Some knowledge may be so deeply embedded in brain processes controlling interactions with the environment that the self will not be even aware of knowing until it finds out that the organism performs correctly. If I am not aware that I know, as in the case of implicit knowledge or memory (Schacter 2001), does it meant that "I" really do not know? A better

understanding of what knowledge is for the brain is to look at the whole organism and its ability to regulate its own behavior and interact with the environment. All processes that in some context are used for this purpose are a form of knowledge. The self, understood as the object of reflective consciousness, should extend beyond the processes of which it has an explicit memory, it should include also potential processes that the organisms may be capable of running. A child may be encouraged to try despite claiming "I can't do it", and see that she/he can do it. For the brain the ability to perform is not very different from the ability to memorize and use abstract knowledge. Knowledge is context dependent, processes able to operate with the help of memory state may be available only in specific situations.

Any brain process that in some context may influence my action is a form of knowledge.

Identifying myself with the model based on introspective illusions I shall know little about my true self, and be surprised in many situations by my own behavior. This model of self is one of many functions that the brain maintains. Self is not monolithic or unitary, as introspection may suggest, but it is rather a collection of various functions for representation, monitoring, evaluation and integration of self-referential stimuli, implemented by distinct regions within the cortical midline structures (Northoff 2004), mainly cingulate cortex. Attribution of personal relevance leads to brain activations that partially overlap with those of the reward system, and depend on the personality type. This can be seen in the activity of anterior cingulate cortex and insular cortex (Enzi, 2009). Various aspect of the self may be selectively damaged leading to personality disorders, or subnetworks of coherent activations may form, leading to dissociative identity disorder (multiple personalities). Cognitive neuroscience expands towards the first-person neuroscience linking mental and neuronal states (Northoff and Heinzel 2006) and towards social neuroscience, linking brain states of people involved in social interactions. Neurophenomenology offered a systematic strategy to explore the links between mind and consciousness, focusing on the structure of human experience (Varela, 1996). Cognitive neuroscience slowly started to address the self-experience of being a cognitive agent (Christoff et al. 2011). Although much remains to be done many aspects of self has already been revealed.

Cognitive psychology perspective also shows different aspects of self (Baumeister 2011). The folk notion of willpower has found some confirmation in experiments showing depletion of resources; making decisions decrease the energy available for self-regulation. Self, as all other brain processes, needs energy to initiate actions and make decisions. Willpower and self-control compete for energy, so also self-control decrease the ability to make decisions, leading to irrational behavior and extreme choices instead of more effortful compromises. Self-regulation, controlling one's own behavior, consumes glucose, while low level of glucose in blood leads to deficits of self-control.

Anthropology and evolutionary psychology supports neuroscience, seeing the self in relationships to others, connections to family and larger social group. "The purpose of the human self is to facilitate being part of a human social group" (Baumeister 2011). Creation of interpersonal connections requires the self to adjust to groups and acquire culture. Learning to communicate with other people required development of language, symbols to categorize percepts, intended actions, needs, leading to conscious thought about oneself and understanding of others. In this process people started to define themselves, distinguishing and anticipating events that were of relevance to their needs, becoming self-aware and developing reflexive consciousness. Multiple inner processes that enable sensory-motor interactions and inner reflections are perceived from a single perspective of the self, enabling coherent social interactions and predictive behavior. "Free energy minimization" principle for optimization of resources, value (expected rewards, expected utility) or its complement,

surprise (prediction errors, expected costs), short and long-term predictions and minimization of surprises, has recently been proposed as a general principle explaining action, perception and learning (Friston, 2010).

