

Research Essay

Interactionism Read Anew: A Proposal Concerning Phenomenal Judgments

Einar L. Halvorsen*

ABSTRACT

From a classical Cartesian perspective, interactionism implies the transfer of thoughts and feelings from a non-physical phenomenal consciousness to the physical brain. Thereby, phenomenal consciousness is thought to control the physical body somehow like a marionette hanging by strings of non-physical thought. Differing from this depiction, a basic premise of the current interactionist hypothesis is that the non-physical phenomenal consciousness reflexively effects accentuation of thoughts, feelings and sensory experiences which already exist as physical brain processes. In this essay, the mentioned interactionist hypothesis is presented and central philosophical problems which pertain to it are discussed. Conclusively, the hypothesis may withstand the initial scrutiny and is thereby rendered coherent. Nonetheless, the feasibility of interactionism is not thereby significantly influenced. That would require a much more extensive treatment of the feasibility of the current hypothesis as well as of coherent solutions to numerous other problems pertaining to interactionism.

Key Words: interactionism, phenomenal judgment, phenomenal consciousness, physical brain.

Introduction

Like Michael Tye and others (Tye, p. 1), I will refer to phenomenal consciousness as “P-consciousness”. The term “P-consciousness” will here provide for an intently undefined meaning, as justifiable in light of Ned Block’s notion that the meaning of that term is linguistically indefinable (Tye, p. 1). Tye, freely re-formulated, sees P-consciousness as a capacity to possess a phenomenal side to experiences (Tye, p. 144-5). If Block is right, however, the phrase “phenomenal side to experiences”, if understood as the indefinable “P-side”, makes that description circular.

The assumption of the linguistic indefinability of P-consciousness carries implications also for an interactionist perspective. While the physical world constitutes one pole of a purported interactionist dualism, the nature of the other, non-physical pole will remain unknown or a mere object of beliefs. Beyond the notion that P-consciousness is intimately linked to experience, furthermore, a consensual nature of such beliefs can only be assumed.

Interactionism, the general theory discussed in this essay, is the idea that a non-physical P-consciousness causally interacts with the physical brain (Chalmers, 2010, p. 126). Given that the doctrine of interactionism is correct, violation of the doctrine of causal closure of the physical

*Correspondence: Einar L. Halvorsen, Independent Researcher, Germany. Email: EinarLautenHalvorsen@gmail.com

world (Stoljar, p. 220-2), including the quantum mechanical version of that doctrine (Papineau & Selina, p. 74-5), is necessary.

Stewart Goetz convincingly argues that no theoretical hindrance exists which makes scientifically incoherent the notion that causal closure may be violated (Goetz, in Baker & Goetz, p. 104-16). Robin Collins, furthermore, explicates that it is not scientifically necessitated that violation of causal closure must imply violation of energy conservation (Collins, in Baker & Goetz, p. 124-33).

Even if the doctrine of causal closure could coherently be violated, however, another problem is that expressed by the question of how quantum level interactionist influences may amount to neuro-functional changes. This would be necessary should interactionism make a relevant difference (Chalmers, 2010, 126-7n). Sir John C. Eccles early described the neuronal processes which would have to be influenced (Popper & Eccles, p. 232).

A multitude of electro-chemical variables causally co-determine the firing of “action potentials”, the single wave-like signals of invariable amplitude which “shoot” through single neurons (Klinke & Silbernagl, p. 51-77). The challenge, if wishing to rule out interactionism, is to prove each such electro-chemical variable as fully determined by reductive processes. David Chalmers suggests that scientific understanding of these matters is still incomplete (Chalmers, 2010, p. 126-7n).

In order that an interactionist solution to the problem of phenomenal judgements may be correct, a necessary premise is the coherence of the notion that *all* further problems with interactionism may be solved. Only two further problems, however, which have direct relevance to the problem of phenomenal judgements as viewed from an interactionist perspective, will here be granted short preliminary treatments.

Functionality and causality

A well known doctrine is that P-consciousness is functionally undefinable (Chalmers, 1996, p. 46-7). A problem is thus whether a non-physical P-consciousness may be the source of functional phenomena within the physical brain. Chalmers argues that if any interactionist influence would exist, that would require non-physical, yet *functional* variables; “psychons” (Chalmers, 1996, p. 156-8), as indeed proposed within Eccles’ interactionist perspective (Eccles, p. 87-8).

Thereby, however, the problem is merely displaced from the physical world to a non-physical sphere. The next question becomes whether the functionally undefinable P-consciousness can be the source of the psychons’ functioning. From an interactionist perspective, the challenge is to explain how P-consciousness *generally* can be the source of functional phenomena without itself being functionally definable.

As initial analogy, P-consciousness may be compared to an invariant light source. This light could be reflected in the shattered mirror surface of a parabolic sphere (here an analogy to a brain). Those mirror shards retaining an “angle” similar to that of an unbroken parabolic mirror will have surfaces being the most saturated with the light being reflected in them. Other shards,

which surfaces have become more or less parallel to the light beams, will see their surfaces less saturated with light.

While the light itself here symbolizes P-consciousness, the optimal saturation of light on shard surfaces symbolizes the interactionist influence upon brain segments. The saturation of light on shards' surfaces, furthermore, will be fully determined by the "angle" of each shard, not by variances or differentiations within the light source. Hence, any potential changes to the interactionist influence, although change is not represented by the analogy, appear definable by the functional processes of the physical brain itself.

All analogies are imperfect. An imperfection with the present one is that even light emanating from an invariant source is functionally definable as waves of electromagnetic energy. The following account is lacking regarding the structure of what argument it constitutes from a scientific perspective. My intention, however, is merely to draw attention to an *apparent* conceivability of a solution to the problem.

According to the theory of relativity, an implication of length contractions and time dilations occurring by motion (Sartory, p. 185-200), is that the spatio-temporal universe appears as lacking actual extension if seen from the "perspective" of light (Haisch, p. 95-7). That perspective is principally impossible to acquire through scientific observation, since no observing, physical object may ever behave like a photon of light (Sartory, p. 210).

From quantum physics, it is known that the spatio-temporal location of single photons may be indefinable thanks to so called "superpositions" (Halvorson, in Baker & Goetz p. 142-6). Without manifest location, furthermore, there is no manifestation at all. When photons *are* localizable, distances between them appear, through so called "entanglement", as causing no hindrance for their instantaneous "contact" (Halvorson, in Baker & Goetz p. 146-9).

