

Psychoanalytic Theory: A Historical Reconstruction

In this paper I sketch a reconstruction of the basic psychoanalytic conception of the mind in terms of two historical resources: the conception of the subject developed in post-Kantian idealism, and Spinoza's laws of the affects in Part Three of the *Ethics*. The former, I suggest, supplies the conceptual basis for the psychoanalytic notion of the unconscious, while the latter defines the type of psychological causality of psychoanalytic explanations. The imperfect fit between these two elements, I claim, is reflected in familiar conceptual difficulties surrounding psychoanalytic theory and explanation.

I

Enquiry in the history of philosophy may serve a variety of purposes and assume a variety of forms. In addition to tracking the actual historical story of the formation of ideas, we may seek to illuminate the thought of historical figures by drawing systematic comparisons in the absence of any relation of historical influence. Longuenesse's account of the relation between Kant's 'I ought' and Freud's superego fits this description, and I am in full agreement that there is insight to be had by drawing out systematic affinities between concepts, even when these are historically remote in the sense of belonging to quite different schools or traditions, or of participating in quite different intellectual projects, as is evidently the case with Kant and Freud.

Another thing which we may legitimately do, with the same end of philosophical illumination rather than strict historical explanation in view, is to offer systematic reconstructions of concepts and theories in terms of resources supplied by the history of philosophy. In this vein, I am going to use the space available to suggest a reconstruction of psychoanalytic theory—more

specifically, of the basic conception of the mind evidenced in Freud's metapsychology—in terms of historical materials. The reconstruction falls into two parts. The first proposes a source for the psychoanalytic conception of the unconscious in post-Kantian idealism. The second attempts to locate a point of affinity of psychoanalytic explanation with Spinoza. Although I regard the first as a defensible historical claim—I think it is plausible to maintain on historical grounds that the legacy of idealism transmits itself into Freud's model of the mind, and I will formulate my account in such terms—for present purposes it can be taken in a purely systematic spirit, which is in any case how my second claim, concerning Spinoza, is intended. What follows may be regarded therefore as a purely analytical exercise cast in the form of a just so story: in a nutshell, I am suggesting that psychoanalytic theory may be viewed as the product of combining a conception of the subject developed in post-Kantian thought with a Spinozist view of the mechanics of human motivation; as if Freud had come up with the metapsychology as a facilitating framework for developing empirical psychological hypotheses through a reading of classical German philosophy and the *Ethics*. In the final section of the paper I will suggest that the idealist and Spinozist dimensions of psychoanalytic theory do not cohere fully, and that the tension between them is responsible for and reflected in certain familiar conceptual difficulties that have been held to confront Freudian theory. The ultimate point of the reconstruction is thus to shed some light on the famously puzzling character of psychoanalytic theory and explanation, in a way which bears on Longuenesse's alignment of Kant with Freud.

II

One component of psychoanalytic theory which is both fundamental to it and, as it seems to me, not likely to be well accounted for by locating Freud's metapsychology in a line of descent solely from developments in nineteenth-century scientific psychology, is the opposition which lies at its core between consciousness and the unconscious. This distinction is not simply equivalent to, even if it

at some point involves, a distinction between mental contents characterized in terms of different degrees of accessibility to introspection or conscious reflection, or of different (e.g. personal *vs.* sub-personal) modes of processing (different types of) information. Rather it invokes as its starting point the ordinary idea of the person as a self-conscious self-determining subject, to which it then counterposes the notion of an enduring mental structure, capable of changing its contents over time, causally disposed to maintain itself in existence, systemically distinct from the network of mental states manifest in the stream of consciousness, and governed by a heterogeneous set of laws: a structure which is nonetheless posited *within* the orbit of the subject and held to stand in an intimate, immediately determining relation to conscious and self-conscious mental life.

This idea is, as Freud appreciated, highly puzzling: common sense psychology accommodates mental states which are in various ways obscure, subliminal, latent, implicit and so on, but nothing in it prepares for the notion that, as it has been put, I am (also) my Other. Freud of course spells out several justifications for the concept of the unconscious, but these are best regarded as essentially warrants for employing an implicitly pre-given concept, not *ab initio* explications of a novel theoretical construction. So we have the question: How, through what sequence of reflections, can it come to be thought that persons as ordinarily conceived harbour counter-selves, abiding configurations organized independently from their self-conscious subjectivity, their relation to which naturally draws comparison with a relation to another person, yet with which they are required to identify themselves?¹

¹ For a clear statement of the point, see Tugendhat 1986, pp. 131–2. The puzzle here may be regarded as a heightened, theoretically elaborated form of the perplexity Saint Augustine reports in contemplating the vastness of memory: ‘Although it is part of my nature, I cannot understand all that I am. This means, then, that the mind is too narrow to contain itself entirely. But where is that part of it which it does not itself contain? Is it somewhere outside itself and not within it? How, then, can it be part of it, if it is not contained in it? / I am lost in wonder when I consider this problem. It bewilders me.’ (2002, p. 216; Book X, 8)

At least the rudiments of this idea can be arrived at through epistemological reflection of the type pursued by Kant's immediate successors.² The usual suspects cited when nineteenth-century antecedents of psychoanalysis are sought include Schelling, Romantic *Naturphilosophen* such as Gotthilf Heinrich von Schubert, Ignaz Paul Vital Troxler and Gustav Carus, Eduard von Hartmann, and, famously, Schopenhauer and Nietzsche. In all of these thinkers an identification is made of the unconscious portion of the subject with Nature and the forces holding sway in it, which is also of course a foundational tenet of psychoanalysis, but in order for that identification to make sense, the concept of the former had to be available initially; and in order to locate that original notion, we need to look further back, to the first generation of thinkers occupied with the post-Kantian transcendental project.

