Merleau-Ponty's Concept of the Unconscious in Perception and the Attempt to Model Cognitive Information Processing in Cognitive Science and AI

Bernhard Irrgang *
irrgangb@msx.tu-dresden.de

ABSTRACT

In *L'institution, la passivité. Notes de cours au Collège de France (1954-1955)*, Maurice Merleau-Ponty develops approaches to a concept of the unconscious that can also be relevant to a mental philosophy from the perspective of cognitive science and AI in the tradition of continental philosophy. I consider this all the more important as previous attempts at an analytical philosophy of AI tend to remain stuck in Cartesian or Platonic dualism, which – as a dream of Silicon Valley — could encourage gnostic thinking (Irrgang 2021).

1. Merleau-Ponty and the Hermeneutics of Psychoanalysis

According to Merleau-Ponty's *L'institution, la passivité*, a field of social sedimentation and structure formation emerges in historical perception. It is a social entity, in which everything is social and a dialectical milieu in full reality. It produces crises, problems, solutions, regressions, transformations, but is never at rest in itself (Merleau-Ponty 2015, 153). The perceptual foundation of the real, which is unconscious in Levi-Strauss' work and the object of psychoanalysis in Freud's, contains a sexual simplification in terms of nomenclature and classification (Merleau-Ponty 2015, 154). Merleau-Ponty points to a weakness in Levi-Strauss' analysis with regard to the treatment of complex forms (Merleau-Ponty 2015, 158): by determining institutions, a remedy is sought for the difficulties posed by philosophy of consciousness (Merleau-Ponty 2015, 161). The fragments of an analysis presented here by Merleau-Ponty enable a revision of Hegelianism and lead to the discovery of

* TU Dresden, Germany.

phenomenology as a living, current, original connection between the elements of the world. But this disappears into the past, as it causes a subordination to the systematic vision of philosophy (Merleau-Ponty 2015, 165).

Far from reducing philosophy to psychology or indeed to sociology, and far from developing a superposition of philosophical existence to ontology and its disciplines, our phenomenal body and the perceived world of the pre-objective allow this same body to come into being and act through the soul. The latter can be understood as an action of the objective body that affects the phenomenal body. In the process, consequences may be produced that are still not intelligible (Merleau-Ponty 2015, 210et seq.). The phenomenal body (*Leib*) presupposes the perceiving body (*Körper*). Now, is this the vegetative soul, and what system of finality does it develop (Merleau-Ponty 2015, 218)?

The unconscious can be seen as sedimentation of the perceptual world in the sense of primal sedimentation. In contrast, there might be a second world of sedimentation, filled with matter of the symbolic. The unconscious represents a world that does not come to reactivation and thus cannot control the dream world. Here, unconscious compromises may occur, a vital drama and variations surrounding it, and crystallisation points may be reported throughout the day's events (Merleau-Ponty 2015, 278). Thus, the unconscious appears as the eternally existential, as the coherence of life through the exuberance of events. This raises the question whether the unconscious is adequately explained by the persistence of symbolic matrices (Merleau-Ponty 2015, 293).

What if the unconscious is only an implicit thought, a second thought in myself, so to speak, like some kind of delirium, an unconscious resurrection of myself? Could there be a causality of the unconscious within a third person? Could it be in the sense that we are only results (Merleau-Ponty 2015, 301et seq.)? Even if there are certain limits with regard to recollection, it seems to me that it is always possible to regain a previous level of my life through regression. But can this existing passivity, that precedes an event, that questions me and demands the proof of my existence in the world, can it be sufficient to make me aware of the burden of my past? Should we interpret the unconscious as a capacity for memory, as a sedimented memory of what we experienced and learned before? Ultimately, memory grows out of our life story, making the passivity of memory between preserving and forgetting the basis of our own life story (Merleau-Ponty 2015, 330et seq.).

In the 20th century, the theory of the unconscious and its hermeneutics was oriented towards psychoanalysis and psychotherapy. The

psychoanalytic tradition of interpreting the unconscious drew on its conception in late Romanticism, for example in the work of Schelling. Alfred Schöpf begins by reconstructing the forms of knowledge of the unconscious in the 19th century. Ultimately, however, these attempts prove unsatisfactory. In order to understand the relationship between language and affect, the concept of the body (*Leib*) needs to be understood as a mediating instance between nature and culture (Schöpf 2001, 49).

The treatment of mental illness was the horizon of experience within which psychoanalysis established itself. In this respect, psychoanalytic experience can be understood as an interpretation of the lifeworld (*Lebenswelt*). For Schöpf, the tension between Enlightenment and Romanticism, between Enlightenment reason and its shadow side, is crucial. He assumes that the demands of ethics cannot be met on a discursive – i.e., rational – level alone. All cognition is rooted in affects. Moral action presupposes a stable function of the self. Schöpf addresses affective development and the basis of self-formation and cognitive development. Another essential element is the effort to ground psychoanalysis in scientific theory. The phenomenology of (inter-)corporeality (*(Inter-)Leiblichkeit*) plays a particularly pivotal role here (Schöpf 2001, 13-23). However, there seem to be certain limits to the objectification of psychoanalytical content: The contents of unconscious desires and fantasies are much more difficult to assess in testing procedures than conscious contents. In addition, therapists are observers that participate more than they objectify – they have to be able to put themselves in the other person's shoes, to assume the role of the other person (Schöpf 2001, 107).

2. Early Romantic Ironic Scepticism as an Alternative Approach to the Bodily Unconscious

An approach that deviates from the standard interpretation of the unconscious in the 20th century dates back to the transitional period between late Enlightenment and early Romanticism in the context of Kantian scholarship and the development of a theory of self-consciousness. It was more sceptical than is generally assumed and was explored systematically as recent as between 1990 and 2005. This is not surprising, as the Age of Enlightenment was more sceptical and sensualistic than is usually suggested by its rational interpretation (Irrgang 1982; Irrgang 1990). Irony and scepticism are attitudes of a sobered

existence. They also express the early Romantic worldview: Life is perceived as an insoluble riddle, and the human being as appropriately boundless.

