

Georg Northoff

„Am I My Brain?“
Personal Identity and Brain Identity –
A Combined Philosophical and Psychological
Investigation in Brain Implants

Abstract

Though the issue of personal identity has been discussed extensively in philosophy, the exact relationship between personal identity and the brain remains unclear. Some philosophers like T. Nagel regard the brain as a necessary and sufficient criterion for personal identity („I am my brain“). In contrast, other philosophers like D. Parfit entirely dissociate personal identity from the brain („I am not my brain“). The purpose of the present paper is to investigate both Nagel's and Parfit's assumption from a neurophilosophical perspective. In a first step, the philosophical theories of Nagel and Parfit are presented. Their criteria served as a guideline for a psychological investigation of personal identity and personality in patients who received brain implants. The methods and results of this study are reported in a second step. Finally, in a third step psychological results and philosophical theories are compared with each other and neurophilosophical questions about the relationship between personal identity and brain identity are raised. The comparison between psychological data and philosophical theories reveals the following. First, Nagel's characterization of the brain as an „additional fact“ and by „physico-mental intimacy“ is not compatible with the subjective experience of the brain implant and the own person in these patients. Second, Parfit's neglect of mental states and the first-person perspective remains empirically implausible and raises the question for the epistemic abilities and inabilities of our own brain thus requiring „neuroepistemology“. Third, Parfit's ontological definition of the brain in a purely mechanical-technical sense, as presupposed in his physical spectrum argument, remains incompatible with empirical results of brain function during implantation. This raises the issue of the ontological definition of the brain thus requiring „neuroontology“. It is concluded that „neuroepistemology“ and „neuroontology“ might contribute to an empirically more plausible philosophical definition of the brain and conceptual clarification of the relationship between personal identity and brain identity.

Zusammenfassung

Obwohl das Problem der personalen Identität extensiv in der Philosophie diskutiert wurde, bleibt die genaue Beziehung zwischen personaler Identität und Gehirn unklar. Einige Philosophen wie zum Beispiel T. Nagel betrachten das Gehirn als notwendige und hinreichende Bedingung der personalen Identität („I am my brain“). Im Unterschied dissoziieren andere, wie zum Beispiel D. Parfit, personale Identität vollständig vom Gehirn („I am not my brain“). Das Ziel des vorliegenden Artikels besteht in einer neurophilosophischen Untersuchung von Nagels und Parfits Ansätzen. In einem ersten Schritt werden die Theorien von Nagel und Parfit kurz vorgestellt. Die hier entwickelten Kriterien zur personalen Identität dienen dann als Richtlinie für eine psychologisch-empirische Untersuchung der personalen Identität bei Parkinson-Patienten mit einer fetalen Hirngewebs transplantation. Die Methode und Resultate der empirischen Ergebnisse werden in einem zweiten Schritt vorgestellt. In einem dritten Schritt werden psychologisch-empirische Ergebnisse und die philosophischen Theorien von Nagel und Parfit direkt miteinander verglichen. Es werden Fragen zum Verhältnis zwischen personaler Identität und Gehirnidentität diskutiert. Der Vergleich zwischen psychologisch-empirischen Ergebnissen und den philosophischen Theorien zeigt die folgenden Resultate. Erstens: Nagels Charakterisierung des Gehirns als ein „additional fact“ mit „physico-mental intimacy“ ist nicht kompatibel mit dem subjektiven Erleben des Gehirnimplantat and der eigenen Person in diesen Patienten. Zweitens: Parfits Vernachlässigung der mentalen Zustände und der Ersten-Person Perspektive ist empirisch unplausibel und wirft die Frage nach den epistemischen Fähigkeiten und Nicht-Fähigkeiten unseres Gehirns und somit einer ‚Neuroepistemologie‘ auf. Drittens: Parfits ontologische Definition des Gehirns in einem rein mechanisch-technischem Sinne, wie in seinem physikalischem Spektrum Argument vorausgesetzt, ist nicht kompatibel mit den empirischen Ergebnissen zur Hirnfunktion bei Implantation. Dies wirft die Frage nach der ontologischen Definition des Gehirns und somit einer ‚Neuroontologie‘ auf. Die Entwicklung von „Neuroepistemology“ und „Neuroontology“ könnte zu einer empirisch plausibleren philosophischen Definition des Gehirns sowie der Beziehung zwischen personaler Identität und Gehirnidentität beitragen.

Introduction

The intimate connection between our brains and our behaviors, as well as the peculiar relationship between our brain and our selves, generates questions that beg for the interplay between philosophical and neuroscientific methods. For example, many of us overtly or covertly believe

in a kind of „neuroessentialism“, that our brains define who we are and what we are (see Roskies 2002). In short, personal identity may be considered as a brain-based notion or, to put it even stronger, personal identity is brain identity. This, however, remains questionable from a philosophical point of view.

Personal identity is a conceptual notion which refers to strict numerical identity through time as an All-or-nothing category. In contrast, the issue of brain identity has not been raised yet, partly because the brain is considered as an empirical notion rather than a conceptual one. Despite these basic differences, the brain, among others, has nevertheless been discussed as a potential candidate for the explanation of the identity of persons (Northoff et al. 2001): Either the brain is regarded as a necessary and sufficient criterion for personal identity („I am my brain“) as it is, for example, suggested by T. Nagel (1986). Or the brain is neither regarded as a necessary nor a sufficient criterion for personal identity („I am not my brain“) as it is, for example, reflected in the account by D. Parfit (1989). In addition to these positions, there are various intermediate positions (like, for example, Shoemaker, Williams, etc.) which relate personal identity to the brain though not exclusively, i. e., as a sufficient criterion (see Northoff 2001 for an overview).

The purpose of the present contribution is to demonstrate that both Nagel's and Parfit's theories of the relationship between personal identity and brain presuppose an empirically implausible definition of the brain. Drawing on own psychological investigations in patients with brain implants, Nagel's and Parfit's philosophical definitions of the brain turn out to be empirically rather implausible. Accordingly, there is a need for an empirically plausible epistemic and ontological definition of the brain and thus for „neuroepistemology“ and „neuroontology“.

