

What we (don't) need the Death drive for: Kant and the philosophy of science in Bion's Theory of Thinking.

Abstract

Kleinian theory controversially identifies the Death drive with those forces in the mind which produce a canonical phenomenology of mindlessness. In the explanation of this phenomenology of psychic self-destructiveness Bion's theory of thinking has been a powerful organiser of psychoanalytic observation, but his use of Kant needs to be more clearly understood. I evaluate the Kantian contribution to the theory in the light of Bion's avowedly critical employment of ideas he drew from contemporary philosophy of science. I argue that Bion uses Kant's account of cognition as an 'as if' model with which to organise his own clinical theorising, in an opening-up of new ways of understanding the mind. This allows us to explain a phenomenology of mindlessness as arising when experience of extreme anxiety challenges the mind's capacity for thought, without reference to a pessimistic thesis of mental self-destructiveness. Instead, the appearance of destructiveness is explicable in terms of Kant's account of sublime judgment by a discursive intellect, an established mind always oriented to continuity of thought and, a fortiori, not self-destructive. Re-positioning Bion's theory helps us understand how the mind grasps the experience of profound anxiety and what prevails when it cannot do so.

Keywords

Broad: History of Psychoanalysis; Comparative Psychoanalysis; Epistemology; Philosophy.

Narrow: Psychic Pain; Psychic Structures; Self; Theory of Thinking; Unconscious; Conceptual Research.

Authors: Bion, Wilfred; Freud, Sigmund; Klein, Melanie.

1. Introduction

Freud's initial positing of the Death drive as both a source and a principle of psychical destructiveness continues to present psychoanalysis with a theoretical and clinical conundrum. The clinical picture in question is of a distinctive, though at first sight disparate, group of mental phenomena: psychoanalytic stasis with absence of affect; passivity in thinking; indecision in acting, disruption and attenuation of thought processes; eidetic and imagistic references to catastrophe contrasting with affective blankness or tranquillity; an active pleasurable destructiveness of thought process and analytic progress. These are some of the manifestations of what has been called a deathly state of mind, attributed to the operation of the Death drive. The precise dynamical explanation informing clinical interpretation of these phenomena depends on whether the Death drive is conceived as: a primary destructive force in the mind opposing the force of life with a return to entropy; a drive towards the goal of mindlessness and pleasure in tranquillity with the rejection of, and destructive attack on the demands of the real world for work and suffering; an object-directed affect of envy or hatred directly opposed to love.¹ Conceived as manifestations of an underlying, primary but essentially unknowable originating force, drive, or affect, these alternatives correspond to different theoretical models

¹ These correspond to three different phases in the Freudian canon: 1. the biological, "organismic" mind seeking stability of a "energy economy", originating in the "Project for a Scientific Psychology" (Freud 1895); 2. the "Two Principles of Mental Functioning" mind that opposes the drive to seek pleasure and avoid pain to the drive for survival in an economy of affect (Freud 1911); 3. the person-level psychology of affective relating inaugurated as object relations theory in "Mourning and Melancholia" (Freud 1917).

of the mind – energetic, functional-teleological, and object-relational – whose metapsychologies predominate at different periods in the ongoing evolution of psychoanalysis.

While contemporary Kleinian psychoanalysts largely eschew biological energetics and functional teleology in favour of object-relational psychology in the explanation of the clinical picture, the terminology of received Kleinian discourse is replete with cross-borrowings from these earlier models with the assumption, and expectation, of conceptual equivalence between three very different modes of explanation. The theoretical conundrum posed by the Death drive is partly due to the ensuing conceptual discontinuities confounding explanatory coherence. On the first, “energetic” explanation in terms of disruptive forces in the mind, the Death drive is (already) explained as the threat to affective stability of a destructive force producing either a direct interruption of thought, or a profound anticipatory anxiety such that, on the second, functional explanation in terms of psychic defence against the un-pleasure of anxiety, mindlessness becomes a defensive outcome against disruption by this anxiety. Turning to the third, psychological explanation in terms of affective object relations such as love, hate, curiosity, or indifference, the destruction of thought is framed object-rationally as an attack motivated by envy or hatred, on the capacity for thought itself.

The upshot from the concept’s evolution through the different phases of psychoanalytic metapsychological thinking is that even those psychoanalysts with a theoretical commitment to the idea of a Death drive do not uniformly agree on what it involves. While opting for the third, object-relational, metapsychology would remove much of the confusion, this would require discarding the other two models, something to which institutional psychoanalysis is resistant. However, my critical target in this paper is not this institutional inertia, whether conceived as psychological or ideological. Nor is it the inconsistency of psychoanalytic theses

surrounding the Death drive or the precise places in which assumed conceptual equivalences fail, although the unwarranted assumption of conceptual equivalence as a mistake in reasoning does contribute to the problem I address in this paper: the Kleinian hypothesis of a primary mental destructiveness that can be turned on the mind itself.

This thesis, which is my critical target, is elaborated by Bion as the mind attacking itself. In that form it has itself become institutionalised as psychoanalytic dogma and so risks meeting precisely the sort of fate that, as we shall see, Bion wished to avoid, for his own theories and for psychoanalysis in general. The attribution to the mind of a primary self-destructive drive, to explain why certain of its states present as forms of mindlessness, has been found clinically powerful; however, I shall argue that the theoretical explanation offered is mistaken. I propose a different explanation bringing together two areas of philosophical thought not usually connected with each other or with psychoanalysis: the theory of cognition psychoanalysis inherits from Kant and the idea of a theoretical model in science.