Aenesidemus Schulze recognises and describes the causes of Carl Leonhard Reinhold's irresolvable difficulties: they are due to the fact that Reinhold interprets all consciousness, especially self-consciousness, according to the model of imagination which states that the subject must be opposed by an object (or itself as an object). This, seemingly inconspicuous, corrective imagination model of self-consciousness represents nothing less than the first serious attempt at radical and representational interpretation of self-consciousness, as it subsequently reappears in Novalis' *Fichte-Studien* (1795/96), and in Johann Gottlieb Fichte's lectures on the *Wissenschaftslehre Nova Methodo* (1796/99). According to Aenesidemus, however, there is also non-representational consciousness (e.g., sensations, pain). In particular, self-consciousness is not a case of imagination, since imagining always entails an intentional accusative *of which* it is imagining. In the case of self-consciousness, however, I do not imagine an object that is distinct from me in object status, but rather - *indistinctly* - myself as subject (Frank 1997, 293-295).

Immanuel Kant identified judgement with thinking. Every judgement functions in service of a synthetic unity, for judging means linking a subject and a predicate term with each other so that, if the sentence formed from it is true, it can be recognised what the object designated by it is about. Friedrich Hölderlin, in *Urtheil und Seyn* (1795), follows not only in Kant's footsteps but also a widespread, although erroneous etymology that interprets the expression "*Urteil*" (judgement) as indicating a division – "*Ur-teilung*". In the act of judgement, something previously unified divides itself into two components or relations. The relationship between the two simultaneously conceals and reveals the original unity. It reveals, for in judgement, two different ideas are connected with each other and thus related to a ground of unity. But it also conceals, because unity does not come about as such, only as difference between two classes of ideas that are intertwined with each other. Hölderlin then applies this fundamental consideration to the particular case of the judgement "*Ich ist Ich*" ("I is I"). Here, too, a differentiation takes place: a judgement divides the relata.

This, though, happens in such a way that the content of the judgement contradicts its own form. What is expressed through the judgement is the non-differentiation of the relata; the form of the judgement, however, consists precisely in distinguishing the non-differentiation. Hölderlin calls this unity – following Spinoza and Jacobi – "*Seyn*" (being). *Seyn* is situated even above the

relative identity that Fichte discusses in his texts. It cannot be thought, because thinking is judgement, and judgement is non-differentiation. For without a basis of unity the reference would only be postulated, the actual and evident experience of an "I am I" as the identity of "I" would have to remain unexplained. Strictly speaking, it is not only the unity of pre-identity that is derivable from judgement-relations. Hölderlin assumes that self-consciousness is only possible by opposing myself, by separating myself from myself, but that despite this separation I recognise myself and the opposite as being one and the same (Frank 1997, 750et seq.).

According to Manfred Frank, this is the initial idea that expresses the fundamental shared conviction of early Romanticism: It consists in the assumption that being – as simple, jointless unity, in contrast to the identity of the Kantian-Fichtean *Cogito* – cannot be made comprehensible from the judgemental and reflexive relations that all original divisions of those that have to be united make, and that can only ever presuppose the original, simple unity. Hölderlin's critique of Fichte thus consists in the fact that he most firmly opposes intellectual perception to the act of division through judgement (*Urteilung*) (Frank 1997, 751). Even though Hölderlin, with his conception of love and harmony, remains in the wake of Heraclitus' search for himself under the heading of opposing harmony, he still prepares a shift from the rational interpretation of the Kantian thought of "I think" as the supreme principle of subjectivity as self-consciousness, and a turn towards the pre-rational, self-feeling and implicit knowledge (Irrgang 2015).

Hölderlin's and Novalis' argumentations work factually towards approaching the phenomenon of our undeniable familiarity with ourselves (Frank 1997, 822). Taking into account the enduring conflicts of interpretations, they suggest that an Enlightenment hermeneutic tradition might be – sceptically – open to its extension in the 21st century. This is against the background of the positions of tacit knowledge, pre-rational consciousness, and bodily sense of self, i.e., the concept of a human corporeality (*Leiblichkeit*) that is ontically and ontologically prior to its inner subjectivity, which could be empirically substantiated within the framework of a cybernetic respectively an evolutionary cognitive anthropology (Irrgang 2015; Irrgang 2020). According to my interpretation, hypermodernity in philosophy thus overcomes the opposition between a transcendental position of reason and the contrasting positivist data position by investigating the empirical preconditions of human bodily mentality (*menschlich-leibliche Mentalität*) with scientific means. But it

has to use the means of exactly the same mentality and will therefore never be fully able to grasp reality as foundation for mentality in its own realistic essence. For me, this is part of the reason for the aporetic constitution of mental self-consciousness, which cannot be grasped by exclusively empirical means.

Thinking no longer grasps being – this is the real crux that early Romantic scepticism argues against Reinhold's philosophy of principles, which is not to be confused with the system philosophy of late Romanticism and the hermeneutic tradition based on it. For Novalis and Hölderlin, the only thing that remains to be demonstrated is how reflection itself captures its own inability to represent the absolute. It is specifically for this capacity of reflection-through-its-own-deletion that Novalis summons intellectual contemplation (*Anschauung*). For him, like for Hölderlin, being (*Seyn*) is more than just aspects of identity and non-relations. Thus, intellectual contemplation no longer reaches being itself, but merely explains our motive, namely the transcendental reasons for holding onto its inconceivability. Like Hölderlin, Novalis does not consider the self to be a principle of philosophy at all, for the foundation of consciousness is being (*Seyn*) (Frank 1997, 529et seq.).

In Novalis' *Fichte-Studien*, under the heading "*Vom empirischen Ich*" ("About the Empirical I"), we encounter for the first time the formula of the "*Trieb, Ich zu seyn*" ("Drive, to Be I"). The fact that this drive appears in connection with the "empirical I" is easily explained by reading the third part of the *Wissenschaftslehre* (Frank 1997, 845et seq.). In contrast to the ideal of "I", only the form of "I-ness" as a respect of action is present in the intellectual view (Frank 1997, 850). At the end of the *Fichte-Studien*, Novalis points out yet again that truth has to be replaced by probability, and ultimately by the criterion of coherence. He draws this conclusion from the impossibility of ultimately establishing the truth of our convictions (Frank 1997, 854). In the same vein, Schlegel explains in 1804 that what is actually contradictory in our "I", is that we feel both finite and infinite at the same time (Frank 1997, 865). Mysticism is the abyss into which everything sinks. Although this is the most sober and solid of all frenzies, it is nevertheless not a viable alternative to critical philosophy. The latter remains pathetically linked to scepticism, insofar as it assumes that no principle is secured by evidence, by virtue of which our fallible creations would become certain (Frank 1997, 867). Thus, the motive for philosophising is not the successful perception of an *omnitudo realitatis*, but, as with Novalis, on the contrary: the feeling of a lack, of an incompleteness, a striving for knowledge, a

tendency towards the absolute, a longing for the infinite, and more of many similar wordings (Frank 1997, 874).

