IS THERE NEUROPSYCHOLOGICAL EVIDENCE FOR A THEORY OF INTENTIONAL STATES?

Helmut Hildebrandt & Eckart Scheerer

'Representation' and 'Intentionality' in current cognitive science

The concepts of representation and intentionality have a long and diversified history in philosophy (see SCHEERER, 1992, for the representation concept), but here we will be concerned with how they are used at present. Despite the protean nature of their historical use, most current philosophers of mind and cognitive scientists will agree with the following working definitions (SCHEERER, 1993). *Representations* are internal states of a system that have some content, i.e., that bear some referential relation to objects different from themselves and prototypically existing outside of the system. The referential relation enables representations to serve as internal 'placeholders' of the objects for which they stand, provided that there is some lawful relation, or at least statistical correlation, between the represented object and the representation. Thus defined, representations are not necessarily mental states but could occur in any system that has internal states which are deterministically or stochastically dependent on its environment. In order to prevent such an inflation of the concept, it is usually stated that there must be some „interpretation" of the internal state as being representational. In the context of this „semantic interpretation", the concept of *intentionality* is invoked. Other than the ethical or judicial use of 'intention', 'intentionality' does not only refer to mental states underlying voluntary, purposeful or planned behavior, but to the much broader class of *semantically interpreted representational states.* Expressed differently, representations are individuated by *reference* and intentional states by *meaning.* Depending on whether the subject of the interpretation is outside or inside the representing system, it is useful to distinguish between *extrinsic* and *intrinsic* intentionality (SEARLE, 1980). For instance, any notational system has extrinsic intentionality once it is deciphered, but for a language user his/her own speech acts have intrinsic intentionality, though this is *derived* from the mental state which is the locus of *primary* intentionality inasmuch as speech acts are linguistically expressed thoughts. Thus, we can define the concept 'mind' as 'system endowed with *intrinsic, primary* intentionality.' For the sake of brevity, our own use of 'intentionality' and 'meaning' will be restricted to the notion of intrinsic, primary intentionality. Readers who find this use puzzling are invited to substitute 'subjective experiential meaning' for 'intentionality'. At any rate,

we want to make it clear that while there is nothing wrong with the notion of a non-mental representation, intentional states are by definition mental states, though depending on one's approach to the mind-body problem they may or may not be brain states as well.

With few exceptions, such as GIBSONian 'ecological realism' (TURVEY et al., 1981), current cognitive science subscribes to an asymmetric relation between representation and intentionality: intentionality presupposes representation, but not the other way round. This is especially clear in the computational theory of mind (e.g., PYLYSHYN, 1984) which regards the mind as a computational system capable of multiple instantiation by different physical systems, including man-made systems. In the philosophical underpinning of the computational theory (e.g., FODOR, 1981), the key concept linking representation and intentionality is '*propositional attitude'*. As the only carriers of meaning within a system, propositional attitudes have two components: a representational state in propositional form, and some relation of the system to the representational/propositional state, belonging to the two broad classes of *belief* and *desire.* For instance, if *p* is the proposition that can be expressed in the ordinary-language sentence 'The sun is shining' and *X* is a subject, then the structure of meaning is captured by expressions such as '*X* believes that *p'*, '*X* hopes that *p*', '*X* wishes that *p'*, '*X* fears that *non-p*', etc.

At first sight, the concept of propositional attitude seems perfectly suited for explaining the occurrence of representation/intentionality dissociations, since it presupposes the ability of the theoretician to „factor out" the constant „propositional core" of propositional attitudes from the various „attitudes" in which they figure, and a similar „factoring out" seems feasible for the attitudes themselves. Nevertheless, such a way cannot be taken. Let us mention some (but not all) reasons for that. First, according to the computational approach intentionality (i.e., meaning) is not a mental state with „absolute", independently verifiable properties but resides exclusively in the *relation* between the attitudinal and the relational component. Thus, fear is individuated neither by experiential qualities nor by physiological indicators, but by a formal (actually, even „computational") structure which itself has propositional format („*X* fears that *p*" is as much a proposition as „*p*", though with some special properties; for instance, it does not support the inference to *p.)* Second, the relation must always be to some *specific* proposition, rather than to a class of propositions or to some state of the world that cannot be captured in a proposition. Thus, „diffuse" or „objectless" emotions cannot find a place in the computational account and in fact they are dismissed as belonging to physiological psychology (e.g., PYLYSHYN, 1984.) Finally, and most importantly, propositions are *types,* or classes, of *sentences* (e.g., 'The sun is shining' and 'Die Sonne scheint' are tokens of the same proposition) and these. in turn, are types of *utterances* (e.g., spoken vs. written,. male vs. female voice etc*.)*. However, according to the

computational theory mental and physical (i.e., brain) states are lawfully related at the token, not at the type level (FODOR, 1981). It follows that while brain states and utterances may be lawfully related, this will not be the case for brain states and propositions, nor -(because attitudes relate to propositions) will there be a lawful relation between brain states and propositional attitudes. As a result, it is unlikely that the attitude-proposition relation will be disrupted by impaired brain functioning, arising as it does from conditions such as arterial distribution, localization of tumors, lesions etc. which cannot be brought „into register“ with an abstract, algorithmic structure as envisaged by the computational theory.

These consequences of the computational approach put rigid constraints on the feasibility of a naturalistic account of representation and meaning, which is indispensable from the perspective of neuroscience. This did not escape the notice of authors who are in pursuit of a naturalistic program in the philosophy of mind. One increasingly popular way to naturalize the mind is to adopt an *externalistic theory of meaning*. The computational theory of the mind takes a frankly internalistic stance to the emergence of meaning; the meaning of representations arises from their causal role within the system, i.e., from their insertion into the „propositional-attitude“ scheme. In contrast, externalistic theories seek to derive meaning from the covariation of representational states with the environment. Externalistic theories provide a reasonable starting point for a neuroscientific approach to intentionality. While internalistic theories link the properties of intentional states in a „top-down“ fashion to the existence of propositional representations and thus to a „language of thought“ (FODOR, 1975). externalistic theories start with the evolutionary and ontogenetic beginnings of meaningful states and work their way „bottom-up“ to higher forms of cognition. Within this general framework, DRETSKE (1988) distinguishes between the mere covariation of environmental events and organismic processes and a covariation which was learned to indicate the environmental event and assigns representational status to the latter only. Thus, representations differ from causal effects in virtue of their *indicator function,* which is based on the *learning history of the individual organism*. Moreover, the fact that the indicator function is acquired by instrumental learning may serve as a starting point for naturalizing the „attitudes“ without tying them to propositions in the first place. According to DRETSKE (1988, ch. 5), all kinds of motivation start with „pure desires“, defined by the receptivity, selected by evolution, of an organism relative to the outcome of its own behavior. In the course of instrumental learning, the past outcomes of an organism’s behavior serve as a structuring cause of the present behavior, and this results in the acquisition of a „cognitive skill“, an „implicit belief“ about what kind of behavior produces the desired result. For DRETSKE (1988, p. 127) desires are not representational states, and yet they share all the properties of intentional states as usually defined. In short,

„pure“ desires are intentional without being representational. Dretske does not deny the existence of explicit beliefs as mediators of purposeful behavior, but derives their emergence from interactions (including conflicts) between multiple beliefs and desires.

Dretske's externalistic theory of meaning has a number of attractive features for an evolutionary and ecological analysis of cognition. (1) It adverts to basic evolutionary principles such as survival value and by this means provides some functional points of view which are needed to complement a purely physicalistic description of the mind/brain (ROTH & SCHWEGLER, 1995). (2) By focusing on the indicational function of representations, it suggests a plausible difference between the representational and the causal mode of explanation. (3) The pivotal concept of the theory - the learning history of *individual* organisms - serves to delimit natural living systems from man-made control systems.

These advantages notwithstanding, we do not believe that Dretske has succeeded in formulating a *comprehensive* theory of intentionality. In our view. his theorizing suffers from insufficient differentiation between the objective and the subjective (experiential) aspect of intentional states. On the one hand, his theory is meant to provide an objective (third-person) account of the external and internal conditions necessary for the emergence of representation and meaning. On the other hand, DRETSKE (1988, p. 117) emphasizes that intentionality is a first-person phenomenon, that the capacity for having genuinely intentional states exists only insofar as the *agent* knows that his/her behavior produces the desired or intended result. These two aspects can be maintained at the same time only under the condition that the objective aspect serves as the *explanans* and the subjective aspect as the *explanandum.* The theory is reductive in that it tends to replace the explanandum by the explanans, i.e., subjective states by their objective conditions. This has two consequences. which are on the face of it opposed to each other but in fact flow from one and the same premiss.

The first consists in *over-extending subjectivity*, i.e., in the attribution of intrinsic intentionality to system states and/or processes that are devoid of it if other criteria are applied. Whenever a system fulfills the prerequisites for „goal-intended behavior“, it is, on Dretske's account, a system possessing first-person intentionality. This criterion is too broad, however. For according to Dretske, intentionality emerges from the process of instrumental, or operant conditioning, with the proviso that what is learned is not some specific „behavior pattern“ to be described in physicalistic terms, but the behavior of the organism as a whole insofar as it is subordinated to goals that are intended by the organism. Now, it is a basic property of operant conditioning that it may function, in principle, without awareness. While the extent of „conditioning without awareness“ in *awake* humans is under

dispute (e.g., BREWER, 1974), the occurrence of aversive conditioning during *sleep* has been demonstrated (ANTROBUS, 1990). Thus, certain environmental properties are coupled with an objective (physiological) emotional response to bring about a specific piece of behavior. To the extent that conditioning could be extended to goal-anticipatory responses while the subject is asleep (whether or not this is the case is an empirical question), the complexity of the underlying conditioning process and its behavioral „product" is not different from operant conditioning in the waking state. In this case, a state would be achieved about which it one could say - from a 3rd-person perspective - that the sleeping person behaves in that particular way *because* he/she believes that this behavior helps to avoid the anticipated aversive consequences. Following DRETSKE (1988, pp. 115 ff.) this would satisfy all criteria for „goal-intended behavior". Yet aversive conditioning while asleep is quite different from a behavioral act motivated by *experienced* emotional arousal in the wakeful state. Conditioned responses acquired during sleep are of no consequence for the behavior of the awake individual. Emotions are not states of the *organism as a whole*, but properties of the *organism in a particular state*, such as sleep. drowsing, wakefulness, hallucination, etc. (HOBSON et al., 1986). The reason for this is that the brain is organized in such a way that different, multilayered structures are responsible for emotional states, and as a consequence emotions can arise independent of, or parallel to, the representational analysis and decoding of environmental properties (LE DOUX et al., 1988). In other words, the Dretskian „pure desires" can function in operant conditioning without giving rise to „cognitive skills" that are available to the organism or subject as a whole. The structuring of behavior by representations and intentions is more „modular" than envisaged by Dretske's holistic account. One consequence of this is that emotions, inclusive of their coupling to environmental conditions, can occur outside of the realm of conscious awareness.