In a first step, some of the main features of the philosophical theories about the relation between personal identity and brain, as by Nagel and Parfit, are briefly summarized. The present account focuses on the theories by Nagel and Parfit since both paradigmatically reflect the most extreme positions in the current debate about the relationship between personal identity and the brain. The characteristics of their philosophical theories serve as an orientation for the development of questionnaires and thus for an empirical study as reported in the second step. These are applied to Parkinson patients who received brain implants. In a third step, the psychological data and the philosophical theories are directly

compared with each other. The neurophilosophical implications for their arguments and the need for an epistemic and ontological definition of the brain are pointed out (see also Figure 1).

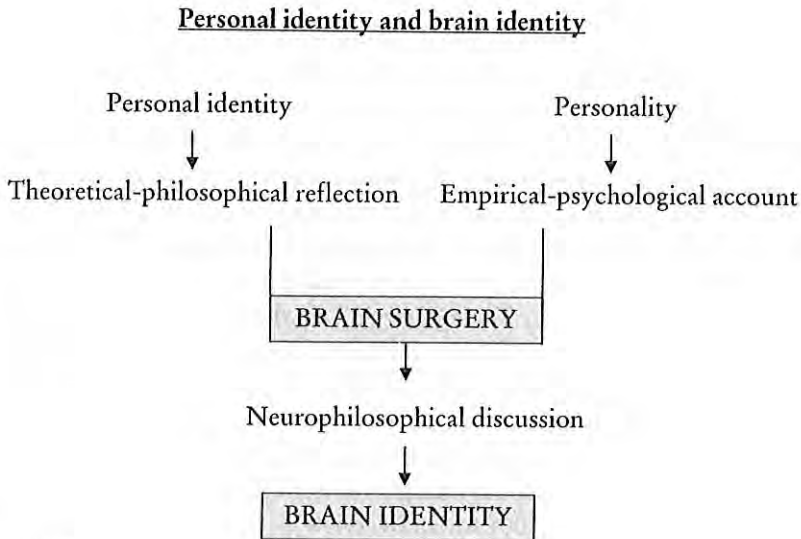


Figure 1

Philosophical Definition of Personal Identity and the Brain: „I Am My Brain“ Or „I Am Not My Brain“?

„Additional fact“

Nagel (1986, p. 98) considers the brain as both a necessary and sufficient criterion for personal identity which results in the statement „I am my brain“. Nagel (1986, p. 64f., p. 69) considers the term „identity“ as a definitive and non-conventional term. The term is definitive in the sense that it necessarily implies a Yes-No, i.e., All-or-Nothing decision with regard to the identity of persons. Either a person is identical or non-identical. There are no intermediaries. The term is non-conventional in the sense that it does not imply its own necessary and sufficient conditions. Since the term does not imply its own necessary and sufficient conditions, it contains an „empty position“ which must be filled by an „additional fact“ (Nagel 1986, p. 71). He compares the term ‚identity‘ with the

term ‚gold‘. Before the chemical formula for ‚gold‘ was detected, the term ‚gold‘ contained an „empty position“ which later was filled by an „additional fact“, i. e., the chemical formula of gold.

Analogous to the situation before the detection of the chemical formula of ‚gold“, we currently have no idea about the „additional fact“ which might potentially fill the „empty position“ in the term ‚identity‘. According to Nagel (1986, p. 72), the „additional fact“ in the case of the term ‚identity‘ must bridge the gap between the subjective experience of the person, i. e., its First-Person Perspective and its „I“ on the one hand and the objective necessary structures, i. e., its body of its existence on the other. Since the brain might eventually bridge this gap between subjective experience and objective structures, it may be considered as a suitable candidate to fill the „empty position“. Accordingly, the brain might be considered as an „additional fact“ in the case of the term „identity“: On the one hand, the brain must be considered as the necessary foundation for the possibility of subjective experience since without a brain we remain unable to experience mental states in First-Person Perspective. On the other hand, the brain is the carrier of psychophysiological processes which remains essential for regulation and maintenance of the body. In contrast to other organs like, for example, liver or kidney, the loss of the brain is accompanied by loss of personal identity. The brain must subsequently be regarded as a necessary and sufficient condition for personal identity. However, due to our current lack of empirical knowledge, this assumption must be considered as a preliminary hypothesis according to Nagel.

Instead of assuming an „additional fact“, Parfit rejects the notion of personal identity altogether and replaces it by the term „survival“. According to Parfit (1989, p. 476f.), „what matters“ is not „personal identity“ but „survival“ of the person („identity is not what matters“). Unlike the term „identity“, the term „survival“ does neither imply a one-to-one relationship nor an All-or-Nothing relationship. For example, in the case of fission in hemispherectomy, the psychological continuity and thus the „Relation R“ (see below) may be preserved so that the person „survives“. However, the person cannot be considered as identical in the numerical sense since in this case there is a one-to-two and more-or-less relationship. „What matters“ for the person itself in this case is therefore not numerical identity and thus personal identity but mere „survival“. Unlike in the case of personal identity, the assumption of an „additional fact“ is

no longer necessary in „survival“. Instead, „survival“ can be defined by the „Relation R“ which he characterizes by „psychological connectedness“ and „psychological continuity“ (Parfit 1989, p. 206): „Psychological connectedness“ points out direct psychological connections („particular direct psychological connections“) as, for example, those between the memories in the sense of Locke. In contrast, „psychological continuity“ is defined by „overlapping chains of direct connectedness“. If there are no direct connections, i. e., „psychological connectedness“, as for example, in the case of amnesia, „psychological continuity“ can nevertheless be maintained. „What matters“ for the „survival“ of the person is thus „psychological continuity“.

Parfit also challenges Nagel's assumption of „identity“ as a non-conventional term. If the term „identity“ can be replaced by the term „survival“, the question for characterization of the latter has to be raised. Persons can survive without reference to something else so that the term „survival“ can neither be characterized by an „empty position“ nor an „additional fact“ as suggested by Nagel (see above). He illustrates his point by an analogous example, the one of nations. Although nations neither refer to a separate existing entity nor to other objects, properties, entities, etc. (governments, territorium, etc.), they nevertheless „survive“. Analogous to nations, the term „survival“ refers exclusively to itself. In contrast to Nagel, the assumption of an „additional fact“ is therefore no longer necessary (Parfit 1989, p. 471f.). Since there is no such „empty position“, the brain is no longer needed for the term „survival“.