As early as in Schlegel's notes of 1796 and 1797, the elucidation with regard to fundamentals occurs genetically, i.e., progressively and, since the principle of the entire theory of science (*Wissenschaftslehre*) ultimately proves to be an idea in the Kantian sense, never finally. The "I" is, in fact, not really given as the result of an act, but only proves to be assigned as a task to be subsequently resolved. Thus, deduction never has an end. Knowledge implies truth, feeling does not. This leaves an enormous gap between truth and feeling (Frank 1997, 876et seq.). Schlegel notes in the second half of 1797: *Archai*, as principles, are always plural and construct themselves among one another as alternating constructions and alternating determinations (Frank 1997, 888-890). Thus, the failure of self-justification as system philosophy by Kant and the representatives of idealistic philosophy leads in early Romanticism to attempts at finding solutions between common sense and scepticism. This produces the first models of a reciprocal constitution of sensuality and mind. 200 years later, once again in a situation of upheaval after the end of modernity, subjectivity can now be transformed within the framework of a sceptical-empirical meta-epistemology as a pattern of self-organisation between sensorimotor and thought processes, and a changing positionality of "I" and of subjectivity through perspectivity in transition to multi-perspectivity.

The elementary bodily (*elementar Leibliches*) is spontaneous, not fundamental or principled. Approaches of developmental thinking, rather than architectural constructivism emerge – or more precisely, a combination of the two. Thus, romantic scepticism paves the way for hypermodernity as a form of civilisation that advances ironic play with the third principle of logic (*tertium non datur*). Self-justification and intersubjectivity of the "I" are formally not possible on a mental level. The necessary, pragmatic reformulation of Kant's transcendental philosophy and the reflections of meta-mathematics on the mathematisation of mathematics clearly show this, and ultimately legitimise a sceptical thinking that carefully tries to rethink itself and its presuppositions. A self-referential, ironic form of scepticism is a possible starting point for meta-epistemologies of hypermodern, technologized science. Real possibilities of self-organisation, for example in biological evolution, have to be distinguished from self-organisation of possibilities of thought and forms of knowledge. Self-consciousness (*Selbstbewusstsein*) is grounded in self-awareness (*Selbstgefühl*), and must not be confused with rationality itself. In this context, classical

scepticism, which was primarily concerned with sensorimotor systems and their mental comprehensibility, is now supplemented by a scepticism of the mental that turns ironically against itself, and is also prepared to doubt logic and mathematics (Irrgang 2015). In Hume's second book of the *Treatise of Human Nature*, we find the famous analysis of reflected feelings, which also allowed for a self-attribution as an "I" with intentional or cognitive character. In this context, two types of self-awareness can be distinguished: (1) a perception of one's own inner state; and (2) an epistemic type of self-ownership (*Selbsthabe*), a specific knowledge, which can also be referred to in terms of a conscience (Frank 2015a, 26-31).

Novalis already distinguishes himself from Fichte's conception of self-awareness by interpreting it as a feeling of dependence on being (Frank 2015a, 34). In this context, it can be assumed that being is related to experience and that reality takes precedence over possibility (Frank 2015a, 41). Feeling can also be understood as a basic sense within the meaning of an overarching sensation (Frank 2015a, 66). For those who believe with Leibniz in the continuing (so to speak, ontogenetic) elucidation of degrees of consciousness, starting from darker animal and plant feelings, the idea of considering palpatory or tactile sensation as a kind of primordial sense cannot seem too farfetched (Frank 2015a, 69). In this respect, Leibniz may be regarded as a forerunner and pioneer of evolutionary thinking in 18th century natural history (Irrgang 1983). As a shared, common feeling (*Gemeingefühl*), initially based in the sense of touch, it is still known in the lectures on psychology (*Vorlesungen über Psychologie*) that Schleiermacher gave repeatedly from 1818 onwards (Frank 2015a, 73). Condillac, too, bases the identification of a general sense with the skin system in terms of life as surface. The Scottish school led by Thomas Reid understands this then as the basis of common sense (Frank 2015a, 74et seq.).

Thus, the feeling as consciousness of reality is correlated with the feeling of self as consciousness of the reality of the self (Frank 2015a, 77). This feeling does not deceive as such, because, as Rousseau remarks, errors are not based on sensation but on the judgement of a possibly false statement (Frank 2015a, 81). The inner sense is not self-observation but an act of turning to oneself, or, as Novalis writes, following Leibniz: what reflection finds already seems to be there. This anticipates Edmund Husserl's approach as well (Frank 2015a, 107). In this context, the theory of inner sense as a theory of how consciousness of our inner life emerges is inadequate, as Franz Brentano and Sydney Shoemaker have shown, because it interprets the act of reflection as

sensory perception, as an inner observation (Frank 2015a, 109). Brentano's attempt to overcome the model of inner observation basically merely envisaged a more modern version of the inner sense model that Kant already grappled with (Frank 2015a, 111). Strictly speaking, this concerns the consciousness of existence of one's own mind, which Locke as well as Descartes and the whole tradition of theorists of the sense of self assumed to be certain (Frank 2015a, 114).

In order to arrive at a comprehensive conception of the feeling or rather awareness of self (*Selbstgefühl*), we have to blend two theoretical hybrids, the reflexive and the pre-reflectivist conceptions of self-awareness, as well as take into account this feeling's dependence on being (Frank 2015a, 176). A suggestion to relate Novalis to Sartre and Schelling with their positions on the sense of self could thus also constitute a proposal to solve the problem of what existential consciousness has to do with pre-reflective self-consciousness, or with the double problem of ontological and epistemological regress (Frank 2015a, 234). Self-consciousness is an ambiguous compound. In German, a distinction is made between "*selbst*" and "*sich*", a distinction that is not possible in the same way in the English "*self*" and the French "*soi*", although Jean-Paul Sartre and Paul Ricoeur made a philosophical effort to distinguish between the two. In their language, this difference could not be represented in a simple word, but had to be outlined and explained. Self-consciousness refers to an awareness of the consciousness of itself (Frank 2015b, 14et seq.).