2. Kleinian theory

While Klein herself equated the Death drive with envy as a primary destructiveness directed at the good as object, I have suggested that the considerable variation in interpretations of clinical material attributed to the operation of the Death drive arises from the way that in Klein's account the destructiveness is theoretically overdetermined by the attempt to accommodate all three metapsychologies. The mind is at the mercy of an unthought force let loose within it; this force is both seen as envy causing anxiety through the threat it poses to mental functioning, and as itself identified with anxiety. The mind is also conceived as home to its own object- and self-directed affect of hate. While, then, the Kleinian clinical literature is replete with case

material presented as exemplifying deathly states of mind, opinions differ over the value and indeed the coherence of interpretations that rest on the posit of the Death drive. Non-Kleinian object relations theorists, notably Bollas and Winnicott, argue that fear of annihilation and other profound anxieties are developmentally acquired, having already been experienced in a trauma whose effects on the capacity to think are unavailable to thought. Other thinkers argue that the clinical presentation of feared internal annihilation is instead the return into the mind of an outwardly directed primary aggression. Such thinkers reject the Kleinian view of envy as a fundamental destructive propensity of the mind reflexively directed onto the mind's capacity for thought.

At the same time, and despite being marginal to the wider psychoanalytic theoretical corpus, the Death drive refuses to die. This is not only because it is institutionally entrenched in post-Kleinian thought. The Death drive in Kleinian metapsychology is a theoretical elaboration which is intended as continuous with the Freudian canon; in it the idea of psychic destructiveness grafts readily onto the duality retrieved for psychoanalysis in "Beyond the Pleasure Principle". David Bell writes that there are "substantial empirical grounds" for a psychic force that opposes life and development and a "compelling theoretical framework for ... a duality governing psychic life" (Bell 2014, 12). The clinical and the theoretical power of the Kleinian thesis thus acquire merit from re-provisioning psychoanalysis with its cultural and historical foundations, newly powerful in clinical application.

However, there is reason to think these foundations are insecure. Bell locates the source of an inherent tension within the Freudian concept of the Death drive in the mysterious and unexplained 'transformation' of the repetition of un-pleasure in "Beyond the Pleasure Principle" into a pull towards inertia. In Freud's metapsychological explanation this is

overdetermined as both energetic (the tendency of the “mental apparatus” to divest itself of tension) and functional (the goal of the organism as an entropic “nirvana” state). This state of mind presents phenomenologically as mindlessness but, as Bell observes, Freud also went on to describe an overt, “noisy” destructiveness, aimed at the processes of thought and bringing with it gratification. Referencing Freud’s treatment of negation, Bell provides a more sophisticated analysis of the connection Freud wishes to make, in which the repetition of negation of particular pleasures is replaced, or displaced by the annihilatory negation of psychotic “negativism”; here “*the negation of all desire*” (Bell 2014, 3) yields a different form of pleasure, the absence of any demand for thought and work.²

Bell maps out the clinical phenomenology of this tension in three theoretical models. All three clinical pictures are taken to support an interpretation in terms of a destructive attack wrought by the mind on itself as object. The first model foregrounds active destructiveness, the Death drive manifesting as a clinical picture of overt sadistic pleasure in the mindlessness and the destruction of life and thought. In Kleinian object-relational terms what is encountered clinically is envy, as a primary object-relation present *ab initio* in the constitution of the mind, whose object-directed aim is the annihilation of the thinking self. The third model is overtly sadistic; here the Death drive aims at paralysis of the analysis itself, keeping it alive so as to continue its sadistic treatment *ad infinitum*. Clinically there is stasis, quietude, and a sadistic

² It should be said that as a plausible interpretation of Freud’s reasoning, Bell’s analysis does not thereby make Freud coherent on this point. If Freud is read as talking about determinate negation (the negation propelling dialectical reasoning) we could read him as saying that “within the domain of desires any (one) desire can be negated, so the domain of all desires together can be negated”. This is to re-state the conundrum of self-annihilation by the mind.

object relation of a never completed, ongoing annihilation of the analytic process as its object. The second model is frankly annihilatory; it is object-relational but with the aim of the object's annihilation; clinically, it presents as a seductively mindless, deadening passivity in which the destructiveness is represented as its completion.

3. The Death drive: a conundrum

In attributing a destructive purpose to psychic force, we encounter a conundrum: either the “energetic” interpretation of the Death drive yields an a-psychical dynamics of brute destructiveness, or the goal of psychical self-destruction is impossible. The conundrum points to the holding pattern between the first two of the three metapsychological framings of the Death drive, noted earlier; those given in terms of economies of energy and affect. Bell's object-relational typology of deathly states of mind offers a way through by making destructiveness an action of the mind upon an object where that object can be itself; following Bion, we understand the deathly state of mind as both depicting, and resulting from, self-directed attack motivated by hatred. The object-relational metapsychology makes the clinical picture intelligible, and the hypothesis of the Death drive thereby clinically powerful for elucidating both the dynamics of these states of mind and their consequences in human relating. For post-Kleinian theorists, the thesis of primary destructiveness takes its theoretical justification from Bion's theory of thinking (Bion 1967) and its empirical confirmation from clinical typologies of the deathly state of mind. I will be challenging this justification in what follows, arguing that the hypothesis of self-attack by the mind is an uncontrolled implication of Bion's theory which cannot provide a secure basis for empirical observation. The thesis's clinical power cannot be predicated on the evidence of the phenomenology since the deathly state of mind is not about death, after all.

This argument for this critical point is made, in the general case, by the philosopher Richard Wollheim, who cautions against reliance on phenomenology as a guide to mental typology.³ While mental states' phenomenology presents as direct and unmediated, this deceptively simple presentation obscures the complex psychological role in the mind that lies behind it. By the same argument, phenomenology is an unreliable basis on which to construct, adjudicate or agree on metapsychological explanations. Rather it is the case that phenomenology follows, and serves, the underlying function so that the philosophical task, as put by Wollheim, is to "exhibit, over as wide a range of mental phenomena as is feasible, how function is served by phenomenology" (1979, 218). To clarify the role of the Death drive in post-Kleinian metapsychology I shall pose this question with respect to the phenomena that it purportedly produces and which, as I shall argue, can be otherwise explained in terms of function.

