Containment and 'Rational Health': Moran and Psychoanalysis

Edward Harcourt

**Faculty of Philosophy, University of Oxford
and
Keble College, Oxford
Oxford OX1 3PG
UK**

edward.harcourt@philosophy.ox.ac.uk

EJOP-Mar-16-3544.R1

0. I take Richard Moran's *Authority and Estrangement* (2001; cp. Moran 2004) as an extended attempt to theorize the distinction between – to borrow a useful label from Bernard Williams – ‘happenings outside one’s moral self’ and happenings inside it (1977: 104). After explaining the inside/outside distinction very briefly, section 1 of this paper presents a second distinction that I want to lay alongside that first one, the distinction from a distinctively post-Kleinian strain within the literature of psychoanalysis between ‘contained’ and ‘uncontained’ states of mind. At first sight, it looks as if Moran’s inside/outside distinction is a much needed theoretical deepening of the distinction from the post-Kleinian psychoanalytic literature. Real though the similarities are, however, sections 2-5 go on to explain why the two distinctions are nonetheless not the same. Section 2, by way of some reflections on disgust, argues for a normative distinction *within* the class of states that fall on the wrong side of Moran’s norm of ‘rational health’, so an adequate normative mapping calls at least for two norms not one; whether it’s ‘rational health’ or this second norm that best corresponds to the post-Kleinian notion of ‘containment’ is left open. By fleshing out what in practice containment involves and what it contrasts with, section 3 then argues that it is indeed the second norm that corresponds to ‘containment’, so Moran’s ‘rational health’ is not identical with containment, but a more exigent notion. That conclusion is compatible with the view that the two norms, though distinct, are complementary and can be fitted together into a more nuanced picture of grades of psychological health or ill-health. Sections 4 and 5 argue that the relation between containment and ‘rational health’ is more complicated than that: some states that meet Moran’s transparency condition are not rationally healthy (section 4), but (section 5) the notion of containment can help us to see what more is needed to render them so.

1. Williams’s phrase ‘happenings outside one’s moral self’ occurs in the course of a discussion of the prospective negative psychological effects on an agent of choosing a course of action, and whether and how these effects might constitute reasons for the agent to alter his choice. Such effects are happenings *inside* the agent’s moral self if they are, in Williams’s words,

emotional expressions of a thought that to [do whatever it might be] would be *wrong* (1977: 103),

or as ‘indications of what he thinks is right and wrong’ (1977: 103); whereas if there is no such thought they express, they are seen by the agent ‘*just* as disagreeable experiences of his’ (1977:103), and are thus ‘outside’. By way of illustration – though the case will turn out to be more complicated than this initial presentation allows - consider someone who recoils in disgust at the sight of slaughtered animals hanging in a butcher’s shop. If, for example, they are vegetarians on moral grounds, their disgust will express this conviction and to accuse such a person of squeamishness is at best a rhetorical device which takes aim at a disposition they don’t have as an easy alternative to taking aim at one – that is, their moral opposition to killing animals for food – which they do (1977: 102). The disgust of *such* a person is a happening ‘inside’ their moral self. If on the other hand one likes eating meat and

sees no reason against it, the disgust is squeamish: there's nothing in the thought department for it to express, and the happening is 'outside'. Moreover, if there are reasons for action in this second case (such as reasons to order meat online, thus avoiding the disagreeable sights at the butcher's), they can stem only from the intrinsic unpleasantness of the disgust. By contrast if the disgust is an 'inside' happening, reasons for action will – by and large: I will leave explaining this qualification till later - be the very same reasons as the reasons for the disgust itself. As these illustrations show, the thoughts which feelings must express in order to count as 'inside' need not be only thoughts of *wrongness*: they could be quite various. Nor need the distinction apply only to feelings: though perhaps it works especially nicely here, it applies to attitudes more generally. Following Moran, I'm going to take it that the touchstone of the inside/outside distinction is (crudely) whether such feelings (or one's dispositions thereto) are or are not, by the agent's lights, *rationalized by their objects*, so

first-person questions [about one's attitudes are] transparent ... to
corresponding world-directed questions

where these are questions 'about the justification or appropriateness of the attitude' (Moran 2004: 423).

Now it's striking how present the inside/outside distinction – or at least, *an* inside/outside distinction – is within some versions of psychoanalysis, that is, within the 'object relations' school and, perhaps especially, where Melanie Klein has been a dominant influence: 'post-Kleinian psychoanalysis' for short.ⁱ An infant's 'instincts' – and I won't put any pressure on the choice of term yet – are, says D.W. Winnicott,

not yet clearly defined as internal to the infant [and can] ... be as much
external as can a clap of thunder or a hit (1990c: 141).

These 'instincts' threaten the infant with 'annihilation' (Winnicott 1977: 303) or 'disintegration' (Winnicott 1990a: 61). If, with maturation, the infant acquires enough psychic structure to be able to meet this threat, the instincts however will be not 'external' but 'contained' in his mind – Bion's word (1984a: 106; 1984b: 118); cp. Levine 2011) - making it possible for them to interact in the right way with other states of mind of his: he 'acquires the capacity to hold ... his own mental state' (Waddell 1998: 208), becoming someone 'capable of experience' (Waddell 1998: 210). If on the other hand this doesn't happen, these states will

lead a split-off, repudiated existence on the edge of the mind, but with an
undiminished, if not actually increased, potential for impact on the life of the
individual (Shuttleworth 1989: 37).ⁱⁱ

For an ego which 'defensively shuts out elements from the id' is likely to be 'driven by unmastered unconscious forces' (Loewald 1978: 25), with 'somatization' – compare Lear's remark that 'vomiting can be thinking' (Lear 1990: 7) - and 'acting out' among its typical modes of self-expression (Levine 2011: 190).

How is the maturing subject supposed to acquire enough psychic structure to 'contain' his own states of mind? The answer from psychoanalysis is by way of a containing other (Bion; Winnicott's expression for what I take to be the same idea is an 'auxiliary ego',

Winnicott 1990b: 75). Good enough caregivers help infants to digest, or to make sense of, difficult experiences by helping to digest these experiences for them ('mak[ing] it possible for [a person] to investigate his own feelings in a personality powerful enough to contain them', Bion 1984a: 106), whether it is a mother who isn't thrown by an infant's initial (and puzzling) unwillingness to settle down to feed (Waddell 1998: 29-31); or one who helps a frustrated child to finish a jigsaw by turning a piece around, enabling it to discover for itself how the piece fits (Waddell 1998: 29); or who suggests explanations of why a child is tearful in such a way that the child comes to be able to say why, and in so saying becomes less so. As Margot Waddell puts it,

[When all is well] there is a 'thinking' ... mother who is able to ... resist being overwhelmed by [the child's feelings], to render them manageable. ... The baby [or child] has been offered, and has been able to receive, a true version of himself (1998: 30-31; cp. Bion 1984b: 114-5).

In a second theoretical moment, this 'digestive' capacity is then *internalized*:

[If the baby has] a sufficient number of experiences of such a containing structure within another person ... [he will] *develop a structure within himself* for dealing with his own distress. ... [He] *internalizes his mother's capacity to think about him*, ... [which] makes possible a degree of tolerance of, and openness to, experience (Shuttleworth 1989: 36, my italics; cp. Levine 2011: 189).

What the good enough caregiver did for us we become able to do for ourselves, that is, to 'tolerate' or 'contain' or genuinely to *have* our own experiences. If this idea of 'genuinely having' seems puzzling, consider this example: a teenager heavily invested in his prowess as a mathematician doesn't make it into the final of a maths Olympiad. Asked how he feels, he says he doesn't feel anything because emotions are a waste of time. But asked to say how an *ordinary* person would feel in a situation like this, he goes on – his voice strained, his body tensed as if trying to hold something together - that they would probably feel angry and upset. The anger and upset that's very obviously present isn't unowned: it is the teenager's. But there is also a clear sense in which he is *not* 'having', and indeed is straining not to have, these emotions which in that first sense are his.