A second consequence is that the theory *does not capture experiential features of subjectivity.* To be sure, Dretske refuses to consider objectively goal-directed behavior as satisfying the criteria for intentionality and instead requires the possession of *knowledge* about the production of intended effects by one's own behavior. In itself, this is a great advance over the computational theory, where meaning could emerge in a well-designed, „intelligent" computer program. Nevertheless, Dretske is not really concerned with „how it feels" to be in an intentional state; his is a theory of structured behavior, not a theory of conscious awareness. Others have been more urgent on this point. SEARLE (1990, p. 585) insists that „we have no notion of intrinsic intentionality except in terms of its accessibility to consciousness". Van GULICK (1988, p. 84) has summed up a fundamental axiom of the theory of meaning as follows: „A capacity for having conscious subjective experience is a necessary condition for having any states with

intrinsic intentional content“. While we are fully in accord with that statement, we want to emphasize the expression ‘conscious subjective experience’ rather than ’capacity for’. That is, we consider it necessary to characterize the phenomenal properties of *intrinsically meaningful states*, over and above the conditions that lead to their emergence, which we take to have been satisfactorily analyzed by Dretske.

In advocating a phenomenological approach to intrinsic intentionality, we shall, for the moment, not be concerned with its emotional/motivational aspect; instead, we shall focus on the relation between the experiencing subject and the intrinsic meanings that he/she experiences. What we wish to emphasize is that representations get their meaning in virtue of their content being integrated, in a particular way, into the conscious experience of the subject, as opposed to the syntactic/formal processing perspective adopted by the computational approach. More specifically, we shall discuss this peculiarity of intrinsic intentionality under the two headings of *transparency* and *dynamic integration* and exemplify them by means of a brief excursion into Gestalt theory.

The Gestalt approach to intentionality: transparency and dynamic integration

First, let us briefly clarify our perspective on Gestalt psychology. In what follows we accept Wolfgang Köhler’s stance that by applying the principle of mind-brain isomorphism it is permissible to infer the dynamics of brain processes from the dynamics of experience. One of us (SCHEERER, 1994) has attempted to demonstrate the contemporary relevance of the isomorphism principle, which rests, in part, on an evaluation of KÖHLER’s ideas on the dynamic self-organization of brain processes that is more positive than their current image as „brain mythology“. However, readers who do not wish to follow us in that direction may cling to the current notion of Gestalt psychology as a phenomenologically oriented approach and still accept (or reject) our argument.

In the propositional-attitude approach to intentionality, its properties are usually illustrated by reference to an „intentional idiom“ (consisting of expressions like „*X* believes that *p*“ etc.; see above) which has the distinguishing mark of *referential opacity* (e.g., DENNETT, 1987, p.240 f.). That is, in propositions assuming the „that p“ place in a statement expressing a propositional attitude, it is not allowed to substitute a given term by another term with exactly the same reference. Thus, while ‘Goethe’ and ‘the author of Werther’ have the same reference, we must not substitute one for the other in a sentence like *„Napoleon knew that Goethe was the author of Werther“*, since this would result in an empty tautology. While we do not deny that intentional statements have this particular logical form, we

think that it is misleading to derive the properties of intrinsic intentionality from them. Rather, in describing everyday experiences, people use a non-intentional idiom (e.g., „this book has a beautiful brownish leather binding“) and the intentionality resides not in the linguistic expression but in the experience itself. It was a historical achievement of the Gestalt psychologists to encourage the use of everyday language in psychological experiments, as opposed to the „analytical introspection“ advocated by the Würzburg school and TITCHENER. For the use of theoretically non-committed „thing language“ brings out a basic property of immediate experiences: their full *transparency* to the experiencing subject. While the subject may err with respect to the *referential object* of the experience (the „leather“ binding may actually be plastic), the *experience itself* (it looks like a brown leather binding) cannot be disputed. To be sure, this property of immediate experience was not „discovered“ by Gestalt psychologists and under the guise of the „incorrigibility thesis“ it figures in very diverse philosophical approaches; but there it is often confounded with notions such as „qualia“ or „sense data“, which force a non-intentional interpretation on immediate experience.

The second aspect was introduced when the Gestalt psychologists began to consider intentionality (HILDEBRANDT, 1991). It concerns the form of integration of a semantic content into an intentional act. Discussing the experience of an overheated radiator, KÖHLER wrote:

> „The direct awareness that the discomfort is „coming from the heat“ is sufficient for this process to be determined by that other brain event which is phenomenally expressed as „heat there“. But the tendency to move away from the heat is also directly perceived as a natural result from the situation. Thus, we have to draw the same physiological conclusion: From, or better, in the joint brain field event which parallels „my discomfort from the heat there“, a ... vector has to arise which phenomenally is expressed as a tendency away from the heat and which physiologically represents, in the brain field, a tendency to increase the distance between ego region and the correlate of „heat there“. In natural science, a tendency to increase the distance between two physical regions is called a force field. So we can say that within the processes involved a force field occurs that tends to increase the distance between the processes corresponding to radiator and heat, and the ego region“ (KÖHLER, 1933, p. 249, translated by the authors; see a similar statement in English in KÖHLER, 1947, p. 356 f.)

Several points come to mind when reading this description.

Most importantly, the relation between intentional act and semantic content is not of the static nature as envisaged by contemporary theories of intentionality; rather,. it has to be described in terms of the violation and restoration of *equilibrium states*. In the example given by KÖHLER, the restoration of equilibrium will be brought about by overt action, but the studies of the POSNER group (e.g., POSNER & ROTHBART, 1992) on covert attention shifts show that new equilibrium states can be attained without external movement of the organism or its sensory systems.

Another point concerns the implication of the Ego as one term of the intention-content relation. Intrinsically meaningful experiences presuppose *integration into the Ego system*. (To a certain extent, this is acknowledged in the computationalists' insistence on „attitudes" being a relation between *system* and representation.) In his example, KÖHLER stresses the Ego as a locally segregated subsystem of the phenomenal (and brain) field. While this is adequate in the case of perceiving external events and acting on them, a more general sense of 'Ego' is given in William JAMES's (1890) metaphor of the *stream of experience,* which encompasses - in the form of the 'Me' - the Gestaltists' notion of the perceived Ego. Leaving aside JAMES's notion of the „pure Ego" (which seems to be an epistemological construct and in this capacity was abandoned in his later work), the stream of experience serves as the subject of single experiences without being a separate subsystem. Rather, it is the (dynamic) structure of experience itself, and single experiences reveal their incorporation into this structure by a feeling of intimate belongingness. In physiological terms, all we can say at present is that intrinsically meaningful experiences never can correspond to local brain processes as such but to the relation they bear to some overall brain state which serves as the correlate of „Egohood". While in a very rough manner the perceived Ego might perhaps be „localized" in the parietal lobes and the Ego as an actor in the frontal lobes, Ego-belongingness must arise from the joint action of (at least) these subsystems, presupposing the functioning of subcortical feedback circles.

Finally, we note, from KÖHLER's description, a discrepancy between the (objective) relational structure of meaningful experiences and their (subjective) immediacy and unity. We are aware of the difficulties inherent in the notion of top-down causality implied by the Gestalt notion of holistic determination of part processes; nevertheless, we are prepared to accept the minimal conclusion that emotionally relevant experiences have a *proprietary representational basis* that is different from, and functions in parallel with, the representation involved in „cold", distant cognition, rather than being a secondary response to the latter. The „attitudes" of the computationalists, in contrast, all have access to the *same* representation.

In a speculative vein, one might even deduce the incorrigibility of intrinsically intentional states from the dynamic integration process. Inasmuch as representations are caused by environmental events, they are certainly the result of local, modular brain processes, and these, while overall reliable, are error-prone with respect to specific cases; as is well known from perceptual illusions, the perceptual system may be „cheated" for a variety of reasons. On the other hand, knowledge that the perceptual system has been cheated with respect to some referential objects or their properties does not affect their experienced properties. The reason would be that once a local, modular process is integrated into an overall, relationally determined organization, it is no longer open to modification from other local

brain processes, as long as the relation remains in force. In other words, incorrigibility would result from the invariance of relations with respect to the local processes that constitute them.

The preceding considerations imply a methodological conclusion. While phenomenological description is adequate and indispensable for revealing the basic properties of experiences, it cannot reveal the objective relational structure that underlies them. In order to „carve nature at its joints“ we need a carving knife. Actually, we have at least two, but in the rest of this paper, we shall be concerned with only one of them, viz., the effects of „natural“ or artificial brain lesions. But we want to emphasize that some if not all of the neuropsychological dissociations can also be produced, in the intact brain, by means of psychological experimentation. Neuropsychology and „normal“ experimental psychology complement each other. However, as a rule the neuropsychological dissociations are more impressive, and though their precise analysis requires experimental techniques, sometimes the clinical symptomatology is almost self-explanatory.

Intermediate summary and preview

Let us sum up some conclusions we have reached in the preceding paragraphs. Representational states are founded on the covariation or correlation between environmental events and states of the CNS and they acquire their representational *function* in virtue of the learning histories of individual organisms. At a biological level, *emotions* arise as a primordial form of intentionality, but they are not necessarily associated with awareness and may subserve organismic subsystems, rather than the organism as a whole. For intentionality at a human level, we have to start from *conscious experience*. In the computational version of cognitive science, intentional states are defined as propositional attitudes and intentionality presupposes representation. However, the reduction of intentionality to representations plus attitudes is open to some major objections. A phenomenological account of intentional states reveals that the relationship between a conscious intentional system and a representational state is not computational or referential, but involves dynamic interaction between the Ego system and the experiential qualities of the representational state.

In the following, we shall defend the thesis that under certain pathological conditions revealed by neuropsychological research there is a *dissociation* between intentionality and representation. That is, intentionality may be disturbed or eliminated in spite of virtually intact representational states, or it may function in parallel to and perhaps even independent of representational states. The first case has two varieties. In the first variety,

exemplified by *blindsight* and some other neuropsychological syndromes, some representational states loose their transparency to the subject altogether. In a second variety, illustrated by cases of *abnormal pain experience,* intact representations are no longer emotionally evaluated and as a consequence they loose their subjective meaning to the patient. The opposite case - intact intentionality in the absence of any kind of representation - is difficult to demonstrate because in order for an emotional evaluation to occur, a minimum level of external stimulation must reach the brain and there be subjected to an initial perceptual analysis. (Otherwise, we would have extrasensory perception!) Accordingly, we confine ourselves to cases where cognitive and emotional processes seem to occur in parallel for different parts of experiential space, as in *unilateral neglect,* or in the absence of conscious recognition, as in *prosopagnosia* and some associated disorders.

Blindsight: Representation without intentionality

In animals, lesioning the primary visual cortex does not result in complete blindness. Though the animals fail to respond to stimuli arriving at the lesioned areas, some visual capabilities can be demonstrated through discrimination learning. Contrary to this, for a long time it has been the majority opinion that with humans comparable injuries always lead to a *complete loss* of visual processing capacities. Not until the 70s, PÖPPEL, HELD & FROST (1973) offered evidence that hemianopics are capable of fixating stimuli projected into their blind visual half-field in an above-chance manner. Despite their subjective blindness, the eye movement of such patients correlated with the position of the presented stimulus.