The ground for the development of a transcendental unconscious was, in effect, prepared by Kant himself: in the first place through his notion (as the first *Critique* allows itself to be and has been read, whether or not correctly) that mental contents are not conscious *per se*, in and of themselves, but need to be *made* conscious, and that cognitive awareness of both self and objects involves conceptually motivated operations, comprising synthetic acts not given in inner sense; and second through his thesis of the necessary limit of self-knowledge which is due, not to the impossibility of theoretically cognizing noumena, but to the consideration that any attempt to cognize the self *qua* subject of thought 'can only revolve in a perpetual circle' (A346/B404), it being 'very evident that I cannot know as an object that which I must presuppose in order to know any object' (A402).

² In what follows I am heavily indebted to Marquard's ground-breaking study, 1987, though my reconstruction does not follow his account in all details. Other illuminating studies of the field include Assoun 1976, Redding 1999, and Völmicke 2005. Detailed discussion of major figures in the development of the concept of the unconscious can be found in Nicholls and Liebscher (eds) 2010.

The notion that self-consciousness contains an aporia—that the ‘I or he or it (the thing) which thinks’ must represent itself as and only as ‘= X’, a something ‘known only through the thoughts which are its predicates’ (A346/B404)—undergoes a radical modification in Fichte’s *Wissenschaftslehre*. The complex theory of mental acts which Fichte provides in its Theoretical Part shares with Kant’s theory of synthesis the aim of establishing the preconditions of empirical consciousness, but has a very different character. Kant’s agencies and acts of synthesis, though non-conscious, can still be said to manifest themselves directly in conscious mental life, in so far as Kantian syntheses are necessarily discernible in—they can be read off from—the conceptually shaped sensible given. By contrast, Fichte’s multi-layered account of the absolute I’s positing and counter-positing, in which elements are cross-related in oppositional structures, identified without reference to recognizable types of mental representation, and form genetic series in which successive forms displace their predecessors, are set at a vastly greater distance from the conscious surface of mental life. Fichte’s speculative structures do not merely shadow what is present in natural consciousness but seek to bring to light its complete sufficient ground, hence their greater depth and remoteness.

What in particular screens Fichtean transcendental grounds from ordinary consciousness is the operation of the faculty which he calls productive imagination. The transition to empirical consciousness involves on his account an abrupt discontinuity, which has no conceptual representation and so must be effected by imagination, which converts purely intelligible structure into the kind of (ultimately, spatio-temporal) mental content that an empirical subject can recognize as its own.³ Experience depends upon an interplay and ‘clash’ of opposed components, but this is possible, Fichte argues, only if the ‘boundary’ between them is positively represented, which is something that no purely intellectual function can do: the role of productive imagination is accordingly to give phenomenologically concrete, intuitable form to a structure which pure thought can grasp only as a relation of irreconcilable mutual exclusion. Fichte draws an analogy with the

³ See Fichte 1982, pp. 187–8, 193–4, 200–1.

way in which instants of light and darkness can be given in experience as alternating only if the boundary between them is extended into a temporal instant (Fichte 1982, pp. 187–208): in similar fashion, the I can grasp itself as distinct from its objects, and figure as an object for itself, only by relating to itself through the medium of imagination.

The upshot of Fichte's account, as regards the theme that we are tracking, is complex. In one sense, the *Wissenschaftslehre* eliminates the aporia that Kant locates in self-consciousness: though the I *qua* subject cannot be given objectively within empirical self-consciousness, it is not unknowable, since philosophical insight into the self-positing I provides knowledge of the thing that for Kant remains a mere 'X'. In another regard, however, Fichte reaffirms and underlines the aporia: the genesis of empirical self-consciousness from absolute I-hood covers its tracks, and the standpoints of life and philosophy are distinct and exclusive, the way things appear from the one being an inversion of how they appear from the other.⁴

Although, for Fichte, the transcendental grounds of ordinary mental life are invisible to the pre-philosophical subject, the relation between them is open to full philosophical comprehension. Such epistemological optimism is open to challenge, however. Two figures are of particular importance in the present context. Kant's early critic Salomon Maimon argued that Kant's theory of empirical knowledge fails to close the gap between the a priori conditions which issue from his transcendental proofs, and the a posteriori contents of empirical judgement, to the advantage of Hume. In order to address skepticism, it is necessary, Maimon argued, to embrace a more rationalistic form of idealism than Kant's. Maimon proposes accordingly a Leibnizian recasting of the Kantian divisions between spontaneity and receptivity, and apriority and aposteriority, in which what he calls 'infinitesimals of sensation', akin to *petites perceptions* and belonging to the subject *qua* passive, are postulated as the elements of synthesis, which proceeds without consciousness and in accordance with rules not given to our understanding. This gives rise to the *appearance* of an a posteriori sensible given. In a cognitively perfected consciousness, Maimon supposes, Kant's

⁴ See Fichte 1982, pp. 207–8.

divisions would be overcome, and such a subject would grasp itself and its power of cognition as parts of an infinite mind or reason. On Maimon's account, the ground of sensation is located therefore within the subject, and derives ultimately from the deficiency of self-understanding which is constitutive of ordinary consciousness.