Now Moran himself refers to Freud throughout *Authority and Estrangement*, and assumes that the goal of psychoanalytic therapy is to restore not only psychic health (few would argue with that) but psychic health as interpreted by his own notion of transparency:

[There is a] crucial therapeutic difference between the merely 'intellectual' acceptance of an interpretation, which will normally be seen as a form of resistance, and the process of working-through that leads to a fully internalized acknowledgment of some attitude which makes a felt difference to the rest of the analysand's mental life. ... [W]hat is to be restored to the person [by psychoanalysis] is not just knowledge of the facts about oneself, but self-knowledge that obeys the condition of transparency (2001: 89-90).ⁱⁱⁱ

This suggests that the arranged marriage I have in mind between Moran and post-Kleinian psychoanalysis could work well for both parties – that is, because it's the distinctively post-Kleinian notion of containment that's needed to explain *why* Moran might think that the psychic health psychoanalysis aims at is properly interpreted by his own notion of transparency - and that if it doesn't work, it is worth exploring why. However, since I will be arguing in the remaining sections of this paper that there are very specific grounds for thinking that this arranged marriage does not work so well, it is worth pausing at this early stage to be sure that there are not after all some obvious and very general grounds for thinking the arranged marriage is doomed from the start: otherwise the argument of the remaining sections of the paper risks being labour in vain.

One might think that the arranged marriage is doomed from the start if one assumed that whatever according to post-Kleinian psychoanalysis is 'split off' or 'uncontained' is also unconscious, since there is no reason to think that all 'outside' states in Moran's scheme must be unconscious. In the early years of psychoanalytic writing, and to the extent that these ideas were deployed at all, there was indeed an assumed equivalence between 'split off' (or 'uncontained') and 'unconscious'. Now, however, there is more of an awareness that this is not so. We can be driven by psychic forces of which we are fully conscious (panic or irrational disgust, for example) (MacIntyre 2004):^{iv} what matters is that though conscious, these states affect the subject's behaviour and other states of mind in ways that may not be apparent to him and which he is poorly able to control. One of the benefits of the contained/uncontained distinction is that it helps to loosen the link between the split off and the unconscious.^v

A second reason for pessimism concerns the nature of the mental items to which Moran's and the post-Kleinian theories respectively apply. It looks, and not just from the *Authority and Estrangement* passage quoted above, as if Moran's theory applies first and foremost to attitudes with objects – beliefs, desires, emotions – whereas in the Winnicottian formulation I quoted earlier, the post-Kleinian theory applies to 'instincts', which sound more primitive and so - perhaps because objectless – like a different sort of creature altogether.^{vi} But the less attitude-like these post-Kleinian mental items are, the less applicable to them is anything like Moran's transparency condition: as Moran himself says, 'because sensations and the like are *not* defined by internal normative demands ... the possibilities of alienation and failure to identify that are characteristic of desires and beliefs simply don't apply to them' (Moran 2004: 2002: 206).

In reply to this worry, several points can be made. First of all, however remote post-Kleinian 'instincts' are from attitudes, it is surely the case that *some* of the states that post-Kleinians see as falling within the range of the concept of (un)containment *are* attitudes: for example emotions such as envy or anger, and memories (or whatever we are to call the mental traces of experiences when they have yet to be metabolized into memories properly so called). Secondly, there is a question as to how remote post-Kleinian 'instincts' can be allowed to be from attitudes. Let us go back to one of our examples of how, in post-Kleinian theory, 'instincts' express themselves when they are

not contained: somatization. If there is to be any sense to the idea that the analyst renders a mental item 'digestible' for the analysand, it must surely be the *very same thing* that is as it were dragged from the category of the perforce only somatizable to the category of the thinkable. So when only somatization *is* possible for the analysand, vomiting – say – or the entertaining of 'inchoate sense-impressions' (Levine 2011: 189) is the inarticulate vehicle of the very same content which, once 'digestion' has happened, is available for expression to the analysand in the form of, as it might be, articulate speech. What changes is not that his mind comes to have an attitude in it where beforehand it had a mere contentless something: on the contrary, on pain of denying that analysts can ever correctly interpret vomitings as the expression of attitudes and by doing so change the role they play in the analysand's psychology – that is, 'make them [amenable] to the normal processes of review and revision that constitute the rational health of belief and other attitudes' (2001: 108; cp. 2008: 7, and 60, 'cognitive isolation'): Moran's description of working through - the change lies not in the content itself but in the resources available to the analysand to express it.^{vii} In enhancing our capacity for symbolic as opposed to concrete thinking, and thereby for communication between the ego and other parts of the self, it is thus reasonable to see post-Kleinian psychoanalysis, as Moran sees psychoanalysis generally, as engaged - in more Aristotelian language – in the expansion of our rational powers.^{viii}

The case is admittedly different, and more complicated, in the case of infants and young children, since presumably part of what makes an attitude ascribable to an adult even when that attitude is very far from being *self*-ascribable by them is that the attitude's articulate expression is in the adult's repertoire (for example, they understand the words used, can form other sentences of comparable complexity and so on), but infants and young children do not have such a repertoire. On the contrary, the analyst's (or caregiver's) re-presenting to them their mental contents in digestible form may constitute an expansion of their repertoire, part of the process by which they progressively acquire more psychic structure. In such cases there is an indeterminacy in the identity of the uncontained state, and a corresponding difficulty in saying what it is for the adult to re-present the infants' or young children's experiences to them truthfully (though we often know falsifying re-presentations of early experience when we see them). Now psychoanalysis has a standing tendency to assimilate the pathological and the immature. Embracing that tendency, the pessimist about the 'arranged marriage' might argue that if very young children's uncontained mental contents aren't attitudes with fully determinate contents, neither are those of adult analysands, thus threatening the reply to the pessimist even for the adult case. In my view this tendency should be resisted, though since I have no space to argue that here, that it *can* be resisted must remain an assumption for the purposes of this paper.^{ix} Supposing that it can, however, even if the case of young children shows there may not be a coincidence between Moran's attitudes and what Winnicott (in the light of the foregoing considerations perhaps incautiously) terms 'instincts', the foregoing considerations show there is a substantial overlap. So the pessimist's very general case for

thinking that the 'arranged marriage' falls at the first hurdle - that there's no type of mental item that falls within the range of application both of Moran's transparency condition and of the post-Kleinian contained/uncontained distinction – can be met. So there is a more than merely *ad hominem* interest in exploring further the relations between the post-Kleinian contained/uncontained distinction and Moran's distinction between 'fully owned' and 'alienated' states of oneself, or between happenings inside the moral self and happenings outside it.

2. I wouldn't have done all this preparatory manoeuvring if I didn't think there was something in the comparison between Moran and the post-Kleinians. Nonetheless I want to spend the rest of this paper arguing that the psychoanalytic distinction between contained and uncontained relates in a more complicated way to what I'll lazily call Moran's inside/outside distinction than these initial moves might lead one to suppose. I hasten to add that it need not be thought of as an objection to Moran if the two distinctions *don't* align perfectly: his evident interest in psychoanalysis is no reason to think he was setting out to theorize *that very* psychoanalytic distinction. Nonetheless to those of us who are attracted by Moran's 'anti-Cartesian' campaign to return the rational mind to the world it thinks about, it would be reassuring if that campaign could be shown to be at least consistent with post-Kleinian psychoanalysis, and if showing that requires us to go beyond the simplicity of the arranged marriage, the attempt to do so would seem to be well worth the effort.

There are two lines of questioning I'd like to pursue. The first concerns a distinction, suggested by the 'containment' terminology, *within* the class of 'outside' or 'alienated' states (this section); the second, a possible complication in the idea that the question what mental states I have becomes a 'purely psychological question about me' only when the states are imperfectly rational.