This observation was supported and confirmed by a series of further studies. Hemianopic patients were presented with stimuli in their blind visual half-field and, though they had no subjective knowledge about the presence of the stimulus, they were asked to guess its location or other properties. In this way, a number of residual visual capacities was demonstrated (MARCEL, 1988; WEISKRANTZ, 1989). The current list comprises the following:

- Localization of stimuli by fixating them or pointing to their position
- Hue discrimination
- Discrimination of movement direction within the blind region,
- Discrimination of simple shapes
- Figure completion within the blind field, provided that the perceived figure actually extends into the blind region
- Reaching for objects in the blind region, where the movement bears an above-chance relation to the size, form and location of the object.

Despite such achievements the patients denied any subjective knowledge of the visual stimuli, and accordingly Weiskrantz called this phenomenon „blindsight“. For the residual visual capacities of destriated animals two types of explanation had been developed (WEISKRANTZ, 1980). The first model postulates a *quantitative* degradation of all primary visual capacities. The second model assumes a *qualitative* alteration of visual processing: some kinds of visual representation are deficient, while others are not affected by the lesions. Transferring these models to humans, blindsight could be explained either in terms of an *overall representation deficit* or of the *failure of a group of intact representations to bring about behavior appropriate to them.*

Empirical decision between these two possibilities is difficult. Other than in the monkey, the human striate cortex is not expanded along the outer surface of the skull but is buried in the medial surface in the cortical fissure dividing the two hemispheres. Consequently, occipital damage will usually affect secondary visual areas in addition to striate cortex. Therefore, the best reports on blindsight come from patients having undergone unilateral surgical ablation of the striate cortex. In patient D.B., who fulfilled this condition, WEISKRANTZ (1986) found that some spatial detection tasks were performed equally well or even better in the „blind“ visual half-field, whereas there was better form discrimination in the „good“ field. There was a *double dissociation* between detection and form discrimination, showing that deficient perception in one part of the visual field does not imply a deficiency in *all* visual representation processes. The fact that blindsight patients do not have visual experiences in their blind field, then, cannot be caused by quantitative degradation or qualitative alteration of visual *representation (*for a similar interpretation see SCHACTER, 1992).

The relevance of blindsight for the intentionality issue becomes clear when we look at the subjective reports which are typically given by patients participating in tests for blindsight (WEISKRANTZ, 1980). Patient K.P., for instance, was able to discriminate the horizontal and vertical orientation of lines with more than 70% accuracy. However, he reported that he actually had seen only a dimly flashed spot. Patient T.F.H. asserted even more strongly that he had recognized no line at all, yet he reached 77 % correct discriminations. Patient E.E.H. produced 80% correct estimates of line orientations. Nevertheless, he believed that he had just reached random values. He stated that everything had been a mere case of speculation. Patient R.U., who was even more seriously impaired than the other subjects, produced good results in grasping and in the discrimination of movement directions; yet she reported that she had not the faintest notion of how she had been able to do this. There are some patients who obviously are capable of similar achievements, but refuse to participate once more in similar experiments, because they do not want to „speculate“ and „invent“

any longer what could be happening in a particular environmental region of which they are not consciously aware (WEISKRANTZ, 1988, p. 188f).

These assertions should be taken seriously. Spontaneously, the patients did not respond to the corresponding stimuli projected into the impaired visual field, and they are often seriously disabled in everyday life. Based on informal observations, MARCEL (1988, p. 146) believes that a thirsty blindsight patient will not reach for a glass of water shown in the blind field, though he/she should be able to distinguish it from other objects in the testing situation. Artifacts such as stray light, remaining striate tissue and criterion shifts have been suspected (CAMPION et al., 1983), but can be excluded (see the discussion papers in CAMPION et al., 1983). Granted, there are representation deficits in blindsight, owing to selective loss of the Pβ system that forms the bulk of the retino-geniculate-striate pathway, with the consequence that owing to transneuronal degeneration the cell loss affects even the corresponding retinal ganglion cells (STOERIG & COWEY, 1993). But concerning those visual capabilities that remain unaffected, blindsight is an example for a dissociation between the *presence* of representational states (defined in terms of the covariation criterion; see above) and their *significance* for a particular patient and her/his behavior. In the blindsight patient's everyday life, intact representations are simply not utilized; not integrated in her/his ongoing behavior; they become relevant only when he/she is „pushed" or „goaded" to respond in the test situation. Blindsight, then ,. is a case of *representation without intentionality*. This, of course, does not mean that the patients lack intentionality altogether, not even in their „higher" visual capabilities; blindsight is not necessarily associated with agnosia, visuo-constructive disorders etc. The existence of blindsight may be taken as empirical evidence for the thesis that representation and intention are situated at two different levels. The covariation of representational states with environmental objects does not guarantee, in itself, that intentional states stand in a similar referential relation to their objects.

Dissociations similar to the blindsight phenomenon are observed in amnesic, dyslexic, and aphasic patients (e.g., SCHACTER, MCANDREWS & MOSCOVITCH, 1988; SCHACTER, 1992). In all of these cases, there is loss of subjective, intentional access to unimpaired cognitive processes (cf. SCHACTER, 1992). Taking our cue from the blindsight phenomenon and the theoretical notions developed above, we suggest that the lack of transparency cannot be a purely local effect of the relevant brain lesions, but must be due to the resultant failure of the relevant brain processes to be integrated with those brain processes or their properties which are responsible for intentionality to occur.

Dissociations of pain experience

In blindsight and the other neuropsychological conditions just mentioned, intentionality is impaired in its *cognitive* aspects: either there is no subjective experience at all, or the characteristic feeling of familiarity (which JAMES considered as the defining mark of an experience belonging to the stream of consciousness) is absent. Now we want to discuss the case where the *emotional* relation to an intact representation is missing or substantially altered. Our examples come from certain dissociations of pain experience to be observed in patients who have undergone „psychosurgery".

Traditionally, pain has been defined as an emotional state opposed to pleasure and playing an important role in learning and the formation of visual images. It was accorded the status of a specific sensory modality only in the last quarter of the 19th century, although this always remained disputed (GRAHEK, 1993; SCHEERER, 1995). Today, most scientists working in the field will probably agree that pain encompasses emotional and cognitive processes in addition to, and sometimes independently of, the input from specific „nociceptors" (GRAHEK, 1991; 1993).

From a philosophical point of view, BIERI (1987) put together the reasons which argue against a reduction of pain to sensation. In particular, he emphasizes that pain perception is non-projectible (i.e., it cannot be referred to external objects) and that talk about perceptual illusions does not make sense in the case of pain perception. Bieri considers pain experience as a unity of emotional, cognitive and perceptual processes. It is this unitary nature and the particular subject-relation experienced in it which DENNETT (1978) offers as an argument that computers can't feel pain. While it may be feasible to simulate the effects of pain on behavior, instantiation of pain *experience* in an automaton or a robot is impossible.

Although unitary, pain experience has cognitive, emotional and sensory *aspects*, and it may be asked whether dissociations between these aspects can occur as a consequence of brain damage. Here, we are interested in a dissociation between pain as a representational state and the intentional pain experience. However, in the literature there is little evidence that the central, representational or intentional processes which are basic for pain could be affected in an isolated manner. To our knowledge, there is no clearly documented case of *pain agnosia,* perhaps because there are many different afferent pain pathways, some of them non-specific, and these project in parallel unto different brain areas (KURTHEN, 1984; MEINHART & McCAFFERY, 1983; LARBIG, 1982). On the other hand, loss of the emotional or motivational component of the pain experience has been established. The brain surgery known as *prefrontal lobotomy*, carried out during the 40s and 50s in the USA, should be mentioned in this context. In these operations the connections of the BRODMANN areas 8, 9, 10, 11, 32,

46, and 47 in the frontal lobe were disconnected from the rest of the brain[1]. Among the operated patients there were also some people who suffered from pain arising from cancer, tabes dorsalis, causalgias, phantom limb sensations, neuralgias etc. FREEMAN & WATTS (1950) describe among other cases of functional pain complaints the following. A woman of „hysterical temperament“ reported for the first time at the age of 16 that she suffered from chronic abdominal pain. In the following time she was operated 12 to 18 times with no fundamental improvement of her condition. After a minor brain injury she reported that now she was suffering from backache instead. Before being operated by FREEMAN and WATTS, she had been in bed for two years, lying most of the time on her left side because her pains did not allow any movement. Then a prefrontal lobotomy was carried out. Two days after the operation, the woman asserted that her backaches had disappeared. She allowed tactile stimuli at the former pain region and reported that this was no longer disagreeable. Though stretching her legs was associated with pain, the woman liked to do that after having been in bed with her legs bent for several months. Ten days later she could be discharged from the hospital and was able to walk by means of an aid. In a medical follow-up she reported that she was fine and had no problems. At the end of the interview the woman got up and left the room without an aid and without any sign of pain. Six months later, she was capable of gainful employment, and she remained employed for the next ten years.

Contrary to what one might expect, the cases documented by FREEMAN & WATTS (1950) are not instances of cortical pain blindness[2]. CHAPMAN, SOLOMON & ROSE (1950) and KING, CLAUSEN & SCARFF (1950) examined several prefrontally lobotomized patients for a possible change in the pain threshold. A thermal stimulus was applied to one cm^2 on the upper arm and it was recorded from which temperature on the patients reacted with a defensive movement. If there was a change at all compared to normals, it was in the wrong direction i.e., a small lowering of the pain threshold as a result of the operation.

Careful clinical examination has revealed a constant pattern: no pain agnosia. The patients reported that their pain was continuing, but only when questioned directly. Nevertheless, the persisting pain was hardly ever expressed in their overt behavior (cf. for instance the dialogue with patient No. 398 in FREEMAN & WATTS, 1950, 362; as well as the accounts in DYNES & POPPEN, 1949; NEMIAH, 1962). Prefrontal lobotomy typically dissociates particular aspects of the experience of meaning from the

[1] It should be noted that these operations were inhuman because of the deep alteration of the operated subjects' personality and that they are no longer carried out.

[2] In view of the different operation techniques which were carried out, it can be supposed that the results were of a global nature and do not allow a reference to localized processing centers.

representational perception of pain. Accordingly FREEMAN & WATTS (1950, p. 354) comment on their findings as follows:

„Prefrontal lobotomy changes the attitude of the individual toward his pain, but does not alter the perception of pain. Whereas previous to operation it occupied the focus of his attention, after lobotomy pain fades into the background".

The fact that some patients with anxiety neurosis were also operated „successfully" lends further support to the assertion that this kind of surgery cuts off the emotional aspects of pain. Administration of morphine seems to have consequences similar to those of prefrontal lobotomy: loss of the emotional experience of pain (see the discussion in DENNETT, 1978).