To make personal decisions, and to integrate them with interpersonal relations, the self must coordinate many brain processes and makes an impressions of a unitary system, but in controlled experiments it is clear that this is more like a collection of functions, something that Minsky called "The Society of Mind" (1988). Unification and coordination of these processes requires a flow of information in the brain, broadcasting available information to all areas that may find it useful and that make the behaviorally salient information explicit, easy to use in decision processes. This mechanism has been called "a global workspace" (Baars 1988) and is now a basis for theory of consciousness that links subjective reports and objective neuroimaging data during conscious perception (Dehaen et al. 2003). Information that can easily be accessed can be commented upon, so every system that creates physical states that carry such information may express it through behavior or (in case of humans) verbal comments (Duch 2005). This mechanism is not all-powerful, does not control all brain processes, and when it fails social and mental health problems arise. Some addictions and urges are "stronger than me", leading to serious problems, but everyone sometimes succumbs to temptations that may expand into major problems. Cognitive neuroscience of self-regulation explains how the prefrontal cortex exerts top-down control over subcortical regions involved in reward and emotion, and why its influence is sometimes not sufficient to subdue strong emotional subcortical impulses, or what are the effects of damage to the prefrontal cortex itself (Heatherton and Wagner 2011).

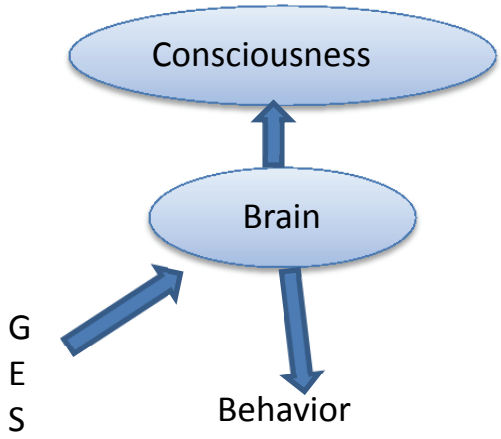
For the first time in the history we have theories that explain internal mechanics of the self, will, consciousness and decisions making process. We start to understand ourselves in a way that just could not be imagined in the past, and this knowledge has many consequences that go beyond neuroscience, extending to clinical psychology and psychiatry, all areas of social sciences, philosophy and even theology. Even if some details are still uncertain there can be no going back to a primitive folk psychology model of all-powerful "I" that makes decisions, survives death and either is reborn or goes to some supernatural place to live forever. But does it all mean that my decisions result from a physical process controlled by genetic, environmental and stochastic factors, and the feeling that I decide and act is just an epiphenomenon, without any causal power?

To understand this crucial point and to counterbalance extreme reductionist views we should look at mental level. It is true that there are physical mechanisms behind each mental event, an alternative is only a belief in magic, as Cashmore has put it (2010). Some philosophers, like John Searle, really believe that electronic elements like some magical causal powers of biological neurons and therefore computers cannot understand the meaning while brains can (Searle, 1992). This type of philosophy created a lot of confusion, pushing the discussion into shallow and muddy waters, without really defining basic concepts, how can we know something, what brain processes are involved, what is self and what is understanding. Knowledge is any process that can influence my decisions, and understanding cannot be granted on other people or artificial systems by without testing their knowledge. From the inner perspective the actual implementation of processes behind mental events, the type of hardware they run on, may determine the mind structure. Percepts obviously depend on our senses, and primary concepts are built extracting categories of percepts and relations between them. Functionalism, assuming that mental states can be reduced to their functional role, tried to build concepts on abstract representations. This has totally missed the richness of experience and thus its influence on cognition.

Suppose that we shall build silicon-based hardware that has identical physical states as the brain matter. We could replicate cognitive processes this way, but because carbon atoms in biological tissue interact in a different way than silicon atoms this replication cannot be perfect. The best we can do is to approximate these processes, and create different types of cognitive systems (artilects), that may nevertheless have some kind of mind-like properties, depending on their embodiment, interactions with the environment and accuracy with which they approximate biological processes. Such artilects will obviously be physical devices, although their complexity, learning in an open environment and stochastic elements may not allow us to understand their behavior in deeper way than we can understand behavior of an individual human. Should we then believe that their actions result from a simple causality in the decisions making process? Knowing that there is no ghost in such machine, can artilects be reduced to a physical level of existence?

