CHAPTER TWELVE

THE HEIDELBERG SCHOOL AND THE LIMITS OF REFLECTION

DAN ZAHAVI
Danish National Research Foundation: Center for Subjectivity Research, University of Copenhagen

Analytical philosophy of mind is currently engaged in a renewed and intensified debate about such issues as subjectivity, phenomenal consciousness, and the nature of selfhood. While it is undeniable that its discussion of these topics has reached a high level of complexity and sophistication, it is however, a discussion that has also remained rather inward looking. Apart from some occasional references to historical figures such as Locke or Kant or James, it has largely been a discussion for and among analytical philosophers. There has been a lack of any real interest in the parallel discussions to be found on the Continent, even though there is a long and rich tradition for discussing and analyzing the very same problems in Austrian, German, and French philosophy.

In this chapter, I wish to present and discuss some of the contributions to a theory of self-awareness that can be found on the Continent. I will focus on the work of a group of German philosophers known as the *Heidelberg School*. However, my point of departure will be contemporary analytical philosophy. I will start out by briefly outlining an account of consciousness that has recently enjoyed great popularity – the so-called higher-order theory. I will then turn to the Heidelberg School, which not only has formulated an incisive criticism of the higher-order theory, but also has developed a position of its own that arguably stands as the most important contribution to a clarification of self-awareness in recent German philosophy. I will analyze this contribution in detail, and then end the chapter by drawing attention to some of its specific limitations.

12.1 THE HIGHER-ORDER THEORY

It is customary to distinguish between two uses of the term "conscious", a transitive and an intransitive use. On the one hand, we can speak of our being conscious of something, be it x, y, or z. On the other, we can speak of

S. Heinämaa, V. Lähteenmäki and P. Remes (eds.), Consciousness: From Perception to Reflection in the History of Philosophy, 267–285.

our being conscious *simpliciter* (rather than non-conscious). For the past two or three decades, a popular way to account for intransitive consciousness in cognitive science and analytical philosophy of mind has been by means of some kind of higher-order theory.[1] According to the higher-order theory, what makes a mental state (intransitively) conscious is the fact that it is taken as an object by a relevant higher-order state. It is the occurrence of the higher-order representation that makes us conscious of the first-order mental state. One way to illustrate the guiding idea is by comparing consciousness to a spotlight. Some mental states are illuminated; others do their work in the dark. Those that are illuminated are intransitively conscious, those that are not, are non-conscious. What makes a mental state conscious (illuminated) is the fact that it is taken as an object by a relevant higher-order state. In short, a conscious state is a state we are conscious of, or as David Rosenthal puts it, "the mental state's being intransitively conscious simply consists in one's being transitively conscious of it".[2] Thus, intransitive consciousness is taken to be a non-intrinsic, relational property,[3] that is, a property that a mental state only has insofar as it stands in the relevant relation to something else. Consciousness has consequently been taken to be a question of the mind directing its intentional aim upon its own states and operations. Self-directedness has been taken to be constitutive of (intransitive) consciousness, or to put it differently, the higher-order theory has typically explained (intransitive) consciousness in terms of self-awareness.[4] As Robert Van Gulick puts it, it is "the addition of the relevant meta-intentional self-awareness that transforms a non-conscious mental state into a conscious one".[5]

One of the clearest articulations of this link between self-awareness and a higher-order account of consciousness can be found in the writings of the

---

[1] See Armstrong, David M., *A Materialist Theory of the Mind* (London: Routledge and Kegan Paul, 1968); Lycan, William G., *Consciousness* (Cambridge, MA: MIT Press, 1987); Carruthers, Peter, *Language, Thoughts and Consciousness: An Essay in Philosophical Psychology* (Cambridge: Cambridge University Press, 1996).

[2] Rosenthal, David M., "A theory of consciousness", in N. Block, O. Flanagan and G. Güzeldere (eds.), *The Nature of Consciousness* (Cambridge, MA: MIT Press, 1997), 739.

[3] Rosenthal 1997, 736–737.

[4] In the following, I will not distinguish terminologically between "self-consciousness" and "self-awareness".

[5] Gulick, Robert van, "Inward and upward: Reflection, introspection, and self-awareness", *Philosophical Topics* 28, 2 (2000), 276.

British philosopher Peter Carruthers. According to Carruthers, the aim of a theory of consciousness is to explain what it is for mental states to be conscious as opposed to non-conscious.[6] This question is different from the question of what it is for an organism or creature to be conscious (i.e., awake) in opposition to non-conscious (i.e., asleep) and it is also a different question from the question of what it is for an organism to be conscious rather than non-conscious of events or objects in the world. Carruthers admits that it might be argued that what he is really after is a theory of self-consciousness rather than simply a theory of consciousness. For in his view, a conscious mental state is one of which the agent is aware, and to that extent it is something that involves self-consciousness. The only reason why he prefers to think of his own theory as a theory of consciousness rather than as a theory of self-consciousness is because the reference to self-consciousness seems to suggest that the subject of a mental state must possess a developed conception of self, a conception of the self as an enduring agent with a determinate past and an open-ended future, in order for the mental state to be conscious. However, because Carruthers finds it highly likely that there are organisms capable of having conscious mental states, but with only the most tenuous conception of themselves as continuing subjects of thought and experience, he regards it as being quite legitimate to maintain a view of creatures which have conscious mental states but which lack self-consciousness.[7] It soon becomes obvious, however, that Carruthers fails to comply with his own admonition. He points out, that the subjective feel of experience presupposes a capacity for higher-order awareness, and he then continues, "such self-awareness is a conceptually necessary condition for an organism to be a subject of phenomenal feelings, or for there to be anything that its experiences are like".[8] To speak of what an experience is like, or of its phenomenal feel, is an attempt to characterize those aspects of experience that are subjective. But the subjective aspects of experience must be aspects that are available to the subject. According to Carruthers, this means that for mental states to be conscious the subject of those states must be capable of discriminating between them; they must be states of which the subject is aware, and this obviously involves a certain amount of self-awareness. In fact, it requires reflective self-awareness.[9] To be more precise, for a creature to be capable of discriminating between its mental states is for