The latter notion potentially even mirrors relativity theory, as also there, it is as if for photons, space and time lacks reality-defining power. Lee Smolin (Smolin, p. 210) predicted that the theory of relativity may be coherent with quantum physical theory only if one assumes that the universe exists according to a holographic principle (Smolin, p. 169-78).

What at any rate appears indicated is that some "factors" of reality exist beyond spatio-temporal parameters. Hence, the idea that P-consciousness could exist in or as some spatio-temporally indefinable state is coherent. P-consciousness could then potentially affect various brain areas non-functionally.

Given that all functionality is definable within spatio-temporal parameters, purported "psychons" could apparently exist only within a non-physical reality having spatio-temporal characteristics. Curiously, however, in light of a *general* solution to how P-consciousness may affect functional phenomena, even the idea of psychons becomes coherent.

What can be assumed transferred to the brain?

By interactionism, one may oft imagine that the non-physical P-consciousness “has” thoughts and feelings which get transferred to the brain. That would require that normal human experience can take place independently of the brain. As well known, however, if a part of the brain is destroyed, the psychological processes which rely on the functions of the damaged brain areas are disrupted. Consequently, I take as premise concerning normal human experience, that all experiential objects; everything we experience through sensing, thinking and feeling, are represented as neuro-functional brain activity.

I assume the same also when specifying experience as phenomenal (P-) experience. From a first person perspective, arguably, no phenomenal character or feature of experience seem barred from becoming motive of directed actions or verbalizations which obviously depend on neuro-functional brain processes. From a third person perspective, it arguably appears as the only rational alternative to trust first hand report and thus assume reports of blindness, example wise, to also imply the phenomenal lack of vision.

In light of the above, a question is how interactionism may be coherent with neuro-science. As an initial consideration, contrary to the position favoured by Tye (Tye, p. 155-82), it appears at the very least coherent that brain processes may be subconsciously accessible to P-consciousness. Not even subconscious experiences, however, will here be assumed to exist independently of brain activity.

Rather, I will assume that a direct effect of an interactionist influence consists of an *accentuation* of experiential objects already existing as patterns of neuro-functional brain activity. This better fits Eccles’ belief concerning inter-neuronal communication, that interactionist influences only may “*modify the probability of vesicular emission of the activated synapses*” (Eccles, p. 77).

Eccles nevertheless favoured the view, contrary to that here favoured, that interactionist influences may transfer discrete experiences (Eccles, p. 71-2). Within both Eccles’ view and the present, however, an interactionist influence should alter the frequency of action potentials in those neurons which combined activity make up the overall pattern of neuro-functional brain activity constituting the given experiential object.

An interactionist influence which direct effect is the accentuation of pre-existing brain activity could be constituted by a non-intentional, reflex-like mechanism, like in the previous “mirror analogy”. It would not have to constitute any consciously experienced effort. Further, it is also not necessitated that the brain in a direct manner should experience any “transmission” which mediates an interactionist influence. Rather, the influence could take place as a fully subconscious process.

In the above, however, no mechanism has been identified whereby the subconscious accentuation of experiential objects should be detectable *as that* by the brain. Another question is then whether the brain may detect the presence of an interactionist influence through a different mechanism. If neither alternative is possible, we may apparently not explain changes to

phenomenal judgements as resulting from interactionist influences. The second alternative, however, may in fact be coherent.

Required for this is a stable correlation between two known variables. Firstly, the amount of absorbed neurotransmitter molecules sufficient for the release of one action potential within a neuron (Klinke & Silberagl, p. 65-6). Secondly, the amount of neurotransmitters absorbed from the first neuron by an adjacent, neuro-functionally subsequent neuron as a consequence of the action potential of the first neuron (Klinke & Silberagl, p. 62-5).

We may assume that the suggested type of correlation, which may be uneven as long as stable, can operate despite input to the receptor neuron from “third neurons”. What is decisive is the amount of neurotransmitters absorbed from the “first neuron” as a consequence of one of its action potentials. Even within networks of interconnected neurons, we may single out, at least as a theoretical construct, the “linear” effect upon single neurons from other single neurons.

A potential reservation is the idea that synaptic transmission from third neurons could affect the very function of the receptor neuron so that the amount of neurotransmitters absorbed from the first neuron is thereby changed. However, such variance should be lawfully determined and thus principally correctable by theoretical models constructed to take height for effects of third neuron influences.

The suggested type of stable correlation will here be called “SLAN-correlation” (Stable Linear Absorption of Neurotransmitters Correlation). Apart from potential effects of interactionist influences, a mode of functioning according to which the SLAN-correlation exists may be ingrained through evolution as a premise for healthy psychological functioning. That notion is supported by the fact that its violation oft appears to be a mechanism of psychopathology (Laruelle, in Hirsch & Weinberger, p. 365-81).

Hypothetically, we can here imagine that some neuro-functional sequences of neurons within the brain are insulated against input from third neurons. Further, we may imagine excitatory synaptic connections between the neurons of such sequences to occur according to a “chain principle”. Further, inhibitory synaptic connections (Klinke & Silberagl, 67) could operate only between neurons separated by one or more intermediate neurons within the “excitatory chain”.

Such “insulated sequences” could potentially be activated exclusively in cases of SLAN-correlation violations. Each neuron of the sequence could be wired for transmissions through both excitatory and inhibitory synapses according to the above outlined principles. The result, by intact SLAN-correlation, could be lacking activation from the first inhibitory neuronal interconnection onward. Inhibitory and excitatory influences might then cancel each other out, as permitted by the principles of neuro-physiology (Klinke & Silberagl, p. 68).

By violation of the SLAN-correlation, however, excitatory influence may outweigh inhibitory influence in receptor cells receiving inhibitory input, thereby causing those cells’ activation. The reason is that in those cases, any accentuating effect of SLAN-correlation violations may accumulate over two or more excitatory synapses but over only one inhibitory synapse.