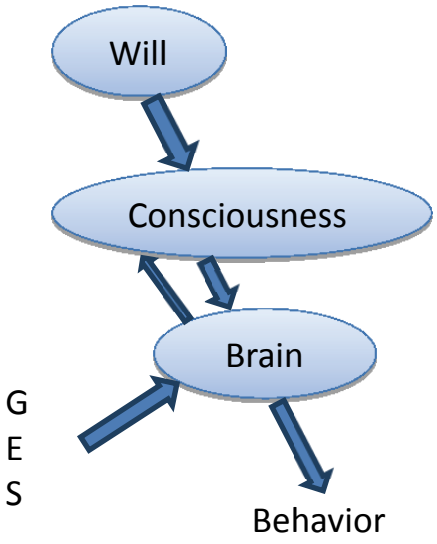
Looking at the electronics and checking the connectivity between artificial neurons will not tell us much about what the system knows. A full description of the human brain, neurons and synapses, will not tell us about the mind. Brains could have been very different if the environment and social conditions in which we live had been different. Dolphins have big brains and some culture but it seems much less sophisticated than human. Brains evolved to fulfill specific needs, enabling sophisticated minds and language, creating conditions for more intricate social organization. At the same time the direction of brain evolution has been constraint by social requirements. Static information describing brains as physical objects in details is not sufficient to understand minds supported by them. This information has to be animated by the working hardware, brains or electronics, recreating the flow of dynamical states. Brain is a substrate in which immaterial mind may develop. Minds do not occupy space, so they are immaterial. The only way to find out what type of the mind brains or artilects have is to interact with them. Testing brains or artilects in many different ways one can probe various behaviors that are capable of, including memory, beliefs, motivations, goals, desires. Learning in an open-ended environment is going to create many memories and associative processes. Immaterial, dynamical states of the brain form a network of events related to each other, activating such processes. Only a small number of dynamical states will be actualized in the lifetime of the system, the number of potentially available states that may be distinguished from each other is in case of human brains almost infinite.

Understanding why certain decisions are made is easy on the mental level, while it may be almost impossible at the physical level. There will be some logic to these decisions, as well as some attempts to rationalize actions that are the result of low-level hidden activity. In case of brains it may be some wetware failure, leading to neuropsychological conditions such as hemispatial neglect, or delusional disorders. Almost anything may have an effect on our decisions, priming the brain in some way. A shocking example has been presented by Danziger et al. (2011): parole decisions made by judges with long experience, according to the legal rules, were favorable in about 65% of cases after meal break, dropping to nearly zero just before the break. The true factors influencing their decisions must have a strong physiological component, although they have been totally unaware of it. In case of artilects there may be a partial hardware failure that will lead such artilects – without access to their low-level processes – to wrong interpretation of the own behavior. However, in healthy brains and in correctly functioning artilects actions will be in agreement with the goals and beliefs of the system and their interpretation should be correct. Self-processes are not always well informed of the precise goals, therefore we do not always fully understand our actions, thinking in vague terms and hoping that selected actions may lead us to general goals. We tend to call it intuition and value it highly, although it may be a failure in internal communication, a process to weak or to diffused to be captured in a precise way and made conscious. Such processes are still claimed

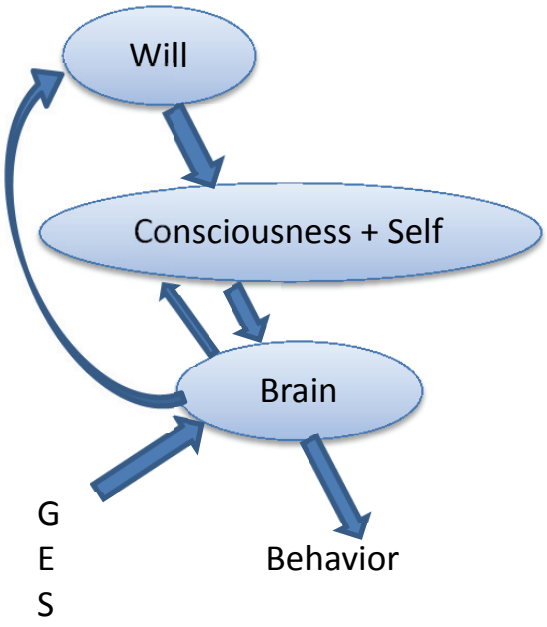


to be our own by the Self. So the problem arises when we identify the "self" with a small subset of self-related brain processes, ideas about oneself that we are conscious off, instead of identifying it with all brain processes or even the whole organism.