Rationality maintains a fundamental relationship to self-knowledge, as is expressed in the thought "I think" (Frank 2015b, 19). The crucial factor in the connection between self-knowledge and rational thought and action lies in what can be called the "motivating power of rational insight". In this context, a close link between self-awareness and motivation for self-preservation can be formulated from a philosophical perspective, for which there is philosophical-historical evidence (Frank 2015b, 22 et seq.). In doing so, we must not ascribe infallible certainty to self-knowledge, but should acknowledge that it may occasionally be mistaken about itself. We must also distinguish fundamentally and in philosophical terms between phenomenal and cognitive self-consciousness, meaning self-consciousness and self-knowledge (Frank 2015b, 25).

In the context of awareness of consciousness of itself, Sartre introduced the term of "irreflexive" or "pre-reflexive" in his 1936 essay *La Transcendance de l'Ego*, and elaborated on it widely in his main work (Frank 2015b, 26). From

Descartes to Leibniz, Kant and German Idealism, the dominant conception of modern philosophers – who might as well be called philosophers of the subject – has been similar to that of most representatives of a Philosophy of Mind, traditionally called in German "*Geistphilosophie*" (Frank 2015b, 35). Leibniz, for example, assumes that every distinct external perception incorporates self-awareness, although he was also aware of the danger of regressive entanglements in the definition of reflection. He tried to defuse it with the concept of "*petits reflexions*", of small perceptions (Frank 2015b, 36et seq.; Irrgang 1982). Returning to Sartre, this makes the differentiation between self-consciousness and self-knowledge plausible and fruitful for further considerations. Sartre justifies the primacy of being over consciousness in this way (Frank 2015b, 92). Moreover, the phenomenon of intersubjectivity could be made intelligible in a pre-reflexive way (Frank 2015b, 95).

If I find myself in a condition of feeling, or sensing, I do not need to know how to identify, name and classify it. But even if I am able to do so, and thus perform a cognitive, high-level intellectual operation, it does not have to correspond to a propositional attitude – at least not if it refers to myself as myself (Frank 2015b, 97). The theory of tacit knowledge by Michael Polanyi and his successors already suggested this and identified it within a concept framework as personal knowledge (Irrgang 2001). The circularity of Hegelian theory as a form of idealist subject philosophy was first exposed in full clarity by Sartre (Frank 2015b, 98). This circularity is, in my nonidealistic interpretation, the antecedent being of being-subject, which can be interpreted as subjectivity of the human and bodily subject. From the perspective of individual scientific disciplines, this being can be considered as having evolved – just like sociobiology, neurodarwinism and evolutionary epistemology put it. This way, a conception of philosophy can be argued that does not want to be located in transdisciplinary scientific endeavours of individual disciplines, but that should be conceived as methodologically integrated into them with reference back to the common sense of life-world. A self-presentation of the properties of consciousness is thus not known by virtue of reflection. Self-consciousness presents itself as a *sui generis* phenomenon. There seems to be a type of consciousness that is not intentionally directed towards something distinct from itself. Those who argue in favour of a kind of self-representationism try to circumvent the consequence of regress because it impedes the intention of naturalistic reduction. Materialists accept the proposed consequence of representational self-knowledge, which presupposes non-representational self-

knowledge, and thus also that not all consciousness is representational (Frank 2015b, 150).

One can identify this, as I do, as the non-representational being of the bodily subjectivity of the human subject, which is, however, in reciprocal justification open to natural as well as mental – or cultural – interpretation (Irrgang 2007; Irrgang 2009; Irrgang 2021). This is an epistemological, not a metaphysical interpretation of the classical-ontological philosophy of *Dasein* (Heidegger). It ties in with the conception of the Heidelberg School, which in ontological neutrality seeks to stop the emergence of intensive recourses. Its conception of emerging intensive regresses is based on the relation of self-representation, which is embedded in itself like a Matryoshka doll, so that it self-represents itself infinitely. Such a regress ought to end by the second loop (Frank 2015b, 158). It distinguishes self-consciousness of a self-referential and representational kind, which coincide in pre-reflexive consciousness. In this way, however, self-consciousness is no longer the fundamental principle of an idealistic philosophy of a subject's subjectivity, but rather its bodily interpretation. In it, pre-rational and rational self-knowledge mutually justify and criticise each other and thus indicate a model for reflection that is no longer fundamentalist (idealistic or materialistic), but in a certain sense pre-dualistic. It does not identify thinking and being, rather it makes human corporeality (*menschliches Leibsein*) accessible to a scientific, anthropological – empirical as well as mental-culturalistic – approach.

3. Slow and Fast Thinking and Decision-Making: How Conscious and Unconscious Dimensions of the Mental Interact – and Why There is Consciousness

Usually, our trust in our intuitive thinking is justified. Humans are excellent intuitive statisticians (Kahnemann 2012, 13-15). Still, reliance on intuitive heuristics with the risk of cognitive bias are sources of error. Similarities thus become stereotypes; associative memory is the source of any interpretation. The experiencing self and the remembering self do not have to be identical to recognise intuition (Kahnemann 2012, 17-27). Two decision-making systems can be distinguished: slow and fast thinking and decision-making. Thinking System I works automatically and quickly, largely effortlessly and without deliberate control. Quick thinking means intuitive comprehension, multiplication with and without arithmetic steps. System II directs attention to

the strenuous mental activities upon which it relies for decision-making, including complex calculations. Thus, System II operations are often accompanied by the subjective experience of agency, freedom of choice and concentration. The automatic operations of System I produce amazingly complex patterns of ideas, but only the slower System II can construct thoughts in an orderly sequence of steps. System I prepare operations of System II, while System II requires attention. These are each individual capacities with interactions between the two systems. Whenever System I encounters difficulties, it demands more detailed and specific processing from System II (Kahnemann 2012, 31-37).