Thus, prefrontal lobotomy supports the idea of a dissociation between emotional and representational aspects of experience. Let us admit, though, that we know of no investigation where changes in physiological parameters (skin resistance, heart frequency etc.) likely to occur in pain have been measured. Also, it is not understood which dimension of emotional experience actually is affected by prefrontal lobotomy. All prefrontal lobotomy patients exhibited profoundly changed personality traits in the direction of lessened emotional irritability with simultaneous flattening of affect. Moreover, under specific conditions there was a distinct pain response; but this behavior was tied to the situational context and did not outlast it.

In an analysis of the case reports by FREEMAN & WATTS, GRAHEK (1993) has tried to distinguish between two emotional reaction patterns for pain. In his opinion, the feeling of unpleasantness remains, while enduring emotional reactions like fear, depression, worry are eliminated. According to Grahek it is not the pain experience which is altered, but the attitude to it (see also ELITHORN, GLITHERO & SLATER, 1958). The brain loci subserving basic emotional responses were not affected by the surgery, and so it is indeed probable that the patients still were able to experience the pain as unpleasant. But it seems not correct to describe the loss of the experience of meaning as to pain states simply as a *different form of emotional reaction*. The downright „efficiency" of the surgery for quite diverse disorders, which hardly are in a logical, not to say an anatomical relation with each other, in our opinion needs to be explained through the loss of a subjective relationship to the content of the experience, which can be *evaluated* only be means of an emotional reaction. Clearly, the patients no longer take the pain perception as an experience which is significant to them (cf. DAMASIO, TRANEL & DAMASIO, 1991, for the role of the orbitofrontal lobe and the basal forebrain - structures which were severed by the surgery - in the emotional evaluation of stimuli). As in blindsight, the pain remains, but it is not integrated into ongoing behavior and into the stream of (intentional) experience. At any rate this is a possible interpretation of FREEMAN & WATTS (1950) who base their explanation on the concept of

the „self“ and think that the „efficiency“ of their method rests on the fact that relating the experience to the self becomes more difficult after the operation:

Unconscious preference in neglect patients

We have interpreted blindsight and altered pain experience in terms of a *deficient intentional stance to ongoing representational processes*. In the next two sections, we shall be concerned with evidence for a dissociation between emotional and representational processing.

In neuropsychology, neglect is defined as the inability to attend to the half of egocentric space which is contralateral to the lesion. Neglect occurs in the most serious manner and most frequently as a result of damage of the right inferior parietal lobe, but it can also be a consequence of left-parietal, frontal, or thalamic unilateral brain damage. Although neglect may be associated with sensory loss (hemianopsia in the visual case), the more typical case is that sensory functions, as assessed by EEG methods, are intact in the neglect patient (VALLAR, 1993).

Earlier neglect theories often presupposed that the relevant phenomena are modality-specific. While this may indeed be the case, more often an in-depth analysis reveals the neglect to be a supramodal phenomenon. Together with evidence on preserved sensory function, this has led to the conclusion that neglect results from an impairment of one or several *encapsulated centers of attention* (e.g., HEILMAN, VALENSTEIN & WATSON, 1985; GAINOTTI, D'ERME & BARTOLOMEO, 1991). The most convincing evidence against this type of attentional theory comes from the well-known studies of BISIACH and his group. BISIACH & LUZZATTI (1978) asked some neglect patients to imagine the cathedral square in Milano and to describe verbally whatever there was. It turned out that the patients - though they were not perceiving but recollecting their premorbid experience in the form of visual images - could describe one side of the place only. However, when they took the reverse perspective, the previously neglected side of the place was described, and the other one was now left out. BISIACH, LUZZATTI & PERANI (1979) presented patients with abstract visual patterns moving behind a narrow slit. This results in a translation of movement into space perception; i.e., when the pattern is moved from left to right, the parts that become visible first are perceived to form the right half of the figure and the other way round with movement to the left. Depending on the localization of the brain lesion and independent of the direction of motion, the patients described only one (typically the right) half of the figure,. thus ruling out an explanation in terms of recency effects. Nor could the effect be due to an unilateral loss of attention, because attention remained focused on the narrow slit behind which the figure appeared. Accordingly, Bisiach proposed a *representational deficit* theory of neglect, according to

which visuospatial hemineglect arises from an impairment in generating visual representations for the half-space contralateral to the lesion.

Current theoretical debates on neglect tend to soften the hard opposition between representational and attentional accounts, owing to a more cognitive interpretation of attention. According to HALLIGAN & MARSHALL (1994), there are two hemispheric contributions to the spatial distribution of attention. Each hemisphere controls attentional deployment to the contralateral half of visual space. In addition, the right hemisphere is tuned to the global „overview“ of the situation and the left hemisphere is concerned with local details within the global frame of reference supplied by the right hemisphere. As noted by HALLIGAN & MARSHALL (1994, p. 183) a purely representational account of BISIACH's results is ruled out by dissociations between imaginal and perceptual neglect. On the other hand, the global processing going on in the right hemisphere is closely akin to the spatial framework BISIACH had postulated for reconstructing visual representations. For our purposes,. it is enough to know that different representational formats are supported by the two hemispheres, and that the type of representation associated with intact right-hemisphere functioning appears perfectly suitable to serve as the basis for assessing and evaluating global situations. The importance of this type of functioning has been stressed by BISIACH. Summing up several decades of research on unilateral neglect, he writes:

> „The data so far reviewed suggest that conscious experience of external stimuli cannot be separated from, and is affected by, the whole situation in which they are perceived, including any response to such stimuli (no matter whether spontaneous or, as in most experimental conditions, complying with instructions)“ (BISIACH, 1992, p. 119 f.).

When investigated under experimental conditions, the neglect syndrome shares certain features with blindsight such that it can interpreted as a dissociation between representation and intentional state of the representation, albeit at a higher level of processing. For instance, a patient with an „extinguished“ visual field (this is a type of neglect that appears only when the two visual half-fields are simultaneously stimulated) was able to give a correct same/different judgment on objects on the left and on the right, though she could name only one of them. Physical identity of the objects was not required for this, they could be different viewpoints of one and the same object, or visually dissimilar members of the same category (BERTI et al., 1992). „Unseen“ objects influence categorial judgments on objects in the „good“ field, again even when they are visually dissimilar but belong to the same category (BERTI & RIZZOLATTI, 1992). Associative priming effects have also been found (LÀDAVAS, PALADINI & CUBELLI, 1993). The effects are similar to those observed in amnesic patients and testify to the fact that neglect patients are able to represent objects up to a categorical level of judgment, though they deny the presence of these very

same objects. Unfortunately, as far as we know there are no experiments in which emotional arousal by „unseen“ objects has been investigated. Thus, direct evidence for a dissociation between representation and emotionally mediated intentionality so far is not available. However, there are two observations which argue for such a dissociation.

WEINSTEIN & FRIEDLAND (1977) and FRIEDLAND & WEINSTEIN (1977) have investigated, using *clinical* methods, the consequences that being anosognosic for one half of the body has for the patients' discourse about this half. Anosognosia turned out to be complete in particular situative contexts only; moreover, in verbalizations about the neglected half, an emotional, rather than a factual speech style proved to be dominant. For instance, one patient denied the existence of his left arm and the fact that it was paralyzed. Nevertheless, under appropriate conditions he described his arm as being shrivelled and like a „yellow claw“. (WEINSTEIN & FRIEDLAND, 1977). BISIACH (1992, p. 131) was motivated by instances of *laevophobia* (a phobia concerning the left side of the body) and *misoplegia (*a feeling of hatred associated with this fear) to expand his representational deficit theory by acknowledging that denial phenomena „reflect the [patient's] needs to come to terms with the pathological representation of one side of their body“; and he entertains the idea that non-cognitive patterns of arousal from the damaged brain region lead to negative emotional processes of evaluation. On the basis of these and similar observations, SCHILDER (1950) and GOLDSTEIN (1939) drew a close parallel between neglect symptoms and psychodynamic processes and considered them as emotionally mediated attempts to cope with an impairment.

Using experimental methods, MARSHALL & HALLIGAN (1988) found evidence for a dissociation between cognitive and meaning-related aspects in neglect. They presented two drawings of the same house to neglect patients. The houses were shown one below the other on the same sheet of paper. In one of the drawings, the house was in flames, but only on the side which was neglected by the patients. At a verbal level they denied to see any difference between the two houses. But when asked to point to the one which they preferred, they preferred the house which was not in flames with significantly above-chance frequency. Apparently the patients were capable of an unconscious evaluation of the state of the house which was not consciously accessible.

BISIACH & RUSCONI (1990) have reported similar results, but given that two of their four patients preferred the „burning“ house, they think that the preferences may not reflect a choice being made on the basis of the meaning of the „flames“, but could be based on low-level sensory features such as complexity and color of the stimuli. At any rate, the reasons given by the patients for their choices had hardly anything to do with the relevant differences between the two drawings. Asking patients to trace the outline of

the stimuli with their fingertips in most cases was not effective in bringing about conscious recognition of their altered left-side parts. This seems to indicate a very high degree of unawareness of the critical features in the neglected half of the houses. Accordingly, BISIACH & RUSCONI (1990, p. 646) state: „The 'confabulations', however, could be induced by left-side features of which the patients seem to be totally unaware".

If this assumption can be confirmed, it could be inferred that an evaluation due to emotional meanings may occur totally outside of conscious awareness but exert unconscious effects on subsequent intentional attitudes.

Intentionality and the representation of faces

The loss of the ability to recognize faces, „prosopagnosia" in medical terms, is one of the neuropsychological phenomena which point to the possibility of a dissociation between representational processes and the generation of meaning. Prosopagnosia occurs rarely as single symptom. Visual field defects are common, but they are not responsible for the inability to recognize faces. The critical lesions are in the occipito-temporal region, with a predilection for the right hemisphere, though lasting cases require bilateral lesions. A rough outline of the subprocesses underlying face recognition was proposed by ELLIS & YOUNG (1988). According to their model, the perceptual analysis of the structural features of faces is followed first by their recognition and then by the activation of knowledge about the perceived person. Since prosopagnosic patients often are able to tell apart faces from other objects, in terms of the model the condition must result from the interruption between intact perceptual analysis of faces and the „recognition units" for faces.

Sensitivity to the figural aspects of faces does not necessarily guarantee the recognition of faces. BRUYER et al. (1983) reported a patient W. who could sort out human faces, even if their hair was covered, from a series of other objects. He had no problems in copying human faces by drawing and he had no particular difficulty recognizing the emotional expression of faces when shown appropriate pictures. Obviously, for him there were hardly any impairments in the perceptual analysis of faces. Yet when shown videos of faces outside of the social context in which he used to meet these persons, he was could rarely if ever recognize them. This was true for photographs of patients he liked to be with, for photographs of the practitioners who treated him and even for a photograph of himself. For the present issue the highly interesting fact was that W. was capable to match faces on the basis of same vs. different emotional expression. His difficulty concerned only an uninteresting neutral facial expression if presented outside of social context

and if additional features of the faces (such as the person's hairdo) were hidden.