The same broad type of development—this time in reaction to Fichte—is found in Novalis, who provides a clear exposition of the shared view of Romantic post-Kantians that the ground of consciousness is necessarily inaccessible to discursive reflection and so must be grasped by other, non-discursive means. Novalis rejects Fichte's claim that the concept of positing can lead us to the absolute ground of self-consciousness and empirical cognition, on the basis that the structures of opposition and non-identity which positing involves are themselves products of and valid only within our limited cognitive structure.⁵ The consequent problem of elucidating being prior to our representation of it in predicative and identity statements, is then provided with an affective-aesthetic solution: if being *qua* ground of consciousness must be considered in abstraction from any objectual character, then its proximate manifestation within us must be similarly non-objectual, i.e., it must comprise a mode of consciousness which is not intentionally directed, which is as much as to say that it must have the character of feeling. Feeling must however stand in some relation to our judgementally articulated consciousness, in order for it to function as a transcendental ground. This relation, Novalis argues, is one of content to form, but it follows what Novalis calls the principle of '*ordo inversus*': reflection reverses the true relations obtaining within the subject, on the analogy with a mirror image, so that when reflection takes up feeling, being's affective self-manifestation is lost from view, and the status of being *something* is attributed instead to what has been formed conceptually.

This provides a rough indication of how the post-Kantian model of the subject can evolve, under the momentum of transcendental reflection, in a psychoanalytic direction. The transcendental project, as Dieter Henrich has argued (2003, Ch. 2), aims to construct a theory of the subject in

⁵ See Novalis 2003, pp. 167–8.

which the subject can recognize itself: it specifies the conception under which the mind operates and which is deployed implicitly in making those operations possible. This subjective fit provides one measure of validation for the theory. The transcendental image of the mind should thus be or correspond to—in a phrase of Richard Wollheim's (1972)—the mind's own image of itself. This conception of the end and method of transcendental enquiry does not however, as we have seen, bind it to representing the subject as fully self-transparent: for a variety of reasons, gaps may be affirmed in the transcendental theory of the mind, and where these exist, they will imply regions of opacity in the subject's apprehension of its own grounds.

The next historical and logical step consists, as indicated earlier, in the insertion of Nature into the place opened up by post-Kantian transcendental reflection. The conception of the empirical subject as constituted by an aporetic self-relation—the mind as a dynamically self-concealing structure, which can open up to itself and the world only on the condition that the facilitating ground of its doing so is excluded from consciousness—is annexed to the philosophy of nature by Schelling, who first formulates explicitly the concept of the Unconscious, identifying it with the infinite activity of *natura naturans*. Schelling's unconscious has no specific role in psychological explanation, but Schubert, Carus and others in his wake ascribe an unconscious to the human subject in an individualized form: the development of human personality is regarded as the effect of a subsisting unconscious core, which expresses in microcosmic form the principles, forces, and order of Nature at large. The final chapters in the story of progressive approximation to the psychoanalytic unconscious are familiar: Schopenhauer and then Nietzsche advance speculative hypotheses concerning the foundations of human motivation which repudiate the original transcendentalist association of the unconscious with rationality.

III

In Part Three of the *Ethics*, ‘Concerning the Origin and Nature of the Emotions’, Spinoza devotes many pages to a systematic catalogue of the psychological laws governing the affects. On the basis of three ‘primary’ emotions of pleasure, pain, and desire, and their immediate derivatives, love and hatred (IIP11 and IIP13S; 2002, pp. 285, 286), Spinoza defines in Propositions 13–59 (2002, pp. 286–311) the complex kinds of emotion which may be formed and the conditions under which they increase or decrease in intensity, and yield to, or are transformed into, one another.

Spinoza’s purpose in giving this account is, at a minimum, to bring home to us the ways in which, and the degree to which, our passions determine our action dispositions independently of anything that could be called reflective endorsement. The account thereby serves the end of increasing our freedom: correct insight into the causation of our passions—above all: appreciation that they cannot be expected to respond to reason in the way that we naively suppose—can spur us to adopt the less direct but truly effective measures for their regulation described by Spinoza in the opening pages of Part Five.

As commentators have observed, there is much in the *Ethics* that can be regarded as proto-psychoanalytic. Spinoza envisages explanation of mental states and changes in a way that parallels and tracks bodily conditions and changes. Spinoza’s strict psychological determinism echoes Freud’s equally categorical statements on the topic. The picture of the mind as moving from one condition to another according to laws of efficient causality—registered by changes of hedonic state but not explained by invoking pleasure as an end—recalls the model that Freud draws up in his pre-psychoanalytic *Project for a Scientific Psychology* and which survives recognizably in the ‘economic’ component of the psychoanalytic metapsychology. The concept of imagination, given very broad scope in the *Ethics* and introduced primarily in order to account for the limitations of human cognition, is employed by Spinoza in the context of the affects in ways that show it doing duty for the panoply of concepts (pleasure as opposed to reality principle, primary as opposed to secondary process, phantasy, screen memory, etc.) employed by Freud to explain the mind’s capacity for operating with indifference to reality. Spinoza, like Freud, regards wish-fulfilment as

having its basis in the fundamental structure of the mind.⁶ Spinoza's assertion of the affective impotence of cognitions, their inability to check emotions directly by virtue of their truth alone,⁷ is mirrored in Freud's account of the insufficiency of intellectual assent to interpretations and of the necessity of working-through. Spinoza anticipates Freud's devaluation of the property of consciousness, if not in fact to have taken it a good deal further, in so far as consciousness barely figures in the *Ethics*, which nowhere treats it as a topic in its own right. Conatus corresponds to Freud's conception of libido. And so on.