„Physico-mental intimacy“

According to Nagel, the brain bridges the gap between subjective experience of mental states and psychophysiological states (see above). The brain can therefore not be considered as a purely physical organ since mental states and subjective experience cannot be reduced to physical properties (Nagel 1986, p. 57, p. 74). In addition to physical properties, the brain must therefore be characterized by mental properties. Nagel speaks of a so-called „physico-mental intimacy“ (in an ontological sense) as an „apparent intimacy between the mental and its physical conditions“ (Nagel 1986, p. 20). Due to these mental properties, the brain shows a special kind of „insiderness“ which accounts for its foundational character for subjective experience:

It (the brain) can be dissected, but it also has the kind of inside that can't be exposed to dissection. There's something it is like from the inside to taste chocolate because there's something it's like from the inside to have your brain in that condition that is produced when you eat a chocolate bar. (Nagel 1987, p. 34f.)

Accordingly, the brain can be described mentally and physically which both may be traced back to a common foundation, i. e., a „fundamental essence“ (Nagel 1979, p. 199). This „fundamental essence“ can be defined by complex forms of organization and combination („unusual chemical and physiological structure“) of matter which shows both proto-physical and proto-mental properties (Nagel 1979, p. 201). However, neither the exact definition of both kinds of properties nor their relation are clearly determined.

In contrast to Nagel, Parfit does not consider the importance of mental states at all since he defines „survival“ by psychological functions exclusively. There is subsequently no need for him to consider mental states and any kind of „physico-mental intimacy“.

Characterization of the brain

Even if the brain is neither needed as an „additional fact“ nor as a sufficient condition, it may nevertheless be regarded as a cause, i. e., necessary condition for the possibility of the „Relation R“ and thus for „survival“. Parfit concedes that the brain is the „normal“ cause of the „Relation R“ which however does not imply that it is the necessary cause. Following Parfit (1989, p. 207f.), there may also be other causes, i. e., „wide“ and „widest“ causes which refer to „any reliable cause“ („wide causes“) and „any cause“ („widest causes“) respectively. For example, a computer which reliably simulates the activity of the brain may be considered as a „wide cause“ for the „Relation R“. The brain can therefore no longer be considered as the necessary cause for the „Relation R“ – „survival“ and brain are not necessarily related to each other („Relation R is not its brain“).

In addition, Parfit concedes that Nagel's definition of the brain carries a certain ambivalence. On the one hand Nagel (1986, p. 74f.) speaks about „a brain“/„intact brain“ and on the other about „a brain in certain (mental) states“. The brain as a „brain/intact brain“ presupposes an anatomical definition of the brain. While the brain as a „brain in certain states“ implies a functional and thus psychological definition. Only the brain as a „brain in certain states“ can therefore be related to Relation R

and thus psychological continuity. However, this relation is not necessary because the brain as a „brain in certain states“ as the „normal cause“ underlying Relation R can be replaced by „any cause“ such as computers. Accordingly, there is no necessary relation between brain and „Relation R“.

Finally, the emphasis of Parfit on „psychological continuity/connect-edness“ suggests that the underlying „physiological or physical continuity“ may be considered as a necessary condition for the „Relation R“ and thus for „survival“. However, Parfit (1989, p. 474ff.) rejects this argument of „physical continuity“ on the basis of his „physical spectrum argument“ where compares two different surgical procedures: In the first procedure, 100 brain cells are exchanged sequentially in 100 steps while in the second procedure the same 100 brain cells are exchanged at once in one step. According to Parfit, the results of both surgical interventions remain the same while they can only be distinguished in their procedures. Moreover, both surgical interventions differ with regard to „physical continuity“ which is preserved only in the first, i.e., the sequential case but not in the second one. If „physical continuity“ is considered as essential for „Relation R“ and thus for „survival“, the person should „survive“ only the first procedure but not the second one. This however is absurd since both procedures lead to the same result with regard to the person which „survives“ in both cases. Accordingly, neither the brain nor „physical continuity“ of the brain can be considered as a necessary condition for the „Relation R“ and thus for „survival“. Moreover, the example underlines „what matters“: In the first procedure, the identity of the person is preserved while this is no longer the case in the second procedure. Here the person only „survives“ but, due to lack of „physical continuity“, remains no longer identical. Since the person „survives“ both surgical procedures, any additional differentiation between „identity“ and „survival“ remains meaningless for the person itself. „What matters“ for the person itself is thus „survival“ and not the distinction between „survival“ and identity.

Psychological Investigation: Subjective Experience of Personal Identity and Personality in Patients with Brain Implants

Methods

Patients: Five patients with severe Parkinson's disease were investigated. All patients showed only motor symptoms (akinetic-rigid type) while no psychological symptoms (depression, dementia) were present (i. e., patients with those symptoms were excluded). Consequently, they received only medication for their motor symptoms but no psychotropic drug so that reliability of reports, self-observation and judgement was assured. Two of these patients received fetal brain tissue transplantation (transplantation of fetal cells into the basal ganglia (i. e., striatum) as a subcortical structure) and three patients received electrode implantation into the basal ganglia (i. e., internal globus pallidum – see Brundin et al. 2000 for exact details about both patient selection and surgical procedures). All five patients were investigated with a standardized interview (see Appendix 1), a semi-quantitative visual analog scale (see Appendix 2) and a personality scale (see below) before (t1) and six months after (t3) implantative surgery. In addition we investigated five age- and sex matched Parkinson patients at the same time points (t1) and six months after (t3), who received only pharmacological treatment (drug-treated), and five healthy controls with the personality scale.

Personal identity: The different items in both the standardized interview (see Appendix 1) and the visual analog scale (see Appendix 2) focused on the various kinds of criteria for personal identity as discussed by Nagel and Parfit.

First, as derived from Nagel (1979), subjective experience of the own „I“, the First-Person Perspective as the „What is it like“, the „centralization“ as the point of view, and the „relation of mineness“ reflecting „my experience“ (Metzinger 1993, p. 177, p. 204, p. 237, p. 265ff.; Northoff 2001, p. 156f.) were asked for with regard to the mental states and the body.

Second, the subjective experience, perception, role and relation of the person to the own brain and the importance of the brain for personal identity, as compared to other organs of the body, was asked for so that the direct relation between brain and personal identity could be explored.

Third, after the implantative surgery, the influence of the surgery on personal identity/personality/individuality, brain/relation to own brain, mental states and subjective experience and body and environment was asked for. This allows for exploration of the direct relationship between changes in the brain and personal identity.