The above account is theoretically rather than empirically based. Any effort, furthermore, to formally demonstrate the theoretical coherence of the account exceeds the spatial confines of this essay. Nonetheless, I hold the coherence of the account as a theoretical premise for the arguments of this essay. For simplicity, activation of neuro-functional “insulated sequences” which lay un-activated except by SLAN-correlation violations will be called “LINS-activation” (Linear Insulated Neuro-functional Sequence Activation).

It is necessary, principally, to allow that LINS-activation may occur without an interactionist influence (positive error) and that it may be absent despite of an interactionist influence (negative error). More generally stated, it is not logically necessitated that SLAN-correlation violations, which may potentially also occur without interactionist influences, should always lead to LINS-activation.

By schizophrenia, as example, deregulation of the amount of neurotransmitters within single neurons leads to what is *here* termed SLAN-correlation violations (Laruelle, in Hirsch & Weinberger, p. 365-81), albeit not extremely abrupt ones. To suggest that this would lead to LINS-activation would appear unfounded, though the idea that it could occasionally do so is not incoherent.

Fluctuations also of other, even subtler and more instantaneously effective electro-chemical variables could cause so called SLAN-correlation violations, however (Eccles, p. 55-69). Those variables, Eccles described, could conceivably be influenced by quantum level changes. Like also Eccles, Chalmers sees the potential existence of quantum level interactionist influences as coherent (Chalmers, 2010, 126-8).

For theoretical reasons, a statistically high degree of correlation between interactionist influences and LINS-activations will here be assumed. Further, the occurrence of LINS-activations could conceivably get registered by the reflectively conscious function of the physical brain as an experiential phenomenon. Pertaining to phenomenal judgements, thus, the present interactionist account could potentially explain how effects of an interactionist influence may become objects of claims and cognitive beliefs.

Phenomenal Judgements & Interactionism

Phenomenal judgements refer to beliefs concerning the phenomenal character of experience (Chalmers, 1996, p. 173-5). Such beliefs can be expressed by claims such as “colours are mysterious” or “I am phenomenally conscious”, example wise. The problem of phenomenal judgements is that of giving an explanation which accounts for the existence and the real nature of such beliefs (Chalmers, 1996, p. 184-6).

The “hard problem” of consciousness, strictly interpreted, is that of explaining the existence of P-consciousness (Chalmers, 2010, p. 3-6), not merely the existence of the *belief* that it exists. The problem of “the explanatory gap”, next, is that of why there is an indescribable (strictly assumed phenomenal) side to discrete experiences (Chalmers, p. 1996, p. 47). With strict interpretations of both problems, the problem of the explanatory gap will be integral to the hard

problem, since all phenomenal experiences require a phenomenal subject (Strawson, in Freeman, p. 189-91).

The hard problem is relevant to phenomenal judgements only insofar as actual phenomenality constitutes a reason for beliefs about phenomenality. Eliminativists have denied the very existence of P-consciousness (Macpherson, in Freeman, p. 75) and logically thus also that of a strict “hard problem”. The problem of phenomenal judgements, then, can potentially (but must not) be answered in concert with an answer to the hard problem.

If remaining “undogmatic” regarding the strictness of the hard problem, P-consciousness appears potentially reducible to any reductively definable property which is capable of solving the problem of phenomenal judgements. Tye holds that there may be a crucial difference between “knowledge by description” and “knowledge by acquaintance”. Furthermore, that no part of the former can be “part and parcel” of the latter (Tye, p. 139). Feelings that something is missing to statements about phenomenal qualia, example wise, such as also implied by the explanatory gap, could thus be explainable (Tye, p. 143).

For Tye, in the sense that he is no eliminativist, P-consciousness exists yet *is* a reductive property, namely the brain state(s) for which knowledge by acquaintance takes place (Tye, p. 144-5). If his theoretical framework is correct, however, it serves his case of defending physicalism only insofar as it defends any substance monist view, including variances of property dualism (Chalmers, 1996, p. 124-5). Conceivably, it could explain why our cognitive beliefs *correspond* to our phenomenal experience. If so, the hard problem would remain, even as the problem of phenomenal judgements could appear solved.

The phenomenal judgement, however, that without the problem of phenomenal judgements, the hard problem could still remain, relies on first person knowledge. If one chooses to take first person perspectives upon consciousness seriously, it may arguably appear that there is something to P-consciousness which is not only intellectually unexplainable (a mystery), but something taking on an existential importance (a captivating mystery). The philosophical question of what extent to which first person knowledge should be taken seriously (Taliaferro, in Baker & Goetz, p. 26-40), however, will be left untouched at this point of the discussion.

An interpretation of Tye

For the sake of the subsequent discussions, a short interpretation of Tye’s perspective on phenomenal judgements will here be presented. Since Tye sees phenomenal experience and knowledge by acquaintance as synonymous, he understands the implicated explanatory gap between knowledge by acquaintance and by description as that which is assumed integral to the hard problem.

Because knowledge by acquaintance has no part or parcel of knowledge by description, Tye sees the character of all phenomenal experiences as linguistically indefinable, like also implied by the well known philosophical notion of inverted qualia (Chalmers, 1996, p. 263-6). Contrary to appearance, this is consistent with Tye’s conviction that P-consciousness is definable as a property of the function of the physical brain.

Using the colour “blue” as example, that colour is descriptively definable as a pattern of brain activity. It appears that also the phenomenal character of blue experience is “fixed” by the neuro-functional characteristics of that very same brain state. This should be assumed even by property dualism and interactionism, given that all normal human experience relies on brain activity.

Tye holds that even from a purely reductive, physicalist perspective, one may *recognize* (pick out) “blue” as a colour separate from other colours regardless of whether one is aware of or able to describe its neuro-functional characteristics (Tye, p. 139). Descriptions of the brain states defining the experience, furthermore, may never convey the character of the experience by acquaintance which allows the recognition. We see, thus, that although Tye sees P-consciousness as neuro-functionally definable, he does not believe the same to concern the subjective *character* of discrete phenomenal experiences.

The character of experiences by acquaintance represents “brute qualities”, identifiable only by expressions such as “one of those” (Chalmers, 1996, p. 288-92). True, they can also be referred to using some random token name, like “blue”. The meaning even of such consensual token names, however, is flexible given the existence of inverted spectrum variances. Further descriptive knowledge will also be unhelpful in specifying the phenomenal character of the experience of a concepts’ referent. Example wise, telling someone that blue is the colour of the sky is unhelpful given inverted spectrum variances.