Let's consider what are the decision making options by the brains (Cashmore 2010). There are 3 main factors: G) genetic, leading to the overall evolutionary brain construction and some variance in individual brains due to inheritance and mutations; E) environmental, influencing the development of the brain through nutrition, exposure to various chemical, and learning; S) stochastic factors, such as the recent history of our activity and interactions (priming effects), physiological factors (glucose level), synaptic noise and other effects that may make some behavior unpredictable. According to Cashmore and other neurobiologists the GES factors determine the state of the brain, and the brain processes lead to actions, while consciousness has no causal effect, it is epiphenomenal, "a mechanism by which we follow unconscious neural activity and behavior" (Cashmore 2010). In this view consciousness has no function. Experiments performed by Libet (Freeman et al. 2000), or more recently Soon et al. (2008) and many others show that there is a time lag between brain activity indicating that a simple decision has been made by the brain, and feeling conscious urge to act (I have discussed that in Duch 2011 in more details). We become conscious when the brain is ready to act and motor cortex is getting active, but obviously there must be prior plans preparing for action, that prefrontal cortex is preparing. One reason we cannot be aware of such plans is that there may be different plans at the same time. Becoming aware of all such activity will change the nature of conscious experience. What I am proposing here is that brains prepare information for conscious approval. Salient features must be made explicit for decisions processes, so they are extracted by separate streams of information processing before reaching prefrontal cortex. Conscious processes are of course also a flow of brain states, but at the mind level they have a different status. GES factors, and in particular social influences, created a network of potential mental states, endowing the system with beliefs and dispositions to act as an individual.



In the traditional view the will to act springs from nowhere, has no causal component, is completely disembodied. Volitional decisions arise without any causes, become conscious and command the brain to act, perhaps with weak influence of the brain on the content of consciousness. Such view based on naïve introspection cannot be reconciled with scientific knowledge, in particular understanding of spontaneous brain processes. It also seems irrelevant from experiential point of view, because there is no reason why the embodied self should accept such whimsical decisions of will as its own. What I will accept as my decision has to be reconciled with my world view, my system of values and beliefs. It is not important at all where is the source of the propositions to act: plans made by prefrontal cortex in cooperation



with the rest of the brain (which are certainly the real source of such plans), random events or some otherworldly noosphere. I may feel an impulse to act violently or do indecent things but such impulse is going to be inhibited confronted with the processes that are part of the self. I will then come with a correct rationalization of my behavior, referring to my beliefs and goals.

Why should I be entitled in this view to claim that I have causal power to act? Mental level has some autonomy. Although all immaterial, mental events have their material support in the form of brain activations, they cannot be analyzed and made sense of only at the neural level. Processes that give rise to the mental events are activated in the brain substrate because it has been prepared by GES factors, reflecting evolutionary history, individual development and recent history. This is what makes me, my self. Given sufficient time I can make an effort and select the action that is in agreement with my goals and beliefs. Playing ping-pong I have no time for reflection, I have to act without thinking.

In this view then brains create mental events that are volitional impulses, but before any action is taken they may become conscious, and thus accessible to self and amenable to further processing by the brain. Conscious selection or at least conscious approval of the proposed decision allows the brain to implement it by acting out.

We can control our behavior in a responsible way only if we can reflect on the consequences of our actions, and that requires conscious reflexive activity that controls spontaneous impulses and thoughts. That activity can only be understood at the mental level.