The qualitative interviews were analyzed according to a qualitative analysis of content with a consecutive phenomenological analysis (see Mayring 1990, p. 46–80 for details and Northhoff 2001, p. 435–466 for the interviews, i. e., the results). The lines of the different items in the semi-structured visual analog scale were divided into centimeters and the respective answers were measured (see Appendix 2).

Personality: In order to investigate the personality, the German version of the Minnesota Multiphasic Personality Inventory (MMPI) was applied. The MMPI includes various dimensions of the personality like hypochondriasis, depression, conversion hysteria, psychopathic deviate, masculinity/femininity, paranoia, psychasthenia, schizophrenia, hypomania, social introversion. Statistically, comparisons between groups (surgically-treated patients versus drug-treated patients versus healthy controls) at the different time points (ANOVA with post-hoc t-tests) as well as between time points (before-after) within the surgical group (paired t-tests) were calculated.

Results

Personal identity: First, all subjects made no principal difference between the terms ‚personal identity‘ and ‚personality‘ and often considered both terms as synonymous. Personal identity was mostly characterized as the personality of the own person. The term ‚personality‘ was defined as the sum total of all psychological functions which, according to the patients, however can neither be reduced to one particular property nor to a mere statistical account. Accordingly, the patients refer rather to a synchronic meaning than to a diachronic meaning as presupposed in philosophical discussion (see above).

Secondly, no patient experienced a change in the „I“, the First-Person Perspective and her/his point of view by implantation of either fetal cells or an electrode. Though the electrode is regarded as a purely technical device, it nevertheless becomes individualized so that, for example, one

patient speaks of „my electrode“ – one may characterize this relation as a „relation of minessness“. Moreover, patients said that the function of the electrode may be influenced by psychological factors including the own emotional status and attitude towards it. One patient reported that „she would feel her brain during the on and off switching of her electrode“ and „could influence the electrode emotionally and mentally“. The fetal cells are regarded as „substitutes for dopamine“ (i. e., the lacking transmitter) and are thus experienced as „natural and physiological“. The psychological integration of these cells is not problematic for the patients. Similar to the electrode, the fetal cells may be influenced by psychological factors. For example, one patient performed meditation before and after surgery in order to „create a good and beneficial mental-moral environment for the growth of the cells“. Moreover, another patient asked whether he would speak Swedish after the implantation (the surgery was performed in Sweden with fetal cells derived from Swedish fetuses). Finally, some patients said that during the course of their illness they „became much more aware of their brain“.

Thirdly, all patients attributed a „special role“ to the brain as compared to other organs. Though implantation of cells and electrodes in the brain does not lead to changes in personal identity, it nevertheless cannot be compared with other organs like, for example, the heart because the „brain constitutes psychological and mental functions“. According to our patients, there is a principal difference between transplantation of the whole brain and transplantation of the whole heart with regard to personal identity and personality.

Fourthly, neither the disease itself nor the implantative therapy was related with any changes in personal identity and personality. Some patients spoke about a „greater freedom and individuality“ in their movements after surgery. The fetal cells were regarded as a „natural and physiological substitute“ for the dopamine deficiency in the own brain. The electrode was considered as „technical device which makes the brain function in the correct way“.

Fifthly, there were no principal differences between the answers given by the patients and those by their partners/spouses.

In summary, analysis of subjective experience revealed that surgically-treated patients did not experience changes in their personality and personal identity. Moreover, they were well able to psychologically integrate the implant within their subjective experience in First-Person Perspective.

Personality: Surgically-treated patients differed in several different subscales, i. e., dimensions of the personality from drug-treated patients and healthy controls at either time point, i. e., t1 and t3 (see Table 1). In contrast, only the dimensions of ‚schizophrenia‘ and ‚social introversion‘ showed significant differences between t1 and t3 within the surgically-treated group which was neither the case in drug-treated patients nor in healthy controls.

Limitations: it should be noted that the investigated sample was quite small. Therefore, further significant effects in a larger sample cannot be excluded. Moreover, the sample was specifically selected excluding, for example, patients with emotional instabilities. This might have biased our data in a certain way. Consequently, our results must be considered exploratory awaiting further support in larger samples.

In summary, analysis of psychometric personality tests revealed that surgically-treated patients showed only some specific changes in their personality after surgery especially in the dimensions of ‚social introversion‘ and ‚schizophrenia‘.

Discussion

„Additional fact“

One should first point out that there is a principal difference between personal identity and personality. Personal identity is a conceptual and philosophical term whereas personality is rather an empirical and psychological notion – this implies that only the latter but not the former can be empirically investigated. Despite these principal differences, there might nevertheless be some implications of the psychological results for the philosophical concept of personal identity.

Nagel, for example, assumes an „additional fact“ in his concept of personal identity (see above). Analogous to the chemical formula, this „additional fact“ itself should be empirically accessible in experience. However, the psychological data show no hints or indications of such an „additional fact“. The patients did not experience anything but their personality which they regarded as the sum total of their psychological functions. In contrast, they did not define (or experience) their personality (and personal identity) by any particular property or fact which could be regarded as an equivalent to Nagel’s „additional fact“. Nagel’s as-

Table 1: Results of scores in personality dimensions in the Minnesota Multiphasic Inventory (MMPI) in surgically- and drug-treated Parkinson patients and healthy controls before (t1) and six months (t3) after surgery/ later

	Surgically-treated patients (n = 5)			Drug-treated patients (n = 5)			Healthy controls (n = 5)			Variance analysis between groups (p)		
	t1	t3	t1-3	t1	t3	t1-3	t1	t3	t1-3	t1	t3	t1-3
Hypochondriasis	67.5b**	64.1b*	n.s.	77.4 c**	81.7c**	n.s.	50.8	50.3	n.s.	0.004**	0.003**	n.s.
Depression	66.8b**	63.8b*	n.s.	68.8c*	73.5c*	n.s.	48.3	48.0	n.s.	0.021*	0.018*	n.s.
Conversion Hysteria	67.0b*	63.4a*	n.s.	72.6c**	82.0c*	n.s.	53.3	52.0	n.s.	0.028*	0.031*	n.s.
Psychopathic deviate	54.0	55.3	n.s.	58.2	55.5	n.s.	51.8	50.2	n.s.	n.s.	n.s.	n.s.
Masculinity/Femininity	50.8	48.2	n.s.	53.6	49.5	n.s.	53.4	50.7	n.s.	n.s.	n.s.	n.s.
Paranoia	51.0b*	52.6	n.s.	60.0c*	54.0c*	n.s.	44.1	45.3	n.s.	0.047	0.041	n.s.
Psychasthenia	60.6b**	52.9	n.s.	69.0c**	76.3	n.s.	48.1	48.2	n.s.	0.049*	n.s.	n.s.
Schizophrenia	59.9b*	53.9	0.042b*	64.8c*	68.7c*	n.s.	49.5	49.4	n.s.	0.042*	0.039*	0.039*
Hypomania	50.8	49.5	n.s.	53.9	49.5	n.s.	51.2	54.3	n.s.	n.s.	n.s.	n.s.
Social introversion	60.0b**	53.8	0.019ab*	58.2	52.5	n.s.	45.0	44.3	n.s.	0.015*	0.049*	0.029*