A “concept”, Tye holds, is deferential (Tye, p. 40-1), meaning that it can be “possessed” even if it is not fully understood (Tye, p. 63-74). A person thus possessing the concept BLUE without full understanding could say about the colour indigo that it is blue. He or she could be colour blind and experience all blue as indigo. Hence, it is not necessary to have undergone acquaintance with the colour blue in order to possess the concept BLUE (Tye, p. 66).

There is, of course, a problem of defining “full understanding of concepts”. True, if a person is either colour blind or experiencing inverted qualia, he or she must logically, at least within an assumed physicalist reality, differ from other people as to how colours are represented in his or her brain. Unless the presence of the relevant type of neuro-functional brain activity is determined in each single case, however, it will remain a matter of mere presumption which phenomenal character is consensually assigned to each colour concept.

The above underlines that, for all practical purposes, Tye presents a physicalist account which respects the lingual indefinability of phenomenal qualia. What might still appear as a remaining question is that of why people commonly fail to realize that the character of experiences by acquaintance cannot be referred to by mere token names of brute qualities.

The principle of inverted qualia might rarely be spontaneously realized, yet this does not make it a scientific mystery. The erroneous belief that concepts may refer to phenomenal qualia appears to rely on a simple ego-centric error, namely the implicit, un-reflected belief that the own subjective perspective is defining of interpersonal reality. In summary, thus, Tye’s perspective allows the notion that people believe to refer to phenomenal qualia using “normal” concepts.

A non-physicalist perspective

Initially in agreement with a non-interactionist, property dualistic perspective (Chalmers, 2010, p. 243-4), I suggest that phenomenal experience is conveyed by the same patterns of brain activity which define experiences in cases of knowledge by acquaintance. The phenomenal side to the experience of any experiential object will thus be definable as the phenomenal experience of the neuro-functional pattern of brain activity defining the experience of that experiential object. We here end up with a double explanatory gap, one between knowledge by acquaintance and knowledge by description, another between phenomenal and non-phenomenal experience of that known by acquaintance.

A comment on the relationship between knowledge by acquaintance, knowledge by description and the meaning of concepts will be useful for some of the following discussions. Using Chalmers' example, the concept WATER can have different secondary intensions (known a posteriori) like " H_2O " or, in a hypothetical other world; "XYZ" (Chalmers, 1996, p. 57). The primary intensions could in both cases be "*the dominant clear, drinkable liquid in the oceans and lakes*" (Chalmers, 1996, p. 57).

However, the brute qualities of water known by acquaintance must possess an a priori nature even relative to its primary intension. Through its descriptive content, the latter is generalizing and classifying of what may (but must not) be a priori "acquaintances" with the brute qualities of the experience of water. Sainsbury and Tye similarly hold that the so called "two-dimensional semantic" with primary and secondary intensions "*fails to connect in a natural way to the notion of a priori knowledge*" (Sainsbury and Tye, p. 36).

From the perspective of an originalist theory of concepts, the acquiring (taking into possession) of a concept can be separate from its origin, since concepts typically are shared (Sainsbury & Tye, p. 40-4). Hence, the acquiring of concepts may oft reflect the adoption of a non-originating use of the given concept, although one may later learn the concept's originating use.

In such cases, acquiring cannot occur based on a priori experience by acquaintance of a non-descriptive referent phenomenon. Even if the originating use of a concept would have such a priori experience as referent, through adopting a non-originating use, the acquiring will not involve such experience. Oft, furthermore, even the originating use of a concept has merely descriptive referents, like in the case of the concept QUARK (Sainsbury & Tye, p. 42-3).

Sometimes, however, the acquiring and the origin of a concept may be synonymous (Sainsbury & Tye, p. 42). This allows the occurrence of discoveries that the given concept has the same originating use as a pre-existing, shared concept (Tye, p. 39-40). It appears that exclusively in such cases, the use of concepts *can* (but must not) reflect the a priori experience of the given concepts' non-descriptive referent phenomena by acquaintance.

Although here assumed that only experiences by acquaintance may have phenomenal sides, even concepts which use involves merely descriptive referents constitute experiential objects existing as patterns of brain activity. Consequently, also these should *secondarily* be possible to experience a priori by acquaintance. Even in the case of a posteriori, descriptive knowledge,

thus, it will be assumed that one may possess a phenomenal side to the experience *of* this knowledge.

The core interactionist hypothesis

The non-physical P-consciousness is assumed only *ideally* to be the source of the previously purported LINS-activation. Furthermore, no identified mechanism appears to allow the brain to decide whether the SLAN-correlation is violated because of an interactionist influence or thanks to some reductively definable cause. Indeed, SLAN-correlation violations may not be detectable in any direct way from the perspective of the psyche.

I previously suggested that they may at least sometimes be detected indirectly through the occurrence of the LINS-activation. Another consequence of SLAN-correlation violations could logically be the abrupt, maybe intrusive appearance of thoughts, feelings or sensory impressions to conscious awareness. As an extreme, but thereby obvious example, we could take schizophrenic hallucinations (Laruelle, in Hirsch & Weinberger, p. 365).

As a potential consequence of all such experience, furthermore, the psyche may cumulatively come to experience a lack of autonomy and integrity of the psyche. The apparent truth; that the psyche does not “master” the physical brain, but can only function adaptively given a fine, homeostasis-like balance within it (the SLAN-correlation) may hypothetically constitute an existential threat to the psyche. It is a known doctrine within psycho-analytically oriented theory that the psyche harbours “unrealistic needs”, such as the preservation of an illusion of solidity, permanence and perfection, which Charles Hanly described as “*a self-image that is distorted by idealization*” (Epstein, p. 24).

Analogically, a need of the psyche not to experience itself as at the mercy of potentially merciless neuro-physiological processes of the physical brain is conceivable. An archaic and subconscious psychic defence mechanism is “projective identification”, the ejection of internal threats in order to perceive (identify) them as external ones (Ogden, p.144-6). Wilfred Bion held that projective identification can account even for many bizarre psychotic experiences of schizophrenics (Ogden, p. 146).