What I want to focus on, with a view to getting a unified perspective on these doctrinal points of contact of Spinoza with Freud, is the *nature* of Spinoza's claims in IIP13–59.

Some of the propositions advanced in IIP13–59 are indisputably true, perhaps even, as we would put it, conceptual truths, while others are far from obvious and require decoding. Spinoza gives no explicit account of their epistemological basis, and their status is not immediately clear. In so far as they deal with observable patterns in the phenomena of psychological life, it would be natural to suppose that those which are not candidates for conceptual truth must be a posteriori, but it is barely conceivable in light of the method and ambition of the work as a whole, and the specific role of Part Three, that Spinoza should intend a mere inventory of psychological tendencies drawn from common observation, amounting to the inferior kind of knowledge that he calls 'casual experience' (IIP40S2; 2002, p. 267). Spinoza's aim must be, rather, to nail down the essential 'logic' of affect—in the context of the *Ethics*, nothing less would do.

This raises the expectation of being able to regard Spinoza's propositions as a priori, and Spinoza does indeed spell out, in the proofs of the propositions, deductive connections with the foundational metaphysical claims about human beings that he has made in Part Two and the earlier portion of Part Three. In these proofs Spinoza's constant reference point is his thesis of the mind's

⁶ See IIP13 (2002, p. 286) and IIP28 (2002, p. 293).

⁷ See IVP1 (2002, pp. 323–4), IVP7 (2002, pp. 325–6), and IVP14 (pp. 328–9).

necessary endeavour 'to think of those things that increase or assist the body's power of activity' or (what is the same) 'the power of thought of our mind' (IIP11–13; 2002, pp. 284–6).

What deserves attention is the nature of the necessity whereby conatus expresses itself in the causation of affect. One type of necessity that we can make good sense of in the analysis of psychological explanation is the normative necessity associated with rational choice theory and the ordinary practical syllogism,⁸ and some of Spinoza's propositions may look as if they can be expressed in such familiar terms: that he 'who hates someone will endeavour to injure him unless he fears that he will suffer a greater injury in return' (IIP39; 2002, p. 298), for instance, might be considered a prudential norm. But this cannot be what Spinoza has in mind, since such rational necessity is manifestly absent from his claims, for example, that we shall love or hate an object which we imagine to be similar to an object that is wont to affect us with pleasure or pain (IIP16; 2002, pp. 287); that we will share the emotion of whatever we imagine to be like ourselves (IIP27; 2002, p. 292); that we love, desire, or hate a thing the more steadfastly, when we think it loved, desired, or hated by somebody else (IIP31; 2002, p. 294); that we will seek to deprive someone of anything they enjoy that only one person can possess (IIP32; 2002, p. 295); that our love for someone will turn to hatred if they form a more intimate bond of friendship (IIP35; 2002, p. 296); that we will desire to possess a thing that gave us pleasure in the same circumstances as when we first took pleasure in it (IIP36; 2002, p. 297); that we will hate someone whom we imagine to hate us without cause (IIP40; 2002, p. 299) or someone imagined to be similar to ourselves who hates something that we love and imagine to be similar to ourselves (IIP45; 2002, p. 301); that someone's hatred of an object that he has previously loved will be proportionate to the strength of his former love, and vice versa (IIP38 and IIP44; 2002, pp. 298 and 301); that we will have higher regard for an object that we have not seen in conjunction with other objects or that has nothing in

⁸ Thus Davidson 2005 seems to identify causal explanation within the attribute of Thought with rationalization. My construal accords with the interpretation of Spinoza in Carriero 2005; see esp. pp. 136–41.

common with them (IIP52; 2002, p. 304). The vectors and tendencies of emotional life to which Spinoza is pointing are familiar and, as we would put it, thoroughly *natural*, but if we choose to honour them on that count with some sort of claim to some kind of rationality, it is simply because of their familiarity, prevalence and embeddedness in human nature, not because we suppose any articulable principle of reasoning or judgement of ‘rightness’ to underlie them.⁹

On a weak reading, of the sort that I offered above as a minimal characterization of his intentions, Spinoza would be taken to have shown only how the affects comport themselves when left to their own devices, information which we could regard as giving helpful forewarning of the ways in which our rational self-determination is prone to being subverted. This, however, would be to picture the affects as situated on the outside, and ourselves—as centres of rational self-determination—on the inside, in the situation of a swimmer calculating the forces of tide and current. A compelling reason for thinking that this view of our situation cannot be Spinoza’s is that it corresponds to that of Descartes, as Spinoza construes—and explicitly rejects—it; it is to ‘conceive man in Nature as a kingdom within a kingdom’ (III, Preface; 2002, p. 277).¹⁰ That a much stronger reading is in order becomes clear when Spinoza says that ‘mental decisions are nothing more than the appetites themselves’, ‘that mental decision on the one hand, and the appetite and physical state of the body on the other hand, are ... one and the same thing’ (IIP2S; 2002, p. 281);¹¹ a claim which coheres with his declaration of intent to ‘consider human actions and appetites just as if it were an investigation into lines, planes, or bodies’ (III, Preface; 2002, p. 278). Spinoza envisages, therefore, the affects and their causation as occupying the full space of

⁹ See IIP9S (2002; p. 284).