In the division of labour between the two systems, conflicts can arise between the automatic response and the will to maintain control. Such conflicts occur frequently in everyday life. Cognitive illusions and deceptions of thinking occur in both systems when mental actors are confronted with individual personalities, abilities and boundaries. Driving a car on a country road is generally left to System I. It is an automatic and unintentional system (Kahnemann 2012, 38-44). The pupils of the human eyes are sensitive indicators of mental strain. In the economy of action, effort is a cost factor; moreover, the brain strives to avoid mental overload. With regard to forms of human work, there are also two systems that can be distinguished: People who are cognitively overloaded are more likely to make selfish decisions, use sexist terms and make superficial judgements in social situations. This is because self-control requires attention and effort. In contrast, there is a "lazier" System II, where intelligence, control and rationality interact. Rationality must be distinguished from intelligence (Kahnemann 2012, 45-67).

Implicit knowledge and automatic, quick decision-making and addition comprise one area of human action and knowledge, another area can be defined as explicit knowledge and deliberate action. Patterns of action and rules of thought may be recorded and documented as rational structures (in spoken word, in writing, in algorithms, in technical rules, in Artificial Intelligence and in robots). The structures of explicit human intelligence emerge from those of implicit intelligence. Implicit mental structures are typically of human bodily nature *(leiblich*) and form the basis for explicit mental competences. Being able to deal with rules and operative structures of explicit rationality presupposes explicit knowledge and consciousness.

Logic and mathematics often confuse structures and schemes of language, of knowledge and of action of human explicit intellectuality with

competence. Implicit knowledge is no knowledge in a conventional sense, but knowledge of a sensorimotor kind in relation to bodily reactions and behaviours that everyone has at their disposal, even if they are not mindful about it. Conscious action, however, presupposes a mental dimensional leap: in evolutionary terms, it constitutes a preliminary stage of consciousness and intelligence. For the most part, it is meant to escape from emergency situations. Implicit knowledge is part of the mental developmental path of human intelligence, human consciousness and self-consciousness, as well as knowledge-capacity. So, given the issue of quick and deliberate decision-making, the question arises whether the mental represents a parallel reality or an aspect of the brain process itself. The experience that enters into consciousness is processed in the frontal lobe. Thus, consciousness seems to be an act of coordination between experience and memory – but the construction of the integral performance of the brain is the self. However, it may be that experience without a human self or rather, without the corresponding subjectivity, does not exist at all and that both are generated in a reciprocal relationship, possibly even in a reciprocal causality.

Alan Fogel researches about the evolutionary genesis of embodied self-awareness, our ability to feel our own movements, perceptions and emotions. This is enabled by neuromotor and neurohormonal pathways between our brains and the rest of our bodies, connections that transmit information about bodily states to enable optimal health and well-being. It is about the psychophysiology of self-awareness and pathways of self-consciousness (Fogel 2013, IX). A bodily embedded sense of self is necessary for survival (Fogel 2013, 2). In the early days of humanity, hunter-gatherers had to have their senses attuned to observe climate and ecological cycles (Fogel 2013, 3). Foetal self-awareness already involves some sense of self in interoception (monitoring the internal milieu of the body in relation to its environment), body scheme as well as in recognising parts of another body (Fogel 2013, 11). This approach is expandable to self-awareness in young and growing humans (Fogel 2013, 12). Self-awareness is fundamentally related and connected to the awareness of others (Fogel 2013, 15). Embodied self-awareness develops through interpersonal relationships during the first three years of life, ultimately grounded in interpersonal encounters with others (Fogel 2013, 29).

Usually, feelings and emotions interlock (Fogel 2013, 39). Interoception is fundamental because almost all people consider this experience, which derives from the vitality of their bodies, to be certain. Their self-

regulation depends on interoception and its neurophysiology. The body is a complex, dynamic system (Fogel 2013, 41). The last decade has seen a rapid growth in neurophysiological research methods, in part due to powerful brain imaging techniques. Prior to this, brain research was mainly conducted on animals (Fogel 2013, 43).

Self-consciousness and self-awareness begin with receptors in spinal cord and brain (Fogel 2013, 46). There are nerve cords that go up and down and regulate each other. They allow us both interoception and exteroception (Fogel 2013, 49). The second level of the formation of self-consciousness is constituted by emotions: the limbic system and the insula (Fogel 2013, 51). At its end are neural integration and spontaneous emergence of interoceptive self-awareness and -consciousness (Fogel 2013, 58). Emotions without self-consciousness exist (Fogel 2013, 66), for instance, within the ability to locate ourselves (Fogel 2013, 71). It manifests in our sense of balance and is important, for example, when dancing (Fogel 2013, 46).

The neurophysiology of bodily self-awareness also includes proprioceptors, nerve pathways that mediate our sensory perception, for example, and integrate into the somatosensory cortex (Fogel 2013, 83). They guarantee, for instance, the organisation of the physical position of individual body parts, classified into subgroups of, say, different muscles (Fogel 2013, 84). Interoception, emotions and body scheme become integrated in self-awareness, leading up to self-consciousness (Fogel 2013, 89). There are neurophysiological connections between self-awareness that is embedded conceptually and bodily. Their material basis manifests itself in the dorsomedial as well as ventromedial nerve cords of the prefrontal cortex (Fogel 2013, 95et seq.), the pathways of fast and of slow thinking as well as of decision-making. There are ergoreceptors and proprioceptors that prevent sensations from entering into embodied self-consciousness. These generate phenomena of suppression and information absorption about the self (Fogel 2013, 102et seq.). Long-lasting suppression of emotions and actions can lead to pathological deviations in behaviour, for example to eating disorders (Fogel 2013, 100-119). Another possible form of reaction is the syndrome of depersonalisation (Fogel 2013, 131). Phenomena of protection, security or threats have an impact on the embodied sense of self (Fogel 2013, 141).

Only gradually do scientists unravel a mystery that has preoccupied philosophers since antiquity (Koch 2019, 46): Not until we understand how our brain generates consciousness can we turn our attention to resolving the

problem at its core – what exactly constitutes and comprises consciousness. What we know about the riddle of conscious experience can be summarised in three theses:

(1) In the posterior areas of the cerebral cortex, researchers are searching for the neuronal correlates of consciousness (NCC).

(2) States of consciousness can be roughly quantified by electroencephalography (EEG).

