

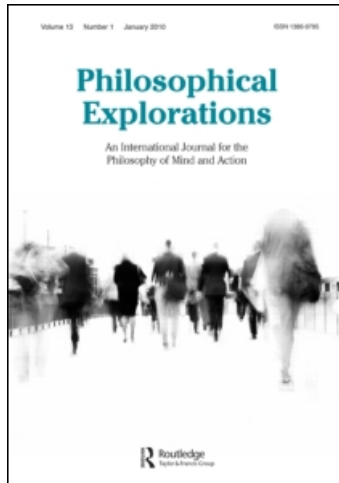
This article was downloaded by: [Swets Content Distribution]

On: 22 April 2011

Access details: Access Details: [subscription number 936475023]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Philosophical Explorations

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713706422>

### Intention and the authority of avowals

Andy Hamilton<sup>a</sup>

<sup>a</sup> Philosophy, Durham University, Durham, UK

**To cite this Article** Hamilton, Andy(2008) 'Intention and the authority of avowals', *Philosophical Explorations*, 11: 1, 23 — 37

**To link to this Article:** DOI: 10.1080/13869790701834599

**URL:** <http://dx.doi.org/10.1080/13869790701834599>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Intention and the authority of avowals

Andy Hamilton\*

*Philosophy, Durham University, Durham, UK*

There is a common assumption that intention is a complex behavioural disposition, or a motivational state underlying such a disposition. Associated with this position is the apparently commonsense view that an avowal of intention is a direct report of an inner motivational state, and indirectly an expression of a belief that it is likely that one will A. A central claim of this article is that the dispositional or motivational model is mistaken since it cannot acknowledge either the *future-direction of intention* or the *authority of avowals of intention*. I argue that avowals of intention – first-person, present-tense ascriptions – express direct knowledge of a future action, knowledge that is not based on examination of one's present introspectible states or dispositions. Such avowals concern a future action, not a present state or disposition; just as self-ascriptions of belief concern the outer not the inner, so self-ascriptions of intention concern the future outer, not the present inner. One way of capturing this future-direction is to say that avowals of intention – and perhaps sense intentions themselves – are a kind of prediction, and not a description of one's present state of mind. This position is suggested by Anscombe in her monograph *Intention* (1963), and treats avowals of intention as judgements about the future, which unlike ordinary predictions are not based on evidence. However, since talk of prediction everywhere suggests an evidence-based stance – that meaningful hypothesis about the likely occurrence of events is being proposed, an hypothesis that can be falsified by evidence – the description *future-outer thesis* is preferred. I defend this thesis against various objections, arguing that it complements Anscombe's characterisation of intentions as based on reasons.

**Keywords:** agency; akrasia; Anscombe, G.E.M.; disposition; epistemic authority; first person authority; intention; self-knowledge; weakness of will

## 1. Authority and decision principles

There is a common assumption that intention is a complex behavioural disposition, or a motivational state underlying such a disposition. Associated with this position is the apparently commonsense view that an avowal of intention is a direct report of an inner motivational state, and indirectly an expression of a belief that it is likely that one will A. A central claim of this article is that the dispositional or motivational model is mistaken since it cannot acknowledge both the *future-direction of intention*, and the *authority of avowals of intention*. On the traditional view, the forward-looking nature of intention consists in the fact that it is a motivational state, and that it involves having a belief about the future – or alternatively that its propositional content concerns the future. A dispositional analysis of intention attempts to capture future-direction by treating intention as a motivational state, one that involves having a belief about the future.<sup>1</sup> I argue in contrast that intentions are

---

\*Email: [a.j.hamilton@dur.ac.uk](mailto:a.j.hamilton@dur.ac.uk)

the products of motivational states, and are not motivational states in themselves. Intention is not a further motivating factor in addition to the agent's desires and means-end beliefs; rather it is the result of practical deliberation concerning those desires and beliefs. Intention is a commitment to act, not a motivation to act, rather as belief is a commitment to the truth of the proposition believed.<sup>2</sup> The future-direction of intention implies that just as avowals of belief concern the outer not the inner, so avowals of intention concern the future outer, not the present inner. Intentions do not merely imply a prediction of future action – they are judgements about the future which are wholly directed towards future action. Avowals of intention – first-person, present-tense ascriptions – express direct knowledge of a future action, knowledge that is not based on examination of one's present introspectible states or dispositions. Such avowals concern a future action, not a present state or disposition. When we ask someone what they intend, we want to know what they are likely to do, not what is going on inside their head.

The received view is that an avowal of intention at most *implies* a prediction of future action – that is, the subject must believe that they have a good chance of performing the action. My more radical proposal is that avowals of intention – and perhaps intentions themselves – constitute a kind of prediction, that is, a judgement about the future and not a description of one's present state of mind, though unlike ordinary predictions it is not based on evidence. The prediction thesis is an analogue of the thesis that expressions of belief function as assertions of the fact believed, which I have defended elsewhere.<sup>3</sup> The position is suggested by Anscombe in her monograph *Intention*, no doubt influenced by various remarks of Wittgenstein – notably 'When people talk about the possibility of foreknowledge of the future they always forget the fact of the prediction of one's own voluntary movements', and his reference to 'the fact that one can predict one's *own* future action by an expression of intention', and to 'the prediction contained in [it]'.<sup>4</sup> However, talk of prediction is everywhere suggestive of an evidence-based stance – that hypothesis about the likely occurrence of events are being proposed, hypothesis that can be falsified by evidence. So the clumsier description *future-outer thesis of intention* is preferable. This thesis faces the objection that it fails to recognise that intentions and beliefs exhibit opposed direction of fit with the world, and a central problem of this article is to justify the claim of future-direction whilst respecting this opposition.