4. Conclusions: the Big Mind

Psychology had once the ambition to be the science of human experience, but now it has more modest goals to be the science of behavior and mental processes. However, with the help of social neuroscience, research on consciousness, neurophenomenology (Varela 1996) and the first person neuroscience methodology (Northoff and Heinzel 2006) there is a chance that closer links between mind and body, subjective perception and objective evaluations will be established.

Neuroscience research has already led to a rather radical, reductionist interpretations supported by eminent neurobiologists. It is clear that traditional views of the human nature cannot be reconciled with neurobiological facts, and common beliefs about the mind and its powers are vestiges of magical thinking (Cashmore 2010). Interpreting results of brain sciences requires understanding the process of developing the self and the theory of mind in childhood, the role of introspection in this process, forming beliefs about oneself, access of the self to internal flow of information, learning from observation of effects of one's own behavior in social environment, emergent nature of mental processes and their relative autonomy. As the discussion in the previous sections has shown we may have indeed serious limitations in access to many internal processes, learning about ourselves, and may even be misguided about our own experience in untypical situations. It may be premature to claim

that mind is powerless, controlled by blind forces of nature. At least studying these questions instead of turning our back on them and putting our head in the sand gives a chance for deeper understanding our own nature and numerous problems that face humanity.

From the social point of view self-regulation and control are much desired traits. Part of my extended self comes from internalized social interactions. Everything that influences brain processes related to self in a strong way may be treated as a part of the same dynamical system, an extended mind (Logan 2007; Menary 2010), or a Big Mind. In this sense some people become partially part of our own Big Mind, not only because we think about them often but and we form a lot of expectations about their behavior. Computers and telephones may form an extension of cognitive abilities of our mind. I may make a choice between actions that arise in my mind, but why do they pop up there in the first place? They have been implanted there by the family, local community, by the society. Plans and ideas that come to my mind result from specific culture I was brought in, life experiences and especially results of reactions to different life situations. But it is also important to realize that not everyone has a choice, it is not easy to break out from the ghetto, from criminal influences or various addictions. Children brought up in the Khmer Rouge camps new only killing and their training reduced their natural empathy. Sometimes one wrong step may determine the rest of one's life.

Seeing myself as the whole organism, not just as a set of ideas or model of myself, allows me to accept all aspects of my nature, all bad and good sides, understand how fragmentary and illusory is the knowledge of myself. It cannot be otherwise, because I cannot test myself in all possible situations, and the self does not have an access to all processes going on in the brain. These processes may be too weak or too diffused to become conscious, but observing my own behavior I can sometimes correct it, bringing it in line with my value system and beliefs, and work on the improvement of my character. Reflection enables conscience, and looking in this light at the social development one can see the value of various cultural and religious practices, contributing to deeper reflection and to building conscience. I may feel empowered to act, accepting as mine all that my brain produces, but choosing to act only when the self accepts it. Darwin's note that men should know that "everything he does is independent of himself" has not been based on deeper understanding of the self and brain processes that create it, as such knowledge was not existing in his times.

In the movie "Of Gods and Men", based on the true story of Trappist monks stationed in Atlas Monastery in Algeria who came under the life threat from fundamentalist terrorists, Christian, the leader of this monastery, is told by one of them "You don't have a choice", and he answers: "I always have a choice". It is impossible to predict one's own behavior in such a stressful situation, but working on one's own character to achieve such stability is a noble goal. Reacting without time for reflection we sometimes do things that we later regret. Education should show children a way how to develop oneself to become in future what we would really like to be, to internalize important values so that they will become a part of our instinctive behavior, without the need for additional reflection. In many religious traditions "cultivation of mind", called in Buddhism *bodhicitta* or the mind of enlightenment, has been an explicit goal. Training was aimed at elimination of selfishness, dropping all personal fears, developing empathy and achieving "goodness for the goodness sake". Such goal does not have to be connected with religion, although religious inspirations may be helpful. Chu Hsi (1130-1200), philosopher in the Neo-Confucian tradition created a secular image of a sage as a role model. In the *Reflections on Things at Hand* he wrote: "The essential training [of government] should be the way of choosing the good and

cultivating the self until the whole world is transformed and brought to perfection so that all people, from the ordinary person up, can become sages" (Chan 1967).