a = significant ($p < 0.05$) difference between surgically- and drug-treated patients
b = significant ($p < 0.05$) difference between surgically-treated patients and healthy controls
c = significant ($p < 0.05$) difference drug-treated patients and healthy controls

* = $p < 0.05$ (corrected)
** = $p < 0.005$ (corrected)

sumption of an „additional fact“ remains therefore empirically implausible.

Though Nagel could argue that this „additional fact“ is epistemically not directly accessible in subjective experience in First-Person Perspective. This leaves him with the following three options to argue: Firstly, the „additional fact“ might not be accessible at all neither in First- nor in Third-Person Perspective. In this case, the „additional fact“ remains principally non-detectable for us. One might consequently consider the „additional fact“ as rather mysterious. Accordingly, this option does not seem to be very attractive. Secondly, the „additional fact“ might be accessible in Third-Person Perspective. Then it should be detectable in empirical and thus neuroscientific investigation which relies on the Third-Person Perspective. However such an „additional fact“ has not been detected yet in neuroscientific investigation of the brain. Accordingly, this option remains empirically rather implausible. Thirdly, the „additional fact“ might be accessible in First-Person Perspective though in a disguised form and thus rather indirectly. In this case, the „additional fact“ might not be accessible as a „fact“ but rather as a particular type of state as distinguished from neuronal states. Instead of looking for an „additional fact“, one should then aim at revealing an „additional state“ and its relation to the brain. What could this „additional state“ be? Is it a mental state? If mental states account indeed for the „additional state“ as the disguised form of the „additional fact“ and thus the brain, mental states must reflect the access to one's own brain in First-Person Perspective. We may perceive our own brain states thus not as brain states in First-Person Perspective but as mental states. Our brain as the „additional fact“ can thus be accessed only indirectly via mental states as „additional states“. Conversely, we remain unable to access our own brain as a brain in First-Person Perspective. This epistemic inability of our own brain to access itself directly as a brain can be called „autoepistemic limitation“ (Northoff 2004). This „autoepistemic limitation“ may be subserved by specific principles of functional brain organisation which prevent the brain from directly perceiving itself as a brain. These principles of functional brain organisation may thus fulfill the same role as the chemical formula did in the case of gold (see above). Accordingly, „autoepistemic limitation“ might account for what Nagel means by „additional fact“. Though his formulation „I am my brain“ should be rephrased as „I am my brain, though, due to „autoepistemic limitation“,

I remain unable to directly access myself as my brain“. Moreover, investigation of the empirical mechanisms underlying „autoepistemic limitation“ requires direct relationships between empirical function and epistemic abilities/inabilities, i. e., so-called „epistemic-empirical relationship“ (Northoff 2004). This in turn makes the development of „neuroepistemology“ (Northoff 2004) as an „epistemology on a neurological basis“ (Kuhlenbeck 1965, p. 137) necessary.

In contrast to Nagel, Parfit's assumption of psychological continuity and psychological connectedness, i. e., the Relation R as crucial for personal identity is well compatible with the empirical results. Our patients experienced and defined their personality as the sum total of the psychological functions which might account for what Parfit calls „psychological connectedness“ and „psychological continuity“. It should however be noted that Parfit refers to the numerical sense when he speaks of psychological connectedness and continuity whereas our patients imply rather a synchronic sense in their meaning of psychological functions. The relationship between the numerical and the synchronic sense remains however unclear. Does the synchronic identity of a person, i. e., its identity at a certain time point (which is often related to the personality of a person) implies its numerical identity, i. e., its identity through time? How are personality as an empirical notion and personal identity as a rather conceptual notion related to each other? Must personality and synchronic identity as the sum total of current psychological functions be regarded as necessary criteria for numerical identity? As reflected in the personality assessment, our patients did not experience any major changes in their personality and thus their synchronic identity (see above). Moreover, they showed no abnormalities in their personality and psychological function through time, i. e., before and after implantation of fetal cells or electrodes (Northoff 2001). Following Parfit, one might consequently speak of „psychological continuity“ and thus preservation of „Relation R“ in our patients; i. e., they „survived“ brain tissue or electrode implantation.

More interesting are those cases (i. e., personality disorders) with major changes in their personality (i. e., synchronic identity). Here, the implications of the synchronic changes in the personality for the numerical concept of personal identity could be discussed. Other discussions might focus on schizophrenic patients with ego-disturbances who experience themselves as different persons through time. The linkage of these expe-

periences to the concept of personal identity and to empirical changes in the brain might be particularly worthwhile to consider (Northhoff 2001). Would Parfit still speak of „psychological continuity“ and consecutively of „survival“ in these cases? Since these patients experience a different identity, one might distinguish between ‚subjective identity‘ and ‚objective identity‘. The ‚subjective identity‘, as experienced from the inside in First-Person Perspective, might be altered in these patients. Whereas their ‚objective identity‘, as observed from the outside in Third-Person Perspective, remains the same. Following Parfit, one might nevertheless speak of ‚survival‘ in these patients because he disregards subjective experience in the First-Person Perspective (see below) and thus ‚subjective identity‘ (see below).