[6] Carruthers 1996, 148.
[7] Carruthers 1996, 149.
[8] Carruthers 1996, 152.
[9] Carruthers 1996, 155, 157.

a creature to be capable of reflecting upon, thinking about, and hence conceptualizing its own mental states. Since mental concepts get their significance from being embedded in a folk-psychological theory of the structure and functioning of the mind, this ultimately means that only creatures in possession of a theory of mind are capable of enjoying conscious experiences.[10] Given this setup, Carruthers draws the obvious conclusion: Creatures who lack a theory of mind – such as most animals, young infants, and autistic patients – will also lack conscious experiences, phenomenal consciousness, and a dimension of subjectivity. In his view, they are blind to the existence of their own mental states; there is in fact nothing it is like for them to feel pain or pleasure.[11] Carruthers concedes that most of us believe that it must be like something to be a young infant or a cat, and that the experiences of these creatures have subjective feels to them, but he considers this common-sense belief to be groundless. In fact, he believes it to be something of a scandal that people's intuitions are given any weight at all in this domain, let alone believed sufficient enough to challenge the higher-order theory of consciousness.[12]

12.2 THE HEIDELBERG SCHOOL

One might share Carruthers' view concerning the close link between consciousness and self-consciousness and still disagree about the nature of the link. In contrast to the higher-order theory, the Heidelberg School explicitly denies that the self-consciousness that is present the moment I consciously experience something is to be understood in terms of some kind of reflection, or introspection, or higher-order monitoring. It does not involve an additional mental state, but is rather to be understood as an intrinsic feature of the primary experience. That is, in contrast to the higher-order account of consciousness that claims that consciousness is an extrinsic property of those mental states that have it, a property bestowed upon them from without by some further states, the Heidelberg School argues that the feature in virtue of which a mental state is conscious is located within the state itself; it is an intrinsic property of those mental states that have it.

According to the higher-order theory, self-awareness is an intentional act; it involves a subject–object relation between two different mental states.

---

[10] Carruthers 1996, 158; Carruthers, Peter, *Phenomenal Consciousness: A Naturalistic Theory* (Cambridge: Cambridge University Press, 2000), 194.

[11] Carruthers 2000, 202–203; Carruthers, Peter, "Natural theories of consciousness", *European Journal of Philosophy* 6, 2 (1998), 216.

[12] Carruthers 2000, 199.

In *An Essay Concerning Human Understanding* (1690) Locke used the term reflection to designate our mind's ability to turn its view inward upon itself, making its own operations the object of its contemplation.[13] It has since been customary to describe such a higher-order account of self-awareness as a reflection theory of self-awareness. But although it might at first sight seem obvious that self-awareness is precisely a question of the mind having itself, rather than anything else, as its object, this approach has also been subjected to severe criticism. In fact, one of the most persistent attacks can be found in the writings of the Heidelberg School.[14] What is wrong with the reflection theory?

The reflection model of self-awareness operates with a duality of moments. In its classical form it is a question of a second-order mental state taking a first-order mental state as its object. Consequently, we have to distinguish the reflecting from the reflected. Of course, the aim of reflection is then to overcome or negate this division or difference and to posit both moments as identical – otherwise we would not have a case of *self-awareness*. However, this strategy is confronted with fundamental problems: How can the identity of the two relata be certified without presupposing that which it is meant to explain, namely, self-awareness; and why should the fact of being the intentional object of a non-conscious second-order mental state confer consciousness or subjectivity on an otherwise non-conscious first-order mental state?

The reflection theory claims that self-awareness is the result of a reflection, i.e., in order to manifest itself phenomenally (and not merely remain non-conscious) a feeling, perception, or thought must await its objectification by a subsequent reflection. However, it is not enough for the reflection theory to explain how a certain state becomes conscious. The theory also has to explain how the state comes to be given as my state, as a state that I am in. Why? Because when one is directly and non-inferentially conscious of an occurrent pain, perception, or thought, the experience in question is characterized by a first-personal givenness that immediately reveals it as being one's own. In this sense, the first-personal givenness of the experience can be said to entail a built-in self-reference, a primitive experiential self-referentiality, which is

[13] Locke, John, *An Essay Concerning Human Understanding* (Oxford: Clarendon Press, 1975), 107, 127.