By schizophrenic psychoses, patients frequently claim to have thoughts that are not their own (AMDP, p. 85-6). As above assumed, such “intrusive” experiences may cause a sense of the psyche’s lacking autonomy or integrity and be experienced as threatening. Beside the mentioned sense of thought insertion, projective identification may lead to delusions of alien influence (AMDP, p. 86-7) and persecution (AMDP, p. 68-9). The latter could more directly account for the projective identification of an experienced, internal threat.

Also a priori experience by acquaintance of the LINS-activation may, according to the same principle as above outlined, be experienced as a threat to the autonomy and integrity of the psyche. This, because LINS-activations, which is thought to exclusively follow SLAN-correlation violations, may be experienced as abrupt, intrusive presences felt by the psyche as something uncontrollable.

However, if a concept LINS-ACTIVATION would be rightly possessed, unlike typically the case among other concepts having referents being experienced by acquaintance, it might represent a phenomenon with no positively identifiable referent. That is, it may logically represent nothing in the *external* physical world and might also constitute no discrete thought, feeling or illusory sensory impression.

Hypothetically, the LINS-activation could be experienced as was it nothing at all, still penetrantly present *as that* in an a priori fashion. Effectively, thus, the LINS-activation could be unique as being impossible to positively describe a posteriori, bar in terms of the very brain activity constituting its neuro-functional referent. Devoid of any positively identifiable, descriptive referent phenomena, the only accessible referent of the LINS-activation might thus be the very character of its experience by acquaintance.

Descriptively identifiable only as negations like “nothing” or “absence”, the LINS-activation may find no physical referent at all. By projective identification, thus, the only accessible referent of the LINS-activation, the character of its experience by acquaintance, might thanks to lacking alternatives be falsely judged as transcendent and ontologically non-physical. Purely hypothetically, illusory beliefs in transcendent qualities to experience may even be “welcomed” by the psyche, since transcendence could serve the “unrealistic needs” of solidity, permanence and perfection.

According to the current hypothesis, beliefs in a non-physical character to experiences would be “falsely founded”, since the LINS-activation actually refers to a physical brain phenomenon. Hence, the presumed, *actually* non-physical nature of P-consciousness would not be the direct reason for beliefs about that very non-physical nature. Still, it could constitute an indirect reason for it in the form of an interactionist influence.

Answer to some criticisms

A potential criticism of the above interactionist hypothesis can be expressed through the question of how we justify the idea that phenomenal experience merely is present as the LINS-activation. Answering this, it should once more be clarified that between phenomenal and non-phenomenal experience, no different functionality in the brain is here assumed, like also not in cases of property dualism.

Presumably, there is always a phenomenal side to experiences by acquaintance. The purported LINS-activation entails, on the one hand, one single and separate phenomenal experience among the multitude of other, phenomenal experiences. On the other hand, it may also constitute the reason for convictions about a quality over and above everything physical to the experience of any experiential object currently accentuated by an interactionist influence

One could still ask how, without adding any new experiential contents, the LINS-activation alone may explain judgements about the entire range of phenomenal experiences. The answer is that the LINS-activation is assumed to ideally occur because of an interactionist accentuation of discrete experiences already present. The decisive factor regarding the character of each

phenomenal experience is then the pre-existing experience *being* accentuated by the interactionist influence.

The above may still be further questioned. In light of the above accounts, it appears that people should believe recognizing a reductively indefinable character to the experience of the LINS-activation rather than to that of the experiential object being accentuated. A confusion may be comprehensible, however, because the LINS-activation and the accentuation of the present experiential object presumably occur simultaneously and co-dependently upon an interactionist influence.

Further and as previously mentioned, the LINS-activation may be thought of as being, in a sense, “object-less”. That is, the only identifiable experiential “contents” could be those of the experiential object *being* accentuated. Given the sum of those circumstances, a fallacy of correlation and causality (Losee, p. 41-8) could conceivably occur. Hence, one might believe the impression of a transcendent, non-physical quality to the character of any present experience (a-phenomenally defined) to be caused by the presently accentuated experiential object rather than by the LINS-activation.

Another question is that of whether the present interactionist account may explain why and how the physical brain can believe in the existence of phenomenality to experience. After all, the brain must produce the claims reflecting the beliefs about phenomenality. The flip side to the same problem is the question of how P-consciousness may believe that functional brain states refer to its own phenomenal experiences.

Even within an interactionist account of reality, however, we must assume that the physical brain alone may know nothing about phenomenal (P-) experience. Like also according to a property dualistic rationale (Chalmers, 1996, p. 124-5), the physical brain and P-consciousness could mutually lack ways to “inform” the other that a discrepancy exists between an experience as functionally definable and its phenomenal side. The discrepancy may simply not matter as long as the neuro-functional determinants of the experience are identical.

Thus, even as one experiences the phenomenal side to an experiential object, one may be “blind” to the fact that there is a difference between an experience as functionally definable and its phenomenal side. This also implies “blindness” to any potential experience that there should be a mystery to the notion that the character of phenomenal experience is reductively definable.

In the absence of any LINS-activation, thus, P-consciousness may experience the phenomenal side to discrete experiences by acquaintance without accompanying beliefs that the character of those experiences is transcending everything physically definable. The specific case of the LINS-activation, however, may constitute an exception. As previously outlined, the character of its experience by acquaintance could hypothetically be experienced as transcendent and ontologically non-physical.

The explanatory power of the present hypothesis

An important question is whether an interactionist influence could result in beliefs differing from those of a phenomenal zombie (Chalmers, 1996, p. 94-9) following “normal” knowledge by acquaintance. In light of the difference between knowledge by acquaintance and by description, zombies’ beliefs about a lingually indefinable character to experiences may take place without interactionist influences. Exactly like Tye holds, this could appear to solve the problem of the explanatory gap. This would also leave the present interactionist hypothesis stripped of any unique explanatory power.

Rightly, the character of the experience by acquaintance of the colour blue is seen as indescribable from Tye’s perspective. However, within any substance monist view, of which physicalism and variances of property dualism are examples, a strong subjective conviction that the indescribable character of blue experience is transcendent should not arise. As above mentioned, P-consciousness would arguably be unable to detect any major mystery to the notion that phenomenality should be reductively definable.