Secular programs aimed at cultivation of mind should be acceptable to everyone. Frequently a false dichotomy is presented: materialist vs. idealist, or atheist versus religious people. Materialists are supposedly only interested in material things and show no spiritual inclinations, while idealists are oriented towards higher goals, shunning material wealth. Numerous examples show that there is not true and that we should optimize our behavior towards achieving more distant goals that will bring true and lasting satisfaction, sharing our happiness with other people. Even animals show many altruistic behaviors. The role of prefrontal cortex and amygdala nuclei in moral behavior has been well documented. Less known is the fact that genetic factors contribute to the reward mechanism that makes some people susceptible to confirmation bias (Doll et al. 2011; Frank et al. 2009). Mutations that lead to changes in the DARPP-32 and COMT genes that are expressed in dopamine receptors in striatum and prefrontal cortex are correlated with the length of time people stick to rules they have been told despite observing that they fail, or in general exploitation of what they know in contrast to own exploration. Many high-ranking members of various religious organizations in the western as well as eastern countries have been accused of terrible crimes, The differences between atheists and religious people may have biological roots, but such differences are of secondary importance. Attitudes toward spirituality, understood as individual, transpersonal pursuit of sense and meaning in life seem to be more important. There seems to be no correlation between selflessness and empathy with idealistic beliefs, although hard data on this issue is hard to find. Certainly there is a negative correlation between religiosity and crime rates (compare for example USA with any Scandinavian countries using statistic information at nationmaster.com).

The "free energy principle" proposed by Friston (2010) should be extended beyond individual brain. An accounts of action, perception and learning without social context cannot be complete. What is optimized is not only internal predictive error of one's own behavior but of other people's behavior as well, and of social reactions to our behavior. Much of what we know about ourselves comes from literature and art, exploring the ways our minds may tune into the environment and interaction with other people. Humanities should join natural sciences in the great enterprise of building a new vision of what it means to be human, the science of human experience. In the end we shall know the truth and (hopefully) it will liberate us.

References

Anderson M (2010) Neural reuse: A fundamental organizational principle of the brain. Behavioral and Brain Sciences 33: 245–313

Aristotle, De Anima (350 B.C), Translated by J. A. Smith, the Internet Classics Archive.

Austin J.H. (1988) Zen and the Brain: Toward an Understanding of Meditation and Consciousness. Cambridge, MA: MIT Press.

Baars, B.J. (1988) A Cognitive Theory of Consciousness (Cambridge, MA: Cambridge University Press)

Baumeister, R.F. (2011) The Unity of Self at the Interface of the Animal Body and the Cultural System. Psychological Studies, 56, 5-11.

Baumeister R.F, Masicampo E.J, de Waal C.N. (2009) Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness. Personality and Social Psychology Bulletin, 35, 260-268.

Berlin B. and Kay P. (1969) Basic Color Terms: Their Universality and Evolution. University of California Press, Berkeley.

Brown W.S, Murphy N, Malony H.N, red. (1998) Whatever Happened to the Soul? Scientific and Theological Portraits of Human Nature. Minneapolis: Fortress Press.

Bunson M.R. (2002). Encyclopedia of Ancient Egypt, Facts on File, New York.

Carruthers P. and Smith P. (1996) Theories of Theories of Mind. Cambridge University Press

Cashmore A.R. (2010) The Lucretian swerve: The biological basis of human behavior and the criminal justice system. Proc. of National Academy of Science 107: 4499-4504.