[14] The first use of the term Heidelberg School is found in Tugendhat, Ernst, *Selbstbewußtsein und Selbstbestimmung* (Frankfurt am Main: Suhrkamp, 1979), 10, 53. Cf. Frank, Manfred, *Die Unhintergehbarkeit von Individualität* (Frankfurt am Main: Suhrkamp, 1986), 35. The name reflects the crucial influence of Dieter Henrich's seminars in Heidelberg.

exactly what the reflection theory has to account for. But in order for the experience to appear as my experience, it is not sufficient that the experience in question (A) is grasped by a reflection (B). If A is to be given as mine, it is not enough that B is de facto about A. B must recognize itself in A. That is, the first-order state must be grasped as being identical with the second-order state (and since a numerical identity is excluded, the identity in question must be that of belonging to the same subject or being part of the same stream of consciousness). This poses a difficulty, however, for what should enable the act of reflection (which according to this model is itself non-conscious) to realize that the first-order state belongs to the same subjectivity as itself? In order to identify something as oneself one has to hold something true of it that one already knows to be true of oneself. Just as I cannot recognize something as mine unless I am already aware of myself, a non-conscious second-order mental state (that per definition lacks consciousness of itself) cannot recognize or identify a first-order mental state as belonging to the same mind as itself. If the second-order state is to encounter something as itself, if it is to recognize or identify something as itself, it needs a prior acquaintance with itself.[15] As Konrad Cramer puts it,

> How should the reflective subject be able to know that it has itself as an object? Obviously only so that the self knows that it is identical with its object. But it is impossible to ascribe this knowledge to *reflection* and to *ground* it in reflection. Because the act of reflection presupposes that the self *already knows* itself, in order to know that the one which it knows when it has itself as an object is identical with the one which accomplishes the act in the reflective thinking (back to itself). The *theory* that tries to make the *origin* of self-awareness understandable through reflection, ends necessarily in a circle which needs to presuppose the knowledge that it wants to explain.[16]

---

[15] Cramer, Konrad, "'Erlebnis': Thesen zu Hegels Theorie des Selbstbewußtseins mit Rücksicht auf die Aporien eines Grundbegriffs nachhegelscher Philosophie", in H. G. Gadamer (ed.), *Stuttgarter Hegel-Tage 1970*, Beiheft 11, Bonn: Hegel-Studien (1974), 563.

[16] "Wie aber soll das Subjekt in der Reflexion wissen können, daß es sich selbst als Objekt hat? Offenbar einzig dadurch, daß das Ich sich mit dem, was es als Objekt hat, identisch weiß. Nun ist es aber unmöglich, dies Wissen der *Reflexion* zuzuschreiben und aus ihr zu *begründen*. Denn für den Akt der Reflexion ist vorausgesetzt, daß das Ich sich *schon kennt*, um zu wissen, daß dasjenige, was es kennt, wenn es sich selbst zum Objekt hat, mit dem identisch ist, was den Akt reflektiver Rückwendung auf sich vollzieht. Die *Theorie*, die den *Ursprung* von Selbstbewußtsein aus der Reflexion verständlich machen will, endet daher notwendig in dem Zirkel, diejenige Kenntnis schon voraussetzen zu müssen, die sie erklären will." (Cramer 1974, 563.)

Any convincing theory of consciousness has to account for the first-personal or egocentric givenness of our conscious states, and has to respect the difference between our consciousness of a foreign object, and our consciousness of our own subjectivity. However, this is exactly what the reflection theory fails to do. Thus, it is highly questionable whether one can account for the first-personal givenness of phenomenal consciousness by sticking to a traditional model of object-consciousness and simply replacing the external object with an internal one. Self-awareness does not come about as the result of a procedure of introspective object-identification. I do not first scrutinize a specific pain and subsequently identify it as being mine, since that kind of criterial identification implies the possibility of misidentification, and self-awareness is not prone to that kind of error. In fact, when one is aware of one's occurrent thoughts, feelings, beliefs, and desires, one does not seem to be confronted with objects of any sort, and this is exactly what the reflection theory overlooks.

When something is given as an object, it is given as something that transcends the merely subjective. For something to be given as an object of experience is for it to differ from the subjective experience itself. But if this is so, if object-awareness always involves a kind of epistemic divide, object-awareness cannot help us understand self-awareness. After all, self-awareness is supposed to consist in some kind of acquaintance with experiential subjectivity; it is not supposed merely to acquaint us with yet another object of experience. It could, perhaps, be objected that there surely are cases where I am confronted with a certain object, and then recognize that the object in question is in fact myself. This is true of course, but this kind of objectified self-recognition can never constitute the most fundamental form of self-awareness. In order for me to recognize a certain object as myself, I need to hold something true of it that I already know to be true of myself, and the only way to avoid an infinite regress is by accepting the existence of a non-objectifying self-acquaintance. In analytical philosophy of mind, a similar line of thought can be found in Sidney Shoemaker:

> The reason one is not presented to oneself "as an object" in self-awareness is that self-awareness is not perceptual awareness, i.e., is not a sort of awareness in which objects are presented. It is awareness of facts unmediated by awareness of objects. But it is worth noting that if one were aware of oneself as an object in such cases (as one is in fact aware of oneself as an object when one sees oneself in a mirror), this would not help to explain one's self-knowledge. For awareness that the presented object was φ, would not tell one that one was oneself φ, unless one had identified the object as oneself; and one could not do this unless one already had some self-knowledge, namely the knowledge that one is the unique possessor of whatever set of properties of the presented object one took to show it to be oneself.