It was suggested, however, that the experience by acquaintance of the LINS-activation could appear as having a non-physical character to it. Obviously, even without any purported LINS-activation one may form *hypotheses* (and incidentally even favour these) that the character of experiences by acquaintance may best be explained as transcendent. This would then work according to the principles of scientific hypotheses in general, which of nature are a posteriori, matter-of-factly and affect-less.

Violations of the matter-of-factly and hypothetical nature of claims as expectable concerning a posteriori descriptive notions could indicate the presence of the direct a priori experience of some phenomenon. If, additionally, such “irrational convictions” would concern notions of transcendent or ontologically non-physical aspects of experience, this would fit that which is expectable within the current interactionist hypothesis.

As previously mentioned, both positive and negative errors might occur regarding the co-occurrence between an interactionist influence and LINS-activation. Thus, irrational convictions must sometimes be granted to be “just” irrational. Given that errors are exceptional, however, the current interactionist hypothesis generally favours the notion that first person knowledge may have a central role when it comes to our judgements about the nature of consciousness.

A question is still why we should prefer the present interactionist hypothesis over new, more complex forms of substance monism. In those alternative scenarios, the LINS-activation always occurs by reductive means, still causing phenomenal judgements about the character of phenomenal experience as being transcendent. Within a property dualistic scenario, as example, one could imagine that beliefs about phenomenality would fully correspond to *actual* phenomenality, yet the beliefs would not even indirectly be caused by that phenomenal reality.

However, the LINS-activation would then have to be seen as resulting from random deregulation of the SLAN-correlation. This, furthermore, does not appear to be the principle according to which our phenomenal judgements are produced. Example wise, we do not hear claims that the

character of the experience of the colour blue *some times* and *without predictability* has a quality transcending everything physically definable.

In summary, although a traditional property dualistic perspective may explain beliefs that one refers to phenomenal experiences, it leaves the subject “blind” to the difference between phenomenal experience and knowledge by acquaintance. Tye’s physicalism explains the same beliefs without that dilemma, since it sees phenomenal experience and knowledge by acquaintance as synonymous. Nonetheless, it seems that only an interactionist alternative may account for the mystery of consciousness as expressed by the first person experience of the hard problem.

The problem of lingual reference to P-consciousness

As evident by the fact that people try to express convictions about P-consciousness in language, certain concepts must frequently be believed to refer to P-consciousness. All thus relevant concepts, however, have reductively definable referents, existing independently of any purported interactionist influence. A fitting example is exactly the concept PHENOMENAL CONSCIOUSNESS, here understood as distinct from P-consciousness.

The word “phenomenal” is a term from philosophy, especially linked to 19th century philosopher Edmund Husserl. It concerns the manner in which experiential objects (phenomena) “appear” to consciousness (Fahey, online). The word “consciousness”, next, can be defined as a mode of functionality which can be shared by the neuro-physiology of biological organisms and artificial computing systems (Chalmers, 1996, p. 275).

Tye argues that a reductively definable mental state (or states) *for which* knowledge by acquaintance takes place likely is the referent of P-consciousness (Tye, p. 144-5). I also judge it as safe to infer that Tye, from his physicalist perspective, sees that (or those) state(s) as the referent phenomenon of the concept PHENOMENAL CONSCIOUSNESS. For reasons of simplicity, such (a) mental state(s) will here be referred to as “the mental state of acquaintance”. In that state, no agency appears involved, only passively “receptive” experience.

Many concepts other than PHENOMENAL CONSCIOUSNESS could have the same, reductively definable referent. Examples are concepts such as PURE BEING and SUBJECTIVITY. Another example is the concept I, which is arguably also used with the belief that it refers to P-consciousness. Accordingly, statements like “my brain registers colours, but only I see their qualia” may follow. According to Mark Epstein (Epstein, p. 47-8), the “I” is a portion of the psychological ego. Further, all psychological phenomena are functionally definable (Chalmers, 1996, p.46-7).

A question is whether the reductive referents of the two concepts PHENOMENAL CONSCIOUSNESS and I could be reduced to one single, neuro-functional brain phenomenon. Given that the two concepts may be believed to refer to the same phenomenon, we would thereby avert a theoretical problem. Examples illustrate, however, that “the mental state of acquaintance” does not appear to be synonymous with the reductive referents of the concept I.

By many sleeping dream states or other states induced by drugs, hypnosis, meditation or traumata, experience by acquaintance appears to take place without being “brought under” the “I” (Howell, p. 19-20). The “I”, on the other hand, can apparently experience itself as an agent behind actions (Epstein, p. 47-8). That violates the assumption that the mental state of acquaintance is purely passive, without agency.

Hypothetically, the reductive referents of the concept *I* and “the mental state of acquaintance” could nonetheless be intertwined phenomena. The “I” naturally appears to contain an “observer-observed duality” due to its self-reflective function. The observed part of the “I” may oft be its agent-part (Epstein, p. 47-8), yet may apparently also be its very observing part. The “observed observer” of each new second appears to require a “pure observer”. The self-reflective process of the “I” may thus appear like a Russian Matryoshka doll with an endless amount of concentric layers. David Chalmers made a humorous note of a similar point (Chalmers, 1996, p. 230).

Sainsbury and Tye hold that we use a specific “I-concept” (presumably the concept *I*) to think about ourselves (Sainsbury & Tye, p. 144-5). Potentially, the observed part of the “I” (including the observed observer) corresponds to the concept *I*, while the *momentarily* observing instance is not part of it. Conceivably, the momentarily observing instance, when understood as detached from the concept *I*, is synonymous with the mental state of acquaintance. Thus, the reductive referent of both the concept PHENOMENAL CONSCIOUSNESS and the concept *I* could ultimately be identical.

A non-physicalist perspective

Concerning P-consciousness as such, the situation appears by first glance to be more complex than concerning a phenomenal character of discrete experiences such as those of sensory objects. Regarding colours, as example, the phenomenal side to a colour as neuro-functionally definable is “*the colour as it is phenomenally experienced*”. Similarly, it appears that the phenomenal side to the mental state of acquaintance simply is “*the mental state of acquaintance as phenomenally experienced*”. This cannot be thought of as being synonymous with P-consciousness. The same goes for the “I”, since the “I” as phenomenally experienced is not the same as P-consciousness.