I will begin by examining the *authority principle*. The authority principle for belief says that I can be wrong in my belief, but not wrong about my belief. The claim defended here is analogous: that while I can fail *in* my intention – that is, fail to carry it out – I cannot be wrong *about* my intention. Avowals of intention exhibit the same authority principle as avowals of belief:

If X truthfully, attentively, comprehendingly and non-akratically avows 'I intend to A', then the truth of 'X intends to A' is guaranteed (where 'A' is an action, and the sense of 'X' and 'A' is constant, appropriate indexical changes being made).<sup>5</sup>

The limited exceptions to the authority principle – post-hypnotic suggestion is the only convincing counter-example – involve some kind of defect in the subject. It should be noted that the formulation 'truthfully avows' is not tautological, since clearly it is not the case that all truthful utterances express a truth.

An agent's avowal may be truthful – that is, it is correct to say that they intend to perform the action – even when they subsequently fail to do so, or do not perform it intentionally. There are a number of circumstances under which this is so: (1) they were prevented; (2) they changed their mind or forgot their intention, or the intention became redundant because the action was performed by someone else or the intended outcome came about without the agent's intervention; (3) they did not recognise that the time or circumstances for performing the action had arrived;

(4) they were incompetent or inattentive and so, although they were not prevented from carrying out the action, its consequence was not as intended (this condition may however undermine the ascription of the intention); or (5) they suffered from weakness of will. Note also (6): if the agent performs the action unintentionally, in their sleep or by accident, it may also be the case that their avowal was truthful. This list may not be exhaustive. But if the agent fails to perform the action, none of these conditions applies, and no other comparable condition is suggested, then the avowal is taken not to be truthful.

The list may be open-ended, but the authority principle asserts that there is one condition which it cannot include. The avowal could not turn out truthful, but false, because the agent was simply mistaken about their intention. That is, the fact that the agent did not perform the action could never show that they had made a mistake about their intention. It is worth elaborating on this claim because the essential thought underlying it is not always grasped. Imagine that the subject's avowal appears inconsistent with the rest of their behaviour, linguistic and non-linguistic, including perhaps – most crucially – a failure subsequently to perform A. The authority principle rules out one possible explanation of this inconsistency: that the subject was simply ignorant of their own intention. The principle therefore implies that intention is analogous with belief – on my account – and disanalogous with ordinary behavioural dispositions to drive over the speed limit or to smoke cigarettes, and with character- and personality-traits such as generosity and obstinacy. In the latter cases, the possibility of self-ignorance cannot be ruled out; the subject may, non-pathologically, lack insight. In such cases there is no authority, and no avowal in the sense in which the term is used here.

The authority principle generates for intention, as it does for belief, a version of what I have termed the dispositional dilemma – namely how can I have what seems to be immediate and certain, or even presumed, knowledge of my intention, given that it is a state extended in time, a state of which others apparently come to know by interpreting my linguistic and non-linguistic behaviour? In the case of intention as with belief, a reliabilist resolution, which views the subject as having merely reliable knowledge of their intention, is generally regarded as inevitable. The present article argues against this position, concluding that for the subject, present intentions are not, properly speaking, objects of knowledge at all.

There is a further principle associated with authority which, as with the case of belief, also conflicts with the dispositional model. I have called this the *decision principle*, and it has two versions, the first of which says:

Having decided the question whether I will A/am going to A, there is no further question for me to decide, concerning whether I intend to A.

The decision principle claims that the outcome of deciding whether I will A – a process that mostly involves considering the pros and cons of A-ing – is the basis for my willingness to avow the intention to A. The avowal is either the result of a process of making up one's mind, or else constitutes a re-affirmation of a prior commitment; it is not based on introspection of a present state, nor on examination of present behaviour. I cannot merely report the presence of an intention: 'I intend to A', or 'I will A' when it functions as an avowal of intention, express a commitment to act. The converse principle also applies:

Having decided the question whether I intend to A, there is no further question for me to decide, concerning whether I will A/am going to A.

The decision principle constitutes one expression of the future-direction of intention. When asked whether I intend to go shopping, it would be absurd for me to respond 'Here's my

provisional answer, but I can't tell you for sure until the verdict is in on my behaviour'. Unless my intention to do something simply expresses my desire to do it, deciding what I intend – that is, what I am going to do – involves a process of practical reasoning, including estimation of practicalities or means-end beliefs. This process will involve an assessment of my chances of carrying out the intention; this assessment can be reconsidered, but such a process would amount to a reconsideration of the intention itself. Consider my avowal 'I intend to mark 50 essays over the weekend'. If someone responds 'Are you sure?', their question concerns the future; it does not require investigation of a present state, but invites me to reconsider the question of what I will do. Insofar as the question seeks a justification of my claim, it asks not for evidence – behavioural or introspective – that I have the intention, but requests my reasons for action. I am being asked whether I have really considered whether I am able to carry out the action.