> Perceptual self-knowledge presupposes non-perceptual self-knowledge, so not all self-knowledge can be perceptual.[17]

This reasoning holds true even for self-knowledge obtained through introspection. That is, it will not do to claim that introspection is distinguished by the fact that its object has a property, which immediately identifies it as being me, since no other self could possibly have it, namely the property of being the private and exclusive object of exactly my introspection. This explanation will not do, since I will be unable to identify an introspected self as myself by the fact that it is introspectively observed by me, unless I know it is the object of *my* introspection, i.e., unless I know that it is in fact *me* that undertakes this introspection. This knowledge cannot itself be based on identification if one is to avoid an infinite regress.[18]

So the basic claim being made is that self-awareness cannot come about as the result of the encounter between two non-conscious experiences. Consequently, the reflection must either await a further reflection in order to become self-aware, in which case we are confronted with a vicious infinite regress, *or* it must be admitted that the reflection is itself already in a state of self-awareness, and that would of course involve us in a circular explanation, presupposing that which was meant to be explained, and implicitly rejecting the thesis of the reflection model of self-awareness, namely, that *all* self-awareness is brought about by reflection.[19]

So far, the Heidelberg School's contribution to a clarification of self-awareness has mainly consisted in its criticism of the reflection theory. If it could offer nothing more than these negative observations, however, it would hardly qualify as an alternative theory of self-awareness. Dieter Henrich readily acknowledges that it is essential to transcend a mere disclosure of what he believes to be the aporetical implications of the reflection theory and offer a more substantial account. However, as he points out one has to realize that the difficulty in interpreting the familiar phenomenon

---

[17] Shoemaker, Sidney and Swinburne, Richard, *Personal Identity* (Oxford: Blackwell, 1984), 105.

[18] Shoemaker, Sidney, "Self-reference and self-awareness", *The Journal of Philosophy* 65 (1968), 561–563.

[19] See Henrich, Dieter, "Selbstbewußtsein, kritische Einleitung in eine Theorie", in R. Bubner, K. Cramer, and R. Wiehl (eds.), *Hermeneutik und Dialektik* (Tübingen: Mohr, 1970), 268; Henrich, Dieter, *Selbstverhältnisse* (Stuttgart: Reclam, 1982b), 64; Frank, Manfred, "Fragmente einer Geschichte der Selbstbewußtseins-Theorie von Kant bis Sartre" (1991b), in M. Frank (ed.), *Selbstbewußtseinstheorien von Fichte bis Sartre* (Frankfurt am Main: Suhrkamp, 1991), 498, 529.

"consciousness" by direct description is so extreme that it is practically impossible to overcome.[20] The difficulty Henrich has in mind has to do with the difference between being self-aware and explaining self-awareness. Whereas the self-givenness of lived consciousness is characterized by immediacy, this is certainly not the case with our philosophical understanding of it. In order to examine (reflect upon) the structure of self-awareness we have to direct our attention to it, and since this inevitably implies its objectification, the original subjective dimension will evade our theoretical gaze and remain inaccessible for direct description and investigation.[21] This does not imply that its existence is merely postulated, since we are after all not only acquainted directly (and non-theoretically) with the original state of being conscious, e.g., we all know the difference between wakefulness and sleep, but we are also in a position to ascertain that we are self-aware through reflection. By analyzing reflection we can regressively infer that it has a more primitive form of self-acquaintance as its condition of possibility. Nevertheless, a direct examination of this dimension seems impossible, and the following four features that constitute the core of Dieter Henrich's own theory of self-awareness have consequently been disclosed indirectly, *ex negativo*, through a criticism of the reflection theory:[22]

1. Consciousness is a dimension that contains knowledge of itself, for there is no consciousness of anything that is not implicitly acquainted with itself. "Implicitly" is here not used in the sense of being a mere potential acquaintance, but in the sense of existing even prior to reflection and explicit thematization.
2. Original self-awareness is not a performance, but an irrelational occurrence (*Ereignis*). That is, self-awareness is not only irrelational, it is also something that is given rather than voluntarily brought about.
3. Self-aware consciousness is an *egoless* dimension within which intentional experiences and mental states take place.
4. It is a private or exclusive dimension, in the sense that each consciousness has special access to itself.

Let me add a few clarifying comments. Henrich denies that *original* self-awareness should be understood either as a relation between two mental states or as a relation between the mental state and itself. The general point seems to be that one should avoid theories describing self-awareness as a kind of *relation*, since every relation – especially the subject–object relation – entails

[20] Henrich 1970, 274.
[21] Henrich, Dieter, *Fluchtlinien* (Frankfurt am Main: Suhrkamp, 1982a), 152.
[22] Henrich 1970, 275, 277, 280, 284.

a *distinction* between two (or more) relata, and this is exactly what generates all the problems. Thus, Henrich argues that self-awareness cannot come about as the result of a criterial self-identification, nor is it a kind of reflection, introspection, object-intentionality, or conceptually mediated propositional attitude – all of which entail the distinction between two or more relata. The basic self-awareness of an experience is not mediated by foreign elements such as concepts or classificatory criteria, nor by any internal difference or distance, but must be treated as an intrinsic quality of experience that is completely *irrelational*.[23]

The criticism directed at the reflection theory has generally not been meant to imply, however, that reflective self-awareness and objectifying self-thematization is impossible, but merely that it always presupposes a prior unthematic and pre-reflective self-awareness as its condition of possibility. Thus, it is necessary to differentiate *pre-reflective* self-awareness, which is an immediate, implicit, irrelational, non-objectifying, non-conceptual, and non-propositional self-acquaintance from *reflective* self-awareness, which is an explicit, relational, mediated, conceptual, and objectifying thematization of consciousness.