At first sight, the ability to recognize facial expressions without being able to identify faces seems most unlikely. Yet the dissociation between the emotional evaluation of a face and its explicit representation as the face of a particular person becomes more plausible on the background of the observation that babies are able to respond, at a very early stage of their development, to the emotional expression of faces, probably a long time before they develop precise and subtle knowledge about the persons belonging to the faces. The great importance of facial expressions for human communication in the nonverbal field has long been acknowledged by investigators in the field of emotion (see EKMAN, 1982; EIBL-EIBESFELDT, 1979; and for a historical account HILDEBRANDT, 1989).

Thus, we may take seriously the possibility that the meaning of a facial expression is processed independently of the representation of faces. BOROD et al. (1986) found that in patients with right hemisphere damage the ability to name facial expressions, as well as to pose them, was definitely reduced, compared to patients with left-hemispheric damage. BOROD et al. take this as an objective indication that observed facial expressions lead to reduced emotional arousal in this group. The difference disappeared when the patients were presented with facial expressions which did not provoke a distinct emotional interpretation.

The experiments by BOROD's group were based on direct inquiry of the patients, i.e. not on indirect methods, as they were used for the investigation of „blindsight" etc. The same applies to BRUYER et al.'s prosopagnosic patient. One of the first studies using indirect methods for investigating face recognition is PREILOWSKI's (1979) investigation of a split-brain patient. Pictures of familiar and unfamiliar faces as well as of the patient's own face served as material. The change in skin resistance resulting from the perceived face was recorded. The pictures of the faces were projected in a lateralized way into both visual half-fields. PREILOWSKI found that the split-brain patient N.G. neither with his right nor with his left visual half-field was capable of identifying faces. Yet the change in the electrical skin resistance was highest when his own face was presented. A similar result was found - in line with earlier investigations (cf. SPERRY, 1968) - when emotionally arousing stimuli were presented. On the basis of similar experiments, SPERRY (1968, p. 732) claims that in split-brain patients the emotional arousal is communicated between the both hemispheres *without* representational content: „Apparently, only the emotional effect gets across, as if the cognitive component of the process cannot be articulated through the brainstem".

Changes in electrical skin resistance - which we take as an indicator for the arousal of an emotional meaning response - in the absence of overt

face recognition were found in prosopagnosic patients. TRANEL & DAMASIO (1985; cf. also BAUER, 1984; DAMASIO, TRANEL & DAMASIO, 1990) reported on two female patients whom they had shown 50 photos of well-known faces, faces from their own families, and of unfamiliar persons. Again the patients responded with a change in skin resistance which was correlated to the level of familiarity though they stated, as a rule, that they had not recognized the faces. Additional neuropsychological investigation excluded the possibility that the two patients' prosopagnosia was caused by a deficit in image generation or by a purely perceptual impairment. The occipital lobe and visual areas at the boundary between parietal and occipital lobe were not compromised. TRANEL & DAMASIO (1988) emphasize that physiological parameters of emotional arousal can be modified in the absence of conscious recognition. This is in line with the view presented here that neuropsychology can reveal a dissociation between the generation of meaning and representational processes. The high emotional valence of perceived faces (especially one's own face, though this is rarely present as a visual stimulus) seems to facilitate the „emancipation" of meaning generation from representation in the case of brain damage.

In order for this interpretation to be correct, it must be shown that prosopagnosia can be restricted to faces and does not necessarily extend to other types of visual objects. It has sometimes been claimed that prosopagnosia reflects a failure of within-category discrimination (as opposed to between-category discrimination) or a more generalized deficit in „configural processing" extending to other types of objects displaying non-combinatorial structure (LEVINE & CALVANIO, 1989). On both accounts, there should be no absolutely „pure" cases of prosopagnosia, i.e., lacking other agnosic phenomena. However, it can be taken for granted that face-specific recognition deficits do in fact exist. BRUYER's et al. (1983) patient, a farmer, was able to recognize the faces of his animals. A patient of WARRINGTON, who initially was thought to suffer from a more generalized deficit (McNEIL & WARRINGTON, 1991), has become a sheep farmer after his stroke and recognizes his sheep by face, though still unable to recognize familiar human faces (McNEIL & WARRINGTON, 1993). A patient studied by FARAH et al. (1995) is able to sort objects within categories but fails the same test when faces are substituted for objects. In itself, the existence of strictly face-specific impairments does not necessarily entail an emotional interpretation of face processing; the effect *could* be a purely cognitive one. However, there is at least one patient (L.H.) who fails all kinds of covert tests for the recognition of familiar faces (ETCOFF et al., 1991) except that he seems to have covert knowledge of their facial expressions (ETCOFF & MAGEE, 1992).

The neuropsychological syndromes reviewed so far can be understood as *simple* dissociations (SHALLICE, 1988) among representation, emotional evaluation and intentional relation towards the representation. In the field of

facial and person perception, however, there is a symptom that may reveal a *double* dissociation between representation and meaning generation. Patients affected by *CAPGRAS' symptom* are able to recognize faces, but they refer them to „doubles" of persons that are familiar to them (ELLIS & YOUNG, 1990). Up to now the symptom has not received adequate objective analysis and it may belong to psychopathological symptoms in the narrow sense. However, neuroimaging methods point to bifrontal involvement and though the symptom is frequently associated with paranoid symptoms, it also occurs as a result of brain damage. Neuropsychological examination of some patients with CAPGRAS' symptom found them not to suffer from any basic impairment in face perception and clearly they were not prosopagnosic. Adequate evaluation is not impaired for *all* faces. Only the most intimately known persons (husbands and spouses, children etc.) are experienced as being completely unchanged in outer appearance but somehow changed in identity. Even the perception and recognition of persons merely close to the patients (such as hospital staff) is not altered. According to ELLIS & YOUNG, the CAPGRAS' symptom is

> „a mirror image of prosopagnosia... This would mean that they receive a veridical image of the person they are looking at, which stimulates all the appropriate overt semantic data held about that person, but they lack another, possibly confirming, set of information, which as LEWIS (1987) and BAUER (1986) have independently suggested, may carry some sort of affective tone" (ELLIS & YOUNG, 1990, p. 244).

In CAPGRAS' symptom emotional evaluation appears to be impaired and representation is intact, while in prosopagnosia representation is impaired but evaluation is functioning properly. In one word, there is a clear *double dissociation.*

Discussion

Before we draw some conclusions, we would like to discuss some alternative interpretations that have been forwarded for the dissociations reviewed by us.

At present, neuropsychological dissociations between conscious und unconscious processing are a very popular research topic, and we could not aspire to cover more than a minute fraction of the facts and of the controversies surrounding them. Nevertheless, relevant mainstream thinking in neuropsychology exhibits some recurring features. First, it is usual to envisage consciousness as just another processing component finding its place within „box-and-arrow" diagrams illustrating the functional relations between processing modules assumed to exist in the mind and/or the brain (see, e.g., the models proposed by SCHACTER, 1989 and MOSCOVITCH, 1992). True, the „consciousness" system may be more central, somehow more important than the others, but as all other subsystems it is causally

affected by input from other subsystems and in turn exerts causal influence on them. Second, the intentionality problem is rarely if ever broached by neuroscientists. Apart from their habit to use 'intention' in a non-technical sense (referring to the volitional antecedent of action) and perhaps a sheer lack of familiarity with 'intentionality' as a philosophical term, the responsible factor seems to be that they do not distinguish between causal and intentional accounts and the associated central concepts of representation and meaning. The latter is just another causal factor which may or may not be present. Finally, emotional processes (if considered at all) as a rule are taken as indicators of perception or recognition without awareness, as caused by the meaning of stimuli, rather than conferring meaning on them.

The standard account, then, is that representation/meaning dissociations arise from the disconnection or disruption of functional or neural pathways between the 'consciousness system' and other subsystems. There are, however, some alternative accounts whose relation to our own we want to discuss briefly.

The *perceptual impairment theory,* as recently revived by Martha FARAH (e.g., FARAH, O'REILLY & VECERA, 1993; FARAH, 1994*)* assumes that purported dissociations between covert recognition and conscious awareness result from a combination of two factors. First, patients exhibiting covert/overt recognition dissociations suffer from perceptual impairments, and second, tests for covert recognition are more sensitive than tests for overt recognition, allowing the patients to „pass them" while they fail on the more demanding tests for overt recognition. A separate consciousness system and its disconnection from non-conscious representation is not required on this account and in fact such an assumption is explicitly denied by FARAH. Reanalyzing data in the literature, FARAH found that at least in some instances of covert/overt dissociations the patients' overt recognition scores were not quite as random as they had been assumed to be, and that their covert recognition level was lower than that of unimpaired control subjects. Theoretically, she has underpinned her argument by reproducing some overt/covert recognition dissociations in artificial neural networks which were first trained to pass overt recognition levels and then „lesioned" by removing a certain percentage of processing units; after this, they were still able to pass covert recognition tests while failing overt tests.

At first sight, FARAH's theory seems to contradict our thesis that representation/meaning dissociations are possible; the theory appeals to representation deficits where we maintain the presence of intact but meaningless representations. However, upon closer examination this is not the case. First, blindsight and abnormal pain experience - where we would like to defend an intact representational basis - have not been reanalyzed or modeled by FARAH. In fact it is difficult to see how, e.g., a „lesioned"

network consisting of homogeneous (i.e., undedicated) processing units could reproduce intact localization with absent detection (the essence of the blindsight phenomenon), given that localization requires more precise discrimination and therefore a richer representational basis than mere detection does. Second, two of the dissociations modeled by FARAH (redirection of attention, knowledge of animate vs. inanimate things) are irrelevant to our account, and her face recognition analyses pertain to cognitive aspects (covert name knowledge), rather than to the emotional aspects stressed by us. For reasons given above, we do not propose the complete absence but the proprietary nature of the representational basis for the emotional evaluation of faces. Most likely, emotional evaluation is based on „physiognomic properties“ as emphasized by the Gestalt school (e.g., KOFFKA 1936, pp. 359 ff.), and such properties are not just a quantitatively degraded version of the structural properties used for „cold“ cognition but qualitatively different from them. This is not to say that blindsight or emotional evaluation cannot be modeled by the network approach but it would have to be done by a type of architecture (e.g., more than one output layer representing different attributes such as ‘presence’ and ‘localization’ in the case of blindsight) that already presupposes the dissociation it wants to explain. Other aspects of FARAH’s endeavor , such as her rejection of the „locality assumption“ and of an encapsulated consciousness system, are fully in accord with the approach outlined here.