„Physico-mental intimacy“

Some of our patients reported close interaction between their mental states and the physical substitutes, i. e., the cells or the electrode. They, for example, describe the feeling that they could influence the brain by their psychological and mental states. Nagel’s assumption of „physico-mental intimacy“ might thus be considered as empirically plausible. The subjective experience of an influence of mental states on both the own brain states and the implant would not be possible without some kind of „physico-mental intimacy“. „Physico-mental intimacy“ is thus best accessible in subjective experience in the First-Person Perspective. Does this epistemic characterization of „physico-mental intimacy“ however justifies the inference of physical and mental properties in an ontological sense? Such an epistemic-ontological inference is, for example, reflected in Nagel’s when he infers from the epistemic description of a special „insideness“ to underlying ontological properties. This type of inference might be called an „epistemic-ontological fallacy“ (Northhoff 2004) where one falsely infers from epistemic characteristics to ontological properties.

However, even if one allows for epistemic-ontological inferences, the ontological assumption of mental brain properties, as distinguished from its physical properties, cannot be justified: Though mental states can be experienced in First-Person Perspective, their experience cannot be directly linked to the own brain because of „autoepistemic limitation“. If however mental states cannot be linked directly to the own brain, any inference of mental brain properties remains implausible. From the non-

experience of the brain in mental states, mental properties of the brain cannot be inferred. Though one may not infer mental brain properties one may be able to at least infer mental properties in general. While the more specific assumption of linking these mental properties to the brain cannot be made. Accordingly, any type of epistemic-ontological inference from mental states to mental brain properties remains impossible. Nagel could however argue that mental brain properties may be inferred from Third-Person Perspective. If mental brain properties cannot be inferred from experience of mental states in First-Person Perspective, they may at least be inferred from Third-Person Perspective. This is problematic too because we do not experience mental states in Third-Person Perspective. Instead, we observe physical states in Third-Person Perspective. One may consequently infer mental brain properties from observation of physical states in Third-Person Perspective. This however remains even more implausible: Inference of mental brain properties from physical states is blocked too because physical states can be entirely accounted for by physical properties. There is no need to assume any mental brain properties. Accordingly, either type of inference of mental brain properties, from First- or Third-Person Perspective, remains implausible.

Though the attribution of mental properties to the brain in an ontological sense remains implausible, the epistemic relevance of mental states in subjective experience in First-Person Perspective cannot be denied. This is reflected in both empirical results and the epistemic characterization of the First-Person Perspective: The empirical results show subjective experience of mental-physical interaction and the First-Person Perspective can be defined by experience of mental states rather than by observation of physical states. Denial of mental states implies therefore denial of the principal epistemic possibility of subjective experience in First-Person Perspective, as distinguished from observation in Third-Person Perspective. This is, for example, the case in Parfit's account who considers only psychological functions which as such are accessible in Third-Person Perspective exclusively. In contrast, he neglects subjective experience of mental states and thus the First-Person Perspective itself in his account of personal identity. This neglect makes it impossible for Parfit to even raise the possibility of „autoepistemic limitation“ as an epistemic inability of the brain in First-Person Perspective: If there is no First-Person Perspective at all, the problem of the perception of one's own brain in First-Person Perspective remains elusive. Moreover, Parfit cannot ac-

count for „physico-mental intimacy“ in an epistemic sense: Since he disregards subjective experience of mental states in First-Person Perspective, the idea of a possible (indirect epistemic) „intimate“ relation between mental states and the (physical states of the) brain cannot even occur to him.

Accordingly, Nagel is right by pointing out the epistemic importance of mental states in First-Person Perspective while he is wrong in inferring mental brain properties. In contrast, Parfit is right in rejecting mental brain properties while he is wrong in neglecting mental states and thus the First-Person Perspective.

Characterization of the brain

Parfit considers the brain as the „normal cause“ but not as a necessary cause for psychological continuity and thus the Relation R because „any reliable cause“ could fulfill the same function (see above). Considering that in our case of brain implants, brain functions are replaced by fetal cells and electrodes without any psychological changes in personal identity and personality, Parfit’s characterization of the brain as a non-necessary cause for personal identity must be considered as plausible from an empirical, i. e., psychological point of view. Accordingly, he presupposes a rather extrinsic, i. e., non-necessary relation between brain and psychological function. This extrinsic relation might however become questionable in the case of subjective experience of mental states in First-Person Perspective. If, for example, the subjective experience of mental states is intrinsically related to the brain itself (see Northoff 2003), as it is suggested by „physico-mental intimacy“ in an epistemic sense (see above), the brain might be regarded not only as the „normal cause“ but also as a necessary cause for mental states. Due to its intrinsic linkage to mental states, the brain can then not be replaced by „any reliable cause“: The own brain remains indispensable for the subjective experience of mental states by that particular person. It can therefore not be replaced by either another brain of another person or even another cause such a computer. One’s own brain cannot be replaced by „any other cause“ – one’s own brain must be characterized as „no other cause“: Firstly, the brain as „no other cause“ is individualized since it is not „any brain“ but „my brain“. Secondly, due to an individual genesis and history with a so-called „biopsychosocial historicity“ (Northoff 2001), „my brain“ is intrinsically linked to my subjective experience of mental states in First-Person Per-

spective. Subsequently, no other brain or any kind of device could take over the role of „my brain“ with respect to mental states. Thirdly, due to its „biopsychosocial historicity“, any implant needs to be integrated within the individual context of „my brain“. The implant itself has to be adjusted, i. e., individualized with regard to the needs and demands of the brain of the respective host. Parfit neglects the individual character of the brain almost completely which is also reflected in his implicit presupposition of the brain as a purely mechanical-technical device (see below). He refers to „any brain“ but not to „my brain“ and can therefore not account for subjective experience of mental states in First-Person Perspective. Therefore Parfit is right when he states that „I am not my brain“ with regard to the brain as „any brain“ since the functions of the latter, psychological functions, can be replaced by „any cause“: „I am not *any* brain“. While he is wrong that „I am not my brain“ with respect to mental states: „I am *my* brain“. Accordingly, Parfit's statement „I am not my brain“ must be reformulated as „I am not my brain with regard to psychological functions, but I am *my* brain with regard to *my* mental states“.