In Moran's view, one reason to associate the transparency condition – that is, a condition met by any attitude such that I can settle the question whether or not I have it by asking the 'corresponding world-directed question' (2004: 423; 2001: 62, 64, 67), as I settle the question whether I believe that it's going to rain by settling the question, simply, whether it's going to rain – with psychic health is that an attitude that is *not* in this sense transparent 'undermines one's position as a deliberator' (2001: 98).^x That's more or less what I said about split-off or – not that this is the same thing - uncontained attitudes too. I now want to question the strength of the link between transparency and the capacity to deliberate. I'm going to do this in the first instance by taking a more careful look at the contrast between the 'inside' disgust of the principled vegetarian and the 'outside' disgust of the squeamish non-vegetarian which, though it served well enough as a thumbnail introduction to a distinction, turns out to be more complicated. Nothing in the overall argument of this section, however – which simply aims to question the connection between transparency and the capacity to deliberate – depends on the details of this particular case, and if the case is too controversial, a simpler

and less controversial example would do as well. With that said, it's perhaps worth trying to tease out some of the complexity of the disgust case for its own sake.

Let me present a straightforward account of the disgust case first of all, and then explain why this account is *too* straightforward. Disgust for the vegetarian isn't in itself – except perhaps marginally – a self-standing reason for action: on the contrary, the vegetarian's reasons for action (for turning his head away when he passes the butcher's, or whatever it is) are the very same as the reasons for the disgust itself. By contrast, if disgust is – as squeamish disgust is – merely a disagreeable experience, it *is* or at least can be a self-standing reason, for the same reason any disagreeable experience can be; indeed, in so far as the subject has any reasons to turn his head away or to order his meat online, these will be supplied by the disgust itself, not by anything (because there isn't anything) that rationalizes it. But thanks to this very fact, squeamish disgust can be something the subject can strategize about. So suppose he loves the taste of meat and can't stand the sight of it when raw, he can easily think up ways to enjoy the one without suffering the other; that is, he can manage his disposition to experience disgust so it doesn't get in the way of other things he wants (e.g., to eat meat). If he does this, he is thinking about his disposition as attaching costs – in the shape of the experience of disgust – to certain courses of action in such a way as to place them very low in the league table of means to his given ends, much as a river's disposition to flood or a road's disposition to attract traffic at rush hour places them low in the league table of means to getting to work. This rather special way in which one's own dispositions can feature in deliberation helps to make sense of the thought that whereas unalienated dispositions we deliberate *from* – they shape the way the world appears to us – alienated ones we deliberate *about* (they are themselves parts of what appears) (Thomas 2005: 27; Harcourt 1998; Dreyfus 1989).

Now according to the terms of the 'arranged marriage' – that is, on the view of the contained/uncontained distinction that maps it perfectly onto Moran's inside/outside distinction – the contained state in the pair under discussion is the vegetarian's, the uncontained state the squeamish meat-eater's. But though this fits the thought that containment goes with 'rational health' – squeamish disgust is not rationalized by its object, whereas principled disgust is – it is a bit too neat. Indeed it is a bit too neat for two kinds of reason, one which has to do with the special complexity of this case in particular, and one which this case has in common with any simpler case.

To take the special complexity first, it does not seem quite right to say that the non-vegetarian's disgust is *merely* a disagreeable experience, at least if this means that it supplies him with reasons for action simply on the grounds of its unpleasantness. For non-transparent as this disgust may be to any reasons *to avoid eating meat*, it still seems to have – or at least to be capable of having – an intentionality that isn't captured by the 'merely disagreeable' label. To be sure, like most states of disgust (though not all: Plato's 'corpses' case is a famous counterexample (Plato 1961: 439e-440a)) it is unpleasant. But, first of all, there are quite tight conceptual limits to the possible objects of such states (connected most strongly

with smell and taste, more with the wet than the dry, and so on): compare pride, contrast pleasure and displeasure. Secondly, the non-vegetarian's disgust surely bespeaks the (real or apparent) *disgustingness* of its object: though there's no thought *of wrongness* it expresses (contrast the vegetarian), there does seem to be a thought ('that's disgusting') it expresses, so the reason for avoiding the offending sight won't – or at least needn't be – the intrinsic unpleasantness of the experience of disgust, but rather the real or apparent disgustingness of the sight; correspondingly, it looks as if something can be said about the non-vegetarian's disgust similar to what we said about the vegetarian's, namely that his reasons for avoiding the sight are the very same as his reasons for being disgusted.

The lesson to draw from this might be that transparency comes in layers: that there's (as it were) an opaque layer to the non-vegetarian's disgust which is absent from the vegetarian's – the layer that's not transparent to distinctively moral thoughts about its object – but which leaves it transparent to thoughts of a different kind. But which thoughts? Suppose one suddenly had to take a job as a hospital porter and was faced with the unaccustomed sight of body parts. Here there's no space as it were left empty by a rationalizing thought about wrongness, because it's far rarer for hospital porters to think the presence of body parts in hospitals is *permissible* than it is for non-vegetarians to think this about eating meat: outside special contexts like Christian Science, the question about the permissibility of amputations is not so much raised and settled in the affirmative as never raised. Nonetheless the novice porter would – I imagine – still think of his disgust as a reaction to the disgustingness of these things, and not as a merely disagreeable experience. Now the rationalizing thought *might* be that the parts are *disease-ridden*, or something similar, and this shows that there is at least a kind of thought for disgust to be 'transparent' to even where a rationalizing thought about wrongness is – as it is not in the novice porter's case – *acknowledged* as absent. This favours the 'layers of transparency' view of the non-vegetarian's case. However, the rationalizing thought about disease is scarcely a thought that's available in the case of the non-vegetarian: he is after all planning to *eat* the stuff. Nor am I convinced that, even in the novice porter's case, disgust need have a 'separable' rationalizing thought^{xi} in order to be experienced as a reaction to the disgustingness of the body parts rather than as a merely unpleasant experience. 'It's disgusting' is, after all, a thought; whether it is properly speaking a rationalizing thought I am less sure. (Consider also the fact that there's no 'separable' thought that distinctively needs to disappear for the disgust to cease: it might be that one just gets used to seeing body parts.) So the lesson might rather be that in place of a two-way distinction between experiences that are rationalized by their objects and experiences that are merely pleasant or unpleasant we need a three-way one, between experiences that are rationalized by their objects; experiences – like the non-vegetarian's or (in one variation) the novice hospital porter's – which are *experienced as* reactions to the real or apparent features of their objects but perhaps not properly speaking rationalized by them; and merely (un)pleasant experiences.

Now what does it take finally to relegate one's experiences to – in one's own eyes – the category of the 'merely disagreeable'? If the 'layers of transparency' story is really correct about the non-vegetarian's disgust – as I suspect it is not – then clearly the thought that meat-eating isn't wrong will not be enough, because disgust will still be transparent to thoughts about disease etc.. But in the case of the novice hospital porter, the 'layers of transparency' story isn't straightforwardly available. Nor is it available in what one may be inclined to think of as a paradigm case in this area, the case of irrational or indeed phobic fear. With fear, the rationalizing thought is something like dangerousness. But even where the subject seemingly believes the feared thing is *not* dangerous, so the fear is irrational, fear is still experienced – like, *mutatis mutandis*, the non-vegetarian's disgust – as a reaction to the frighteningness of its object, not as a merely unpleasant mental state of one's own. What's needed to convert the phobic's experience of fear or the non-vegetarian's experience of disgust into the experience merely of an intensely unpleasant mental state of theirs is not merely the *absence* of a rationalizing thought (that meat-eating is wrong or that spiders are dangerous), but the *presence* of the thought that his disgust (fear) is squeamish, i.e. doesn't speak for any rationalizing features of its object. Of course we, as commentators on the situation, will apply the word 'squeamish' (or equivalent) as soon as we note the absence of the rationalizing thought, but it's only the capacity to apply that word *to oneself* that shifts the subject's focus from the object ('I know they aren't dangerous but I'm still frightened of them') to his experience of it.