At the opposite end of the perceptual impairment theory is the *meta-representational account.* On this view, consciousness is a „commentary system“ (WEISKRANTZ, 1986) defined not so much through the presence of sensory qualities but in terms of the ability to refer to them, predominantly by linguistic means. For instance, blindsight patients would suffer from an inability to report or otherwise elaborate on visual experiences in their blind field rather than from their absence. If we believe in the ineffability of qualia, the theory is difficult to reject and provides the only means to „anchor“ the „ineffable“ in publicly observable behavior. If considered from a less principled angle, the theory is subject to various rejoinders. Specifically as blindsight is concerned, it seems not to be true that the patients exhibit an absolute loss of metarepresentational reference to their visual abilities. Patient D.B. (the subject of WEISKRANTZ’s, 1986, monograph) was referred to the attention of psychologists because he himself had noticed certain residual functions of his blind field. Other blindsight patients refer to their residual visual abilities by stating that they „feel“ the objects they cannot see (PÖPPEL et al., 1973). HARTMANN et al. (1991) have described a cortically blind patient with a small intact portion of his visual field that mediated considerable visual abilities including the reading of words (within the spatial limits defined by his intact visual field). The patient - who led a completely independent life despite his insistence on being totally blind - denied the *visual* nature of the experiences corresponding to his intact

abilities objectively mediated by vision; instead, he consistently referred to them as „feeling“. In other words, metarepresentational reference to visual abilities is intact but their *experiential quality* has been lost, or rather transferred to the amodal plane of the „mind“.

This is not to deny that brain damage can profoundly affect the meta-representational level of mental functioning. However, the classical case in which this occurs seems to be *anosognosia*, the „unawareness of deficits“ such as hemiplegia, cortical blindness and other neuropsychological conditions. Theoretically, anosognosia should be distinguished from the psychodynamically motivated *denial* of the deficit, although the dividing line may be difficult to draw in individual cases. Anosognosic patients do not simply deny their deficit at a verbal level, but consistently behave as if no deficit were present. Based on a careful review of the literature, McGLYNN & SCHACTER (1989) have concluded that anosognosia arises from (typically right-hemispheric) lesions in the parietal lobe, in the frontal lobe, or both. The parietal lobes are thought to subserve the *monitoring,* and the frontal lobes the *execution* of sensorimotor and/or cognitive systems, as evidenced by the fact that ansognosias resulting from parietal lesions are often specific for single deficits while frontal lesions involve more generalized anosognosia and anosognosia for higher-order deficits such as amnesia or personality alterations. The monitoring system is supposed to be a c*onscious awareness system* that normally receives input from specific modules, but becomes disconnected from them if they are not sufficiently activated, resulting in specific anosognosias, and/or fails to send output to the frontal executive system, resulting in generalized anosognosia.

As far as the neuroanatomical correlates of anosognosia and the distinction between two types of anosognosia are concerned we have no objections against Schacter’s formulation. We find it problematical, though, to identify the monitoring system with conscious awareness. In the case of undisturbed functioning, at any rate, the monitoring function is automatic and outside of the focus of conscious awareness. In view of this, we prefer to assign the anosognosias not to the consciousness concept but to the Ego concept. Specific (parietal-type) anosognosias, we submit, are a result of the disconnection between the representational level and the *Ego as experienced,* and generalized (frontal-type) agnosias arise from dis-connections involving the *Ego as an actor*. In terms of our intentionality/ meaning account, we suggest to conceptualize the anosognosias in terms of disturbances of *second-order intentionality*. Let us briefly explain.

As outlined above, intentional (intrinsically meaningful) experiences are at once relational and absolute; relational because they presuppose integration into the Ego conceived as stream of experience, and absolute because the *relation* between experience and Ego normally is not thematized but only conveys a „coloring“ of the experience by a feeling of intimate familiarity.

This is the case of first-order intentionality, obtaining when our mental life is functioning smoothly. Consider now what may happen as a result of damage to the neural substrate of the mind. If a defect in sensorimotor or cognitive functioning ensues, the two „poles of intention“ (Ego and experience) will have to be thematized separately („I can no longer move my arm“, etc.), until readjustment to the defect has taken place or, hopefully, the defect subsides or is alleviated in the course of recovery and/or rehabilitation. Separately focusing on Ego and experience, and thematizing their relation, is the defining mark of second-order intentionality, preserved in „ordinary“ acknowledgment of deficit and disrupted in the case of anosognosia. In terms of second-order intentionality, anosogonosic patients act and experience on a representational basis corresponding to their premorbid abilities (*cum grano salis*, somewhat similar to the phantom limb experience after amputation), but at the level of primary intentionality, a more „realistic“ and often emotional evaluation of their defect is discernible, at least from the third-person perspective. Thus, the „explicit“ Ego of second-order intentionality becomes dissociated from the „implicit“ Ego of first-order intentionality, insofar as they do no longer use one and the same representational basis. Denial (in the sense of a psychodynamic complication often consistent with premorbid personality traits) will occur when the emotional evaluation is shifted to the level of second-order intentionality but the cognitive content of the explicit Ego remains tied to the premorbid level of functioning, in the service of maintaining the integrity of the Ego, or more properly the Self. Returning to SCHACTER's (and many others') conceptualization of anosognosia in terms of a dissociation of the representation of the deficit from „conscious awareness“, we think that the juxtaposition of 'conscious' and 'awareness' may be justified when it is meant to convey the metacognitive component in the adjective 'conscious'; otherwise, if meant as a simple reduplication of terms, it may provoke confusion by conflating ordinary, first-order intentionality („awareness“) with second-order, „reflective“ intentionality („consciousness“).

Quite recently, Ned BLOCK also deplores a „confusion about a function of consciousness“. Briefly, BLOCK (1995, p. 227) distinguishes between „phenomenal consciousness“, (P-consciousness) which he defines (as we do) in terms of experience, and „access consciousness“ (A-consciousness) i.e., the availability of representations for „use in reasoning and rationally guided speech and action“. The „conflation“ Block notices among most authors in the field consists in attributing the information processing „machinery“ underlying A-consciousness to P-consciousness; with the result that the latter has the function of enabling representations to guide action and speech. Thus, the fact that blindsight patients lack P-consciousness for stimuli in their „blind“ field is taken to explain their failure of using visual representations (whose presence in their brain is taken for granted by BLOCK) in action and speech. However, according to BLOCK blindsight

patients are also „A-unconscious“ for stimuli in their blind field, and it is fallacious to infer from the joint absence of P- and A- consciousness that one is the cause of the other; both could be caused by a common underlying factor. On the other hand, while BLOCK denies a dissociation between A- and P-consciousness *in blindsight patients,* he thinks that double dissociations (A-consciousness without P-consciousness, P-consciousness without A-consciousness) are conceptually possible. An instance of A-consciousness without P-consciousness would be a „super-blindsight“ patient who uses representations in his blind field in everyday behavior and conversation and still denies to experience them. (HARTMANN et al.'s, 1991, patient mentioned above nearly fulfills this criterion except that the visual representations whose experience he denies come from his perimetrically determined seeing field. He was not mentioned by Block himself or his commentators.) P-consciousness without A-consciousness would be exemplified by cases where we are aware of some object or event but do not realize its meaning, or by an animal deprived of the dorsal system underlying visually guided behavior but still in possession of the ventral recognition system. (Actually, there are patients of this sort who experience and recognize objects but cannot localize them, and the opposite dissociation also has been documented; see MILNER 1992). In sum, BLOCK asks us to consider the relevant neuropsychological syndromes as instances of *dissociations between two kinds of consciousness,* rather than as dissociations between conscious and non-conscious processing as proposed by the standard account.

Given that BLOCK defines (1995, p. 232) access-conscious content as being representational, he maintains the possibility of a dissociation between experience and representation and to that extent we agree with his conclusion. Nevertheless, there is an important difference. In the interest of letting A-consciousness support inferential capacities underlying rational thought and action, BLOCK describes the „prototypical case“ of an A-conscious state as „propositional attitude“, and though on the surface he is not concerned with the role of meaning, it is quite clear from his examples that meaning is conferred to P-conscious states via the propositional attitudes that are a property of the A-consciousness mechanism. Here, of course, we disagree. Meaning, at least the type of intrinsic, experiential meaning we are concerned with, does not accrue to „meaningless“, neutral experience via insertion into an inferential network. As a terminological correlate of our dissenting opinion, we would prefer not call the access mechanism „consciousness“. Why not restrict it to a purely functional notion, with the added convenience of avoiding a dissociation between representation as such and representation as content of A-consciousness? According to BLOCK, 'consciousness' is a „mongrel concept“, and some of its traditional attributes are better saved in the notion of A-consciousness than in the notion of P-consciousness. One such attribute is *rationality.* In his

reply to the commentators, BLOCK (1995, p. 277) has somewhat toned down this requirement; now it is not good or valid reasoning that is needed for A-consciousness, but „the appeal to the use of a representation in reasoning“, even if the reasoning is poor. It seems to us that this requirement would define the notion of intelligence (including artificial intelligence) rather than the notion of consciousness, and at any rate (see DRETSKE, above) it is not goal-directed but goal-intended behavior that counts. The critical difference between the two, we reiterate, is not the absence or presence of rationality, but the absence or presence of individual learning under the (at least initial) control of emotional evaluation.

Conclusion

Based on a brief *tour d'horizon* of current debates in cognitive science, we have discussed some explanatory approaches to the problem of the emergence of subjective meaning. There is a strong tendency among cognitive scientists to identify intentional states with representational states, including meta-representations. This tendency was criticized and it was argued that a distinction between intentionality and representation is conceptually possible and empirically valid. Concerning awareness of representational states, various neuropsychological syndromes could be adduced where the representational level remains largely intact, but its outcome is neither acknowledged consciously by the patients nor is it expressed in the patients' spontaneous behavior. Brief as it had to be, our review has provided evidence for intentionality as *a state of transparency for the subject* which is not situated at the same level as the representational processes covarying with environmental properties.

The evidence of a representation-independent emotional evaluation of environmental properties is not quite so distinct, mainly for the methodological reason that the experimental verification of this possibility requires the establishment of some minimal level of representation. Even so, experimental and clinical findings have been presented for patients with neglect and prosopagnosia which can be interpreted in terms of emotional evaluation of environmental events occurring *before or at least parallel to* cognitive processing. At the very minimum, we hope to have shown that the emotional evaluation of faces rests on a representational basis that is proprietary to it and not reducible to the cognitive processing of the same faces. The often negatively colored emotional relation of neglect patients to the neglected half of their own body or of surrounding space also points to a representational basis that is different from that in the intact field.

To bring about an intrinsically meaningful intentional state, a three-dimensional process joining emotional, ego-related, and cognitive processes has to take place; a viewpoint which still has received little attention in

cognitive neurosciences. Representational and meaning states in themselves are not intrinsically intentional states when the brain structures subserving the emotional component are lesioned. It is in this respect that pain experience in lobotomized patients differs from that of healthy subjects. On the other hand the emotional component alone is not sufficient for the development of a state of intrinsic intentionality when there is no concurrent representation of the object of this emotional process. Though prosopagnosic patients may experience the emotional expression of a familiar face, this does not in itself lead to recognition of the face's personal identity and an appropriate intentional relationship to it. Delusionary misidentifications, in this paper discussed under the heading of CAPGRAS' symptom, show that a significant integration of quasi-representational (i.e., subjectively representational, but objectively misrepresentational) processes may take place and that this integration is an independent aspect of intentional states. Another example is the unawareness of neurological deficit, which accentuates the role of secondary intentionality with its explicit focusing on the Ego, and in the case of psychodynamically elaborated denial, on the self as the integral „core region" of the Ego.