¹⁰ Key points in Spinoza’s extended argument against this conception are IIP48–9 (2002, pp. 272–7), IVP1–7 (2002, 323–6), and Part Five, Preface (2002, 363–5).

¹¹ A mental decision, Spinoza continues, ‘is not distinct from imagination and memory, and is nothing but the affirmation which an idea, insofar as it is an idea, necessarily involves’ (IIP2S; 2002, p. 282).

psychological explanation: they are not what we *qua* rational beings are set *over and against*, rather they are what we *are*, whence their *metaphysical* necessity;¹² the question of why we find ourselves subject to immediate affective determination receives its answer alongside the question of why we conceive physical bodies as divisible and numerable. Spinoza's logic of affect is designed to show that psychological life proceeds in ways that are not driven by any power of practical reason: explanations with the form of practical syllogisms do not express falsehoods—human action is not 'blind' in that sense—but they do not display the causal reality of the mind, since what makes them true, in so far as they are true, is not a grasping of normative relations. Recognition of the true, non-normative character of psychological reality is the first step in escaping from human bondage: only when we have grasped that our 'ought' thoughts are nothing but further expressions of conatus, without *sui generis* reality, are we in a position to steer ourselves in the appropriate therapeutic direction.¹³ There is therefore a good sense in which the ordinary concept of a reason for action does not survive the *Ethics*.¹⁴ Spinoza's interchangeable use of 'reason' and 'cause' gives notice of this.¹⁵

What should accordingly strike us when we consider IIIIP13–59 with Freud in mind is the remarkable parallel of Spinoza's explanations by and of emotion with psychoanalytic explanations,

¹² Which is why achieving freedom is associated with the metaphysical modification of our identity—identification with the immortal part of the mind—described in Part Five.

¹³ Whence Nietzsche's praise of Spinoza for restoring the world's 'innocence' (1994, p. 50; Essay II, §15). Spinoza describes himself as seeking to 'understand' the emotions and actions of men rather than 'abuse or deride' them (III, Preface; 2002, p. 277).

¹⁴ This is of course very far from a complete analysis, which leaves aside the special case where reason has become determining; my intention is simply to highlight points of relevance to psychoanalysis. On Spinoza's complex theory of agency, see Koistinen 2009.

¹⁵ See IP11, Second Proof (2002, p. 222). Also relevant is the famous comparison of a person's consciousness of their agency with the thinking of a stone in motion (Letter 58; 2002, pp. 908–10).

which similarly appear—from the standpoint of common sense psychology—to be poised awkwardly on the border separating reasons from mere causes. Just as it would make no sense to cite the inherent tendency of love to convert itself into hatred in the circumstances detailed in IIP35 as a *reason* for regarding a person with hatred, so similarly unconscious determinants as such cannot be incorporated into our self-reflection as justifications for our actions or propositional attitudes. And yet, nor are these forms of causation correctly described as *merely mechanical*: the relations between cause and effect are *internal*, and the effects are *intelligible* in light of their postulated causes.¹⁶ There is no space here to go into detail, but I think it would not be hard to show, with reference to Freud's own texts and to typical psychoanalytic case histories, that the explanatory base-line in psychoanalytic explanation consists in the interaction of elements in psychic reality in ways which fuse (in a way that at once illuminates and puzzles) activity of thought with forms of activity conceptualized in terms of movements of organic and even inorganic bodies.

The shared nature of Spinozistic and psychoanalytic explanation—their fusion of thought and mechanism into a single form of efficient causality—is reflected in the points of doctrinal convergence listed above. Other revisionary claims concerning the human subject defended in the *Ethics* with conspicuous psychoanalytic resonance include Spinoza's reductive conception of the self, his rejection of free will, his use of the concept of expression as a linchpin of psychological explanation, his overhauling of the common sense taxonomy of mental states, and his argument that the very fact of failures of reason suffices to falsify the kingdom-within-a-kingdom conception of the subject.¹⁷

¹⁶ Of interest here is the sketch of a psychoanalytic logic of affect in Alexander 1935, which attempts to coordinate talk of 'emotional syllogisms' with the modelling of mental processes on bodily functions.

¹⁷ 'But my argument is this: in Nature nothing happens which can be attributed to its defectiveness, for Nature is always the same, and its force and power of acting is everywhere one and the same;

IV

My answer to the question posed earlier—whence the psychoanalytic concept of the unconscious?—is therefore that Spinoza’s logic of affect can be thought to mesh with the idealist conception of self-concealing subjectivity, yielding a model of the mind which approximates to that of psychoanalysis: in rough outline, the psychoanalytic conception of the unconscious is the product of substituting for the synthetic acts of transcendental subjectivity, postulated by the idealist/transcendentalist with a view to solving problems of epistemology and metaphysics, the sorts of contents and operations which Spinoza postulates with a view to explaining human action and affect. Freud has accordingly a foot in two distinct camps: the humanistic tradition of classical German philosophy, in which cognitive and ethical tasks are conceived and pursued from the internal standpoint of the subject, and the alternative naturalistic programme of modern philosophy.