Chan Wing-tsit (1967) Reflections On Things at Hand, Columbia University Press, New York

Christoff K, Cosmelli D, Legrand D and Thompson E. (2011), Specifying the self for cognitive neuroscience. Trends in Cognitive Sciences 15(3), 104-112

Churchland P.M. (1981) Eliminative materialism and the propositional attitudes. Journal of Philosophy 78(2): 67–90

Crick F. (1995) The Astonishing Hypothesis: The Scientific Search for the Soul. Scribner

Damasio A.S. (2010). Self Comes to Mind. Constructing the Conscious Brain. Pantheon Books.

Danziger S, Levav J, Avnaim-Pesso L. (2011) Extraneous factors in judicial decisions. Proceedings of the National Academy of Sciences, DOI: 10.1073/pnas.1018033108

Davidson R. (2010) Empirical explorations of mindfulness: Conceptual and methodological conundrums. Emotion 10(1): 8-11

Dehaene, S., Sergent, C. and Changeux, J.-P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. Proc. National Academy of Science (USA) 100. 14: 8520-8525.

Dehaene S, Changeux J.-P, Naccache L, Sackur J. and Sergent C. (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in Cognitive Sciences 10(5): 204-211

Demertzi A, Liew C, Ledoux D, Bruno M.A, Sharpe M, Laureys S, Zeman A. (2009) Dualism persists in the science of mind. Annals of the New York Academy of Sciences 1157: 1-9

Dennett, D. (1995) The Unimagined Preposterousness of Zombies, Journal of Consciousness Studies, vol. 2, no. 4, pp. 322–326

Depraz P.V.N, Varela F. (2003) On Becoming Aware: A Pragmatics of Experiencing. John Benjamins, Amsterdam.

Deutscher F. (2010) Through the language glass: why the world looks different in other languages Metropolitan Books.

Dobosz K and Duch W. (2011) Visualization for Understanding of Neurodynamical Systems. Cognitive Neurodynamics 5(2): 145-160.

Doll B.B, Hutchison K.E, Frank M.J. (2011) Dopaminergic Genes Predict Individual Differences in Susceptibility to Confirmation Bias. Journal of Neuroscience 31(16): 6188-6198

Duch W. (1999). Soul and spirit, or prehistory of cognitive science (in Polish). Kognitywistyka i Media w Edukacji 1, 7-38.

Duch W. (2005) Brain-inspired conscious computing architecture. Journal of Mind and Behavior, Vol. 26(1-2): 1-22

Duch W. (2009) Imagery agnosia: what goes on in my head? In Coma and Consciousness, Clinical, Societal and Ethical Implications. Berlin, June 2009, p. 46.

Duch W. (2010) Free Will and the Brain: Are we automata? In: Ethics and Science for the Environment, 3rd European Forum. Ed. M.Jaskula and B.Buszewski, Societas Humboldtiana Polonorum, 2011, pp. xxx.

Duch W. (2011) Neurodynamics and the mind. Proc. of the International Joint Conference on Neural Networks, San Jose, CA, IEEE Press, pp. 3227-3234

Edelman G. (2006) Second Nature: Brain Science and Human Knowledge. Yale University Press.

Enzi B, de Greck M, Prosch U, Tempelmann C, Northoff G (2009) Is Our Self Nothing but Reward? Neuronal Overlap and Distinction between Reward and Personal Relevance and Its Relation to Human Personality. PLoS ONE 4(12): e8429. doi:10.1371/journal.pone.0008429

Faw B. (2009) Conflicting intuitions may be based on differing abilities, Journal of Consciousness Studies 16(4): 45-68

Frank, M.J., Doll, B.B., Oas-Terpstra, J. & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nature Neuroscience 12: 1062-1068.

Friston K. (2010) The free-energy principle: a unified brain theory? Nature Reviews Neuroscience 11, 127-138 (February 2010) | doi:10.1038/nrn2787

Frazer, J.G, The golden bough: a study in magic and religion. New York: The Macmillan Co., 1922.