In both formulations, the decision principle perhaps has an odd ring to it. This is because 'I will ...' or 'I am going to ...' can be read not just as an avowal of intention, but also as an evidence-based prediction of an involuntary action – 'I'm going to be sick' for instance – and perhaps even an intentional action. On any of these readings, however, having decided whether I am going to A, there is no further question concerning whether I intend to A. Arising from the decision principle is a Moorean paradox for intention, namely: 'X intends to A, but they will not A' is quite intelligible, while 'I intend to A, but I will not A' is paradoxical, as is 'I intend to A, but I doubt that I will A'. The paradoxes which these utterances express arise from the mistaken assumption that what I intend and what in fact I will do are separate matters.

## 2. Objections to authority and decision principles

The authority principle asserts that intention is a paradigm of the conscious.<sup>6</sup> Although I may act intentionally without prior deliberation, an intentional basic action can never be a surprise to me when I perform it. Only a reflex or involuntary movement – if these count as actions at all – could surprise the agent. Impulsive actions or instinctive reactions such as bravely tackling a gunman are not fully intentional, and may surprise the agent when they reflect on them after the event – as other performances may do too, for instance giving a dazzling interview or crumpling under questioning. But the basic actions by means of which these actions were performed could not be a surprise to the subject. As Wittgenstein (1958, para. 628) noted: 'one might say: voluntary movement is marked by the absence of surprise'. One could not define intentional actions as bodily movements which I am not surprised by – my hair blowing in the wind may not be a surprise – and so lack of surprise is a necessary but not sufficient condition. But the authority principle and the conscious status of intention suggest that the class of intentional actions may be defined by the agent's readiness to claim them as intentional, a criterion unjustly neglected in recent debates in favour of a causal analysis of intentional action. If intentional actions are those which are defined by the agent's readiness to claim them as intentional, then intention may be more paradigmatically conscious than belief.<sup>7</sup>

As in the case of belief, critics will argue that unconscious intentions, and cases where the subject mistakenly believes that they have the intention when they do not, serve to undermine the decision and authority principles, respectively. My response is (1) that there is no interesting sense in which intentions can be unconscious, and (2) that intentions are self-intimating – if I intend to A, I cannot doubt that that is my intention. Akrasia is the first of three candidates for unconscious intention which should be considered:

### **(1) *Akrasia and 'repressed intentions'***

Akrasia, or weakness of will, is normally taken to be one kind of explanation – or label for a non-explanation – of why, having formed the intention, the subject fails to perform the action. It is usually assumed either that the akratic subject has a repressed intention of which they are ignorant, and which is contrary to the intention which they avow; or that they avow an intention which they do not really have. I would argue against these assumptions that it is indeterminate whether or not the subject has the avowed intention; and moreover, that whether or not they do, there is no threat to authority. This is because if the subject does have the intention, the avowal is correct; and if they do not clearly have the avowed intention, then they are not simply mistaken in their avowal. As in the case of self-deception, motivated irrationality is implicated. The akratic avowal is not the product of a process of practical inference; rather it is caused by a wish, desire or fear which the subject is unable consciously to acknowledge. The 'mistakes' of the akrates, like those of the self-deceptive subject, are motivated and not clear-cut. The akratic agent fails to recognise the conflict which they are undergoing, or the pattern of action in which they are engaged. The authority principle as stated above specifies that the avowal has to be non-akratic – a qualification that does not import circularity, since the akratic subject does not make a simple mistake in their avowal of intention. But like the requirement that the avowal of belief should be non-self-deceptive, the qualification 'non-akratic' may not really be required.

### **(2) *Intention to A, where 'Shall I A?' is a question I have never considered***

As in the case of belief, it is often assumed that one can have the intention to A unconsciously, without having considered the question of A-ing. This assumption is unwarranted. Clearly it is of the nature of intentions, as it is of beliefs, that sometimes they are considered or publicly expressed, and at other times not. But it does not follow that there are two categories of intention, conscious and unconscious. As in the case of belief, it is a version of this dubious categorisation which generates the view that intention is closed under entailment – that if I intend to A, and A requires B, it follows that I intend to B (a version of the claim that willing the end implies willing the means). But again as with belief, plausible examples of conflict between avowed intentions and non-avowed products of inference are hard to come by. One suggestion is that one can intend to do whatever is in one's diary. Without checking my diary, I might express an intention to attend a meeting on Monday and then discover that the meeting has been entered for Tuesday – so apparently, the avowal turns out mistaken. I avow the intention to go to a meeting on Monday, but unconsciously intend to go to a meeting on Tuesday. The example is unconvincing from the outset, however. The general intention to do whatever is in one's diary is a peculiar one; diaries do not normally function so mechanically. Moreover, it is more plausible to say that I intend to go the meeting on Monday, and change my intention when I check my diary. While it is too strong to say that one cannot be said to believe something which one has never considered, it does seem that one cannot be said to intend something which one has not thought about doing.