At a more general level, the gap between the description of representational processes in the functional terms of information processing psychology and the development of a theory of intrinsic intentionality poses a major problem for psychology and cognitive science. Two very different perspectives are currently employed in them: a general account of functional processes for all subjects and a specific theory of the reasons for actions of individual subjects (HILDEBRANDT, 1994). The gap is at the core of the debate between *Geisteswissenschaften* (humanities) and natural sciences and may turn out to be an „immortal" (NAGEL, 1979) question, owing to necessarily different general research programs and objectives. Another reason for this gap may well be the wrong or over-generalized choice of models for explaining intentional states. Functional description of general processes is forced to look for separate and invariant processing structures. A more phenomenologically based approach reveals that intentionality involves force-like interactions between some content and a subject which can be defined as a system in a specific emotional and ego-related goal state. The dissociative neuropsychological syndromes reviewed by us all depend on *local* brain damage; presumably leaving intact the functioning of other local processes and thus fulfilling an important postulate of the functionalist research program. The phenomenological approach requires that the coupling of such local processes for conscious awareness depends on their dynamic interrelation. This is not just a personal intuition but supported by the neuropsychological syndromes that have been reviewed. Consequently, in trying to elucidate the brain correlates of conscious processes, cognitive neuroscience should investigate the dynamic *binding* of different neuronal assemblies and in this way look for Gestalt-like properties

of states of the whole brain. Adding more and more representation, and throwing in meta-representation for good measure, will not do the job.

Wolfgang Köhler expected his Gestalt physics of the brain to bridge the gulf between natural science and Geisteswissenschaften (KÖHLER, 1938). On this score, too, we are glad to follow his footsteps.

Summary

Motivated by ideas inherent in Gestalt Theory, we explore the contribution of neuropsychology to an unsolved problem in current cognitive science, viz., the relation between 'representation' and 'intentionality'. We first provide working definitions of these two basic concepts in the philosophy of mind and go on to show that the problem of intentionality usually is assimilated to the problem of representation: intentional states are taken as a subclass of representational states, and representation is taken as a necessary condition for intentionality. An advance over the standard approach is made be Dretske's externalistic theory of meaning, which derives the representational function from the learning history of individual organisms and stresses the role of emotions in the process. However, a purely objective theory of meaning is inadequate at the level of human psychology. Accordingly. we confront the „mainstream" approach in cognitive science with a richer account of intentionality, derived from the thought of Wolfgang Köhler, that makes intentional states dependent on global states of the mind/brain encompassing emotional and self-evaluative aspects over and above „pure" representation. A corollary of this approach is that representational and intentional states,. though closely cooperating under normal conditions, can be dissociated from each other under pathological conditions. A survey of some relevant neuropsychological results shows this to be the case. Intentionality may either be eliminated altogether, or its emotional and self-referential aspects may be profoundly affected, despite virtually intact representational capacities. On the other hand, stimuli which are not consciously represented may still evoke physiological responses indicative of an altered emotional state. These dissociations cannot be resolved by appealing to a hierarchy of meta-representational levels or by reducing all of them to representational deficits.

Zusammenfassung

Von der Gestalttheorie angeregt, erörtern wir den Beitrag der Neuropsychologie zu einem ungelösten Problem der heutigen Kognitionswissenschaft: die Beziehung zwischen 'Repräsentation' und 'Intentionalität'. Nachdem wir Arbeitsdefinitionen dieser beiden Grundbegriffe der Philosophie des Geistes aufgestellt haben, zeigen wir, daß das Problem der Intentionalität in aller Regel dem Problem der Repräsentation untergeordnet wird: intentionale Zustände werden als Teilmenge repräsentationaler Zustände aufgefaßt, und Repräsentation wird als notwendige Voraussetzung für Intentionalität angesehen. Einen Fortschritt gegenüber dieser Standardauffassung bedeutet die externalistische Bedeutungstheorie von F. DRETSKE, in der die Repräsentationsfunktion aus der Lerngeschichte individueller Organismen abgeleitet und die Wichtigkeit von Emotionen in diesem Vorgang betont wird. Jedoch ist eine rein objektive Theorie dieser Art für die Humanpsychologie unzureichend. Daher konfrontieren wir die Standardauffaßung mit einer inhaltsreicheren, durch W. KÖHLER angeregten Konzeption der Intentionalität, wonach intentionale Zustände von globalen Zuständen des Geist/Gehirns

abhängig sind und neben der "reinen" Repräsentation emotionale und ichbezogene Aspekte umfassen. Aus dieser Konzeption folgt, daß repräsentationale und intentionale Zustände, obwohl sie bei ungestörtem Funktionieren eng miteinander verbunden sind, unter pathologischen Umständen dissoziiert sein können. Ein Überblick über einige einschlägige neuropsychologische Daten zeigt, daß dies in der Tat der Fall ist. Bei prinzipiell intakter Repräsentationsebene kann die intentionale Beziehung entweder ganz wegfallen oder in ihren emotionalen und selbstbewertenden Aspekten gestört sein. Umgekehrt können Reize, die zumindest nicht bewußt repräsentiert sind, physiologische Reaktionen hervorrufen, die auf einen geänderten emotionalen Zustand verweisen. Diese Dissoziationen können weder durch eine Theorie repräsentationaler Defizite noch durch die Einführung metarepräsentationaler Ebenen erklärt werden.

References

ANTROBUS, J. (1990). The neurocognition of sleep mentation: Rapid eye movements, visual imagery, and dreaming. In R. R. BOOTZIN, J. F. KIHLSTROM & D. L. SCHACTER (eds.), *Sleep and Cognition* (pp. 3-24). Washington, D.C.: APA.

BAUER, R. M. (1984). Autonomic recognition of names and faces in prosoagnosia: A neuropsychological application of the guilty knowledge test. *Neuropsychologia, 22*, 457-469.

BAUER, R. M. (1986). The cognitive psychophysiology of prosopagnosia. *Bibliotheca Psychiatrica, 164*, 253-267.

BERTI, A. & RIZZOLATTI, G. (1992). Visual processing without awareness: Evidence from unilateral neglect. *Journal of Cognitive Neuroscience, 4*, 345-351.

BIERI, P. (1987). Pain: A case study for the mind-body problem. *Acta Neurochirurgica, Suppl. 38*, 157-164.

BISIACH, E. (1992). Understanding consciousness: Clues from unilateral neglect and related disorders. In A.D. Milner & M.D. Rugg (eds.), *The Neuropsychology of Consciousness* (pp. 113-137). London: Academic Press.

BISIACH, E. & LUZZATTI, C. (1978). Unilateral neglect of representational space. *Cortex, 14,* 129-133.

BISIACH, E., LUZZATTI, C.& PERANI, D. (1979). Unilateral neglect, representational schema and consciousness. *Brain, 102*, 609-618.

BISIACH, E. & RUSCONI, M.L. (1990). Break-down of perceptual awareness in unilateral neglect. *Cortex, 26,* 643-649.

BLOCK, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18,* 227-287.

BOROD, J.C., KOFF, E., LORCH, M.P. & NICHOLAS, M. (1986). The expression and perception of facial emotion in brain-damaged patients. *Neuropsychologia, 24*, 169-180.

BREWER, W.F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W.B. WEIMER & D.S. PALERMO (eds.), *Cognition and the symbolic processes*
(pp. 1-42). New York: Wiley.

BRUYER, R., LATERRE, C., SERON, X., FEYEREISEN, P., STRYPSTEIN, E., PIERRARD, E. & RECTEM, D. (1983). A case of prosopagnosia with some preserved covert remembrance of familiar faces. *Brain and Cognition, 2*, 257-284.

CHAPMAN, W.P., SOLOMON, H.C. & ROSE, A.S. (1950). Measurement of motor withdrawal reaction in patients following frontal lobotomy. In M. GREENBLATT et al. (ed.), *Studies in Lobotomy* (pp. 386-392). New York: Grune & Stratton.

DAMASIO, A., TRANEL, D. & DAMASIO, H. C. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience, 13*, 89-109.

DAMASIO, A., TRANEL, D. & DAMASIO, H. C. (1991). Somatic markers and the guidance of behavior: Theory and preliminary testing. In H.S. LEVIN, A.M. EISENBERG & A.L. BENTON (eds.), *Frontal Lobe Function and Dysfunction* (pp. 217-229). New York: Oxford University Press.

DENNETT, D.C. (1978). Why you can't make a computer that feels pain. *Synthese, 38*, 415-456.

DENNETT, D.C. (1987). *Brainstorms: Philosophical Essay on Mind and Psychology*. Cambridge4, Ma.: MIT Press.

DRETSKE, F.I. (1988). *Explaining Behavior. Reasons in a World of Causes*. Cambridge: MIT Press.

DYNES, J.B. & POPPEN, J.L. (1949). Lobotomy for intractable pain. *Journal of the American Medical Association, 140*, 15-19.

EIBL-EIBESFELDT, I. (1979). Human Ethology: concepts and implications for the sciences of man. *Behavioral and Brain Sciences, 2*, 1-57.

EKMAN, P. (ed.) (1982). *Emotion in the Human Face*. Cambridge: Cambridge University Press.

ELITHORN, A., GLITHERO, E., SLATER, E. (1958). Leucotomy for pain. *Journal of Neurology*, *Neurosurgery and Psychiatry, 21*, 249-261.

ELLIS, A.W. & YOUNG, A.W. (1988). *Human Cognitive Neuropsychology*. Hillsdale, N.J.: Lawrence Erlbaum.

ELLIS, H.D. & YOUNG, W.A. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry, 157*, 239-248.

ETCOFF, N.L., FREEMAN, R. & CAVE, K.R. (1991). Can we loose memories of faces? Content specificity and awareness in prosopagnosia. *Journal of Cognitive Neuroscience, 3,* 25-41.

ETCOFF, N.L. & MAGEE, J.J. (1992). Covert recognition of emotional expressions. *Journal of Clinical and Experimental Neuropsychology, 14*, 95-96.

FARAH, M. (1994). Neuropsychological evidence with an interactive brain: A critique of the „locality“ assumption. *Behavioral and Brain Sciences, 17*, 43-104.

FARAH, M., LEVINSON, K.L. & KLEIN, K.L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia, 33*, 661-674.

FARAH, M., O'REILLY, R.C. & VECERA, S.P.(1993). Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review, 100*, 571-588.

FODOR, J.A. (1975). *The Language of Thought*. New York: Crowell.

FODOR, J.A. (1981). *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, Ma.: MIT Press.

FREEMAN, W. & WATTS, J.W. (first ed. 1942, sec.ed. 1950). *Psychosurgery*. Oxford: Blackwell.

FRIEDLAND, R.P. & WEINSTEIN, E.A. (1977). Hemi-inattention and hemispheric specialization: Introduction and history. In E.A. WEINSTEIN & R.P. FRIEDLAND (eds.), *Hemi-Inattention and Hemispheric Specialization* (pp. 1-31). New York: Raven Press.