Bell allows that "a question naturally arises as to whether hatred of thought is a kind of irreducible datum or a manifestation of a deeper process" (2014, 5). The interpretation of

³ Wollheim argues that appeals to phenomenology to ground functional differences are unsatisfactory in relying on antecedently convergent intuitions, or on conventions, for describing the phenomena. While eliciting "reverberations" between speakers may work "within general limits" for a mental phenomenon, Wollheim maintains this will not support discrimination of mental states, such as wish-fulfilling thoughts, or screen memories, whose function is to mimic another type of state. Writing of the phenomenology of memory, Wollheim argues that memory state types are grounded causally in their functional mental activity, of which phenomenology is at best an indicator (1979, 218); he writes that "At least in our philosophising about them ... phenomenology follows function".

mindlessness in terms of destructive mental processes can be challenged with material that shows the absence of thought to be defensive against pain, such that the deadening of thought is in fact “in the service of life”.⁴ Bergstein (2015) suggests that a clinical picture of extreme incoherence of thinking may be understood not as the result of self-attack but has a functional role as performative; the patient’s state of mind shows what she cannot tell, the state of her mind under attack. Bell himself adopts both explanations: if we “focus on the way destructive activity functions” (2014, 7), we may see it originating in defence against unbearable affect but equally, attempts to preserve thought can be recruited into a manic defence that is itself anti-thought, an extremity of destructiveness in which even the mind’s attempt at self-protection is attacked and psychic defence itself falls victim to the Death drive. Although the point cannot be pursued here, it seems that under Bell’s pessimistic interpretation of the Death drive the theory comes to perform itself; the thesis of primary self-attack self-attributes its power performatively through its negation of the life-preserving nature of the mind.

Pursuing my critical agenda, however, I return to Bion’s own proffered explanation of the phenomenology to see how far the thesis of mental attack and its attendant object-relational metapsychology might be given theoretical support. Is the clinical picture of mindlessness really evidence for a self-destructive force in the mind? Bion posited hate as the affect behind the clinical manifestations attributed to the Death drive, the force of hate driving an internal attack on the subject’s own mind. What his theory needs to explain is not just why the mind, through hatred, should be provoked to attack and disable itself from thinking; we also need an

⁴ Bell cites Alexander 2000 and Feldman 2000. It is a further question what makes an experience so extreme that it can only be represented in the form of a deathly state of mind; Louise Gyler has insisted on this point.

explanation of how this is possible. Positing hate here introduces two considerations; one merely clarifies the way the Death drive forces work (as they are supposed to do in conformity with Freud's binary "Two Principles" model of mental functioning) by opposing the life force, in the mind as part of its equilibrium-maintaining apparatus. Just as love is the driving affect of Eros, so hate is the driving affect of Thanatos. But a further consideration accompanies Bion's positing of hate since like love, hate is object-relational and so, can be reflexive. Klein (1948) had already postulated an "epistemophilic instinct" in which the mind loves its own knowledge-seeking capability. If the mind can be the object of its own love, so too it could be assumed to be the object of its own hate. Bion formalises this line of thought in his grid of positive and negative attitudes to knowledge and to knowing ("K"); that is to say, hate is directed to the capacity for knowledge; more precisely, to the capacity for mental activities which deliver knowledge and which are the subject of his theory of thinking.

But the parallel with love does not hold up; the hypothesis of the mind's attack on itself is self-negating, as we shall see. Bion's model of the mind gives no clue as to how the mind "discovers" self-attack as a route to mental self-regulation when the attack is precisely directed at, and if successful will destroy, its capacity to discover anything.⁵ Part of the clinical power of the self-attack hypothesis is that it provides a suggestive metaphor for a mental activity that is not possible in the terms of Bion's theory of the mind; but what lies behind the metaphor is a self-stultifying assumption that the mind can do something impossible. The mental activity

⁵ To put the matter in a Hegelian way, what is needed is a conceptualisation of thinking such that the negation of mind's capacity for knowledge is a determinate negation; psychologically a (reversible) suspension of cognitive activity allowing some further, different form of consciousness to emerge; this is one description of the path that I shall pursue.

in question is the interminable ongoing annihilation of the object; in clinical experience this is phenomenologically apt but as a theoretical claim it is incoherent since the object continually under annihilation is the activity of annihilation. Phenomenology here is not a guide to function.

To understand what is involved in the claim that the capacity for thought at all, and *a fortiori* for knowledge in particular, can come under attack in the mind, we should return to Bion's theory of thinking. It is recognised that there is a Kantian inspiration behind his theory and my claim is that, properly understood, this can provide an answer to the question of self-attack and, too, with illumination of how the mind represents and "thinks" the annihilation anxiety that supposedly both triggers and responds to mental attack. To do this we have to understand the use Bion makes of Kant, taking as the key the first chapter of his *Learning from Experience*.

4. Bion's theory of thinking

Bion is supposed to owe a debt to Kant but I shall argue that its significance is not what it is thought to be. Psychoanalysts have sometimes been tempted to think of Bion in his clinical observation and theorising as empirically validating Kant's own psychology, but this is certainly not the relation between Kant's thought and Bion's theorising. Kant was not concerned with empirical questions about how the mind actually works, but with the prior question of how the mind must be structured so as to do what we see it as doing. Moreover, although as an undergraduate Bion had some exposure to Kant's work, and owned a number of his works, he is held either to misunderstand or disregard the detail of Kant's work (Noel-Smith 2013, 132). In any case Bion does not import a Kantian theory of the mind wholesale, but most likely picked up the elements of Kant's critical philosophy present in Freud's thought

(Stänike, Zachrisson and Vetlesen 2020; Tauber 2009). Freud's use of Kantian ideas is not systematic and neither is Bion's, at least at first sight.⁶ Nevertheless, we should not dismiss Bion as doing bad philosophy of mind by loosely appropriating Kantian terms to refer to psychological phenomena. Instead, I suggest that Bion understood what he took from Kant well enough to put it to a particular use, not as a template for a psychoanalytic psychology but as a theoretical model for his theory of thinking. I shall then argue that we can extend that model to dismantle the idea of mental self-destructiveness.