### **(3) *Previously formed intentions***

Again intention mirrors belief. If I have deliberated but then forgotten what I intended, I cannot any longer be said to intend it. If someone asks me whether I intend to A, and I recall having formed the intention, then either I will self-ascribe it unreflectively or habitually, or else I will reconsider, and then either revise or restate it. A restatement of one's intention is not simply a

description of an intention already formed, however. I cannot adopt a neutral stance towards the intention, since it involves a commitment to act. So I cannot make a simple error concerning the presence of a previously formed intention. Again there is no justification for postulating a distinct category of unconscious intentions, revealed by introspection, or by observation of one's behaviour.

### 3. The prediction or 'future-outer' thesis for intention

The authority and decision principles constitute a denial of reliabilism – the view that the subject has merely reliable knowledge of their intentions. In opposition to reliabilism, I propose instead that as in the case of belief, the dispositional dilemma should be resolved by a non-cognitive standpoint which undercuts the traditional debate. On this view, 'I wonder whether I intend to A?' and 'I believe that I intend to A' have no meaningful use. The non-cognitive standpoint rejects a self-scanning, descriptive or reporting model of self-ascription; avowals are regarded as expressions and not descriptions of one's belief or intention. Corresponding arguments to those concerning belief carry over to intention.<sup>8</sup> In brief, it is not possible for the subject merely to report their intention to A, without expressing a commitment to act. To observe my present intentions from a third-person viewpoint, merely reporting that I have them, suggests a dissociated personality or psychotic state. Neither intentions nor beliefs are things that just come into the subject's head, which they can observe with concern, fear or approval; forming an intention involves making a judgement. Intentions, unlike desires, are never things which the agent just has. They are formed, not given.

Disanalogies between intention and desire, glossed over by the dispositional model of intention associated with reliabilism, reinforce non-cognitivism concerning avowals of intention. Desire is a disposition, and the paradigm motivational state, and therefore a non-cognitive thesis of desire would be quite implausible. One can have alien desires, but not alien intentions. One may be concerned about one's desire for drugs, pornography or unfeasible quantities of jazz CDs – regarding these desires as alien – but it would be nonsensical to avow concern about one's intention to acquire these things. If I do something reluctantly, my regret is not that I have the intention to A, but that I have to do A. In contrast, one can observe one's desires in a third-personal behavioural manner, and errors in self-ascription are possible. I can try, want and maybe intend to want, but I cannot want, try or intend to intend. Intention standardly involves a rational commitment to act, desire does not. In respect of the dispositional dilemma, therefore, intention parallels belief rather than desire. Desire is not future-directed in the same way as intention. Self-ascriptions of desire are not, properly speaking, avowals which exhibit authority, nor are they predictions – predictions that I will get what I want, for instance – in the way that I will argue those of intention are. Desire is a motivational state, and its propositional content might appear to concern the future; but neither of these features, I believe, suffices for the property of being forward-looking in the sense in which I have defined it.

A non-cognitive thesis of avowals of belief is associated with the *assertion thesis* which says that such avowals are assertions of the fact believed; beliefs are not inner states of which one can have introspective knowledge. A non-cognitive thesis of avowals of intention may be associated with the analogous *prediction thesis* which says that such avowals are predictions of a future action – judgements about the future not based on evidence – and not descriptions of one's present state; intentions are not inner states of which one can have introspective knowledge. In each case, there is no inner state to be known; the subject's knowledge concerns only the outer (or future outer). Just as it is a mistake to say that 'I believe that p' is primarily a report



of a mental state, and only indirectly an assertion of the fact believed, so it is wrong to say that 'I intend to A' is primarily a report of a mental state, and only indirectly a prediction that I will A. The prediction thesis constitutes a stronger claim than the standard view that an avowal of intention implies a prediction. Like the assertion thesis for belief, the prediction thesis highlights a first/third-person asymmetry. The third-person attribution of an intention is a prediction about what the agent will do only in the most qualified sense. The observer may attribute the intention, knowing or believing that the agent will be prevented, or will change their mind, etc.: 'He intends to catch the train to London, but unfortunately the line is flooded so he won't be able to'. So in contrast to the first-person case, the prediction is at best a conditional one: that the agent will take steps to catch the train, or would catch it if they were able. Analogously, an observer will make a third-person attribution of a belief that p, knowing or believing that p is false; the third-person ascription of belief is in no way an assertion of p.