GAINOTTI, G., D'ERME, P. & BARTOLOMEO, P. (1991). Early orientation of attention toward the half space ipsilateral to the lesion in patients with unilateral brain damage. *Journal of Neurology, Neurosurgery, and Psychiatry, 54*, 1082-1089.

GOLDSTEIN, K. (1939*). The Organism, a Holistic Approach to Biology Derived from Pathological Data in Man*. New York: American Book Co.

GRAHEK, N. (1992). Schmerz. In J. RITTER & K. GRÜNDER (Hsg.*), Historisches Wörterbuch der Philosophie*, Bd. 8. Basel: Schwabe.

GRAHEK, N. (1993). Objektive und subjektive Aspekte der Schmerzwahrnehmung. In H. HILDEBRANDT & E. SCHEERER (Hsg.), *Interdisziplinäre Perspektiven der Kognitionsforschung* (S. 109-132). Frankfurt/M.: Peter Lang.

HALLIGAN, P.W. & MARSHALL, J. (1994). Toward a principled explanation of unilateral neglect. *Cognitive Neuropsychology, 11*, 167-216.

HARTMANN, J.A., WOLZ, W.A., ROELTGEN, E.P. & LOVERSO, F.L. (1991). Denial of visual perception. *Brain and Cognition, 16*, 29-40.

HEILMAN, K.M., VALENSTEIN, E., WATSON, R.T. (1985). The neglect syndrome. In J.A.M. FREDERICKS (ed.), *Handbook of Clinical Neurology*. *Vol 1: Clinical Neuropsychology* (pp. 153-183). Amsterdam: Elsevier.

HILDEBRANDT, H. (1989). Piderits "Wissenschaftliches System der Mimik..": wissenschafts-geschichtliche Einordnung. In T. PIDERIT, *Wissenschaftliches System der Mimik und Physiognomik* (S. 14 -47), (hrsg. von W. Prinz & N. Bulst). Göttingen: Hogrefe.

HILDEBRANDT, H. (1991). Zur Bedeutung des Begriffs der Alltagspsychologie in Theorie und Geschichte der Psychologie. Frankfurt: Lang

HILDEBRANDT, H. (1994). Gehirn - Psyche - Subjekt. In H.W. HOEFERT & C. KLOTTERT (Hrsg.), *Neue Wege der Psychologie* (S. 43-54). Heidelberg: Asanger,

HOBSON, J.A., LYDIC, R. & BAGHDAYAN, H.A. (1986). Evolving concepts of sleep-cycle generation: From brain centers to neuronal populations. *Behavioral and Brain Sciences, 9*, 371.448.

JAMES, W. (1890). *Principles of Psychology*. New York: Appleton.

KING, H.E., CLAUSEN, J. & SCARFF, J.E. (1950). Cutaneous thresholds for pain before and after unilateral prefrontal lobotomy. *Journal of Nervous and Mental Diseases, 112*, 93-96.

KOFFKA, K. (1936). *Principles of Gestalt Psychology*. London: Kegan Paul.

KÖHLER, W. (1933). *Psychologische Probleme*. Berlin: Springer.

KÖHLER, W. (1938). *The Place of Value in a World of Facts*. New York: Liveright.

KÖHLER, W. (1947). *Gestalt Psychology*. New York: Liveright.

KURTHEN, M. (1984). *Der Schmerz als medizinisches und philosophisches Problem*. Würzburg: Königshausen & Neumann.

LÀDAVAS, E., PALADINI, R. & CUBELLI, R. (1993). Implicit associative priming in a patient with left visual neglect. *Neuropsychologia, 31*, 1307-1320.

LARBIG, W. (1982). *Schmerz*. Stuttgart: Kohlhammer.

LEDOUX, J.E., ROMANSKI, L., XAGORARIS, A. (1988). Indelibility of subcortical emotional memories. *Journal of Cognitive Neuroscience, 1*, 238-243.

LEVINE, D.N. & CALVANIO, R. (1989). Prosopagnosia: A defect in visual configurative processing. *Brain and Cognition, 10*, 149-170.

LEWIS, S.W. (1987). Brain imaging in a case of Capgras' syndrome. *British Journal of Psychiatry, 150*, 117-121.

MARCEL, A.J. (1988). Phenomenal experience and functionalism. In A.J. MARCEL & E. BISIACH (eds.), *Consciousness in Contemporary Science* (pp. 121-158). Oxford: Clarendon Press.

MARSHALL, J.C., HALLIGAN, P.W. (1988). Blindsight and insight in visuo-spatial neglect. *Nature, 336*, 766-767.

McGLYNN, S.M. & SCHACTER, D.L. (1989). Unawareness of deficits in neuropsychological syndromes. *Journal of Experimental and Clinical Neuropsychology, 11*, 143-205.

McNEIL, J.E. & WARRINGTON, E.K. (1991). Prosopagnosia: A reclassification. *Quarterly Journal of Experimental Psychology, 43A*, 267-287.

McNEIL, J.E. & WARRINGTON, E.K. (1993). Prosopagnosia: A face-specific disorder. *Quarterly Journal of Experimental Psychology*, *46A*, 1-10.

MEINHART, N.T., MCCAFFERY, M. (1983). *Pain*. New York: Appleton-Century-Crofts.

MILNER, A.D. (1992). Disorders of perceptual awareness - Commentary. In A.D. MILNER & M.D. RUGG (eds.), *The Neuropsychology of Consciousness* (pp. 139-158). London: Academic Press.

MOSCOVITCH, M. (1992). A neuropsychological model of memory and consciousness. In L.R. SQUIRE & N. BUTTERS (eds.) The Neuropsychology of Memory (2nd ed.) (pp. 5-22). New York: Guilford Press.

NAGEL, T. (1979). *Mortal questions*. London: Routledge.

NEMIAH, J.C. (1962). The effect of leukotomy on pain. *Psychosomatic Medicine, 24*, 75-80.

PÖPPEL, E., HELD, R. & FROST, D. (1973). Residual visual function after brain wounds involving the central visual pathways in man. *Nature, 243*, 295-296.

POSNER, M.L. & ROTHBART, M.K. (1992). Attentional mechanisms and conscious experience. In A.D. MILNER & M.D. RUGG (eds.), *The Neuropsychology of Consciousness* (pp. 91-112). London: Academic Press.

PREILOWSKI, B. (1979). Consciousness after complete surgical section of the forebrain commissures in man. In L.S. RUSSELL et al. (ed.), *Structure and Function of Cerebral Commissures* (pp. 411-420). London: McMillan Press.

PYLYSHYN, Z. (1984). *Computation and Cognition*. Cambridge, Ma.: MIT Press.

ROTH, G. & SCHWEGLER, H. (1995). Das Geist-Gehirn-Problem aus der Sicht der Hirnforschung und eines nicht-reduktionistischen Physikalismus. *Ethik und Sozialwissenschaften, 6*(1), 69-156.

SCHACTER, D.L. (1989). On the relation between memory and consciousness: Dissociable interactions and conscious experience. In H. ROEDIGER & F. CRAIK (eds.), *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving*. Hillsdale, N.J.: Erlbaum.

SCHACTER, D.L. (1992). Consciousness and awareness in memory and amnesia: Critical issues. In: A.D. MILNER & M.D. RUGG (eds.), *The Neuropsychology of Consciousness* (pp. 179-199). London: Academic Press.

SCHACTER, D.L., MCANDREWS, M.P., MOSCOVITCH, M. (1988). Access to consciousness: dissociations between implicit and explicit knowledge in neuropsychological syndromes. In L. WEISKRANTZ (ed.), *Thought without language* (pp. 242-278). Oxford: Clarendon Press.

SCHEERER, E. (1992). Repräsentation. In J. Ritter & K. Gründer (Hrsg.*), Historisches Wörterbuch der Philosophie*, Bd. 8. Basel. Schwabe.

SCHEERER, E. (1993). Mentale Repräsentation in interdisziplinärer Perspektive. *Zeitschrift für Psychologie, 201*, 136-166.

SCHEERER, E. (1994). Psychoneural isomorphism: Historical background and current relevance. *Philosophical Psychology, 7,* 183-210.

SCHEERER, E. (1995) Die Sinne. In J. Ritter & K. Gründer (Hsg.*), Historisches Wörterbuch der Philosophie*, Bd. 9. Basel: Schwabe.

SCHILDER, P. (1950). *The image and appearance of the human body*. London: Kegan Paul.

SEARLE, J.R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*, 417-457.

SEARLE, J.R. (1990). Consciousness, explanatory inversion, and cognitive science. *Behavioral and Brain Sciences, 13*, 585-642.

SHALLICE, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.

SPERRY, R.W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23*, 723-733.

TRANEL, D. & DAMASIO, A.R. (1985). Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science, 228*, 1453-1454.

TRANEL, D. & DAMASIO, A.R. (1988). Non-conscious recognition in patients with face agnosia. *Behavioral Brain Research, 30*, 235-249.

TURVEY, M.T., SHAW, R.E., REED, E.S. & MACE, W.M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition, 9*, 237-304.

VALLAR, G. (1993). The anatomical basis of spatial hemineglect in humans. In I.H. ROBERTSON & J.C. MARSHALL (eds.), *Unilateral Neglect: Clinical and Experimental Studies* (pp. 27-62). Hillsdale, N.J.: Erlbaum.

VAN GULICK, R. (1988). Consciousness, intrinsic intentionality, and self-understanding machines. In A.J. MARCEL & E. BISIACH (eds.), *Consciousness in Contemporary Science* (pp. 78-100). Oxford: Clarendon Press.

WEINSTEIN, E.A., FRIEDLAND, R.P. (1977). Behavioral disorders associated with hemi-inattention. In E.A. Weinstein & R.P. Friedland (eds.), *Hemi-Inattention and Hemispheric Specialization* (pp. 51-62). New York: Raven Press.

WEISKRANTZ, L. (1980). Varieties of residual experience. *Quarterly Journal of Experimental Psychology, 32*, 365-386.

WEISKRANTZ, L. (1986). *Blindsight*. Oxford: Oxford University Press.

WEISKRANTZ, L. (1988). Some contributions of neuropsychology of vision and memory to the problem of consciousness. In A.J. MARCEL & E. BISIACH (eds.), *Consciousness in Contemporary Science* (pp. 183-199). Oxford: Clarendon Press.

WEISKRANTZ, L. (1989). Blindsight. In F. BOLLER & J. GRAFMAN (eds.), *Handbook of Neuropsychology*, Vol. 1, (pp. 375-385). Amsterdam: Elsevier.

Anschrift der Verfasser:

Dr. Helmut Hildebrandt

AE „Psychologie im Gesundheitswesen“ und

Projekt „Interdisziplinäre Kognitionsforschung“

Universität Oldenburg, FB 5

email: hildebrandt@psychologie.uni-oldenburg.de or

Prof. Dr. Eckart Scheerer

Institut für Kognitionsforschung

Universität Oldenburg, FB 5

D-26111 Oldenburg

email: scheerer@psychologie.uni-oldenburg.de