This seems to me a fair representation of the intellectual forces which come together in Freud’s notion of the unconscious, and it accords well with the salient facts of psychoanalysis’ historical reception—to wit, the huge appeal which psychoanalytic ideas have had across the board in the humanities for over a century, alongside the determined attempt, to the fore in the North American psychoanalytic tradition, to hold psychoanalysis fast (for better or worse) to the agenda of the natural sciences.¹⁸ It is difficult to see that any more economical hypothesis could account for the extraordinary depth of disagreement separating different schools of psychoanalytic thought,

that is, the laws and rules of Nature according to which all things happen and change from one form to another are everywhere and always the same’ (III, Preface; 2002, p. 278). Also relevant are IIP2S (2002, pp. 281–2), on the cognitive illusion of the distinctness of mental decision, and the deflationary remarks on perfection in the Preface to Part Four (2002, pp. 102–3).

¹⁸ Kitcher 1992 provides a sophisticated version of this approach.

ranging as it does from deconstructionist and neo-Nietzschean appropriations of Freud to the interpretation of psychoanalysis as proto-neurophysiology.

Representing psychoanalysis as an attempt to unite in a single concept two major and distinct trajectories in modern philosophy throws light on the theory and accounts for the double character of Freud's notion of the unconscious, but by the same token makes clear the difficulty which it faces. On the one hand, the legacy of idealism showing itself, Freud asks that we *avow* the unconscious, that we regard it as *our own*, while on the other he insists, in line with Spinoza, that the real forms of motivational causation are incongruent with the image of action as proceeding from a centre of self-conscious self-determination. This combination of claims is what forces out the characterization of the unconscious as the 'Other' of consciousness. The extended debate in the philosophy of psychoanalysis concerning whether psychoanalysis is a form of causal explanation or of reason explanation, reflects this duality.¹⁹ So too does the puzzle, which has also given rise to an inconclusive literature, concerning the coherence of supposing that *bona fide* mental states, of the sort attributed in psychoanalytic explanation, can subsist without consciousness.

The problem can be put in sharper focus if we contrast the relation to common sense psychology required for psychoanalytic explanation with the situation of Spinoza. As my earlier discussion emphasized, Spinoza's intentions are uncompromisingly *revisionary*. If Spinoza is not an eliminativist regarding common sense psychology, he comes within a hair's breadth of being one. Psychoanalytic theory cannot, however, mean to follow Spinoza's hard line, because its own distinctive forms of explanation are interwoven with those of common sense psychology, and its own concept of the unconscious is formulated with reference to and as the counterpart of the ordinary conception of the subject.

¹⁹ As Ricoeur puts it, the 'aporia' of psychoanalysis, 'the whole problem of the Freudian epistemology', may be 'centralized in a single question' concerning the integration of 'interpretation that necessarily moves among meanings' with explanation by forces, and 'Freudianism exists only on the basis of its refusal of that disjunction' (1977, pp. 66–7).

The consequences of this conceptual compromise—psychoanalysis' adoption of Spinozistic explanation, without acceptance of its revisionary implications—are reflected in Sartre's critique of Freud, the nub of which is contained in the following passage:

By rejecting the conscious unity of the psyche, Freud is obliged to imply everywhere a magic unity linking distant phenomena across obstacles, just as sympathetic magic unites the spellbound person and the wax image fashioned in his like. The unconscious drive (*Trieb*) through magic is endowed with the character 'repressed' or 'condemned', which completely pervades it, colours it, and magically provokes its symbolism. Similarly the conscious phenomenon is entirely coloured by its symbolic meaning although it can not apprehend this meaning by itself in clear consciousness. (Sartre 1958, pp. 53–4)

This claim is explained more fully in Sartre's essay on the emotions:

The psychoanalytic interpretation [of an emotion] conceives the conscious phenomenon as the symbolic realization of a desire repressed by the censor. Note that, for consciousness, the desire is *not involved in its symbolic realisation*. In so far as it exists by and in our consciousness it is only what it gives itself out to be: emotion, desire for sleep, theft, laurel-phobia, etc. If it were otherwise, if we had any consciousness, even *only implicit*, of the real desire, we should be in bad faith, and that is not what the psychoanalyst means. (Sartre 2004, p. 30)

So the relation between an emotion and the signification conferred on it under a psychoanalytic interpretation must, to some degree, resemble the relation of the ashes of a fire extinct upon a mountain to those who lit it. And yet, Sartre points out, it cannot be *merely* external. For the psychoanalyst 'there is always an internal analogy between the conscious fact and the desire it

expresses, since the *conscious fact is symbolical of the expressed complex*: the symbolic character of the conscious fact 'is *constitutive* of it'. This, however, leads us back in a circle, to the identification of unconscious motivation with mere bad faith: 'if symbolization is constitutive it is legitimate to see an immanent bond of *comprehension* between the symbolization and the symbol'—'the relation between symbol, symbolized and symbolization is an intra-structural bond of consciousness' (Sartre 2004, p. 32).