One must address the most obvious objection to the prediction thesis, an objection that is often based on a misunderstanding of what the thesis is claiming. This is the objection that predictions about what I will do are normally based on evidence, while avowals of intention are based on reasons for action. Now certainly one can make predictions – in the ordinary sense of 'prediction' – about one's own future behaviour, but these are quite distinct from the category of avowals. The distinction has a grammatical mark, now fallen into disuse: 'I will' is a prediction, 'I shall' an expression of intention. Such predictions are normally evidence-based, and do not concern intentional actions, but rather, behaviour or outcomes over which the agent lacks direct control: 'I'm going to be sick'; 'I'll probably trip on my way to the podium'; 'I know I won't be able to stop myself telling him what I think of his behaviour'. 'I'm going to fail my driving test' could be ambiguous; an expression of lack of confidence, or of a deliberate intention. If someone asks 'How do you know?', usual responses would be 'I haven't practised enough' or 'I get very nervous in tests'. But it would be conceivable, if bizarre, for the subject to respond 'No, you don't understand, I *intend* to fail' – they are aiming for the test failures record, or appearing in a documentary about appalling drivers. Where the prediction concerns one's own intentional actions, it is unlikely to be entirely evidence-based. If someone asks me where I am going for my next summer holiday, my response 'I'll probably go to Italy, though I haven't yet decided' might appear to be an ordinary prediction, based on evidence of my past Italian holidays. If I had not thought about the question and have no plans, possibly I could regard it third-personally and estimate my likely behaviour. But it would be difficult to divorce this reasoning from a process of first-personal practical reasoning concerning where I will go on holiday.

Clearly avowals of intention are not predictions of future action in the preceding sense. To reiterate, the prediction thesis says that an avowal of intention is a judgement about the future outer, not the present inner. The point is clearly understood by Anscombe, when she proposes a prediction thesis in her book *Intention* (1963). Anscombe wishes to broaden the notion of prediction to encompass commands and avowals of intention as well as evidence-based estimates of the future. A prediction, she writes, is when something is said with one inflection of the verb, and later that same thing, only with a changed inflection, can be called true (or false) in the light of what has happened. On this criterion, she argues, commands and expressions of intention will also be predictions, since they are all 'descriptions of something future', as opposed to something present and internal to the agent. (She evidently regards 'I shall A', and not 'I intend to A', as the characteristic expression of intention.) Anscombe suggests that in contrast to an evidence-based estimate of the future, 'an expression of intention is a description of something future in which the speaker is some sort of agent, which description he justifies (if he does justify it) by reasons for acting'. I believe that this is a salutary statement of a prediction



thesis, and it captures the view that I wish to defend: that avowals of intention – and perhaps in some sense intentions themselves – are predictions based on, or the product of, reason-providing ends. However, Anscombe (1963, 2–6) pronounces the line of approach mystifying, and presents instead her well-known account of intentional action as based on reasons for acting. This response is itself puzzling, because the two accounts seem to be compatible and indeed complementary – as I will show in my concluding section.

The prediction thesis has several advantages. It helps to elucidate the future-direction of intention, and thereby avoid a Cartesian explanation of the authority of avowals – ‘Cartesian’ in the historically inaccurate textbook sense of introspectionist and mentalistic. If avowals of intention are regarded as non-evidence-based predictions of future action, the tendency to regard them as reports of an inner state or disposition is resisted, and an explanation of their authority, and of the associated decision principle, becomes possible. The Cartesian model of an infallibly known inner state is undermined when one recognises that, from the first-person present viewpoint, the question of what I intend is not distinct from the question of what I will do – just as the question of what I believe is not distinct from the question of what is the case. Hence the Moorean paradoxes for belief and intention which arise when this fact is ignored. The fact that ‘I intend to A, but I will not A’ is paradoxical, suggests that an avowal of intention is a prediction.<sup>9</sup> A structural analogy with belief supports the prediction thesis: ‘I believe that p’ says ‘P unless I am much mistaken’, while ‘I intend to A’ says ‘A unless I am prevented, or . . . etc.’

As Anscombe recognised, a prediction thesis has to regard avowals of intention as part of a broadened class of predictions; beliefs about the future based on evidence, avowals of intention, and commands, are regarded as equally future-directed. Avowals of intention differ from evidence-based predictions of future behaviour in resting unambiguously on reason-providing ends. However, it is not ideal to define a philosophical term of art in a way that conflicts with its everyday connotations, in this case, an evidential basis. As noted earlier, the prediction thesis could more clumsily be called the *future-outer thesis*, and this is how I will henceforth generally refer to it. The thesis – however it is termed – says that avowals of intention concern future and outer states, and are not descriptions of something present and inner. The thesis faces three other principal objections:

**(1) Predictions that I will A are falsified simply by my not doing A; avowals of intention are not so falsified (the agent may be prevented, change their mind, etc.)**

The assertion thesis for belief is often rejected on the grounds that the assertion that p and the avowal of belief that p have different truth-conditions. Anscombe (1963, 4) presents an analogous objection to the prediction thesis: ‘If I do not do what I said I would, I am not supposed to have made a mistake, or even necessarily to have lied. So it seems that the truth of a statement of intention is not a matter of my doing what I said’. Her response is that ‘this only shows that there are other ways of saying what is not true, besides lying and being mistaken’. The suggestion is that ‘I’m going for a walk’, in its normal use as an avowal of intention, may be false not just when the subject is lying, but also when they are prevented, change their mind, etc. and so do not A. But Anscombe’s response is plausible only for avowals of intention which are, grammatically, simple future-tense claims: ‘I will A’ or ‘I’m going to A’. ‘I intend to A’ could not count as a prediction on her criterion, namely: something is said with one inflection of the verb, and later that same thing, with a changed inflection of the verb, is called true (or false) in the light of what has happened.