In the *Critique of Pure Reason* Kant considers how it is possible for us to come to know the world through the activity of cognising. Cognition produces knowledge through the activity of judging how best to fit together our experiences, and all our experience of the world comes to us through our senses as what he calls "intuitions"; when we acquire knowledge of the world it is our sense experience that is judged. Judging brings order to our experience by organising intuitions under the abstract rules we call concepts. Finally, in a more theoretically convoluted step to which I return later, these organising concepts are produced in a collaboration between the imagination which continually assembles sense experience into patterns which Kant calls "schemata", and what he calls the "legislative", that is the law-making, activity of understanding which selects and ratifies those schemata most apt for organising experience into representations of the world. This formal structure of cognition, bringing together the two streams of sense and concept, is what Kant calls a "discursive intellect". It clearly specifies the mind as active in structuring the knowledge it acquires rather than passively receiving knowledge ready-made. Mind's activity in creating and using concepts is essential to it;

⁶ But see Civitarese (2017) on Bion's "matrix" of theoretical terms, including negative capability.

actively thinking is what Kant calls the mind's "whole vocation" (die ganze Bestimmung des Gemüts) (1790, 142). When Kant writes, "Thoughts without content are empty, intuitions without concepts are blind" (1781, 193-94), he means that to have thoughts about the world with true content that we can reason about, we need this cognitive apparatus to supply that content and also that our experience of the world remains unuseable by the mind without concepts to organise it.⁷ It is a further question (to which I return when I deal with Kant on sublime judgment) as to what other types and sources of content thoughts could have, but it already follows from this conception of the mind that any content supplied will be the product of mind's activity; mind as the organ of thought is necessarily active.

If, to situate ourselves for what follows, we now look at how these Kantian ideas are deployed in Bion's theory of thinking, we see that it contains the notion of a discursive intellect.⁸ This borrowed structure is clearly seen in his theory of functions where, in particular, what is done in 'alpha-function' is the processing of sensory beta elements into conceptual thoughts or alpha elements. In Bion's account, the production of thoughts by alpha-function is, if only approximately (see Noel-Smith 2013), the work done by judgment in supplying and applying

⁷ Kant's target here is metaphysical thought; knowledge claims that go beyond experience and so, beyond what we are warranted in claiming to be true.

⁸ The operation of judgment as bringing the intuitions of sense under the concepts of the understanding, which is a formal description of the working of cognition, is unhelpfully parlayed by Bion into the metaphor of "mating", presumably intended to carry the implication of something new being generated.

the concepts or rules of the understanding to the intuitions of sense which, without concepts, are “blind” beta elements of experience.⁹

We should also consider the following argument. It is an assumption of Bion’s theory that thinking as the active handling of thoughts defines the mind as its essential function. It is a logical consequence of this that eliminating that essential function will eliminate the mind. Were Bion to make this argument it would take him beyond Kant for whom the elimination of mind’s activity is of no theoretical interest since without it no discursive intellect is conceivable; we are outside the realm of what can be thought.¹⁰ When therefore Bion does go on to posit the idea of a self-destructive attack on mental activity which disables thinking and which has therefore the potential to destroy the mind, this is not a coherent possibility within the transcendental psychology of Kant’s First Critique. Consequently, this step in Bion’s theory of thinking, even if responding to clinical intuition and in line with psychoanalytic theory, has no basis in his Kantian theoretical model. Bion’s further step remains to be theoretically explained and to do this we will need to extend the model base for Bion’s theory of thinking beyond what can be discerned in his own writing.

5. Bion and philosophy of science

⁹ Bion uses “function” as placeholder for a range of psychological activities; most systematically for the object-relational attitude to knowledge in terms of the active relations of love or hate in the “grid” notation.

¹⁰ For more on discursive intellect see Maddy (1999).

When I said that the nature of Bion's debt to Kant is not well understood, I meant that it resides not in the borrowed structure for cognition itself but in the methodological role the structure is, if only implicitly, called on to play in Bion's theorising. In arguing this, my approach is more interpretive than demonstrative, but I will hope to show from what Bion says that he is committed to a systematic approach to psychoanalytic theorising. For this he turns to philosophy of science. I base my remarks principally on the Introduction and first chapter of *Learning from Experience*.¹¹ This should be seen as a theoretical work setting out an intellectual programme for psychoanalysis, and in it Bion is doing several things. His motivating idea is that much psychoanalytic theorising is unsatisfactory, and he refers to the need to avoid "the elaboration of new and possibly misguided theories" (1962, 2). He is critiquing existing theories and aiming for the development of psychoanalytic thinking and technique through clear and correct theorising, suitably guided by philosophy. He describes a methodology intended to make his own theorising exemplary in contrast to the practices of theory building he condemns. He rejects theory-building on already-interpreted particular clinical experience, on the ground that its explanations are self-fulfilling and wrong theory can never be refuted or, consequently, discarded.¹² His intention is to provide a clear theoretical basis for psychoanalytic generalisation, for which the methodology of the natural sciences as presented in the then current philosophy of science was to provide the guidance and standards.

¹¹ Reasons of space prevent a detailed exposition of Bion's substantive theorising, or of other things he says about thought and its disorders.

¹² Although Bion doesn't reference this Popperian critique of psychoanalysis directly, he will have been aware of it, since he references Popper.