On the whole, if one gets as far as to apply that word to oneself, the experience of disgust fades rather fast: as it is no longer thought of as a reaction even to the real or apparent disgustingness of the thing, it is hard to take it seriously as a reason for anything.^{xii} But of course it may not fade, and it's tempting to say that it's only at *that* point – and not before – that one starts encountering one's disposition to experience it as an 'environmental' feature, like the flooded road. On the straightforward account of disgust we started out with, the non-vegetarian who's nonetheless disgusted by the sight of raw meat encounters his dispositions to disgust as an environmental feature and – like the flooded road – adjusts his behaviour in the light of it by buying ready-meals online. But perhaps that's not a true case of encountering his dispositions as environmental features after all, and he deliberates *from* his disgust at raw meat in just the way that another person deliberates *from* his fear of highwaymen when he avoids the short-cut through the forest. We could, if we insisted, redescribe either as 'strategizing in order to avoid a mental state of one's own', but the *point* to the agent of avoiding the mental state would be supplied by his reasons for avoiding the raw meat (or the highwaymen) namely that they're disgusting (or dangerous). It doesn't matter, of course, whether this or that case ends up being better described as a case of real self-management or not. What matters is that we only have a real case of encountering one's dispositions to react as environmental – of real self-management – when one has gone as far as it is possible to go in detaching one's dispositions to react from *any* evaluatively relevant feature of the situation, as has the man who orders ready-meals online not because he finds

the sight of raw meat disgusting but because he finds his avowedly squeamish disgust at it unpleasant or upsetting. (Only the latter, not the former, would seek hypnosis as a cure for the disposition.)

What I'd like to emphasize – and this is the strategic point of the disgust example, whether or not you buy my various local observations about its complexity - is that the possibility of a 'managerial' attitude to one's psychological dispositions *is itself a substantive psychological achievement*. To be sure, it contrasts unfavourably with the person whose unalienated disposition to disgust is something he deliberates from (whether that's just the case of the vegetarian, or of some non-vegetarians too). But it contrasts favourably with the person who is simply and uncontrollably disposed to be filled – as we can and they cannot see, groundlessly - with disgust. Compare phobic fear: part of what it is for it to be phobic^{xiii} is for it to be unresponsive to thoughts about the harmlessness of its object. But for all that, the phobic's starting point is not self-management but simply terror, which is no less irrational for the fact that the phobic is quite incapable of seeing it as unresponsive to thoughts about harmlessness. If only he could *articulate* the unresponsiveness to these thoughts of his disposition to experience fear, indeed – come to see it as related merely causally to certain sorts of environmental item - he could begin to manage it.

So alongside attitudes that are responsive to thoughts about their objects, we now have a distinction within the class of attitudes that are not so responsive: attitudes that aren't and which, because they aren't and one can understand this about them, can be managed; and attitudes which simply aren't. Which of these three types of attitude are 'contained'? If the mark of containment is 'rational health', the possibility of proper interaction with other attitudes, then only the first counts as contained. That is the answer the 'arranged marriage' – that is, the perfect alignment of the contained/uncontained with the inside/outside distinctions – predicts. But there are reasons to doubt this answer. First, as I've just argued, there's *something* healthier about 'outside' dispositions – dispositions to attitudes that fail the transparency condition – which one can manage which is lacked by 'outside' dispositions which one cannot. Mightn't the norm expressed by the contained/uncontained distinction mark *this* contrast, and therefore classify manageable 'outside' dispositions along with 'rationally healthy' ones, rather than marking the rational health/ill-health contrast itself? Of course simply to introduce a second norm into the discussion is not to offer an argument as to why this norm, rather than rational health, is the better interpretation of the post-Kleinian psychoanalytic notion of containment. In the next section, therefore, I want to set out some further psychoanalytic reflections which suggest a reason why the class of contained states of mind include manageable 'outside' dispositions, and thus why containment doesn't map simply on to Moran's rational health, but captures some of Moran's rational ill-health too.

3. I now want to introduce some empirical work intended to flesh out the psychoanalytic notion of containment, and the psychoanalytic distinction between contained and uncontained states of mind. Work by Peter Fonagy, Mary Target and others explores the idea that by

offering a child^{xiv} words, or non-verbal expression, for its states of mind – which of course requires us to be sufficiently sensitive to them to know which states of mind they are - we render those states less terrifying or unsettling, and so less likely to be dealt with by splitting (Fonagy and Target 1997: 684; Fonagy et al., 2004).^{xv} Though they do not always use the word ‘containment’ – their key terms are ‘reflective function’ and ‘mentalization’ – it’s clear that they see themselves as offering an observationally controlled rounding out of the ideas from Bion, Winnicott and others I have already expounded.^{xvi}

Though it scarcely confirms a theory on its own, the following experiment is a powerful illustration of what Fonagy and Target have in mind. They were interested in the kind of caregiver response which most effectively calms down 8 month-old infants when taken to the doctor for an inoculation (Fonagy and Target 1997: 684; cp. Fonagy et al. 1995). Some caregivers of course don’t respond at all, but for our purposes especially interesting is the difference between caregivers whose responses are – in the jargon – ‘contingent’, i.e. which show in content and timing that they are keeping track of what’s going on, and those whose responses are not only contingent but also ‘marked’ (Gergely and Watson 1996). An example of a contingent but unmarked response would be ‘catching’ the baby’s incipient panic at the sight of the needle and panicking oneself: that certainly shows you are following what’s happening rather than, say, daydreaming. ‘Markedness’, on the other hand, is explained as

the caregiver’s capacity to incorporate into her expression a clear indication that she is not expressing her own feelings, but those of the baby (Bateman and Fonagy 2003: 193).

Marked responses, that is, involve reflecting the infant’s mental state not by duplicating it but by play-acting it. The caregiver ‘combines a “mirror” with a display incompatible with the child’s affect’, e.g. ‘smiling, questioning, mocking display’ (Fonagy and Target 1997: 684). The significance of the distinction between marked and merely contingent responses is that the child is soothed most effectively when the caregiver’s responses are marked, and Fonagy and Target’s thought is that the child is soothed because marked responsiveness on the caregiver’s part enables it to contain or to mentalize its experience. There must be enough in the caregiver’s response that is *foreign* to the state expressed by the baby for the baby to think – possibly falsely, since I don’t see why only adults who aren’t to any degree afraid of needles can be effective ‘containers’^{xvii} - that the caregiver is not herself in that same state, and thus to understand her expression not as the expression of her own state, but as a representation of the baby’s.^{xviii} Otherwise the infant sees not its *own* state in the caregiver’s expression - and so gets no further capacity to mentalize its state than it had before - but merely *that the caregiver herself is in that state* (here, panic), which amplifies the infant’s panic (‘if the mirroring is too accurate, the perception itself becomes a source of fear and it loses its symbolic potential’ (Fonagy and Target 1997: 683)). Again, if the caregiver is herself sufficiently on top of what’s going on to offer a ‘marked’ response - that is, if *she* can ‘contain’ the infant’s experience rather than be overwhelmed by it – she can offer the infant the means of containing *its own* experience (that is, her representation of it, which enables the infant to

represent, and not simply to express without representing, its own state: compare vomiting-as-thinking with saying what one thinks). Without the containment which a marked response makes available, on the other hand, it's as if the panic that the infant might have succeeded in localizing to its mind is now everywhere. In Kleinian terms, this is typical of one moment of the paranoid-schizoid position (the paranoid moment, i.e. what splitting (hence 'schizoid') is supposed to deal with), where the uncontained state that is *in fact* a state of the infant's own mind appears to be a state of the world.^{xix}