There is a more convincing response than Anscombe's to the objection that the truth of the avowal is not a matter of the truth that the agent will A. This response exploits the parallels with belief. To describe someone as denying the claim 'I believe that the Conservative Party is finished' is to represent them as saying 'No, it's not' rather than 'No, you don't'. Thus an assertion thesis for 'I believe that p' does not imply that the avowal is falsified when p turns out false; likewise, the claim that 'I intend to A', which functions as a prediction that I will A, does not imply that the avowal is falsified when I do not A. 'I'm going shopping' is not falsified by my subsequently not going shopping, not even where this failure shows that the avowal was insincere – because then it would be untruthful but not false. If I am prevented from going, it would be a feeble joke for someone to say 'Then it wasn't true that you were going shopping'; I would justly reply 'I was going shopping, but I was prevented'. It is true that the intelligibility of 'I was going shopping, but I didn't go' appears to support Anscombe's position; but it is countered by the senselessness of 'She is going shopping later, but she'll change her mind in a minute'.

What appear at first sight to be relatively radical measures are called for here. A future-outer thesis should be accompanied by a truth-valueless thesis of avowals, according to which avowals are not true or false – made true or falsified – at all. Rather they are truthful or untruthful, fulfilled or unfulfilled. This is not counterintuitive when one considers that a truth-valueless thesis for commands is uncontroversial. Commands, like avowals of intention, are also predictions in the broader sense – they may be expressed as 'You will do this', as opposed to 'Do this!' Commands, promises and requests, like avowals of intention, all say something about the future – that is, they are predictions broadly conceived. What they say about the future – the propositional content – can be true or false; the corresponding third-person statement will be true or false; but the avowal, command or promise is not itself true or false.

## ***(2) The future-outer thesis cannot accommodate the fact that avowals of intention are expressions of a motivational state***

Earlier I stated that my target was the standard picture of intentions as inner, motivational states; the latter aspect should now be considered. Despite appearances, I would argue, the picture of intention as a motivational state is not sufficient to capture its future-direction; intentions are the rational products of motivational states, and not motivational states in themselves. Intention is not a further motivating factor to be added to the agent's desires and means-end beliefs; rather, when there is practical deliberation concerning those desires and beliefs, intention is normally the result. (Though as we will see, not all intentions are the product of practical deliberation.) The fact that intentions constitute reasons for action – one can say 'His intention to get arrested was the reason he knocked off the policeman's helmet' – does not mean that they are motivational states. It might be argued a settled intention can motivate. Although 'I've put a lot of work into this process of decision so that's what I'm going to do' does not count as a reason for action, it may motivate without being a reason for action.<sup>10</sup> However, it is doubtful that it is the intention which motivates. Generally, intention is a rational commitment to act, and a commitment is not a motivation. There might be a consequentialist justification for the policy 'always act on it once you've formulated the intention'; but the motivation then is the policy, not the intention.

An intention is not, in contrast to motivational states, a cause of the action. Consider the question 'What caused him to jump out and make that funny face at me?' To say that his intention to jump out and make a funny face was the cause is not explanatory, even though the existence of a prior intention is not presupposed by the description of the action as intentional – 'She acted

intentionally' does not imply 'She intended to act'. However, 'jumped out . . .' is a datum, not an interpretation. The answer 'His intention was to make a funny face' may show that the action was premeditated, but in this case, in contrast to an habitual or unpremeditated intentional action such as opening a window, that was likely anyway.

In light of these considerations, objection (3) should be reformulated as follows: avowals of intention express the results of practical deliberation concerning motivational states, whereas predictions express the result of evidence-gathering. As noted earlier, wishful thinking and self-deception are irrational precisely because the avowal of belief is motivated by a desire. Objection (3) then becomes a version of the objection that predictions are based on evidence.

### ***(3) The future-outer thesis gives intention the wrong direction of fit with the world***

Predictions in the ordinary sense are beliefs, and beliefs – it is often said – aim at the truth; they change to reflect the state of the world, while intentions aim to change the world to reflect their own state. The future-outer thesis, it may be argued, fails to respect this distinction. Of course, the thesis is a claim about avowals of intention, and not intentions as such. However, if an avowal of intention says something about the future – rather than describing a disposition to act – then, it may be argued, that surely makes an intention into a belief. So the problem remains. Now the standard picture of direction of fit is certainly correct in pointing to a fatal implausibility in Velleman's (1989, 2000) treatment of intention as a kind of self-referential belief. However, the metaphor of direction of fit may be too elusive for it to carry out much useful philosophical work. Certainly it has to be cashed out in terms of the features of intention under discussion – that avowals of intention are not based on evidence, that they express a motivational state, and so on. Moreover, since intentions and beliefs do not literally have aims, it is preferable to say that in forming a belief, subjects aim at the truth; and that in forming an intention, they aim to change the world. Finally, insofar as we persist with the metaphor, intention is more equivocal in its direction of fit than desire and belief, since intentions also change to reflect the state of the world. One cannot intend to do something that one believes to be impossible, and so an intention can go out of existence because of the way the world is – when the agent realises that they cannot in fact A after all. In contrast, one can perhaps desire what one believes to be impossible, though desires also go out of existence because of the way the world is. Hopes and wishes are the most resistant, bearing the weakest relation to how the world is, but even they can vanish – I can no longer wish to be a millionaire if I discover that I already am one. So it is necessary to qualify the metaphor and say: intentions aim to change the world to reflect their own state up to certain reasonable or rational limits. In contrast, beliefs in no way aim, or should – rationally – in no way aim, to change the world to reflect their own state; hence a belief caused by a wish or desire is criticised for its motivated irrationality.