> Reflections can mediately connect to immediate awareness and elevate it to the status of knowledge. The original givenness is however the awareness itself which obviously appears as a single unit and not as an object pole of a conscious subject which directs itself toward it.[24]

Henrich's third feature also calls for a clarification. The question whether it makes sense to speak of a subjectless or egoless self-awareness, i.e., of self-awareness without anybody being self-aware, ultimately depends upon whether one opts for an *egological* or a *non-egological* theory of consciousness. An egological theory would claim that whenever I taste a single malt whiskey then I am not only intentionally directed at the *whiskey*, nor merely aware of the whiskey being *tasted*, but also aware that it is being tasted by *me*, i.e., that *I* am *tasting* a *whiskey*. Thus, the egological theory would claim

---

[23] Henrich 1970, 266, 273; Henrich 1982a, 142. Cf. Pothast, Ulrich, *Über einige Fragen der Selbstbeziehung* (Frankfurt am Main: Vittorio Klostermann, 1971), 76–78; Frank 1986, 34, 61; Frank, Manfred, *Selbstbewußtsein und Selbsterkenntnis* (Stuttgart: Reclam, 1991a), 71, 405; Frank 1991b, 597.

[24] "Reflexionen können mittelbar an unmittelbares Bewußtsein anschließen und es in den Status eines Wissens erheben. Die ursprüngliche Gegebenheit ist aber das Bewußtsein selbst, welches offensichtlich einstellig auftritt und nicht als Objektpol eines sich darauf richtenden Bewußtseinssubjektes." (Frank 1991b, 438. Cf. Frank 1991a, 7, 161.)

that it is a conceptual and experiential truth that any episode of experiencing necessarily includes a subject of experience.[25] This account, which identifies self-awareness with *I-consciousness*, is however regarded by the Heidelberg School as having fallen victim to the language of reflection – *the use of "I" seems exactly to articulate a self-reflection* – and is rejected for the following reasons: whereas reflection is described as the accomplishment of an active principle, as something that is initiated by a subject, pre-reflective self-awareness must precede all performances, and can consequently not be attributed to the ego, but must be characterized as a subjectless or egoless awareness.[26] Moreover, an egological theory claiming that self-awareness is properly speaking an original awareness of my*self*, as a self, subject, or ego seems in an eminent way to take self-awareness as a kind of object-awareness, and thus to be prone to all the problems confronting this approach.[27] Finally, if one conceives of the ego qua subject of experience as that which *has* the experience, one obviously makes a distinction between the ego and the experience. They are not identical. In this case, however, it is difficult to understand why the ego's awareness of the experience should be classified as a case of *self*-awareness.[28]

This criticism does not imply, however, that the ego is a superfluous and dispensable notion. Henrich argues that it is impossible to understand phenomena such as making a decision, solving a problem, expecting an event, or initiating a reflection, without assuming the existence of an active principle of organization in the field of consciousness, i.e., without accepting the existence of an ego or a self. But this egological structure is not a fundamental feature of consciousness; rather, it is merely a mode of its organization. Originally, consciousness is egoless and anonymous.[29]

Is the position of the Heidelberg School convincing? Somewhat surprisingly, both Dieter Henrich and Manfred Frank have later expressed reservations about the adequacy of their central claim, namely, that original self-awareness is strictly *irrelational*. Both explicitly acknowledge that the phenomenon of self-awareness has an internal structural *complexity* that manifests itself in a plurality of ways. More specifically, they have started speaking of *three* moments that together make up the unity of self-awareness: the anonymous dimension of subjectivity, the epistemic self-acquaintance, and the (founded) egological organization. All of these features have to

[25] Cf. Shoemaker 1968, 563–564.
[26] Henrich 1970, 276.
[27] Cramer 1974, 573; Frank 1991a, 252.
[28] Pothast 1971, 64.
[29] Henrich 1970, 276, 279.

co-exist in a structural unity, and this seems to contradict the claim that pre-reflective self-awareness per se lacks internal differentiation and structural complexity.[30]

In short, it seems as if it is too hasty to ban every kind of internal differentiation and structure from pre-reflective self-awareness. This is not to say that the arguments presented against the reflection theory and against the attempt to understand self-awareness as a kind of relation have suddenly lost their validity – one must still display utmost caution not to become vulnerable once more to that criticism. But as Frank suggests, it is possible to escape the previously outlined difficulties if one conceives of the moments as conceptually differentiable, but factually inseparable.[31]

Thus, when all is said and done, self-awareness is primitive in the sense of being irreducible, but it is neither simple nor unstructured. We are ultimately dealing with a unitary phenomenon composed of connected elements that can neither be subsumed under nor deduced from a higher principle. Frank speaks of a unity of identity and difference, in the sense that each of the elements is irreducible, but nevertheless unable to exist in separation from the others.[32]

At this point, however, the clarification and analysis terminate. According to Henrich, we do not possess an *adequate* understanding of the connection between the different elements of self-awareness. Why the elements are inseparable, and how they manage to constitute the unity of self-awareness, are questions that cannot be answered:

> So it is necessary to accept both that self-awareness in itself is complex and that we cannot unravel this complex or understand it in its inner constitution.[33]

In the end, it is consequently claimed that the unitary phenomenon of self-awareness resists comprehension.[34] This conclusion is hardly satisfying. Although Frank admits that it conceals rather than solves the problem – if the different moments are not only to be different, but in fact moments of

---

[30] Henrich 1970, 280; Henrich 1982a, 145–146. Cf. Frank, Manfred, *Zeitbewußtsein* (Pfullingen: Neske, 1990), 113; Frank 1991a, 16–17. These reflections are developed by Henrich in an unpublished manuscript, which Frank summarizes in Frank 1991b, 590–599.