The problem that Sartre describes, expressed in terms of my historical reconstruction, is therefore that psychoanalytic explanation invokes the conception of the unity of the mind and of the nature of relations between mental elements contained in our common sense self-conception, elaborated in the Kantian tradition, *and also* models the mind in terms of Spinoza's patterns of causation, from which the same degree and type of reflexive unity is absent. The psychoanalytic concept of symbolization is a bridge, not just from unconscious to conscious facts, but also from the one philosophical model of the subject to the other, and the strain which it is under reflects the way in which psychoanalytic theory merely *superimposes*—grafts onto one another, without integrating—the two different models of the mind.²⁰

There are, to be sure, ways of softening the impact of Sartre's criticism. We may, for instance, postulate some species of mental connection mid-way between immanent bonds of comprehension, and merely external causal linkages of the fire & ashes sort. And it is clear furthermore that in doing so we may claim a degree of support from common sense psychology (which does not, after all, assimilate all 'bonds of comprehension' to the clear-cut case of knowing the meaning of a sign, as Sartre perhaps tends to imply). Moves of this sort can be given plausibility through being inter-related with one another in a systematic way. Psychoanalysis can thus be presented as an extension of common sense psychology. What I think nonetheless needs to be

²⁰ Sartre does not, however, reject symbolic interpretation as such, and towards the end of *Being and Nothingness* he proposes what he considers a superior conceptualization: see Sartre 1958, pp. 602–15.

recognized as the substantial truth in Sartre's criticism—and which in no way depends upon his professed Cartesianism—is his insight that, to the extent that we extend common sense psychology by proposing a new species of mental state or relation, even though each individual theoretical move in the construction of the extension may have an adequate local justification, there is a sense in which the cumulative effect of the incremental changes is a *philosophical* loss of bearings; in historical shorthand, we wonder how, starting with Kant, we could have ended up with Spinoza. This is quite compatible with the resulting theory's working perfectly well as an explanatory system in accordance with criteria appropriate to empirical theories; what it means is only that we cannot render the theory philosophically perspicuous.²¹

One way in which the problem might be resolved decisively is through an expansion of the boundaries of the self, in a way that permits reconception of the unconscious in terms that make full metaphysical sense of the identification with it demanded by psychoanalysis. If the Nature that feeds into my unconscious is in reality more of myself, if it is a whole with which I am in some sense identical, then it is false that, as Sartre puts it, 'I *am* ego but I am not *id*' (1958, p. 50): rather, the unconscious is a *continuation* of me. This view of the metaphysics of human personality is, in

²¹ Deconstructionist and neo-Nietzschean readings of Freud and applications of psychoanalytic ideas, referred to earlier, construe psychoanalytic theory as engaged in (negative) metaphysical assertion. This removes the problem under discussion, since if at the heart of Freud's theory of the mind *stands the claim* that the concept of the person or subject is incoherent and/or that the reality in which persons and subjects consist does not permit coherent conceptualization, then the theory is freed from the tension I have been exploring and the inconsistency alleged by Sartre: the tension and inconsistency are relocated in the concept which is the theory's object. What this approach may be right about is that psychoanalysis counts among its data the difficulty which subjects encounter in conceiving themselves in the terms that psychoanalysis recommends, in other words, in relating themselves to their unconscious states, but it is not plausible to construe Freudian theory, as a whole, as a consistently skeptical meta-theory.

fact, the picture proposed (for epistemological and metaphysical reasons) by Schelling. It amounts of course to a revision of the common sense view of the borders of the self, but it conserves, unlike Spinoza's metaphysics, the notion of self-conscious self-determination. However, it also, plainly, requires attributing to Nature properties that disenchanted modern natural science denies it, and though a hint of a preparedness to retrieve nineteenth-century Romantic idealism can be found in a very few psychoanalytic thinkers²²—and arguably receives encouragement from Freud's late remarks on Eros and Thanatos—it is in flat contradiction with Freud's official account of the psychoanalytic *Weltanschauung* and not likely to be regarded as an acceptable condition for securing the conceptual integrity of psychoanalysis.

Where does all this leave psychoanalysis, as regards its systematic evaluation? The larger issues which arise concerning the long-running debate concerning psychoanalysis' claim to truth obviously cannot be pursued here, but I have a limited suggestion, arising from the preceding discussion.

One reaction to my account—which will recommend itself to those unimpressed by psychoanalysis' claim to psychological insight—is to take the diagnosis implied by the historical reconstruction as the basis for a quick conviction. The empirical defensibility of psychoanalytic claims is a matter for another occasion, but if the roots of psychoanalysis' conceptual difficulties lie as deep as I have suggested, namely in a collision of transcendentalist with naturalist commitments, then it may reasonably be held that the reemergence of this overarching philosophical antinomy in the psychoanalytic context—albeit in an especially overt, acute, aggravated form—cannot be a reason for impugning the theory's claim to offer genuine psychological explanation. Psychoanalysis ought not to be made a casualty of our inability to resolve what is at the end of the day a wholly

²² Loewald concludes, after consideration of the 'reconciliation of the subject-object dichotomy' involved in sublimation (1988, p. 20), that psychoanalysis refers us to a *natura naturans* characterized by a subjectivity vaster than 'human individual mentation' (1988, pp. 79–80). Also of high interest is the discussion in Lear 1990, Ch. 5.

general philosophical problem. Acknowledging the ambiguity concerning the personal vs.