An important reference for Bion is the work of the Cambridge philosopher of science R. B. Braithwaite. Almost Bion's exact contemporary, Braithwaite belonged to the tradition of logical positivism, where the methods of formal logic were deployed to reveal what was thought of as "the structure of scientific explanation". Within that structure science was defined as conforming to a "deductive-nomological" schema where events were explained in terms of general laws from which they followed deductively as instances. Despite the prominence of this approach, Bion is not to be seen as importing a positivist deductive-nomological structure of science into psychoanalytic theory; it was not his ambition to regularise psychoanalysis into conformity with the positivist canon. Although he takes his inspiration from Braithwaite's logical modelling of scientific theories using a mathematical type of algebraic notation, together with formalisation and abstraction, and logical forms of inference, he gives notice that his use of terminology will depart from its received usage. But he was struck with the rigour of the approach and the usefulness of its methods for putting psychoanalytic explanation on a firm footing, equivalent to scientific explanation in generality and reliability. In his theory of thinking Bion is aiming for a methodological propriety in psychoanalysis for which the positivistic logical methods of the time provided a frame of reference. He is concerned to systematise psychoanalytic explanation of clinical phenomena by using robust generalisations as an equivalent to empirical laws in science, and he explores how to make their role in psychoanalysis systematic. His theory of functions is explicitly intended to supply psychoanalysis with the generality enjoyed by science. Here the "function" is a placeholder, a logical variable term which provides for generality by accepting different values. Bion says, that the term alpha-function is "intentionally devoid of meaning" and compares it to a

“mathematician’s variable” and that it can take different values (1962, 3).¹³ While he adapts this and other terms freely for his own purposes, his emphasis is squarely on methodological propriety and he recognises the value of the formal side of philosophy for philosophically analysing psychoanalytic investigation, particularly through the idea of a model.

6. Thinking with models

How then should Bion’s use of Kantian theory be understood? I have suggested that his likely aim was to formalise psychoanalytic explanations based on observations in the clinic into theories conceived on the lines of the logical-positivist analysis of science he found in Braithwaite. This, on the lines I’ve indicated, would make psychoanalytic theories more precise, unambiguous, and methodologically respectable, with the aid of logical devices such as the factorisable “function”. No more than seeing Bion as doing bad philosophy of mind by trying to get Kantian terms to refer to psychological phenomena should we see him as doing over-ambitious philosophy of science. But he is insistent on the need to improve psychoanalytic

¹³ Bion also foresees the term alpha-function as “re-invested” with extraneous meanings, remarking, “Constant vigilance must therefore be exerted to prevent this development” (1962, 3). Presciently for his own case, he references Popper’s remark that, “When I invented the new name ‘basic statement’ (‘or basic proposition’) I did so only because I needed a term not burdened with the connotation of a perception statement. But unfortunately it was soon adopted by others, and used to convey precisely the kind of meaning which I wished to avoid” (Popper 1935, 12 n2). Bion’s terminology, intended to ensure methodological neutrality, is frequently misused in just this way.

theorising and though his psychoanalytic theories are not logico-deductive models of a Braithwaitian sort, they are another sort of theoretical model.

The idea of a theoretical model for a scientific theory is one of the logical methods discussed by Braithwaite. In the then prevailing positivist canon, a scientific theory was a “deductive [scientific] system ... [a set of hypotheses] which is arranged in such a way that from some of the hypotheses as premises all the other hypotheses logically follow” (Braithwaite 1953, 12) and Bion actually quotes this definition verbatim in his *Cogitations* (1992, 2-3).¹⁴ The provision of a more concise, logical model would thus render a scientific theory more graspable. Braithwaite distinguishes scientific theories from their models, writing: “there are great advantages in thinking about a scientific theory through the medium of thinking about a model for it; to do this avoids the complications and difficulties involved in having to think explicitly about the language or other form of symbolism by which the theory is represented” (1953, 92). A scientific theory is not easy to grasp when in the form of an extensive and unwieldy system of theoretical statements in the technical vocabulary of, say, biology or psychology. Bringing it under a generic theory such as systems theory or functional analysis, or an abstract representation using symbols, analogously to an algebra, makes it more wieldy and accessible.

I shall develop the claim that Kant inspired Bion with a model for his own theory of thinking; moreover Bion’s Kant-inspired theory is to be thought of as a theoretical model, although not for the formal sort of scientific theory that Braithwaite is talking about. For Braithwaite, as a philosopher of science, one starts with the original scientific theory and then finds a logical-

¹⁴ Thanks to Niall Gildea for pointing this out.

mathematical model to make thinking about the theory easier, as I've indicated.¹⁵ The sort of original theory that Bion is starting from is not a fully worked-out scientific theory but his own contributions to the generic psychoanalytic corpus which developed after Freud and over a long period to organise clinical observations and, as his critical remarks indicate, became overburdened with the accumulation of redundant and mistaken claims. Bion's own theories, certainly the theory of thinking, can be seen as theoretical models for aspects of the original Freudian theory, particularly the nature of thought and its disorders which Bion was concerned with. And as models these themselves draw on, or borrow from, the Kantian ideas Bion had encountered, for their structure; as I shall argue, Kant supplies Bion's own theory with a model.

Max Black, a philosopher of science in the USA contemporary with Braithwaite in the UK, calls the Braithwaite type of model an "existential" theoretical model; as we have seen, it applies to an original theory which is a scientific theory of the real world. Such a model couches the original theory in a different vocabulary, for example modelling blood flow with the mathematical theory of fluid mechanics. Advances in understanding from exploring the model are checked out in terms of what the real-world applications and confirmations will be. Black contrasts this with the "as-if" theoretical model, which resembles "a sustained and systematic metaphor" (1962, 236); Freud's "mental apparatus" in the "Project for a Scientific Psychology" is an example of this sort. Both types of theoretical model are heuristic; they work by bringing

¹⁵ Remarking that "Constant vigilance must therefore be exerted to prevent" misuse of terms like alpha-function, Bion is re-quoting Braithwaite's "the price of the employment of models is eternal vigilance" (Braithwaite 1953, 93). Bion concurs with Popper, noting that what are intended as purely theoretical terms become burdened with extraneous meaning leading to "myths" or, "new and misguided theories".

two fields of enquiry or two bodies of theory into relation in such a way as to explore the one (the original theory) through the lens of the other (the theoretical model). The difference lies in the fact that in the existential model the theoretical claims are about things that exist, while the as-if model's terms are not. Instead, like a metaphor, an as-if theoretical model provides fictions or scenarios which are elaborated by the imagination to produce new features generated by the model. Indeed, the metapsychological dimension of psychoanalytic theories often works in just this semi-metaphorical way to suggest new concepts and connections which can be productive for psychoanalysis.