(3) Two main theories seek to explain the origin of consciousness: one by means of a global neuronal workspace, the other as the sum of information integrated in the system (Koch 2019, 49).

Regarding the origin of consciousness, two main theories exist: First, the global neuronal workspace (GNW), or global workspace theory (GWT), formulated by psychologist Bernard Baars and neuroscientists Stanislas Dehaene and Jean-Pierre Changeaux. It is based on the observation that several brain areas have access to information we perceive consciously. But when you type very fast, you do it automatically. If someone asks you how exactly you manage to do that, you are unable to explain it, because you lack conscious access to this information, which is only available to the cerebral circuits in the sensorimotor system. According to GWT, consciousness arises from a certain kind of information processing. The brain thus contains a sort of information board that can be accessed by various brain processes. Part of the incoming sensory impressions make it onto this platform and are available to other cognitive processes for a short time. They can process data stored there and react to it: formulate a response, retrieve or store a memory, start a movement. Only with the help of a model can we infer which systems are capable of conscious experience.

Integrated information theory (IIT), developed by Giulio Tononi and Christof Koch, chooses a different starting point: that what is experienced. Every lived experience has certain fundamental properties: It is intrinsic, that is, it exists only for the one who experiences it; it follows a chronology in time; and it is specific. It differs from other conscious perceptions, like scenes in a cinema film. Furthermore, all impressions form an inseparable unity. We now find two main challenges ahead of us: First, we need to further elaborate the neuronal traces of consciousness. Better and better tools help us with this, because they allow us to examine the multi-layered interactions of neurons in our brain more closely. Given the immense complexity of our central nervous system, however,

this endeavour will take decades. Secondly, we need to confirm or refute the two competing theories of consciousness with data. It may even be that a better theory will arise from the shambles of the former two, that satisfactorily explains the great enigma of our existence (Koch 2019, 50).

Consciousness and the pyramid concept of information processing need to be distinguished. Recent findings pushed consciousness off the top of the information processing pyramid and moved it further down. It has been found that our most sophisticated cognitive abilities, such as thinking or creativity, are not even directly accessible to conscious experience. Our everyday thoughts are shaped by an astonishing abundance of linguistic imagery. For example, it is an inner voice that speculates, plans, admonishes and comments on events. Only intense physical exertion, acute danger, meditation or deep sleep silence these constant companions. Francis Crick called this sphere the "unconscious homunculus". It is responsible for creativity, intelligence and planning. A considerable part of this happens unconsciously. The transition from the "unconscious homunculus" to consciousness is attention. Visual experience can thus depend crucially on selective attention. However, there is also attention without conscious experience, for instance during learned routine activities (Koch 2020, 34-37).

Conscious experience is an additional fact that goes beyond the scope of today's natural science. We need another way to explain experience, and it should recognise the existence of certain bridging principles, empirical observations that link material processes to the phenomenal world. Against this background IIT formulates: systems need to be separate, structured, have to have their specific way of being, need to be a unity and definite (Koch 2020, 69-72). According to IIT, consciousness is determined by the causal properties of each physical system acting back on itself. The theory reveals or sets out the intrinsic causal powers of any system that obeys all of these five postulates. According to IIT, these causal powers are identical to conscious experience, with every aspect of any possible conscious experience being mapped one-to-one into aspects of this causal structure. All these causal powers can be evaluated systematically, using algorithms (Koch 2020, 77). The postulate of information states that a mechanism only contributes to experience by specifying differences within the system itself. A mechanism specifies information in its present state to the extent that it highlights its cause and its effect within the system. A system generates information in its present state to the extent that it specifies the state of a system that may be a possible cause in the past and its effect in the future.

Information about cause and effect is defined as the overlap of cause information and effect information. In this context, the term "information" is used in a quite different meaning than Claude Shannon's term, which is common in technology and science (Koch 2020, 82et seq.).

Information in the sense of IIT reflects a much older, Aristotelian use of the term: the word derives from the Latin "*informare*", which means to form, shape, depict. Integrated information produces a structure of cause and effect, that is, a form. Integrated information is causal, intrinsic and qualitative: it is used from the system's internal perspective, based on how its mechanisms and its present state shape its past and its future. The way the system determines its own past and future states dictates whether the experience feels like a shade of blue, or like the smell of wet dog (Koch 2020, 83-85). Computationalism is the dominant belief of the information age (Koch 2020, 125et seq.). Artificial intelligence is a descendant of functionalism (Koch 2020, 128). Even if we reject the view that the brain is a computer, there is no doubt that computers have the remarkable ability to simulate the brain. Could this eventually lead to a conscious mind (Koch 2020, 133)? Computers have no experience (Koch 2020, 137), because they have very little intrinsic existence (Koch 2020, 140). The difference between what is real and what is simulated is rooted in the respective causal power. Hence, it does not get wet inside a computer that simulates heavy rain: The software can be functionally identical to certain aspects of reality, but will not have the same causal powers as the real thing. Therefore, the concept of mind uploading is meaningless (Koch 2020, 144et seq.). The phenomenology of experience is determined by its intrinsic causal forces (Koch 2020, 149).

4. Conclusion: Merleau-Ponty's Primacy of Perception, Unconscious Information Processing and Progressive Processing in Cognitive Science

A draft on the nature of perception from Merleau-Ponty, written 1933 (in French: *Projet de travail sur la nature de la perception*), was published amongst other writings in a small volume, containing his earliest conceptions about the ability to perceive (Merleau-Ponty 2012). Merleau-Ponty conceived this project outline in order to obtain a scholarship for his dissertation. In it, he intends to deal with the problem of human perception in general, including the perception of the body in particular. He understands epistemology in terms of perception as an operation of the mind, as the linking of data to an objective world. In doing

so, he draws on the findings of the psychology school of Gestalt theory. According to the epistemology of Gestalt theory, we perceive forms and shapes as such, not as disconnected sensations without context. Nerves do nothing more than enable the transmission of impulses. One has to consider cognition as an activity of the nervous system itself and understand it as a continuous association of visual, tactile, and muscle sense perception. In particular, the perception of one's own body should be analysed, including the illusions of perception that some amputees experience. The results of experimental psychology and neurology on perception are decisive for such an analysis (Merleau-Ponty 2012, 7-9).