Spinozistic status of the unconscious, and accepting that the theory does not explicate itself in a fully coherent manner, is an acceptable price to pay for a picture of human beings that makes sense of stretches of human experience which will otherwise be left completely opaque.

In so far as Longuenesse and I differ, our difference lies in our respective attitudes to the proposed unification of Kant's 'I ought' and Freud's superego. This Longuenesse regards as an unproblematic theoretical unification, which promises to help out Kant's ethical theory, by transposing a questionable transcendentalist conception into a secure naturalistic context. In my view, waiving all questions about the capacity of Kant's conception of morality to survive the move intact, the proposal draws attention to an underlying bifurcation in psychoanalytic thought, which I have modelled in terms of its derivation from incongruent historical sources. One of these, the Spinozistic, is antithetical to Kantian ways of thinking. This alone does not exclude Longuenesse's proposal, since, I have argued, psychoanalysis is not consistently Spinozistic. It may be true, therefore, that there are aspects of Freud's metapsychology which accord with Kant's concepts. The problem, however, is that psychoanalysis is not consistently Kantian, either, and that its ambiguity cannot be resolved in either the one direction or the other. This should not, I have urged, be made an objection to psychoanalysis. But if correct, it means that psychoanalysis does not offer a philosophically safe home for Kant's 'I ought' to the extent that Longuenesse supposes.

Department of Philosophy

UCL

Gower Street

London WC1E 6BT

sebastian.gardner@ucl.ac.uk

REFERENCES

Alexander, Franz 1935: 'The Logic of Emotions and its Dynamic Background'. *International Journal of Psycho-Analysis*, 16, pp. 399–413.

Assoun, Paul-Laurent 1976: *Freud: la philosophie et les philosophes*. Paris: Presses Universitaires de France.

Saint Augustine 2002: *Confessions*. Translated by R. S. Pine-Coffin. Harmondsworth: Penguin.

Carriero, John 2005: 'Spinoza on Final Causality'. In *Oxford Studies in Early Modern Philosophy*, Vol. II, pp. 105–48. Edited by Daniel Garber and Steven Nadler. New York: Clarendon Press.

Davidson, Donald 2005: 'Spinoza's Causal Theory of the Affects'. In *Truth, Language, and History: Philosophical Essays, Volume 5*, pp. 295–314. Oxford: Clarendon.

Fichte, Johann Gottlieb 1982: *The Science of Knowledge, with the First and Second Introductions*. Edited and translated by Peter Heath and John Lachs. Cambridge: Cambridge University Press.

Henrich, Dieter 2003: *Between Kant and Hegel: Lectures on German Idealism*. Edited by David S. Pacini. Cambridge, Mass.: Harvard University Press.

Kant, Immanuel 1999: *Critique of Pure Reason*. Translated and edited by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

Kitcher, Patricia 1992: *Freud's Dream: A Complete Interdisciplinary Science of Mind*. Cambridge, Mass.: MIT Press.

Koistinen, Olli 2009: 'Spinoza on Action'. In Olli Koistinen (ed.), *The Cambridge Companion to Spinoza's Ethics*, pp. 167–87. Cambridge: Cambridge University Press.

Lear, Jonathan 1990: *Love and its Place in Nature: A Philosophical Reconstruction of Psychoanalysis*. London: Faber.

Loewald, Hans 1988: *Sublimation: Inquiries into Theoretical Psychoanalysis*. New Haven: Yale University Press.

Marquard, Odo 1987: *Transzendentaler Idealismus, romantische Naturphilosophie, Psychoanalyse*. Köln: Dinter.

Nicholls, Angus, and Martin Liebscher (eds) 2010: *Thinking the Unconscious: Nineteenth-Century German Thought*. Cambridge: Cambridge University Press.

Nietzsche, Friedrich 1994: *On the Genealogy of Morals*. Translated by Carol Diethe, edited by Keith Ansell-Pearson. Cambridge: Cambridge University Press.

Novalis 2003: *Fichte Studies*. Translated and edited by Jane Kneller. Cambridge: Cambridge University Press.

Redding, Paul 1999: *The Logic of Affect*. Ithaca: Cornell University Press.

Ricœur, Paul 1977: *Freud and Philosophy: An Essay on Interpretation*. Translated by Denis Savage. New Haven: Yale University Press.

Sartre, Jean-Paul 1958: *Being and Nothingness: An Essay on Phenomenological Ontology*. Translated by Hazel E. Barnes. London: Methuen.

— 2004: *Sketch for a Theory of the Emotions*. Translated by Philip Mairet. London: Routledge.

Spinoza, Baruch 2002: *Complete Works*. Translated by Samuel Shirley, edited by Michael L. Morgan. Indianapolis: Hackett.

Tugendhat, Ernst 1986: *Self-Consciousness and Self-Determination*. Translated by Paul Stern. Cambridge, Mass.: MIT Press.

Völmicke, Elke 2005: *Das Unbewusste im Deutschen Idealismus*. Würzburg: Königshausen & Neumann.

Wollheim, Richard 1972: 'The Mind and the Mind's Image of Itself'. In *On Art and the Mind: Essays and Lectures*, pp. 31–53. Cambridge, Mass.: Harvard University Press.