Bion's theory of thinking differs from other psychoanalytic metapsychologies if we count as available to it the organising and heuristic power of the Kantian theoretical model it adumbrates, which my remarks about Kant's epistemology in the First Critique were intended to make visible. But without adequate theoretical control Bion's theory is, in the way Popper points out, prone to produce unwarranted claims which become myths. I shall suggest that the idea of self-attack by the mind is such a claim. However, there is a feature of an as-if model that distinguishes it from metaphor; the model's systematicity provides for what Black calls the "control of implications" (1962, 233). Not every speculation or apparent implication heuristically derived from a model counts as a justifiable inference from it. What can allowably be inferred is controlled by reference to the model's theoretical character as a systematic, even if not formally logical, structure. Applying it limits the imagination to inferences which, when applied to the original psychoanalytic theory, can be justified within the model theory. Bion himself recognised the need to control implications to avoid "new and misguided theories", and it is on this ground that we could require him to accept the onus of seeking a satisfactory explanation for the phenomenological presentations attributed to the Death drive. My conjecture, based on Bion's focus on the formal aspects of methodology of enquiry, is that

Kantian theory appealed to him at least partly for this reason. As a highly structured argument-based justification of our necessary cognitive activity, it is both its own control and offers to control both Bion's own theory and others' speculations. Bion borrows from Kant because the structure of Kant's theory lends itself to correct theory-making, limiting the implications and speculations to those that follow correctly from its heuristic use. This structure, as Bion encountered it, belongs with Kant's epistemological project of the First Critique and the question of our knowledge of the world.

It is by referring back to the model of the Kantian epistemology that the theory of thinking can rightly contribute to psychoanalytic understanding. Taking this view of Bion's theory may therefore allow us to make further progress on the vicissitudes of the mind in experiencing profound anxiety and on whether the Kleinian posit of self-attack has any theoretical place in explaining these. To do this I shall look further at how we might augment the implicit Kantian model behind Bion's theory of thinking with a new theme, one not present in the First Critique, beyond which we must now move. I will suggest a way to expand the model, specifically to suggest that alpha-function, as a variable, can be substituted by more than one type of judgment and that the type of judgment we should now consider, in order to extend his theory so as to include extreme anxiety, is to be found in Kant's third work on the nature of judgment, his *Critique of Judgment* itself.

7. Extending Bion's theoretical model: what psychoanalysis might further learn from Kant

I said earlier that the idea the mind could reflexively attack itself was self-negating within the terms of Bion's object-relational approach. It is also an uncontrolled implication within Bion's

theory of thinking; not only is there no justification for it in the Kantian model of cognition but it is self-negating anyway in the terms of Kant's theory of mind. However if we wish to regard clinical observation as both reliable and informative we should continue to seek a theoretical explanation for it within the model supplied by Kant's thought. We should do this not by inference from the phenomenological presentation but as Wollheim recommends, in terms of the processes behind it. Such an explanation is indeed available in what Kant says about judgments of the sublime in the *Critique of Judgment*. We encounter the sublime as what exceeds the power of cognition, when we are presented with an experience or perception of overwhelming force or magnitude, in nature or in objects (real, and mathematical) of vast scale.¹⁶ What is encountered in nature Kant calls the dynamically sublime, but it is in his treatment of the mathematically sublime that we have a formal analysis of the mind's ability to produce such judgments.

¹⁶ The forces of nature are described by Burke as inspiring awe, terror, and a sense of uplift. Kant, while familiar with Burke's psychological account, has the different goal of showing how sublime experience can be thought by a discursive intellect. He distinguishes between the dynamical sublime (approximating to Burke's psychological picture) and a prior, more general form of sublime judgment in which magnitude itself comes to have a representation, which he calls the mathematically sublime. This presents when we attempt to grasp, in the thought of infinity, the endless progression of numbers emanating from the productive imagination and is also that which we (mistakenly) attribute to an object when we attempt to contemplate and represent vastness of scale and complexity in artefacts such as the Pyramids or St Peter's in Rome.

Here I follow Kirk Pillow's account of sublime judgment; for the mind to contemplate the overwhelming under some thinkable representation a necessary cooperation between the faculties of reason and imagination must be reached.¹⁷ In this negotiation the imagination is required to organise its own intuitions into a graspable whole to meet reason's demand for a concept with which to comprehend the magnitude of the mathematically sublime.¹⁸ Kant conceives the mathematically sublime as an endless stream of intuitions or images produced by the imagination, challenging the mind to bring its faculties to bear on the task of representing a boundless excess. Reason's role is to require this un-ruly heterogeneity to be brought together as a thinkable whole. Imagination as the productive faculty is required to find some conjectural

¹⁷ For scholarly exposition I draw on Pillow (2000). In "determinate" judgment, sense-based concepts are formed by the schematising activity of the productive imagination and "legislated" by the faculty of the understanding for their aptness to experience. In Pillow's account of sublime thought the productive imagination comes into its own since Kant allots it a new role in the untrammelled production of intuitions from its own resources. In sublime judgment, which is reflective and not sense-based, the concepts it needs in order to produce thoughts come from reason and imagination. The discursive structure is thus maintained, being reproduced at the level of reflective thought.