The second text in the volume, from 1934 (French: *La nature de la perception*), addresses the nature of perception further. It is the project outline for the dissertation extension request by Merleau-Ponty. In it, he refers to developmental psychology, neurology, Gestalt psychology, physiology and pathology of perception. His programme amounts to an elucidation of the psychology of perception by means of brain physiology. Perception is not a mere fact, but a sensorimotor process with its own pathology of higher sensory perception processes. It is about orientation reflexes, about disturbances, localisation of sensory perceptions and the psychology of the normal (Merleau-Ponty 2012, 10-13). He confronts the empirical aspects with the critique of psychologism: Psychology has to be renewed in its own field.

It is remarkable that his analyses approach the threshold of Gestalt psychology, but do not cross it. Merleau-Ponty's philosophy of perception can be read as a psychological application of Husserl's theory of the intentionality of consciousness. Gestalt is a spontaneous organisation of the sensorimotor field. There are more or less stable organisations. The creation of a form presupposes a structure: namely of form and background. Primitive perception is directed towards the identification of entireties rather than isolated elements. Thus, there is depth perception, field lines and anchor points, with movement being perceived in perceptual space. Merleau-Ponty also explores children's perception, synthetic and analytical perception, perception of wholeness. Knowledge is not gained through adding up individual, separate elements, but through reorganising perception. Objectivity is a characteristic of adulthood (Merleau-Ponty 2012, 14-24).

The third writing in the volume about the primacy of perception and its philosophical consequences, from 1946 (French: *Le Primat de la perception et ses conséquences philosophiques*), consists of a speech to the Academy of

Sciences in Paris and comes to the following conclusion: an unprejudiced investigation of perception shows that it is indeed not objective. The perceived world is only background for cognition. Objects themselves are not grasped, only what of them is within our field of vision. Our own back, i.e., is not part of this. The non-visible sides of an object, for example, are not seen by us, but are imagined and completed by our cognition. Cognition therefore is a synthesis of perception and reason, which takes into account the context of perceiving. It is hence impossible to dissect a perception. It presupposes horizon synthesis and an indeterminate set of perspectival aspects. Merleau-Ponty understands perception to be paradoxical. The world, in addition, is not an object of perception. The field of perception and the behaviour of the other are essential components of perception (Merleau-Ponty 2012, 26-36).

Perception means to realise something through the body (*Leib*) (Merleau-Ponty 2012, 83). In this respect, the accusation of relativism against Merleau-Ponty is not justified (Merleau-Ponty 2012, 56). The starting point of Merleau-Ponty's phenomenology of perception can be seen in the two writings from 1933 and 1934. They deal with the analysis of experience of the body as it perceives. In Gestalt psychology, Merleau-Ponty finds decisive preliminary works for his critique regarding the model of unconscious activity, as well as the dualism of given data and subjective interpretations. Cognitive psychology remains positivist. Gestalt psychology only acquires its significance in overcoming psychological naturalism with the help of Husserl's phenomenology in the work of Merleau-Ponty (Merleau-Ponty 2012, 89-99).

What Merleau-Ponty still rejected in the 1930s has been sought to make comprehensible in non-symbolic AI for about 30 years now, using algorithmic models of information processing through progressive processing within the framework of cognitive science. In his book on dealing with uncertainty, Andy Clark describes how people gain knowledge about the world in order to act in it. He identifies Predictive Processing (PP) as a fundamental characteristic against the methodological background of a connectionist network theory that grounds perception in self-organisation of human knowledge. This model provides a summarising insight into perception, action and imagination of this simulation. It develops new approaches to nature and structure of human experience. And, if one can even speak of a predictive brain, then it shows that our mind is principally not made up of inferring activity, but is rather a praxis-machine, oriented towards action. Strictly speaking, it is a machine perfectly equipped to develop the most parsimonious, action-based

routines possible that lower the demands of neural processing in order to elaborate agile forms of adaptive success. The strong emphasis on prediction has a long tradition in the sciences of mind. To clarify further: the foundation of knowledge is the learning process driven by prediction that develops a new neural framework for the interplay of computational and experimental methods and their results. The latter develop a neural economy with respect to the uncertainty of sensorimotor cognition that connects prediction with signals of error detection, to provide an estimate of what this game of handling uncertainty consists of (Clark 2019, 1et seq.).

In this context, hermeneutic logic, incorporating an epistemology of human corporeality of perception, could mediate (Lipps 1976). Propositional logic was classically identified with judgement, but it means arguments or descriptions of facts, which then, according to hermeneutic pre-understanding, need to be examined as to their validity. The various ways of understanding, comprehending and grasping stem from the sensorimotor ground of the ability to handle (*umgehen können*), not only muscular performance, but also brain processes and the performances of the nervous system. These categories are structures of disclosure, of reveal (*Entbergungsstrukturen*), that emerge. In my interpretation, such basic structures of categorical unveiling (*Entbergen*) are fundamental perspectives of the human ability to handle. In this connection, the achievement of hermeneutic logic as the art of interpretation is the ability to relate fundamental ways of handling (*umgehen können*; Irrgang 2014: Irrgang 2022). A distinction must be made between everyday speech, language and logical formal isolation. As I see it, epistemic preconditions for a hermeneutic logic can be distinguished as follows: (1) sensomotoric handling; (2) emotional handling; (3) instrumental handling and modelling; (4) linguistic-logical handling in the sense of generalisation; and (5) mathematical handling and modelling in the sense of objectification. These categorical forms of handling competence involving sensorimotor cognition and its further developments contain different types of objectification, generalisation and de-individualisation. These are stages of objectification, clarification and testing.

The logic of innovation is based on the logic of inquiry, questioning and doubt, the refusal to accept the paradoxical as the final word of rational argumentation. Hence, the logic of innovation is the opposite of formal logic. However, it seems that reality consists of an interplay of formal and innovative logics. In a logic of the epistemic, what matters is not whether something is true or false, but whether it is right or wrong, whether it succeeds or fails, and

whether it understands or is mistaken. For a long time, formal rational thinking has obstructed a regionalisation of the culturally specific in human beings. Handling ability can be right or wrong, beyond that, success is what determines behaviour – intellectual behaviour as well. Perhaps correctness is the ultimate we can achieve. In hermeneutic logic, then, linguistic logic is applied to sensory experience of a sensorimotor kind in various forms of handling ability (Lipps 1976).