This little inoculation scene serves to explain why some states that fall on the wrong side of Moran's norm of rational health fall on the right side of the norm expressed by the distinction between contained and uncontained. To inhabit that paranoid moment is precisely to be in what I described in the previous section as the phobic's or the squeamish non-vegetarian's basic condition – terror, or disgust. Now relative to Moran's ideal of rational health, coming to see terror as merely a mental state of one's own – so one can, for example, strategize in order not to be disabled by it – counts as failure: since it 'persists independently of any supporting reasons' (2001: 67), the question whether I am in it or not becomes 'a purely psychological question about me'. But relative either to continuing to be overwhelmed by it, or – as a defence - splitting it off, the step from its persisting independently of any supporting reasons (the paranoid moment) to my *seeing* it as so persisting (the stance of the self-consciously deficiently rational agent who can become a self-manager) is a big success. But to take that step is part of what it means – perhaps the bare minimum it can mean – to become a container of one's own mental states. In the rational agent, to come to see (e.g.) disgust as merely a mental state of his is often enough to do away with it altogether; in the properly phobic subject that will not happen – but at least terror will be understood for what it is, a stubbornly persisting irrational mental state which one can (perhaps) circumvent in the manner of the flooded road.^{xx} But if this is part of what containment means, then some states – the manageable 'outside' states, or dispositions thereto – fall on the right side of the contained/uncontained distinction but fall on the wrong side of the norm of rational health. So rational health is not after all the theoretical deepening of the psychoanalytic notion of containment, but a more exigent norm.

4. So far I have done no more than try to show that the contained includes more than Moran's 'transparent' states of mind. This is consistent with the idea that the rationally healthy states, still understood as coextensive with the 'transparent' ones, are a proper subset of the contained. Now, however, I want to challenge Moran's assumption that transparency defines rational health, by considering the possibility that there are transparent states that are not rationally healthy, and indeed that the difference between those that are and those that aren't is precisely containment. To make this challenge, I am going to consider two questions and try to relate them. One is raised by the 'paranoid moment' (as described in the previous section), the other by knowledge of one's own mind that is possessed by others.

Consider to begin with the well-known ‘false belief test’ (Wimmer and Perner 1983; Baron-Cohen et al. 1985; Perner et al. 1987; Gopnik *et al.*, 2001). A child is shown a miniature room with puppets who enter and exit. The puppet ‘sees’ a Smartie (M&M) being put into its distinctive Smarties tube. Then the puppet goes out and, while it’s out, the Smartie is removed and replaced with coloured pencils. The puppet re-enters and the child is asked what the puppet believes is in the tube. The less mature child says ‘pencils’; the more mature child – I think the turning point is usually around 4 years old – says ‘Smarties’.

It is sometimes said that the test is designed to get at whether test subjects do or do not have a ‘theory of mind’. For example, Baron-Cohen *et al.* hypothesize that lack of a ‘theory of mind’ in the sense of the ability ‘to impute beliefs to others’, or ‘to conceive of mental states; that is, knowing that other people know, want, feel or believe things’ (1985: 38) would explain a deficit in respect of social skills. Since such a deficit is a signature of autistic children, Baron-Cohen *et al.*’s claim that autistic children’s difficulties with the test would support their hypothesis implies that what these authors see the test as a test *of* is precisely a child’s possession or non-possession of a ‘theory of mind’ in the sense explained. However, this understanding of the test does not accord with Wimmer and Perner’s original paper. Recognizing the importance of a ‘theory of mind’ in the sense explained, they claim that ‘the majority of children [even as young as 2.5 years] spontaneously use a substantial vocabulary about perception, volition, major emotions and knowledge ... for self and others’ (1983: 105): in other words both the children who pass the test and the substantially younger children who fail it have a ‘theory of mind’. What the test is designed to elicit is not that, but rather ‘a more complicated metarepresentational problem [which] arises when one has to explicitly represent the *difference* between one’s own and someone else’s relation to the *same* propositional content’ (1983: 105); the three-year-olds who don’t pass the test ‘are not yet capable of representing alternative models of the world, one describing the true story world and another, incompatible one, describing the protagonist’s [sc. the puppet’s] false belief about it’ (Perner et al. 1987: 125). Indeed, notwithstanding their (to my mind misleading) headline description of what the test tests for, Baron-Cohen *et al.* say – in line with Wimmer and Perner – that ‘in order to succeed on the task, the child has to be aware *that different people can have different beliefs about a situation*’ (1985: 38, my italics).^{xxi} Properly speaking, then, the conclusion to draw from the test is that the immature child has not yet got hold of the idea of multiple perspectives on a single reality, because it doesn’t occur to him that saying how things appear to him doesn’t also give the way they appear to the puppet.

I mention this test because there is an analogy between the less mature child in the test and the ‘paranoid moment’ I’ve already touched on, or – in Fonagy and Target’s terminology – the ‘psychic equivalence mode’, which like the paranoid-schizoid position in some Kleinian thinking they conceive as both characteristic of an early stage of development and as a response to later trauma.^{xxii} (Whether the paranoid-schizoid position, or the psychic equivalence mode, in later life is really always a matter of regression to a normal infantile state should not affect the use I want to make of the analogy.) In this mode, the trauma victim,

for example, seems to lack an awareness that their own 'thoughts, feelings and wishes are *part of the mind*' so, like the three-year-old who is 'not yet able to conceive of [what's in fact] internal experience as mental', he 'assumes that his thoughts have a reality equivalent to that of objects in the physical world' (Fonagy and Target 2008: 19-20; cp. Jurist, 2005: 427). Manifestations of this mode in adults include an 'omnipotence of subjectivity', a 'you can't tell me anything' attitude which leaves no room for alternative perspectives because, as there's no gap between what the world contains and what's already in one's mind, there can be nothing to tell; an inability to understand that one's own thoughts are not immediately evident to the analyst, but need to be communicated (Fonagy 1989: 101); an incapacity to think about traumatic experience because allowing it into the mind involves re-living it rather than remembering it (i.e. what is in fact a memory, a mental content, is experienced as an event happening in the present);^{xxiii} and 'concrete thinking', that is, a mode of thought in which words, mental imagery etc. are felt to be the things they stand for.^{xxiv} Indeed Fonagy and his collaborators explicitly compare the psychic equivalence mode to the mode of thought of the immature children who fail the false belief test. 'In the psychic equivalence mode, the very young child equates the internal world with the external world: the world *is* how the mind represents it, alternative perspectives cannot be generated. ... If they *think* there is a monster in the closet, a monster *is* in the closet' (Allen *et al.* 2008: 90). By contrast the child who exits psychic equivalence mode in learning to 'mentalize' becomes capable of 'adopting multiple perspectives on the same interpersonal situation ... [as] Perner described' (*ibid.*; my italics).

In Moran's terminology, the less mature child in the false belief test addresses the question 'what does the puppet believe is in the tube?' by asking the world-directed question - 'what's in the tube?' – which *would* be adequate to settle the 'corresponding mind-directed question' about the puppet. What this child doesn't have, and which the more mature one does, is a conception of a distinction between, as we might put it, mind-directed questions that are, and those that are not, transparent to questions about the world. The mind-directed questions that are not so transparent are – as we can see - questions about minds that are not his own, but until he's got hold of 'mind' as a concept that's multiply instantiated, world-directed questions simply settle questions about *mind* (about, as far as he's concerned, the only one there is, though he is not in a position to put it like that, and if he were, he wouldn't).^{xxv}

'Psychic equivalence' thus raises a question about the relation between rational health and the transparency condition. On Moran's view, the latter defines the former. But now consider three types of case where transparency is present and rational health surely absent. First of all, it looks as if the transparency condition is met by the immature child in the false belief test: he settles questions about belief by asking world-directed questions. Secondly, the transparency condition is met by the adult in the 'paranoid moment', and – according to Fonagy and Target – by the infant unaided by 'marked' responsiveness from a caregiver: although an infant is not in a position to ask any 'questions', it is for him as if what is in fact only a content of his own mind is a condition of the world around him, so *if* he asked himself a

world-directed question, he would – though he couldn't see why – get the right answer to a question as to what's in his mind. ('What's the world full of around here?' 'Terror.')