¹⁸ The psychoanalytic application of the dynamically sublime is explored by Weiskel in a non-object relational framework; the clinical picture we are concerned to explain does not figure in his account. Civitarese (2017) maintains that the aesthetics of the sublime in Romanticism provided Bion's with a "secret model", but without noting Kant's distinction of two types. My account therefore precedes any question of relating the dynamical sublime to psychoanalysis, being concerned with Kant's answer to the prior problem of how judgment of the sublime is possible at all.

way of assembling what is intuited as endless into a graspable thought content. This demand is met by the ability of the imagination to schematise its own productions into concepts that can organise, arrest and assemble the progression of intuitions it produces.

This augments our model of the mind with the resources to explain the phenomenology of deathly states of mind as representations which achieve this accommodation of reason and imagination. Instead of mis-interpreting such states' phenomenology as representing the activity of the mind, we should see their phenomenology mis-representing as active destructiveness what the model shows to be an integrative achievement. However, before proceeding to set out this final step in my argument we should note that Kant does in fact see this achievement as involving a certain self-violence within the mind.¹⁹ For in producing schemata that will self-organise its endless productions into a thinkable unity the mind does indeed do violence to itself. Kant describes the sublime as "contrapurposive", meaning that it runs counter to the mind's "vocation" to think.²⁰ We saw that in its accommodation with reason

¹⁹ It is worth noting a likely degree of familiarity with Kant's Third Critique within the early, Germanophone intellectual context of psychoanalysis (all but lost in the analytic diaspora produced by Nazi-ism), and the context of Klein's own analysis with Abraham and Ferenczi.

²⁰ Of the mathematically sublime Kant writes: "the comprehension in one moment of that which is successively apprehended, is a regression, which in turn cancels out the time-condition of the mind in the progression of the imagination and makes simultaneity intuitable. It is thus (since temporal succession is a condition of inner sense and of an intuition) a subjective movement of the imagination by which it does violence to the inner sense which must be all the more marked the greater the quantum is which the imagination comprehends in one intuition" (1790, 142).

the imagination has to rein in its productions under some totalising schema; to do this it must misrepresent endless, temporal succession as simultaneity within a whole and this does violence by opposing or negating the intuition of temporal succession which is an *a priori* intuition structuring the mind. This, Kant says, “subjectively considered, is contrapurposive” but “objectively, for the estimation of magnitude, is necessary, hence purposive; ... the very same violence that is inflicted on the subject by the imagination is judged purposive for the whole vocation of the mind” (1790, 142).

8. Applying the model: sublime judgment and extreme experience

In formulating his theory of alpha-function Bion wished to retain the term itself as a neutral, content-free variable standing for more than one way of processing the beta elements of experience into thought, with the Kantian model of cognitive judgment in the background. The extended model we now have suggests that alpha-function as thought-production can also be conceived as sublime judgment, where thoughts are produced by the accommodation of reason and imagination. One of the canonical clinical descriptions from which the operation of the Death drive is inferred is the fragmentation of thought back into beta elements when the mind’s thinking capacity is attacked. On the extended model it is the imagination that produces intuitions which, being without concepts, are “blind”; we might see these as elements of thought that remain as fragments until assembled under the totality demanded by reason. Setting aside whether the model can accommodate the hypothesis that the attack on thought turns it back into fragments equivalent to the beta elements out of which a thought could be reconstituted, what is posited by the model is a violent opposition to mental integrity; we have just seen that in Kant’s theory the imagination responds to the demand of reason by an attempt

at its own self-regulation which runs counter and does violence to the temporal, reality-oriented activity of mind.

In the clinical picture the self-violence of the imagination's attempt at its own self-regulation is clearly represented by the different depictions of negation: omnipotently misrepresented as total self-annihilation; the endless production of intuitions unified as its own negation in a Nirvana-state of total undifferentiatedness; as a reversal of the endless streaming into the stasis of a living deathliness; or a negation of excess into pleasurable emptiness. We see this in Bell's three clinical pictures: the explosion as a drive that aims at destroying life; the continuous fall-out of dust representing the pull to mindlessness, both as dead (particle) objects and as deadening the mind's function of object-relating; and the fall-out as a continuous paralysing of the mind's objects kept in a suspended live state with sadistic pleasure. Under the model these are different ways the imagination can schematically self-regulate its endless stream of intuitions under an over-arching representation of interminable annihilation.²¹

Thus far, the model explains the phenomenological presentation. It also tells us that such self-violence, as contrapurposiveness, is intrinsic to the achievement of sublime judgment when imagination accedes to the demand of reason to make possible the continuation of thought which is the mind's "whole vocation". This is a demand which may be made in response to

²¹ It may be noted that the perverse pleasurableness of these "sublime" solutions accords with a feature of the dynamical sublime not mentioned in my compressed account, the combination of un-pleasure provoked by overwhelmingness and the pleasure attendant on the mind's mastery of it into thought; the displeasurable experience is at the same time made pleasurable.

I owe this point to Damien Freeman.

extreme experiences from other causes, that are traumatic or affective and which threaten to overwhelm the mind. In following its vocation reason persists in holding everything into a thinkable whole through attempting an accommodation with what imagination produces. To accommodate reason the imagination then produces further defensive transformations, such as the displacement of the representation of extremity and excess into an originating catastrophic event “elsewhere”. Bell reports a patient dreaming that “*there had been a nuclear explosion. The dust, fall-out, fell all over me. It was pleasant and peaceful*” (2014, 4). Equally, reason may adopt some integrative but false idea: a paranoid representation of attack, inside or outside, causing the overwhelmingness; negation emptying the mind by displacement of all mental content “elsewhere”; omnipotence as endless annihilation.