Freeman A, Libet B, Sutherland K, eds (2000) The Volitional Brain: Towards a Neuroscience of Free Will. Imprint Academic

Gibson J.J. (1986) The Ecological Approach to Visual Perception. Lawrence Erlbaum Associates. Hillsdale, N.J.

Heatherton T.F, Wagner D.D. (2011) Cognitive neuroscience of self-regulation failure. Trends in Cognitive Sciences 15(3): 132-139

LeDoux, J. (2003) Synaptic self: how our brains become who we are. New York: Penguin.

Liu C, Tardif T, Mai X, Gehring W, Simms N and Luo Y. (2010) What's in a name? Brain activity reveals categorization processes differ across languages. Human Brain Mapping}, vol. 31(11): 1786-180.

Logan R.K. (2007) The Extended Mind: The Emergence of Language, the Human Mind and Culture. University of Toronto Press.

Hurlburt R. and Schwitzgebel E. (2007) Describing Inner Experience? Proponent Meets Skeptic. MIT Press.

Janata P. (2010) Brain electrical activity evoked by mental formation of auditory expectations and images. Brain Topography 13(3): 169-193

Menary R. (2010) The extended Mind. MIT Press.

Minsky, M. (1988) The Society of Mind. Simon and Schuster, New York

Northoff G, Bermpohl F. (2004) Cortical midline structures and the self. Trends Cogn Sci. 2004 Mar;8(3):102-7.

Northoff G, Heinzel A. (2006) First-Person Neuroscience: A new methodological approach for linking mental and neuronal states. Philosophy, Ethics and Humanities in Medicine 2006 Mar 17;1(1):E3.

O'Regan, J.K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences, 24(5), 939-1031.

Pascal B (2006). Pensées, Project Gutenberg EBook, http://www.gutenberg.org/etext/18269

Pinker, S. (2002). The Blank Slate: The Modern Denial of Human Nature, Penguin Putnam

Ploux S. and Ji H. (2003) A model for matching semantic maps between languages. Comp. Linguistics 29(2): 155-178.

Pronin E. (2009) The Introspection Illusion. In Mark P. Zanna, editor: Advances in Experimental Social Psychology, Vol. 41, Burlington: Academic Press, pp. 1-67

Schacter, D.L. (2001). The seven sins of memory: How the mind forgets and remembers. Boston: Houghton Mifflin.

Schwitzgebel E. (2011) Perplexities of Consciousness. MIT Press.

Searle, J. (1992) The Rediscovery of the Mind, MIT Press.

Soon, C.S, Brass M, Hans-Jochen H, Haynes J-D. (2008) Unconscious determinants of free decisions in the human brain. Nature Neuroscience 11: 543-545.

Tani J. (2008) Objectifying the subjective self: An account from a synthetic robotics approach, Constructivist Foundations 4(1): 28-30.

Ullman S. (1996) High-level vision: Object recognition and visual cognition. MIT Press: Cambridge, MA,

Varela F. (1996) Neurophenomenology: A Methodological Remedy for the Hard Problem „Journal of Consciousness Studies" 3 (4): 330-349 (polish translation: Avant. The Journal of the Philosophical-Interdisciplinary Vanguard 1/2010, 31-76.

de Wall, F. (2005) Our Inner Ape. New York: Riverhead Books.

de Wall, F. (2009) The Age of Empathy: Nature's Lessons for a Kinder Society. Crown.

Weber B, Depew D, eds. (2003) Evolution and Learning: The Baldwin Effect Reconsidered; MIT Press, Cambridge, Mass 2003

Weiskrantz L. (1986) Blindsight: A Case Study and Implications. Oxford University Press

White A.D, A History of the Warfare of Science with Theology in Christendom, 2 vols. (1896), dostępne w projekcie Gutenberg http://www.gutenberg.org

Wright R (1994) The Moral Animal (Pantheon Books, New York).

Zacks J.M, Speer N.K, Swallow K.M, Maley C.J. (2010) The brain's cutting-room floor: segmen-tation of narrative cinema. Frontiers in Human Neuroscience Oct 1;4. pii: 168.