That answer *in fact* gives information about the paranoid adult's or the pre-mentalizing child's state of mind, but it's precisely characteristic of psychic equivalence mode that one cannot appreciate that one's state of mind may be no more than just that, a state of oneself. And, finally, it looks as if the transparency condition is met by Fonagy's patients with borderline personality disorder (Fonagy 1989: 101): they don't understand that if they want Fonagy to know what is in their minds, they need to *tell* him (because they precisely don't grasp that their minds are opaque to others in a way they are not to themselves).^{xxvi} However, the psychic equivalence mode – whether it results from immaturity or from trauma or from some other cause – surely falls short of 'rational health'. So transparency cannot define rational health.

If these rationally unhealthy states meet the transparency condition, what is it that they or their possessors *don't* have? I'd like to say *localization*: the capacity to localize what one discovers *by* asking world-directed questions to *one's own* mind; or the capacity to recognize that, because it's not just any old mind such that questions as to its contents can be settled in the world-directed way, one's own mind is just one among others.

But localization in this sense is at least brought about by *containment by another*: it's the presence of the containing other in the inoculation case that enables the infant to see that it is not the whole world that is full of terror, but only him, and in seeing that, to become less terrified. But this takes place because the containing other offers the infant something – the 'marked' response – that he can see as a representation of what's in his mind. So when the infant comes to make use of this representation himself, he understands that what it represents is *in* his mind.

5. Let me now set on one side the 'paranoid moment', or psychic equivalence mode, and argue somewhat more slowly for the connection between localization and containment, that thoughts and feelings we can 'localize' are those we can contain.

First of all there's at least one difference between the inoculation case and the false belief test. In the latter case, the step forward involved in localization – the becoming available of the conception that an attitude is local to a mind rather than a general condition of the world – is taken by (as it were) my ascribing to myself and to another (the puppet) incompatible attitudes to the same content: that's the step that ushers in the idea that the concept 'mind' is multiply instantiated. In the former case, the step forward in localization is taken first of all by (as it were) another's ascribing a state to me (the caregiver to the infant), and then the infant (or somewhat older child) ascribing it to himself. In this difference there is, I think, an important clue. The latter case tells us transparency may be insufficient for rational health; the former gives us an idealized picture of how what's missing – what I'm calling localization – gets added in. This brings me to the second theme of this section, others' knowledge of our own mental states.

Central to *Authority and Estrangement* is the difference between the way others know us – by observation, by interpretation, by gathering evidence in various ways – and the way in which, when all is rationally as it should be, we know ourselves (i.e., by asking the corresponding world-directed questions). Even though *some* of the knowledge even the most mentally healthy of us has of him- or herself is ‘based on the same kinds of consideration available to others’ (2001: xxx), self-knowledge in any remotely mentally healthy person couldn’t – Moran argues - *only* be ‘mind-reading applied to oneself’, since there would be something seriously lacking in someone *most* of whose knowledge about himself was ‘interpretive and theoretically grounded’ but such that he could not avow it – that is (minimally), express it *without* relying on evidence or interpretation (2001: 90, 89).

All this is surely right. I am more doubtful, however, about Moran’s continuing emphasis on the *difficulty* of converting knowledge of oneself gathered on the basis of such considerations – as it might but need not be, by a psychoanalyst – into the stuff of avowal (2001: 89, on ‘the aim of psychoanalysis’). To be sure, there *are* cases where one accepts another’s interpretation ‘merely intellectually’, and the inability to move beyond that to avowal of the same content is a mark of ‘resistance’, and requires ‘working through’; compare Ron Britton’s patients who ‘regard the unacceptable facts of their daily life as “only theoretically true”’ (Moran 2001: 89; Levine 2011: 207; Britton 1998: 60). However, by far the more usual case is surely of an *effortless traffic* between evidence-based (or interpretative, or whatever) mental-state ascriptions to myself by others on the one hand and, on the other, avowal – *not* mere intellectual acceptance^{xxvii} - by myself of those same states. This is exactly what seems to be happening in the inoculation case:^{xxviii} to offer a sufficiently ‘marked’ response that the infant can understand it as a *representation* of its mental state *is* a piece of (possibly non-verbal) other-ascription by the caregiver, but it has to be effortlessly converted into an avowal by the infant – ‘Does it hurt?’ ‘It hurts!’ - for the marked response to work its magic, that is, for it to bring it about that the infant ‘mentalizes’ its own state and calms down. The same could be said of any stage of life in which we are helped to contain our own feelings, reactions, attitudes (etc.) by the presence of an appropriately endowed other: here ‘working through’ is surely not the norm.

Now the link between localization – between what’s missing where rationally unhealthy states nonetheless meet the transparency condition – and others’ knowledge of our minds is just this, that this effortless traffic is a condition of effective localization. What helps me to understand that an attitude is local merely to my own mind, and not a condition of the world or of mind in general, is *in the first place* that another ascribes the attitude to me – another, that is, more accomplished than I and who can therefore already see that the only bit of the world containing terror or disgust (or whatever) is my mind – and *then* that I have the ability for the most part effortlessly to convert that (observation- or interpretation-based) other-ascription *to* me into an (evidenceless) avowal *about* me.^{xxix}

But now consider this passage from Margot Waddell:

Fortunately [for the infant], there is a 'thinking' ... mother available who is able to take in the projections [i.e. his anger, panic etc.], to resist being overwhelmed by them, to render them manageable and, in a sense, *to hand back to him* a quality of experience which makes him feel divested of terror (Waddell 1998: 30).

'Localization' is just the 'containment' of post-Kleinian theory; Waddell's 'handing back' is what I have been calling 'effortless traffic'. I get to be a container of my own states of mind because I can effortlessly convert into avowals ascriptions to me by others who, happily, are competent in just the way that the less mature child in the false belief test case is not.

In Moran's terms, then, these more accomplished others can already see world-directed questions as limited in the range of minds about which they can deliver answers. For them, questions about my mind are indeed 'purely psychological'; nonetheless I become a self-knower by (on the whole) effortlessly taking on the 'purely psychological' knowledge more accomplished others have of me. So here, I suggest, is another reason for thinking the 'arranged marriage' between Moran and post-Kleinian psychoanalysis makes things too simple. For all that I need to be able to settle questions about my own mind by settling the relevant world-directed questions, I can also do so by avowing without more ado the contents which others arrive at when they answer 'merely psychological' questions about me. It's for this reason that I want to suggest not only (as I did in sec. 3) that because some contained states are non-transparent, containment is not identical to rational health as Moran understands it, but also that because some states are transparent but *uncontained*, transparency together with containment constitutes a more complete picture of rational health than transparency alone.