Does the *future-outer thesis* respect this demarcation? I have emphasised that the thesis does not assimilate intention to belief – indeed, as we have seen, Anscombe originated both the prediction thesis, and the metaphor of direction of fit, so she at least is aware of the need to respect the latter demarcation!<sup>11</sup> Thus a future-outer thesis of the Anscombian variety is compatible with a distinct direction of fit, in contrast to accounts such as Velleman's which set out consciously to undermine it.

## **4. Defending the future-outer thesis**

I believe that the future-outer thesis should be persevered with in the face of the preceding objections. In this final section I attempt to vindicate it by redeeming a pledge made earlier,

and showing how the thesis complements the Anscombian account of intentional action as based on reasons for acting.<sup>12</sup> An interesting development of this account is presented by Gilbert Harman (1976, 453), who regards intentions as psychological states with a dual function – perhaps it would be better to say a dual aspect – serving as ideas of certain aspects of the future and therefore like beliefs, but also as reason-providing ends, leading ultimately to actions: ‘An intention is an idea of the future for which one has one sort of reason; a prediction is normally an idea for which one has another sort of reason . . . an intention is an idea arrived at and maintained by practical reasoning, whereas a prediction is normally an idea arrived at and maintained by theoretical reasoning’.<sup>13</sup>

The Anscombian account requires qualification and development in various ways that draw on the future-outer thesis. Firstly, a qualification, concerning the claim that an intention is a reason-providing end. Intentions, I would argue, are the products of ‘reason-providing ends’ and not such ends in themselves. A more important qualification is that intention is not merely an ‘idea’ of the future. Harman says this, no doubt, because the expression is neutral concerning which kind of basis is present, theoretical or practical reasoning; the problem, however, is that ‘idea of the future’ is also unacceptably neutral between commitments to action or truth, and non-commitments such as imaginings of the future. What is missing from Harman’s account is the understanding implicit in the future-outer thesis that an avowal of intention expresses a *judgement or decision concerning the future* – concerning one’s own future action, not one’s present disposition or motivational state. An intention is a judgement that I will A (unless I am prevented. . .); a prediction, in the sense of estimate about the future, is also a judgement concerning the future. As Harman notes, in response to the question ‘How do you know it will happen?’, the maker of the prediction will reply ‘Because of such and such evidence’, while the avower of intention will say ‘Because I’ve decided, for such and such reasons’. Finally, a development rather than a qualification, and a vital one: the distinction between practical deliberation and evidence-gathering is not more fundamental than the contrast between intention and estimate of the future. There is a conceptual holism here, an explanatory interdependence or mutual presupposition of concepts, of a kind which I elaborate on elsewhere (Hamilton forthcoming). One cannot define, grasp or manifest an understanding of the distinction between practical deliberation and evidence-gathering, without defining, grasping or manifesting an understanding of the distinction between intention and estimate of the future.

We will see shortly how appeal to this conceptual holism is essential in dealing with the objection that the Anscombian account involves an unacceptable rationalism. First, however, a defence of the interpretation of intentions as judgements. It may be objected that judgements, or their makers, unlike intentions, say something that is either true or false. This is incorrect, however. The verdicts of an umpire, judge or adjudicator are paradigm examples of judgements, and they do not say that something is true. Nor do judgements have to be expressed verbally. Unlike commands and ordinary predictions, intentions are not essentially linguistic. Nonetheless, there is the worry that the attribution of intentions to animals is inconsistent with the interpretation of intentions as judgements. Is the attribution of belief to animals coupled with the attribution of intention – that is, does one ascribe beliefs to just those creatures to which one ascribes intentions, and vice versa? W.C. Fields famously remarked that ‘Anyone who hates children and animals can’t be all bad’, and my feeling is that the same applies to any philosophy which avoids the question of children and animals. But I will comment that when the cat jumps up to the door handle and the door opens, it is not clear whether we ask, ‘Did it intend to do that?’ or merely ‘Did it desire that outcome?’ Wittgenstein (1958, para. 647) was evidently in no doubt: ‘What is the natural expression of intention? – Look at a cat when it stalks a bird; or a beast when it wants

to escape'. Perhaps it is perverse to allow, with commonsense, that the dog barking up the wrong tree believes that the cat is up the tree, yet to deny that it intended to catch the cat. I am not convinced, but at this point will simply say that the claim that intentions are judgements is not in tension with the attribution of intentions to animals, since in the latter case it is a more primitive case of intention, parasitic on the human case.