Shannon's theory of information exchange and Andy Clark's PP are thus, from Merleau-Ponty's perspective, not philosophical-epistemological theories of human-level cognition and communication, but theories of machine signal exchange, suited for mathematical respectively cybernetic modelling of the execution of simple procedures. Philosophically, the conceptualisation of perception presupposes a phenomenological-transcendental-philosophical theory of subjectivity of an experiencing, bodily, human subject (*leiblich menschliches Subjekt*), which is derived from all three sources of exteroception, interoception and proprioception. It thus supports the insight that, from a human perspective, we have to infer and reconstruct the unconscious, and thereby can only access the automated forms of information processing from a level of conscious perception, i.e., based on experience. Simplistic technocratic models, like Shannon's and Turing's, are not sufficient as explanatory models, even if they might fool certain subjects. At any rate, I am convinced that a dog would never confuse a Turing machine with its master. Shannon's and Turing's models, as are utilised for the test, are too simple for the complex cognitive world of the bodily embedded human mental, including communication beyond the mere exchange of information (Irrgang 2022).

REFERENCES

Clark, A. (2019) *Surfing Uncertainty. Prediction, Action and the Embodied Mind*. First Edition 2016. Oxford: Oxford University Press.

Fogel, A. (2013) *Body Sense. The Science and Practice of Embodied Self-Awareness*. First Edition 2009. New York/London: W.W. Norton & Company.

Frank, M. (1997) *"Unendliche Annäherung". Die Anfänge der philosophischen Frühromantik*. Frankfurt am Main: Suhrkamp Taschenbuch Wissenschaft.

Frank, M. (2015a) *Selbstgefühl. Eine historisch-systematische Erkundung*. First Edition 2002. Frankfurt am Main: Suhrkamp Taschenbuch Wissenschaft.

Frank, M. (2015b) *Präreflexives Selbstbewusstsein. Vier Vorlesungen*. Stuttgart: Reclam Verlag.

Irrgang, B. (1982) *Skepsis in der Aufklärung. Zur Rekonstruktion der Bedeutung skeptischer Argumentation und ihrer Widerlegung in Versuchen der Rechtfertigung ihres Anspruchs als "Siècle Philosophique"*. Frankfurt am Main: Haag+Herchen Verlag.

Irrgang, B. (1983) "‚Evolution' im 17. und 18. Jahrhundert – Fallstudien zur methodologischen Vorgeschichte von Darwins Theorie." *Conceptus* (42), pp. 3-28.

Irrgang, B.(1990) Metamorphosen eines ‚satirischen Realismus'. Zweifelsexperimente zwischen Spätaufklärung und Frühromantik in Deutschland, in Hans Körner et al. (eds.) *Die Trauben des Zeuxis. Formen künstlerischer Wirklichkeitsaneignung*. Hildesheim: Georg Olms Verlag, pp. 201-233.

Irrgang, B. (2001) *Technische Kultur. Instrumentelles Verstehen und Technisches Handeln (Philosophie der Technik 1)*. Paderborn: Brill/Schöningh.

Irrgang, B. (2007) *Gehirn und leiblicher Geist. Phänomenologisch-hermeneutische Philosophie des Geistes*. Stuttgart: Franz Steiner Verlag.

Irrgang, B. (2009) *Der Leib des Menschen. Grundriss einer phänomenologisch-hermeneutischen Anthropologie*. Stuttgart: Franz Steiner Verlag.

Irrgang, B. (2014) *Handling Technical Power. Philosophy of Technology*. Übersetzung von G. Somasekharan. Stuttgart: Franz Steiner Verlag.

Irrgang, B. (2015) "Skeptisch-kritische Epistemologie, kontextbezogene Selbstorganisation des Verstehens und positional-perspektivische Metahermeneutik", in Michael Funk (ed.) *Transdisziplinär und Interkulturell. Technikphilosophie nach dem Zeitalter philosophischer Kleinstaaterei*. Würzburg: Königshausen und Neumann, pp. 463-532.

Irrgang, B. (2020) *Roboterbewusstsein, automatisiertes Entscheiden und Transhumanismus. Anthropomorphisierungen von KI im Licht evolutionär-phänomenologischer Leib-Anthropologie*. Würzburg: Königshausen und Neumann.

Irrgang, B. (2021) ‚Homo digitalis'. Künstliche Intelligenz und christlicher Glaube. Ein Versuch", in Johannes Brantl (ed.) *Personen gestalten Institutionen – Institutionen prägen Personen. Leben im Spannungsfeld individueller und sozialer Verantwortung*: Regensburg: Verlag Friedrich Pustet, pp. 210-223.

Irrgang, B. (2022) *Vom impliziten technischen Wissen zum technologischen Design von KI und Biotechnologie. Cyberphilosophie des Designkonzeptes der Tüftler und reflektierten Praktiker*. Würzburg: Königshausen und Neumann.

Kahnemann, D. (2012) *Schnelles Denken, Langsames Denken*. First Edition 2011. Hamburg: Siedler Verlag.

Koch, C.(2019) "Kognition. Was ist Bewusstsein?" *Spektrum der Wissenschaft. Gehirn & Geist* (Dec.), pp. 45-50.

Koch, C. (2020) *Bewusstsein. Warum es weit verbreitet ist, aber nicht digitalisiert werden kann*. Translated by Monika Niehaus-Osterloh and Jorunn Wissmann. Berlin: Springer Verlag.

Lipps, H. (1976) *Werke. Band 2. Untersuchungen zu einer hermeneutischen Logik*. First Edition 1938, Fourth Edition 1976. Frankfurt am Main: Vittorio Klostermann.

Merleau-Ponty, M. (2012) *Das Primat der Wahrnehmung*. Edited by Lambert Wiesing, translated by Jürgen Schröder. Frankfurt am Main: Suhrkamp Taschenbuch Wissenschaft.

Merleau-Ponty, M.(2015) *L'institution, la passivité. Notes de cours au Collège de France (1954-1955)*. First Edition 2003. Paris: Edition Belin.

Schöpf, A. (2001) *Unbewusste Kommunikation. Der interne Diskurs des Gewissens und der externe der Gesellschaft*. Wien: Passagen Verlag.