References

- Allen, J.F., P. Fonagy and A. Bateman (2008). *Mentalizing in Clinical Practice*. Washington DC/London: American Psychiatric Publishing Inc..
- Anscombe, G.E.M. (1981). On Sensations of Position. In Anscombe, *Collected Philosophical Papers vol. II: Metaphysics and the Philosophy of Mind* (pp. 71-4). Oxford: Blackwell.
- Baron-Cohen, S. et al. (1985). Do Autistic Children have a 'Theory of Mind'?. *Cognition* 21(1), 37-46.
- Bateman, A.W., and P. Fonagy (2003). The development of an attachment-based treatment program for borderline personality disorder. *Bulletin of the Menninger Clinic* 67(3), 187-211.
- Bion, W.R. (1984a). 'Attacks on Linking'. In Bion, *Second Thoughts* (pp. 93-109). London: Maresfield Reprints.
- (1984b). A Theory of Thinking. In Bion, *Second Thoughts* (pp. 110-119). London: Maresfield Reprints.
- (1984c). *Learning from Experience*. London: Maresfield Reprints.
- Britton, Ron (1998). The suspension of belief and the 'as-if' syndrome. In Britton, *Belief and Imagination* (pp. 59-68). Hove: Routledge.
- Dreyfus, H. (1989). Alternative Philosophical Conceptualizations of Psychopathology. In H.A. Durfee and D.F.T. Rodier (Eds.), *Phenomenology and Beyond: The Self and Its Language* (pp. 41-50). Dordrecht: Kluwer.
- Fonagy, P. (1989). On Tolerating Mental States: Theory of Mind in Borderline Personality. *Bulletin of the Anna Freud Centre* 12(2), 91-115.
- Fonagy, P. et al. (1995). Attachment, the Reflective Self, and Borderline States. In S.

- Goldberg, R. Muir and J. Kerr (Eds.), *Attachment Theory: Social, Developmental and Clinical Perspectives* (pp. 233-278). New York: Analytic Press.
- Fonagy, P., et al. (2004). *Mentalization, Affect-Regulation and the Self*. London/New York, Karnac.
- Fonagy, P. and M. Target (1996). Playing with reality I: Theory of mind and the normal development of psychic reality. *Int. J. of Psycho-Anal.* 77(2), 217-234.
- (1997). Attachment and reflective function: their role in self-organization. *Development and Psychopathology* 9, 679-700.
- (2000). Playing with Reality III: The persistence of dual psychic reality in borderline patients. *Int. J. Psycho-Anal.* 81(5), 853-73.
- (2008). Attachment, Trauma and Psychoanalysis. In E. Jurist, A. Slade and S. Bergner (Eds.), *Mind to Mind: Infant Research, Neuroscience and Psychoanalysis* (pp. 15-49). New York: Other Press.
- Frankfurt, Harry (1998). Identification and Externality. In Frankfurt, *The Importance of What We Care About* (pp. 59-68). Cambridge: Cambridge University Press.
- Gergely G., and J. Watson (1996). The Social Biofeedback Model of Parental Affect-Mirroring. *Int. J. Psycho-Anal.* 77(6), 1181-1212.
- Goldie, Peter (2012). *The Mess Inside*. Oxford: Oxford University Press.
- Gopnik A. et al. (2001). *How Babies Think*. London: Phoenix.
- Harcourt, Edward (1998). Mill's 'Sanctions', Internalization and the Self. *Eur. J. Phil.* 6(3), 318-334.
- Jurist, E.L. (2005). Mentalized Affectivity. *Psychoanalytic Psychology* 22, 426-44.
- Klein, Melanie (1988). *Envy and Gratitude*. London: Virago.
- Lear, Jonathan (1990). *Love and Its Place in Nature*. London: Faber.
- Levine, Howard (2011). 'The Consolation which is drawn from Truth': the analysis of a patient unable to suffer experience. In C. Mawson (ed.), *Bion Today* (pp. 188-211). London/New York: Routledge.
- Loewald, Hans (1978). *Psychoanalysis and the History of the Individual*. New Haven/London, Yale UP.
- MacIntyre, A. (2004). *The Unconscious* (revised edn.). London/New York; Routledge.
- Meins, E. et al. (1998). Security of attachment as a predictor of symbolic and mentalizing abilities. *Social Development* 7, 1-24.
- Mitchell, Stephen (1988). *Relational Concepts in Psychoanalysis*. Cambridge MA: Harvard University Press.
- Moran, Richard (2001). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- (2002). Frankfurt on Identification: Ambiguities of Activity in Mental Life. In S. Buss and L. Overton (Eds.), *Contours of Agency* (pp. 189-217). Cambridge MA/London: MIT Press.
- (2004). *Précis of Authority and Estrangement: An Essay on Self-Knowledge. Philosophy and Phenomenological Research* 69(2) 423-426.
- O'Shaughnessy, Edna (1988). Explanatory Notes. In M. Klein, *Envy and Gratitude* (pp. 324-36). London: Virago.
- Perner J. et al. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *Brit. J. Dev. Psych.* 5(2), 125-37.
- Plato (1961). *Republic*. In E. Hamilton and H. Cairns (Eds.), *Plato: The Collected Dialogues*. Bollingen Foundation. Princeton, N.J.: Princeton University Press.
- Rayner, Eric (1990.). *The Independent Mind in British Psychoanalysis*. London: Free Association.
- Segal, H. (1957). Notes on symbol formation. *International Journal of Psychoanalysis* 38, 391-397.
- Shuttleworth, J. (1989). Psychoanalytic Theory and Infant Development. In L. Miller et al. (Eds.), *Closely Observed Infants* (pp. 22-51). London: Duckworth.
- Thomas, Alan (2005). Reasonable Partiality and the Agent's Personal Point of View. *Ethical Theory and Moral Practice* 8(1-2), 25-43.
- Waddell, Margot (1998). *Inside Lives*. London: Duckworth.
- Williams, Bernard (1977). A Critique of Utilitarianism. In J.J.C. Smart and B. Williams, *Utilitarianism: For and Against* (pp. 75-150). Cambridge: Cambridge University Press.
- Wimmer, H. and J. Perner (1983). Beliefs about Beliefs: Representation and Constraining function of wrong beliefs in younger children's understanding of deception. *Cognition* 13, 103-28.

- Winnicott, D.W. (1977). Primary maternal preoccupation. In Winnicott, *Through Paediatrics to Psychoanalysis* (pp. 300-305). London: The Hogarth Press.
- (1990a). Ego Integration in Child Development. In Winnicott, *The Maturational Processes and the Facilitating Environment* (pp. 56-63). London: Karnac.
- . (1990b). The Development of the Capacity for Concern. In Winnicott, *The Maturational Processes and the Facilitating Environment* (pp. 73-82). London: Karnac.
- (1990c). Ego distortion in terms of true and false self. In Winnicott, *The Maturational Processes and the Facilitating Environment* (pp. 140-52). London: Karnac.
- Wittgenstein, L. (1953). *Philosophical Investigations*, tr. G.E.M. Anscombe. Oxford: Blackwell.

ⁱ The same is true, by the way, of Harry Frankfurt's internal/external distinction in, for example, 'Identification and Externality', in Frankfurt (1998).

ⁱⁱ For the idea of splitting in Klein, see for example Klein (1988: passim), and O'Shaughnessy (1998: 324-6). The psychopathology of splitting is also one of the *idées maîtres* of 'object relations' psychoanalysis: see Rayner (1990: 22), who attributes the origin of the idea to Fairbairn, and says (rightly or wrongly) that it was then taken up without acknowledgement by Klein.

ⁱⁱⁱ Cp. also 2001: xxix-xxx, 7, 35, 85, 108 ('fixation'), 137.

^{iv} Conversely, as Levine argues (looking back in turn to Bion), if psychic health *depends* in part on 'creating structure within the mind by separating conscious from unconscious thoughts' (2011: 189), it can hardly be that in consigning some states to the unconscious we are splitting them off. I have been helped here also by Dreyfus (1989).

^v And indeed between the id itself and the unconscious: what Winnicott refers to as the (infant) id, i.e. the place of infant 'instincts', Bion calls 'the rudimentary infant consciousness' which is 'not associated with an unconscious' (1984b: 116).

^{vi} Bion's term for its range of application – 'beta-elements' (1984b: 117) – is I think deliberately uninformative.

^{vii} There is room here also for a third line of reply. In work that postdates *Authority and Estrangement*, Moran explores the possibility that the transparency condition also applies to pleasure, in other words to contents that are not purely attitude-like (2002: 207-15). But I don't have the space to assess that extension of the notion here.