In addition to the different types of causes, Parfit points out the distinction between the brain as an „intact brain“ and the brain as a „brain in certain states“. The „intact brain“ might refer, for example, to the anatomical structures of the brain. In this case, the replacement of cells, as in fetal brain tissue transplantation, can be considered as a restoration of the brain as an „intact brain“. However, neither the death of the cells, as caused by the disease itself, nor their replacement by fetal cells was accompanied by experience of changes in (personal identity and) the personality. The anatomical brain might subsequently not be regarded as a necessary condition for personal identity (and personality) since otherwise any change in the former should have led to alterations in the latter. It is this sense of the brain as an „anatomical brain“ which is implicitly presupposed by Parfit himself. This is, for example, reflected in his „physical spectrum argument“ where different surgical procedures of replacing cells are supposed to result in similar changes in personal identity (see above). Parfit presupposes here a purely mechanical sense of the brain. This however must be regarded as empirically implausible for the following reasons: Firstly, the importance of interactions between the brain cells and the implanted cells is not considered. This is, for example, reflected in the fact that the implanted cells can exert an influence on motor function only if they interact with the other, i. e., the preexisting

nerve cells (Brundin et al. 2000). Otherwise, the implanted cells remain without any functional and thus therapeutic effects: Without interaction between implanted and existing cells no result of the respective surgical procedure. The interaction might differ in prompt and gradual replacement of cells which then might also exert an impact on the result. Both surgical procedures may thus no longer lead to the same result as it is presupposed by Parfit. Secondly, the importance of integration is not considered by Parfit. He assumes that both surgical procedures lead to the same results despite either gradual or prompt integration of the novel nerve cells. He subsequently does not consider that different types of integration may have an impact on the result. The gradual replacement may lead to better integration and thus to better results than the prompt replacement where a higher amount of cells has to be integrated at one particular time point: The better (or less) integration of the implanted cells into the preexisting brain cells of the respective host, the less (or more) likely any changes in its personal identity. For that reason, brain tissue transplantation has often been performed into two stages, first the one side of the brain and second the respective other side. Accordingly, both surgical procedures, as presupposed in Parfit's Spectrum argument may not lead to the same result. Thirdly, the importance of the distinction between „critical and non-critical nerve cells“ (Stone 1988, p. 521f.) is not considered by Parfit. This is, for example, reflected in Parkinson's disease, where only the dopaminergic cells are critical. In contrast implantation of non-dopaminergic cells does not lead to any improvement in motor function – non-dopaminergic cells are thus non-critical. Analogously, there might be certain brain cells and regions like, for example, the frontal lobe, which can be regarded as „critical“ for personality and personal identity. While the cells in the basal ganglia, as targeted in Parkinson's disease, must apparently be regarded as „non-critical“ since our patients did not show major changes in personality and personal identity.

Due to these shortcomings, Parfit conception of the brain must be characterized as mechanical and static. In order to make it empirically more plausible, it should be replaced by biological and dynamic conception of the brain (see Northoff 2004) which can account for interaction, integration, and „critical and non-critical cells“. In addition, such a conception might allow for clear-cut determination of brain identity. Determination of a clear border („sharp borderline“) as a „critical threshold“ between identity and non-identity of the brain might then be possible

even in the middle cases of the physical spectrum. Parfit's „physical spectrum argument“ should then no longer be regarded as an argument for impossible determination of personal identity. Instead, its modified version as biological and dynamic, should be considered as an argument for the possible determination of brain identity.

In contrast to brain tissue implantation, the electrode implantation concerns more the „brain in certain states“ rather than the „intact brain“ as an „anatomical brain“ (see above). Rather than with anatomical structures, the electrode interferes with the function of the brain and thus its states since, as one patient put it, „it makes the brain function in the correct way“. However, the question for the type of states arises. The „brain in certain states“ might concern physiological states, psychological states, or mental states. If it concerns, as in Parfit's „psychological spectrum argument“, psychological states, the question for the relation between psychological and physiological states arises. If it concerns, as in Nagel's case, mental states, the question of the relation between mental states and psychological/physiological states arises. In our case of electrode implantation, one might ask which types of states is influenced and manipulated by the electrode, physiological, psychological, or mental states? Furthermore, the question for the relationship between physiological, psychological, and mental states arises. Do physiological states, which are probably manipulated and „normalized“ by the electrode, influence psychological and mental states? The definition and characterization of the states of the brain remains unclear. This may however be crucial in determining the ontological properties of the brain. In order to do so a so-called „neuroontology“ will be required (Northoff 2004).

The ontological definition of the brain might provide the basis for the determination of brain identity and its criteria. Brain identity might reflect continuity of anatomical structures (i. e., cells, etc.), physiological states (i. e., action potentials, etc.), psychological functions (i. e., memory, etc.), or mental states (i. e., subjective experience). The „psychological spectrum argument“ might thus not only be applied to psychological states, as it has been demonstrated by Parfit, but in addition to anatomical structures and physiological and mental states. One might consequently speak of an „anatomical spectrum argument“, a „physiological spectrum argument“, a „psychological spectrum argument“, and a „mental spectrum argument“. This extended version of Parfit's spectrum argument might then be regarded as an argument for brain identity rather

than for/against personal identity. Personal identity is then conceived a criterion for brain identity instead the brain being a criterion for personal identity as it is often discussed in the philosophical debate. If the relation between person and brain can indeed be reversed, Parfit might be right by assuming that the particular notion of personal identity does not matter for „survival“ of the person and its brain. While he remains wrong by rejecting the more general notion of identity altogether because the identity of the brain, i.e., brain identity, seems to matter for „survival“ of the person and its brain.

Conclusion

I investigated the relationship between personal identity and the brain focusing on the empirical plausibility of the epistemic and ontological definition of the brain. Either personal identity might be equated with the brain („I am my brain“), as suggested by T. Nagel. Or personal identity is dissociated from the brain („I am not my brain“), as suggested by D. Parfit. However, relying on own psychological results of patients with brain implants, both Nagel's and Parfit's definition of the brain must be regarded as implausible from an empirical point of view. Nagel's assumption of the brain as an „additional fact“ with mental properties remains empirically problematic. This raises questions about our epistemic abilities and inabilities with respect to one's own brain thus requiring „neuroepistemology“. Parfit's neglect of mental states and the First-Person Perspective is not compatible with the subjective experience in these patients. Moreover, Parfit's definition of the brain as a purely mechanical and technical device remains incompatible with empirical data from brain implantation. This raises the question for the type of states by means of which the brain can be defined in ontological regard thus requiring „neuroontology“.

„Neuroepistemology“ and „neuroontology“ might provide us with an empirically plausible epistemic and ontological definition of the brain. Both disciplines will allow to define the notion of brain identity and to develop criteria for it. Explication of brain identity will in turn contribute to empirical and conceptual clarification of the relationship between personal identity and the brain.