Still, the “I” and “the mental state of acquaintance” may display certain idiosyncratic properties worthy of further exploration. Firstly, P-consciousness and the mental state of acquaintance may conceivably “map onto” each other, meaning that experiential objects, features and qualities available to P-consciousness could be exclusively and exhaustively those accessible by acquaintance. This notion is also adaptable to Tye’s view that P-consciousness simply *is* what is here called “the mental state of acquaintance” (Tye, p. 144-5). From a property dualistic perspective, furthermore, the mental state of acquaintance could be experienced as a state within which P-consciousness is solely present.

When it comes to the “I”, next, P-consciousness could, within the present interactionist account, observe the observed portion of the “I” (including the observed observer) through the intimately experienced lens of the *momentarily* observing instance. The momentarily observing instance and the observed observer of the “I” could be experienced as synonymous (this will be further outlined). Next, given that the momentarily observing instance is synonymous with the mental

state of acquaintance, the observed observer of the “I” could also be experienced as synonymous with the mental state of acquaintance.

The mental state of acquaintance could thus possess properties which allow a phenomenal illusion that it is synonymous with P-consciousness. We do not here talk about an illusion in the form of a cognitive belief. Rather, as according to the previously described principles, the person could simply be “blind” to the difference between phenomenal experience and experience by acquaintance. Here, specifically, he or she would be blind to the difference between P-consciousness as such and the mental state of acquaintance. If we accept a standard property dualistic rationale, this could in fact appear to be the natural end point of the current hypothesis.

Scrutinizing the interactionist hypothesis

Hypothetically, the brain activity which constitute the reductive referents of “the mental state of acquaintance” could become accentuated by an interactionist influence. This could cause LINS-activations in the same manner as assumed if, as example, the reductive referents of an experiential object like “vision of the sky” would be accentuated.

Since all experiences must be intimately linked to the presence of an experiencing subject, the interactionist accentuation of discrete experiential objects and of the mental state of acquaintance could potentially always be two sides of the very same event. If so, *all* interactionist influence could be thought to influence phenomenal judgements about both the present experiential objects as well as P-consciousness. We may call that scenario “unitary accentuation of object and subject”.

As previously suggested, the experience of the LINS-activation could give rise to the illusion that there is something transcendent to the experience of that which is being accentuated. Concerning discrete experiential objects, simultaneous and co-dependent a priori experiences of LINS-activations and of experiential objects accentuated by an interactionist influence was assumed. Further, that only the discrete experiential objects have positively identifiable referents, so that LINS-activations are not sensed as independent experiential phenomena.

When trying to picture a similar rationale for cases of an interactionist accentuation of the mental state of acquaintance, we encounter more difficulty. Tye thinks the “realization” that (a) mental state(s) must exist which is “having” experiences by acquaintance (the mental state of acquaintance) only takes place a posteriori (Tye, p. 144-5). If correct, we must here assume the same to be the case also for P-consciousness.

Initially, Tye’s view fits the notion that the mental state of acquaintance may be phenomenally *felt* as synonymous with P-consciousness. This, exactly because there is nothing about it which can be experienced a priori. Rather, it would operate more like an intimate lens for P-consciousness’ experience of other phenomena. However, by unitary accentuation of object and subject, any phenomenal judgements resulting from confusion of that being accentuated and the LINS-activation, would thus appear to concern discrete experiential objects to the exclusion of concerning the mental state of acquaintance.

By separate accentuation of object and subject, on the other hand, LINS-activation following accentuation of the mental state of acquaintance would be experienced a priori without any a priori experience of anything at all being accentuated. This could, by analogy, potentially explain false possession of the concept PHENOMENAL CONSCIOUSNESS. That is, one could imagine that any reference to the LINS-activation as “phenomenal consciousness” or as “the mental state of acquaintance” would be more or less analogical to the naming of indigo as blue in the previous example of colour blindness.

The latter account would imply, however, that since the person does not know the actual mental state of acquaintance a priori, he or she would have no clear conception of what it really means. This, furthermore, does not do justice to what was assumed previously in this essay. It was then assumed that first person, a priori knowledge by acquaintance is both predominantly reliable and decisive for understanding phenomenal judgements.

From a non-interactionist perspective, it appears feasible that the a posteriori realization that a mental state of acquaintance exists is facilitated by awareness about the existence of discrete experiences by acquaintance. This may largely explain temporal co-occurrence of experiences by acquaintance and beliefs about a corresponding mental state having those experiences. It could also explain how one may realize that that mental state is ontologically linked to the character of experiences by acquaintance.

The challenge with the latter alternative is to explain how the insight about the mental state of acquaintance may suffice for the formation and use of the concept *i*. A requirement must be that the actual mental state of acquaintance, having the experiences by acquaintance which facilitate the a posteriori insight about that same mental state, will experience the object of that insight as a “solid” reference to itself.

It appears unclear, at best, whether the above is feasible, given that any a posteriori insight refers to an abstract, merely inferred object. Next, I will loosely outline an account which, in the best case, could aid both non-interactionist perspectives as well as interactionism past the above identified problems. Specifically, I suggest that the mental state of acquaintance may be experienced a priori, despite of Tye’s conviction to the contrary.

We may imagine the mental state of acquaintance as “emitting” neuronal signals in a manner analogical to how light signals were emitted by P-consciousness within the previous “mirror analogy”. As result of (non-interactionist) interaction with discrete experiential objects defined as neuro-functional patterns of brain activity, those neuronal signals could be altered, thereby coming to register and represent present experiential objects.

The mental state of acquaintance might have the capacity of possessing an underlying, intrinsic focus. This focus must not be understood as a reflectively conscious experience. By mere analogy to the visual mode of experience, the suggested focus might conceivably be compared to a direct experience of the sum total of brightness as was light an independent experiential phenomenon. That is, independently of the capacity of light, as medium, to define discrete experiential objects through colours and shades.

Since we talk about an intrinsic self-experience of the mental state of acquaintance, we must not assume that P-consciousness experiences the mental state of acquaintance a priori. This would have contradicted what was previously assumed. We must only assume that P-consciousness may experience, like may also the mental state of acquaintance, the a priori self-experience of the latter as some sort of brute quality.