^{viii} In particular perhaps our metacognitive powers, but these appear to be related to our powers of self-monitoring and affect-regulation: see for example Fonagy and Target (1997; 1996). For the importance of symbolism in communication between parts of ourselves, for symbolic vs. concrete thinking, and also for the oblique relationship of that distinction to the distinction between conscious and unconscious, see for example Segal (1957). For commonalities between developmental trajectories in Aristotelian and post-Kleinian thinking, see Waddell (1998: 3): psychoanalysis is concerned (she says) with 'development, the moral and emotional growth of the self, the character'.

^{ix} For criticism of the tendency, see Mitchell (1988), esp. chs. 5 and 6.

^x Cp. also pp. 7, transparency is 'crucial to the conduct of life', while lacking it is 'disturbing and debilitating'; 35 'the health of the person'; 60 'mutual responsiveness' of our mental states to our beliefs about their objects 'is just what psychological health should involve'.

^{xi} By 'separable' – a term I borrow from Anscombe (1981) – I mean that the rationalizing thought is expressed in words distinct from those used to characterize the reaction itself, as with 'disgust' and 'disease-ridden' but not as with 'disgust' and 'disgusting'.

^{xii} Compare Mark Zuckerberg who, unsure whether his disgust was squeamish or not, set himself the test of slaughtering animals himself, and discovered that it wasn't (because his disposition to experience disgust faded quite quickly): <http://www.guardian.co.uk/environment/green-living-blog/2011/may/27/mark-zuckerberg-kill-animals-meat>.

^{xiii} I say 'part of', since various fears we wouldn't describe as phobic are unresponsive to thoughts about the harmlessness of their objects, but are occurrently rather than stably so, don't exhibit any pattern in the nature of their causes, etc..

^{xiv} The same could be said of adults too, but Fonagy and Target's work on this has been mostly with children.

^{xv} By the way, excitement too can be unsettling and need containment – the need is not confined to bad feelings like fear and sadness.

^{xvi} They say, for example, that their formulation of maternal sensitivity 'has much in common with the British psychoanalyst Wilfred Bion's notion of the role of the mother's capacity to mentally "contain" the affect state intolerable for the baby' (1997: 686; cp. 2000: 855).

^{xvii} Yet another aspect of this story is that the felt need to be a container for another may help one to be a self-container. I am grateful to Margaret Rustin for this point.

^{xviii} Is that difference really just the difference between the caregiver containing the state, and her simply having it without containing it? It could be, in that panic is often 'uncontained' by the person who catches it, and Fonagy and Target certainly suppose it to be; but maybe it needn't be – e.g. if one catches fear, there would seem to be an intermediate state between failure to contain it and the ability to represent it

to another. But maybe what's critical is that the caregiver contains the infant's state – its 'projections' – not just her own reaction to the same object.

^{xix} The comparison between the pathological use of projective identification (Klein) and the reaction of the baby to merely contingent but unmarked mirroring is made by Gergely and Watson (1996: 1202-3): 'he/she will experience is/her on negative affect 'out there'.

^{xx} And in the case of the immature, presumably coming to see terror as a state of their own minds won't have the two possibilities – *merely* such a state vs. a state rationalized by its object – that it has in the case of an adult.

^{xxi} Similarly, Gopnik *et al.* introduce the topic of the false belief test saying that it tests for possession or non-possession of concepts of mental states that are 'representational'. That sounds like it tests for a 'theory of mind'. But quite apart from the fact that desire and emotion are (on any ordinary understanding of the word) representational and they agree that even the three-year-olds who fail the test possess *those* concepts, it is clear from the subsequent discussion that what's really of interest to Gopnik and her co-authors is grasp of whether 'people ... have different beliefs about the same thing' (2001: 44-5), i.e. grasp of the possibility of multiple perspectives.

^{xxii} '[In trauma] modes of representing the internal world *reemerge that developmentally precede* the awareness that thoughts, feelings and wishes are part of the mind' (Fonagy and Target 2008: 19, my italics). Unsurprisingly perhaps, it has been argued that the patterns of nurture like 'marked' responsiveness that foster 'mentalization' also appear to foster precocious good performance on the false belief test (Meins *et al.*, 1998).

^{xxiii} Compare Peter Goldie's discussion of the memories of PTSD sufferers, who either relive the traumatic experience in 'field memory' or affectlessly report it in 'observer memory' (2012: 70-1).

^{xxiv} An oft-cited example is Hanna Segal's schizophrenic patient who, when asked why he had stopped playing the violin since he became ill, 'replied with some violence "Why? Do you expect me to masturbate in public?"'. The point is that to him 'fiddling' was not a word for, but *was* masturbating (Segal, 1986: 49). Fonagy and Target (2008: 20) note the overlap between their 'psychic equivalence mode' and Segal's 'symbolic equation'.

^{xxv} Cp. Gopnik *et al.* 2001: 47: 'It's as if [three-year-olds] have a single theory about the mind, which they apply both to themselves and to others'.

^{xxvi} Cp. Gopnik *et al.* on three-year-olds: 'It's as if the children think that since there is only one world out there, a single reality, everyone will understand it the same way' (2001: 45).

^{xxvii} I wonder how much sense there really is to the idea of 'mere intellectual acceptance' of first-person contents which others ascribe to me on the basis of evidence, interpretation, and so on, at least where this is understood as a step on the way to (but which nonetheless falls short of) avowal of the same contents. Let's distinguish two cases, neither of which seems to me to fit this description. First, there's a good deal of knowledge of myself (for example of my early childhood (as it might be 'he never settled down to feed properly'), or of my physiology ('he has a large cyst on one kidney')) which I have acquired on the basis of others' observations and which I have now confidently incorporated into my autobiography. None of the hesitation involved in 'mere intellectual acceptance' here, then; but unlike first-person memory or experience reports, self-ascriptions of these contents remain answerable to the same kinds of evidence which formed the basis of others' ascriptions of them to me, so they're no more *avowed* than 'it's raining outside'. Because they are not contents to which I *could* have evidenceless access, it seems wrong to describe my acceptance of them as 'merely intellectual': mere in comparison to what? The second case I have in mind is someone who – like Britton's patients (1998: 60) – tells their analyst that some 'unacceptable facts' about themselves are only 'true in theory'. Much of the time at least, I suspect that what's going on here is *not* that they have reached a staging post on the way to avowal, i.e. that on the basis of considerations presented by the analyst, they accept the truth of a proposition about themselves but cannot yet avow it (as it were, avowal minus); but rather the more complex phenomenon (as it were, avowal plus) of avowal combined with denial that they are avowing, an avowal *presented by them* as something they can only acknowledge as 'true in theory'. (Compare puzzles about self-deception: if they hadn't already in some sense made this knowledge their own, why would they be so busy keeping it at arm's length?) This seems an especially attractive account in cases where it's not clear what the evidence for the knowledge-claim might be. Compare the affectless memories of PTSD sufferers (Goldie 2012: 70-1): since there are often no observations by others from which these memories derive, it's pretty clear that the *source* of the reports is (as it would be in avowal) their own past experience, with affectlessness itself playing the role of denial-that-one-is-avowing. These two cases do not show there is no such thing as 'mere intellectual acceptance' (short of avowal) of propositions about oneself, but they do perhaps draw attention to how strange the 'ideal symptomatic stance' would be if we came across it.

^{xxviii} It is also, by the way, what happens in the miniature interaction in Wittgenstein (1953: §244): '[W]ords are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences'.

^{xxix} This effortless traffic also has to be the norm, it seems to me, if the pitfall (of which Moran is well aware (2001: xxxi)) attending any account of mental ascription which makes first- and second- or third-person ascriptions answerable to different sources of evidence (including, on a more or less

Wittgensteinian view of the first person, *no evidence*) is to be avoided, namely that the senses of the relevant predicates fragment into distinct predicates ascribable to self and to others respectively.