Acknowledgements

I am indebted to W. Oertel and A. Kupsch for cooperation in the empirical investigation of the Parkinson patients. I thank all patients who revealed me more about personal identity, personality and the brain than any textbook could ever provide. The study was financially supported by grants from the German Research Foundation (Habilitationgrant, 304/1-1; Heisenberg grant 304/1-4).

References

- Beauchamp, T. L.; Childress, J.F., 1994: Principles of Biomedical Ethics. Oxford: Oxford University Press
- Birnbacher, D., 1995: Identität der Persönlichkeit und Identität der Person. Philosophische Fragen im Zusammenhang mit der Transplantation von Hirngewebe. In: Zentralblatt für Neurochirurgie 56, p. 180-185
- Birnbacher, D., 1998: Hirngewebestransplantationen und neurobionische Eingriffe – anthropologische und ethische Fragen. In: Jahrbuch für Wissenschaft und Ethik 3, p. 79-96
- Birnbacher, D., 2002: Der künstliche Mensch – ein Angriff auf die menschliche Würde? Aachener Studien zur Technikphilosophie.
- Brundin, P.; Pogarell, O.; Hagell, P.; Piccini, P.; Widner, H.; Schrag, A.; Kupsch, A.; Crabb, L.; Odin, P.; Bjoerklund, A.; Brooks, D.; Marsden, C.; Oertel, W.; Quinn, N.; Rehncrona, S.; Lindvall, O., 2000: Bilateral caudate and putamen grafts of embryonic mesencephalic tissue treated with lazarooids in Parkinson's disease. In: Brain 123, p. 1380-90
- Mayring, P., 1990: Qualitative Untersuchung. Heidelberg: Beltz
- Metzinger, T., 2003: Being no one. Cambridge/MA: MIT Press
- Nagel, T., 1979 : Mortal questions. New York: Cambridge University Press
- Nagel, T., 1986: The view from nowhere. New York: Oxford University Press
- Nagel, T., 1987: What does it all mean? New York: Oxford University Press
- Northoff, G., 1995: Hirngewebestransplantationen. Medizinische und ethische Probleme. In: Ethik in der Medizin 7, p. 87-92

- Northhoff, G., 1995: „Ethics of the brain„, in brain tissue transplantation. Preservation or alteration of personal identity. In: Bulletin of the ESPMH, 3–4
- Northhoff, G., 1996: Brain identity in brain tissue transplantation. Preservation or alteration of personal identity? In: Journal of Medical Ethics 2, p. 5–18
- Northhoff, G., 2001: Personale Identität und operative Eingriffe in das Gehirn. Paderborn: Mentis
- Northhoff, G., 2003: Are we our brains? A neurophilosophical account of personal identity. In: C. Kanzian, J. Quitterer, E. Runggaldier (eds.): Wittgenstein Proceedings
- Northhoff, G., 2004: Philosophy of the brain – The Brain Problem. New York: Benjamins
- Parfit, D., 1989: Reasons and persons. Oxford: Oxford University Press
- Roskies, A., 2002: Neuroethics for the new millennium. In: Neuron 35, p. 21–3
- Stone, J., 1988: Parfit and the Buddha. Why there are no people. In: Philosophy and Phenomenological Research 47, p. 519–532

Appendix 1

Questions for the Semistructured Interview

1. How do you define the terms ‚personality‘ and ‚personal identity‘?
2. Did your personal identity, personality and individuality change after surgery? Are you another i. e. different person?
3. Did your relation to your brain and your body changed in the course of the disease and after surgery?
4. Do you have the feeling that you have a personal relationship with the fetal cells/electrode? Is the implant ‚part of yourself‘ and your brain/body?
5. Do you think that the fetal cells/electrode need some kind of ‚psychological integration‘?
6. Do you experience changes in your psychological functions after surgery?
7. Do you experience changes your motor functions after surgery?
8. Has your relation to the environment changed during the disease and after surgery?

9. Do you think that your mental states changed after surgery? Do you have a different consciousness of yourself and others after surgery?
10. How do you compare the brain with other organs of the body? Do you attribute a special status to brain implants as compared to those in other organs?

Appendix 2

Visual-Analog Scale for Subjective Experience of Personal Identity and Personality

1. My movements improved after surgery. _____
_____ My movements did not improve after surgery.
2. My perceptions did not change after surgery. _____
_____ My perceptions changed after surgery.
3. The experience of myself did not change after surgery. _____
_____ The experience of myself changed after surgery.
4. My emotions and my mood did not change after surgery. _____
_____ My emotions and my mood changed after surgery.
5. My will did not change after surgery. _____
_____ My will changed after surgery.
6. After surgery I can command my body according to my will. _____
_____ After surgery I cannot command my body according to my will.
7. I feel no longer locked within my own body after surgery. _____
_____ I still feel locked within my own body after surgery.
8. My psyche did not change after surgery. _____
_____ My psyche changed after surgery.
9. My individuality did not change after surgery. _____
_____ My individuality changed after surgery.
10. My personality did not change after surgery. _____
_____ My personality changed after surgery.
11. I did not become another person after surgery. _____
_____ I became another person after surgery.
12. I do not have a different awareness of myself after surgery. _____
_____ I have a different awareness of myself after surgery.
13. I do not have a different awareness of my body after surgery. _____
_____ I have a different awareness of my body after surgery.

14. I do not feel isolated from the environment after surgery. ———
————— I feel isolated from the environment after surgery.
15. Despite the surgical therapy I remain the same person. ———
————— Due to surgical therapy I became a different person.

Subjects had to make a cross on the line between the two contradictory statements according to the degree of the appropriateness of either statement.

Philosophia naturalis

Band 41 (2004) Heft 2

MICHAEL DRIESCHNER
Die Quantenmechanik – eine Revolution der
Naturphilosophie?

PETER MITTELSTAEDT
Interpretationsprobleme der Quantenmechanik

CORD FRIEBE
Zur Bewegung der Galaxien und des Raumes

GEORG NORTHOFF
„Am I My Brain?“ Personal Identity and Brain Identity

EDUARD KAESER
Der personale Körper – ein erkenntnistheoretischer
Leitfaden

THOMAS SUKOPP
Was ist und was leistet Menschenwürde?



Sonderdruck

VITTORIO KLOSTERMANN · FRANKFURT AM MAIN