[31] Frank 1991a, 10, 83; Frank 1991b, 589, 591.

[32] Frank 1991b, 595.

[33] "So ist also beides anzunehmen notwendig: Daß Selbstbewußtsein in sich komplex ist und daß der Komplex von uns nicht aufgelöst oder in seiner inneren Konstitution verstanden werden kann." (Henrich 1982a, 150. Cf. Henrich 1982a, 152, 157; Henrich 1982b, 102.)

[34] Henrich 1982a, 169. Cf. Henrich 1982a, 155, 162–163.

one phenomenon, it is essential to explain and clarify their connection and interaction – he is ultimately unable to contribute with a more satisfying solution himself, and he as well must in the end admit that the way in which the elements of self-awareness are united remains obscure.[35]

12.3 TUGENDHAT'S CRITICISM

The Heidelberg School's contribution to a clarification of self-awareness has not been met with approval by everybody. One prominent critic is Ernst Tugendhat, who has argued that the Heidelberg School represents the culmination and termination of the traditional discussion of self-awareness. After having pointed to the aporias of previous theories of self-awareness, it fails to provide a less aporetical account itself. Thus, Tugendhat claims that Henrich, in particular, has unwittingly led the traditional concept of self-awareness ad absurdum, and that it is consequently necessary to undertake a fundamental revision of the notion of consciousness which the entire classical tradition has uncritically made use of.

Tugendhat's own alternative is based upon more general language-philosophical reflections. According to Tugendhat, one cannot know or be conscious of an object, one can only be intentionally related to states of affairs. I do not know a table; I know that a table has such and such properties. Self-awareness should be interpreted in a similar way:

> I suggested that we should first make the general structure of consciousness of something clear; on this basis we were to acquire a concept of what consciousness of oneself means by replacing the variable 'something' accordingly.[36]

Thus, self-awareness is taken to be a kind of knowledge. It is not knowledge about an (internal) object, about a self or an experience; rather it is propositional knowledge expressed in the form "I know, that I φ", where φ stands for a mental or psychic state.[37] In contrast to Henrich and Frank, Tugendhat consequently takes immediate self-awareness to be an epistemic relation between an empirical person and a proposition. Self-awareness is a *propositional attitude*.[38]

It is against this background, that Tugendhat claims that the problem discussed by the Heidelberg School is a pseudo-problem. In the phrase

---

[35] Frank 1990, 125, 135; Frank 1991b, 599; Cramer 1974, 591.

[36] Tugendhat 1979, 21, in English Tugendhat, Ernst, *Self-Consciousness and Self-Determination* (Cambridge, MA: MIT Press, 1986), 12–13.

[37] Tugendhat 1979, 22.

[38] Tugendhat 1979, 10–11, 45, 50–51, 54, 57, 66–67.

"I know, that I φ" the word "I" appears twice, and one could then wonder how we know that both uses refer to the same subject. How do we account for the identity between the one who knows and the one who is in the mental state? It is true that I cannot be aware of, for example, that I am in pain or that I am seeing a canary, and be mistaken about who the subject of that experience is. The fact that first-person experience ascriptions are not subject to the error of misidentification is not in need of any further explanation and is in particularly not due to some mysterious self-transparency or self-acquaintance, since no infallible identification or informative reference has taken place. The identity in question is of the purely tautological sort. Thus, that my awareness of an experience does not leave it open whose experience it is, is just as unproblematically true as that A = A or I = I.[39]

Tugendhat attempts to transform the problem of self-awareness into a semantic problem. But rather than clarifying and solving the problem, this transformation merely covers it up. Despite his criticism of the traditional subject–object model, Tugendhat remains convinced that self-awareness is to be understood as a relation between two different entities, a person and a proposition. But he never explains why such a relation should establish *self*-awareness. Nor does he seem to realize that the principal task facing a clarification of immediate epistemic self-awareness is to account for the unique first-personal givenness of our experiences rather than to explain the identity between the knower and the known. Moreover, given that Tugendhat claims that self-awareness is a propositional attitude, he is confronted with an obvious question. Does self-awareness presuppose language-use? Is a person only in possession of self-awareness when it has acquired a sufficient mastery of language to be able to refer to itself with "I"? If it does, are we then to deny self-awareness to children and animals? Tugendhat's reply is remarkably vague. He says that it remains unclear whether we can refer to propositions non-linguistically, but suggests that self-awareness only becomes conscious when it is linguistically articulated.[40] However, not only is it rather unclear what a non-conscious self-awareness should amount to, furthermore many developmental psychologists currently argue that infants are in possession of various forms of pre-linguistic self-experience already from birth.[41]

---

[39] Tugendhat 1979, 55–61, 68–70, 83.

[40] Tugendhat 1979, 21, 26.