We see now how extending the Kantian model explains the range of clinical presentations Bell describes. Unifying these deathly states of mind under a rubric of primary destructiveness had the merit of drawing attention to ways that imagination’s “subjectively contrapurposive” disruption of the mind’s activity can be represented in, and enacted through, phantasy. It also re-situates Klein’s explanation of the Death drive as fear of annihilation in terms of a secondary, paranoid response to the imagination’s violent self-(mis)representation of a mind empty of content or paralysed in its activity. My question about the organising power of this mistaken metapsychology can now be addressed. The Death drive and its attendant concepts are theoretical placeholder terms; they make no explanatory contribution. But together they are a clinically powerful instrument; they provide interpretive guidance for identifying the variety of object-relational defences that misrepresent as destructive a fundamentally life-oriented activity of mind under extreme challenge to its capacities. The extended as-if model thus opens a different perspective on these clinical phenomena and a warranted theoretical explanation for how the mind deals with extreme experience. And it absolves us from postulating a quasi-

energetic and inexplicable primary destructiveness that gets turned onto the mind itself. Instead it provides a model for organising further clinical observation, and focuses on the complexity, fragility, and adaptability of the mental apparatus.

Disclosure of interest

The author reports no conflict of interest.

References

Alexander, R. P. 2000. "On Feldman's 'Some views on the manifestation of the death instinct in clinical work.'" *International Journal of Psychoanalysis* 81 (5): 1007-1009. PMID: 11109583.

Bell, David. 2014. "The death drive: phenomenological perspectives in contemporary Kleinian theory." *International Journal of Psychoanalysis* 96 (2): 411-423. DOI: 10.1111/1745-8315.12212.

Bergstein, Avner. 2015. "Attacks on linking or a drive to communicate? Tolerating the paradox." *Psychoanalytic Quarterly* 84 (4): 921-942. DOI: 10.1002/psaq.12043.

Bion, Wilfred. (1992) 2005. *Cogitations. New Extended Edition*. Edited by Francesca Bion. Reprint, London: Karnac.

Bion, Wilfred. 1962. *Learning from Experience*. London: Karnac.

Black, Max. 1962. *Models and Metaphors: Studies in Language and Philosophy*. Ithaca, NY: Cornell University Press.

Braithwaite, Richard Bevan. (1953) 1968. *Scientific Explanation. A Study of the Function of Theory, Probability and Law in Science*. Reprint, Cambridge: Cambridge University Press.

Civitaresse Giuseppe. 2017. *Sublime Subjects: Aesthetic Experience and Intersubjectivity in Psychoanalysis*. New York: Routledge.

Feldman, Michael. 2000. "Some views on the manifestation of the death instinct in clinical work." *International Journal of Psychoanalysis* 81 (1): 53-65. DOI: 10.1516/0020757001599528.

Freud, Sigmund. (1920) 1981. "Beyond the Pleasure Principle." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. XVIII (1920-1922): Beyond the Pleasure Principle, Group Psychology and Other Works*. Translated by James Strachey, Anna Freud, Alix Strachey, Alan Tyson. Reprint, London: The Hogarth Press and the Institute of Psycho-Analysis.

Freud, Sigmund (1911) 1981. "Formulations on the Two Principles of Mental Functioning." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. XII (1911-1913): The Case of Schreber, Papers on Technique and Other Works*. Translated by James Strachey, Anna Freud, Alix Strachey, Alan Tyson. Reprint, London: The Hogarth Press and the Institute of Psycho-Analysis.

Freud, Sigmund. (1917) 1981. "Mourning and Melancholia." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works*. Translated by James Strachey, Anna Freud, Alix Strachey, Alan Tyson. Reprint, London: The Hogarth Press and the Institute of Psycho-Analysis.

Freud, Sigmund (1895) 1981. "Project for a Scientific Psychology." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. I (1886-1899): Pre-psycho-analytic Publications and Unpublished Drafts*. Translated by James Strachey, Anna Freud, Alix Strachey, Alan Tyson. Reprint, London: The Hogarth Press and the Institute of Psycho-Analysis.

Kant, Immanuel (1781) 2000. *Critique of Pure Reason*. Translated and edited by Paul Guyer and Allen W. Wood. Reprint, Cambridge: Cambridge University Press.

Kant, Immanuel. (1790) 2002. *Critique of the Power of Judgment*. Translated by Paul Guyer and Eric Matthews. Edited by Paul Guyer. Reprint, Cambridge: Cambridge University Press.

Klein, Melanie. (1948) 1997. "On the Theory of Anxiety and Guilt." Chap. 2 in *Envy and Gratitude and Other Works 1946-1963: The Writings of Melanie Klein, Vol. 4*. Reprint, London: Vintage.

Maddy, Penelope. 1999. "Logic and the discursive intellect." *Notre Dame Journal of Formal Logic* 40 (1): 94-115. DOI: 10.1305/ndjfl/1039096307.

Noel-Smith, Kelly. 2013. "Thoughts, Thinking and the Thinker: Bion's Philosophical Encounter with Kant." In *Bion's Sources: The shaping of his paradigms*, edited by. Nuno Torres and R. D. Hinshelwood, 124-136. Hove: Routledge.

Pillow, Kirk. 2000. *Sublime Understanding: Aesthetic Reflection in Kant and Hegel*. Cambridge, MA: MIT Press.

Popper, Karl. (1935) 2002. *The Logic of Scientific Discovery*. Reprint, London: Routledge.

Stänicke, Erik, Anders Zachrisson and Arne Johan Vetlesen. 2020. "The epistemological stance of psychoanalysis: revisiting the Kantian legacy." *Psychoanalytic Quarterly* 89 (2): 281-301. DOI: 10.1080/00332828.2020.1717229.

Tauber, Alfred I. 2009. "Freud's dreams of reason: the Kantian structure of psychoanalysis." *History of the Human Sciences* 22 (4): 1-29.

Weiskel, Thomas. 1976. *The Romantic Sublime: Studies in the Structure and Psychology of Transcendence*. Baltimore, MD: The Johns Hopkins University Press.

Wollheim, Richard. 1979. "Memory, Experiential Memory and Personal Identity." In *Perception and Identity: Essays Presented to A. J. Ayer with his Replies to them*. Edited by Graham F. Macdonald, 186-234. London: The Macmillan Press.