A priori self-experience as here suggested should not be taken note of by the brain as something mysterious. Thus, the present account is fully adaptable to substance monist views like physicalism and versions of property dualism. Given interactionism, however, the neuro-functional manifestation of the a priori self-experience of the mental state of acquaintance may get accentuated by an interactionist influence. Further, the purported general confusion between any experiential object accentuated by an interactionist influence and the LINS-activation may then lead to the experience that there is a transcendent quality to the mental state of acquaintance. This, according to the exact same principles as outlined concerning discrete experiential objects.

Lastly, an "object" experienced a priori, such as here the mental state of acquaintance, appears sufficiently "solid" for the formation and use of the concept *i*. The reflective awareness about the self-experience held by the mental state of acquaintance could bring the mental state of acquaintance under observation and the concept *i*. Thereby, we might potentially also better understand how the momentarily observing instance and the observed observer of the "I" may be experienced as identical. Further, reflective awareness may logically only occur with a minimal delay relative to any referent, a priori experience. Since the a priori experience of the mental state of acquaintance may take place anew each second, we might thereby also understand the mentioned "Matryoshka-doll effect".

Conclusion

Initially in this essay, some problems were mentioned, the coherent solutions to which are instrumental for the notion that discussions of interactionism are warranted at all from a scientific perspective. The problems include those of causal closure of the physical universe, neuro-functional manifestation of potentially quantum level interactionist influences and the notion that causality requires functionality. Lastly, whether from a psychological perspective, thoughts and feelings could coherently be assumed transferred from a non-physical sphere to the physical brain.

In the central part of the essay, comparison was made to Michael Tye's physicalist perspective. I accepted Tye's division between knowledge by description and knowledge by acquaintance and that this represents an "explanatory gap". Distinctly, I suggested the existence of a "double" explanatory gap. I presumed a "phenomenal side" to all experience by acquaintance. This allows, also from the perspective of a non-physicalist hypothesis, that all discrete phenomenal experience is defined by physical brain processes. In isolation, this reflects a variance of a property dualistic rationale.

Because of theoretical complexity and space constraints, many aspects of the distinctly interactionist parts of the essay's hypothesis will remain treated exclusively in the main text. What should be mentioned, however, is that this essay purports that interactionist influences from P-consciousness upon the physical brain are reflex-like, content-less and subconscious. No discrete experiential objects are assumed "transferred" to the brain from a pre-existing state within a non-physical sphere. Importantly, however, the interactionist influence is still assumed to effect a process through which a distinct neuro-functional consequence, the so called "LINS-activation", may be experienced by the physical brain.

Two further features of the present hypothesis are worth emphasizing. It was assumed that "irrational convictions" about any transcendent or non-physical character to phenomenal experience may follow an interactionist influence. The present hypothesis necessitates the allowance of error, so that irrational convictions sometimes may be "just" irrational. Simultaneously, it accounts for the idea that the mentioned irrational convictions about phenomenal experience typically may result from first person experience of the neuro-functional consequence of the interactionist influence in the physical brain. Hence, the present hypothesis grants a very significant, albeit not necessarily all-decisive role to first person knowledge when it comes to explaining phenomenal judgements.

Lastly, the present hypothesis accounts not only for phenomenal judgements about the phenomenal character of discrete experiential objects such as thoughts or the sensory representation of physical objects, but also for ones about P-consciousness as such. Tye achieves the same by assuming that a reductively definable mental state (or states) is synonymous with P-consciousness. He thereby disputes the known doctrine that P-consciousness is functionally indefinable. The present interactionist hypothesis, on the other hand, accounts for beliefs that one may lingually refer to P-consciousness while also accepting its functional indefinability.

References

- Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (2007) *Das AMDP-System. Manual zur Dokumentation Psychiatrischer Befunde* (8th Ed), Göttingen, Germany: Hogrefe.
- Baker, M. C. & Goetz, S. (eds.) (2011) *The Soul Hypothesis. Investigations into the Existence of the Soul*, New York: Continuum.
- Chalmers, D. J. (1996) *The Conscious Mind. In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- Chalmers, D. J. (2010) *The Character of Consciousness*, Oxford: Oxford University Press.
- Eccles, J. C. (1994) *How the Self Controls Its Brain*, Berlin, Germany: Springer-Verlag.
- Epstein, M. (2007) *Psychotherapy Without the Self. A Buddhist Perspective*, New Haven, Connecticut: Yale University Press.

Fahey, T. (2008) *Edmund Husserl*, [Online], <http://www.tonyfahey.com/2008/12/edmund-husserl.html> [03 March 2013].

Freeman, A. (ed.) (2006) *Consciousness and its Place in Nature. Does Physicalism entail Panpsychism?*, Exeter, UK: Imprint Academic.

Haisch, B. (2006) *The God Theory. Universes, Zero-Point Fields, and What's Behind It All*, San Francisco: Weiser Books.

Hirsch, S. R. & Weinberger, D. (eds.) (2003) *Schizophrenia Part Two Biological Aspects*, Oxford: Blackwell Publishing.

Howell, E. H. (2005) *The Dissociative Mind*, New York: Routledge.

Klinke, R. & Silbernagl, S. (1996) *Lehrbuch der Physiologie* (2^d Ed), Stuttgart, Germany: Thieme.

Losee, J. (2011) *Theories of Causality*, New Brunswick, New Jersey: Transaction Publishers.

Ogden, T. H. (1986) *The Matrix of the Mind. Object Relations and the Psychoanalytic Dialogue*, Lanham, Maryland: Rowman and Littlefield, 2004.

Papineau, D. & Selina, H. (2000) *Introducing Consciousness*, New York: Totem Books.

Popper, K. & Eccles, J. C. (1977) *The Self and Its Brain. An Argument for Interactionism*, Abingdon, UK: Routledge, 2003.

Sainsbury, R. M. & Tye, M. (2012) *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concept*, Oxford: Oxford University Press.

Sartory, L. (1996) *Understanding Relativity. A Simplified Approach to Einstein's Theories*, Berkeley: University of California Press.

Smolin, L. (2001) *Three Roads to Quantum Gravity*, New York: Basic Books.

Stoljar, D. (2010) *Physicalism. New Problems of Philosophy*, Abingdon, UK: Routledge.

Tye, M. (2009) *Consciousness Revisited. Materialism without Phenomenal Concepts*, Cambridge, Massachusetts: The MIT Press.