[41] Cf. Zahavi, Dan, "The embodied self-awareness of the infant: A challenge to the theory-theory of mind?", in D. Zahavi, T. Grünbaum and J. Parnas (eds.), *The Structure and Development of Self-Consciousness: Interdisciplinary Perspectives* (Amsterdam: John Benjamins, 2004), 35–63.

12.4 CONCLUSION

In my view, Tugendhat's criticism of the Heidelberg School is misplaced. But that does not mean that the account offered by the Heidelberg School escapes criticism. I find its account significant and illuminating because of its focus upon the aporetical character of the reflection theory of self-awareness, and because of its systematic and instructive analysis of how not to conceive of self-awareness. Despite its insights, however, it basically remains a critical introduction.[42] Although both Henrich and Frank acknowledge that pre-reflective, irrelational self-awareness is characterized by a certain internal differentiation and complexity; they never offer a more detailed analysis of this complex structure. That is, when it comes to a positive description of the structure of original pre-reflective self-awareness they are remarkably silent, either claiming in turn that it is unanalyzable, or that the unity of its complex structure is incomprehensible. This is hardly satisfactory, and in addition, the account offered by the Heidelberg School is also overly negative and formalistic. Moreover, the problem of self-awareness has numerous essential facets which the Heidelberg account either remains silent about, or only analyzes inadequately. Let me conclude by specifying some problems that I believe a convincing theory of self-awareness would have to tackle, but which the Heidelberg School has failed to take into sufficient consideration.[43]

1. The methodological problem

To what extent is it at all possible to investigate subjectivity? If subjectivity rather than being an object that we can encounter in the world is the very perspective that permits any such encounters, to what extent can it then be made accessible for direct examination? Will any examination necessarily take the subject of experience as an object of experience, and thereby transform and distort it? In other words, can subjectivity actually be grasped and described, or is it only approachable ex negativo?

2. The problem of reflection

Although the Heidelberg School has offered a criticism of the reflection theory, it has in fact said rather little about reflection itself. Moreover, even if one concedes that reflective self-awareness rather than being the most basic type of self-awareness, is in fact a more complex form, this still leaves it open how exactly pre-reflective self-awareness is supposed to give rise to reflective self-awareness. This is in particular a problem if one, as it is customary in some of the discussions on the Continent, takes reflection to entail some kind of internal

---

[42] Cf. Henrich 1970.

[43] The list is not meant to be exhaustive.

self-division or self-detachment. For how is pre-reflective self-awareness, which is supposedly simple and irrelational, to give rise to such a fracture? Thus, it will not do to conceive of pre-reflective self-awareness in such a manner that the transition to reflective self-awareness becomes incomprehensible.

3. The problem of temporality

Any convincing theory of self-awareness should not only be able to account for the pre-reflective self-awareness of a single experience, but also explain how I can have self-awareness across temporal distance, that is, it should be able to explain why I can remember a past experience as mine. Thus, the temporality of consciousness has to be accounted for, and in more detail than the Heidelberg School has done. Given the temporal character of the stream of consciousness, even something as apparently synchronic as the conscious givenness of a present experience might not be comprehensible without taking temporality (or as Edmund Husserl would call it: inner time-consciousness) into consideration.

4. The problem of the self

The question concerning the egological or non-egological character of self-awareness also has to be clarified. Does self-awareness necessarily have an egocentric structure, or is self-awareness rather the anonymous acquaintance of consciousness with itself? Since an answer to this question can only be given after it has been established what exactly a self is, this must also be done, and ultimately it will prove necessary to determine the relation between a single experience, the stream of consciousness, and the self. However, the analysis of the self or ego offered by the Heidelberg School is clearly inadequate. The validity of their rejection of an egological theory of consciousness is tied to their very narrow definition of the ego. It is understood either as a principle of activity or as something that must necessarily be conceived as standing opposed to consciousness "having" it.[44] But there are certainly other ways to conceive of the self.

5. The problem of the body

The difference between a first-person and a third-person perspective does not coincide with the traditional difference between mind and body. As an analysis of proprioception reveals, the body itself can appear from a first-person perspective, and an investigation of the different types of bodily self-experience must be integrated into a general analysis of self-awareness. Moreover, this investigation of the body is indispensable if one is eventually to understand how one can appear to oneself as a worldly object, that is, if one is to understand the relation between one's awareness of oneself as

[44] Henrich 1970, 279; Pothast 1971, 66.

an elusive subjective dimension, and one's awareness of oneself as an intersubjectively accessible entity in the world. Thus, a convincing theory of self-awareness cannot allow itself to ignore the body.

6. The problem of intersubjectivity

Not only can I be aware of my own subjectivity, I can also be aware of other subjects, and an analysis of self-awareness must also deal with the problem of intersubjectivity. It must do so not because every type of self-awareness is intersubjectively mediated, nor because the analysis must necessarily account for those types of self-awareness that are in fact intersubjectively constituted, but because a theory of self-awareness must avoid conceiving of self-awareness in such a fashion that intersubjectivity becomes impossible. That is, it will not do to conceive of self-awareness in such private and exclusive terms that it becomes incomprehensible how I should ever be able to recognize another embodied subjectivity. To quote Maurice Merleau-Ponty: "If the sole experience of the subject is the one which I gain by coinciding with it, if the mind, by definition, eludes 'the outside spectator' and can be recognized only from within, my cogito is necessarily unique, and cannot be 'shared in' by another. Perhaps we can say that it is 'transferable' to others. But then how could such a transfer ever be brought about? What spectacle can ever validly induce me to posit outside myself that mode of existence the whole significance of which demands that it be grasped from within? Unless I learn within myself to recognize the junction of the for itself and the in itself, none of those mechanisms called other bodies will ever be able to come to life; unless I have an exterior others have no interior. The plurality of consciousness is impossible if I have an absolute consciousness of myself".[45]

7. The problem of intentionality

Pre-reflective self-awareness might not be a type of object-consciousness, but this does not entail that an analysis of self-awareness can dispense with the problem of intentionality. As Erwin Straus once put it: "In sensory experience I always experience myself and the world at the same time, not myself directly and the Other by inference, not myself before the Other, not myself without the Other, nor the Other without myself".[46] Henrich has himself acknowledged that consciousness is simultaneously

---

[45] Merleau-Ponty, Maurice, *Phenomenology of Perception* (London: Routledge & Kegan Paul, 1962), 373.

[46] Straus, Erwin, "Aesthesiology and hallucinations", in R. May et al. (eds.), *Existence: A New Dimension in Psychiatry and Psychology* (New York: Basic Books, 1958), 148.

and co-originally aware of itself and related to the world.[47] But this connection has to be explored in greater detail than done by Henrich.

8. The problem of language
Even if one can reject the claim that self-awareness is a linguistic phenomenon, it is hardly possible to deny that language can transform our self-acquaintance and make possible new and far more complex forms of self-consciousness. A better understanding of how this is possible is a clear desideratum.

9. The problem of the unconscious
A theory of self-awareness will eventually have to confront the problem of the unconscious. The basic question is whether all of our experiences are characterized by a primitive self-awareness and whether the notion of an unconscious consciousness is therefore a contradiction in terms, or whether it is actually possible to reconcile a thesis concerning a primitive but pervasive self-awareness with a recognition of the unconscious?

It is important not to misunderstand this criticism. I am not claiming that a theory of self-awareness, in order to be convincing, must necessarily account for intentionality, intersubjectivity, temporality, etc., as well. Although a full and comprehensive theory of consciousness would have to tackle all of these issues, it is certainly possible and legitimate to focus on and isolate certain specific topics, including the nature of self-awareness. The point I wish to make is simply that the account offered by the Heidelberg School is problematic because it focuses on self-awareness in abstracto rather than accounting for the self-awareness of the temporal, intentional, reflexive, corporeal, and intersubjective experiences.

The nine problems outlined concern aspects of self-awareness in need of further elaboration and clarification. Part of this clarification can be found in another philosophical tradition from the Continent, that I haven't mentioned so far, namely phenomenology. This is a tradition that the Heidelberg School has regarded with considerable reservation.[48] But as any in-depth

[47] Henrich 1982a, 149.

[48] Although Frank occasionally speaks favourably of Sartre, he and Henrich ultimately argue that Sartre's account is vulnerable to their own arguments against the reflection theory (cf. Henrich 1970, 261; Frank 1990, 83). They criticize Husserl and Heidegger along similar lines (Henrich, Dieter, "Fichtes ursprüngliche Einsicht", in D. Henrich and H. Wagner (eds.), *Subjektivität und Metaphysik* (Frankfurt am Main: Klostermann, 1966), 231; Frank 1986, 43–45; Frank 1990, 53–57; Frank 1991b, 530–531, 536; Cramer 1974, 583–584, 592–593). For a critical appraisal of this criticism, cf. Zahavi, Dan, *Self-Awareness and Alterity* (Evanston: Northwestern University Press, 1999); Zahavi, Dan, "How to investigate subjectivity: Heidegger and Natorp on reflection" (2003b), *Continental Philosophy Review* 36, 2 (2003), 155–176.

study of the writings of Husserl, Scheler, Heidegger, Sartre, Merleau-Ponty, etc. will disclose, their analyses of pre-reflective self-awareness are integrated into and can be found in the context of an examination of a number of related issues, such as the nature of intentionality, spatiality, embodiment, selfhood, temporality, attention, sociality, etc. This is one of the reasons why the phenomenological analyses of self-awareness can easily complement the incisive but rather formal analyses offered by the Heidelberg School. A more extensive presentation of the phenomenological take on self-awareness, however, would exceed the limits of this chapter.[49,50]

[49] See, however, Zahavi, Dan, "The fracture in self-awareness" (1998a), in D. Zahavi (ed.), *Self-Awareness, Temporality, and Alterity* (Dordrecht: Kluwer Academic Publishers, 1998), 21–40; Zahavi, Dan, "Self-awareness and affection" (1998b), in N. Depraz and D. Zahavi (eds.), *Alterity and Facticity* (Dordrecht: Kluwer Academic Publishers, 1998), 205–228; Zahavi 1999; Zahavi, Dan, "First-person thoughts and embodied self-awareness", *Phenomenology and the Cognitive Sciences* 1, 1 (2002), 7–26; Zahavi, Dan, "Internal time-consciousness and pre-reflective self-awareness" (2003a), in D. Welton (ed.), *The New Husserl: A Critical Reader* (Bloomington: Indiana University Press, 2003), 157–180; Zahavi, Dan, *Subjectivity and Selfhood: Investigating the First-Person Perspective* (Cambridge, MA: The MIT Press, 2005).

[50] This study has been funded by the Danish National Research Foundation.