The centrality of the human case is underlined in some concluding comments I will now make on the Anscombian characterisation, concerning the charge of rationalism. The objection is that having a prior intention does not always involve deliberation; so should intentions be described as arrived at and maintained by practical reasoning? An initial move in allaying the charge of rationalism is to note that intentional actions need not involve prior intention; one can A intentionally without having formed the intention to A. However, it may be objected that it is possible to decide to do something spontaneously, without reasoning or deliberation; and that therefore some intentions, lacking this element of deliberation, are not judgements based on reasons. 'I'd like a coffee', I say, and go and make it. Going down the stairs, my intention is to make a cup of coffee. But what were my reasons for making it? 'Because I felt like it', the objection continues, is a non-reason; it means 'Don't ask me for reasons'.

In response to this line of objection, one could argue that since intentions do not just pop into one's head, but have to be formed, then in the coffee-making case there is no pre-formed intention. Compare beliefs, which unlike hunches or suspicions, also do not just pop into one's head. The coffee-making behaviour is habitual, or maybe impulsive. If I really have no reason for going downstairs to make a coffee other than that I feel like having one, then perhaps this is a case of end-directed, intentional action, but one which falls short of being based on a reasoned commitment and so lacks a pre-formed intention. It is true that things we do habitually need not be done out of habit – there is no reason we should not form and announce an intention to do them. This fact leads to the alternative suggestion that just as one can believe unthinkingly, and hold beliefs without any rational basis, so one can have intentions without any rational basis also. So I can be said to intend to do something – in contrast simply to doing it intentionally – without formulating an intention. For instance, I always intended to go for a swim when I got to the beach – my having packed a towel shows this – but I did not formulate the intention; if you had asked me, however, I would probably have said yes, that is my intention. A final option is to argue that the notion of deliberation does not have to be psychological. The suggestion is that I can have a prior intention as long there are reasons to be given for the action – that I could give a public rationalisation of it.

Each of these options has its advantages. I would want to insist that even if not all of our commitments need be based on reasons, the latter are philosophically more central. Intentions must, in core cases, be judgements based on reasons. This is because one could not acquire the concept of intention without acquiring the concept of a reason for action. This is the conceptual holism stated earlier. One acquires the concept of intention at the same time as one acquires the concept of reasons, and the concept of factual claim or evidence-based statement. Animals do not acquire such concepts, and they therefore have intentions in a more primitive sense than humans do. The conceptual holism of intentions and reasons for action is shown by the fact that end-providing reasons can always be expressed as intentions. Consider Anscombe's famous example of the man operating a pump. What is his reason for operating the pump? To poison the people in the house. Does he intend to poison the people? Yes. What is his reason for poisoning the people? To overthrow the party. Does he intend to overthrow the party? Yes. What is his reason for that? To institute democracy . . . and so on. In the presence of a conceptual holism, the dual aspect account which I have developed from Harman can still serve as an elucidation of the

concept of intention, though not as an explanation of intention in more basic terms. But there is no such more basic explanation. Hence my position that an intention is a judgement about the future which in central cases is arrived at and maintained by practical reasoning, while an estimate of the future is a judgement about the future which in central cases is arrived at and maintained by theoretical reasoning.

Intention, like belief, belongs to a range of psychological concepts, each exhibiting different biases towards one or other end of the episode-disposition continuum, but each possessing its own version of the authority and decision principles. Yet its forward-looking nature is often neglected. In traditional presentations of the Toxin Puzzle, for instance, failure properly to acknowledge it leads to such absurdities as postulating a contrast between 'reasons for intending' and 'reasons for acting' (Kavka 1983). Such cases show that I cannot 'merely intend'. Intention looks forward, and reaches all the way to the performance of the action. It is this forward-looking feature which the future-outer thesis is concerned, and ideally placed, to capture. Very deep issues about nature of intention remains, which centre on the question of what it is to be the author of one's own actions. But that is material for another occasion.

## Acknowledgements

I am grateful for comments from Robin Hendry, E.J. Lowe, Al Mele, Roger Squires, Sunny Yang, and the audience at Nottingham University Philosophy research seminar in 2001.

## Notes

1. An example of such an analysis is Pears (1980).
2. As argued in Hamilton (2000).
3. See Hamilton (2000).
4. See Anscombe (1963, 2–6); Wittgenstein (1958, paras 629–30, 191, 223). Wittgenstein is attacking the empiricist view that all knowledge of the future is based on present experience. He also refers to the kinship but fundamental difference between two language-games of 'prediction', one based on evidence, the other not – though here the one not based on evidence is commanding rather than intending (Wittgenstein 1958, 191, 223).
5. This view has been defended by Hampshire (1959, 134).
6. The claim is endorsed by Hampshire, who writes that intention exhibits that kind of knowledge of what one is doing, and of what one is inclined to do, that is fully conscious . . . (Hampshire 1963, 59–68). Compare Wittgenstein's reported remark: 'The case of voluntary and involuntary action is like the case of lying and telling the truth. "A man knows when he is lying" is a grammatical and not an experiential statement. It makes no sense to ask "Am I lying?" except in extraordinary circumstances' Wittgenstein (1988, 205). See Scott (1996).
7. Davis (in Guttenplan 1994) defends a view which is close to that proposed here, except for its cognitivist reference to the agent's 'non-perceptual awareness' of an action as their own.
8. These arguments were presented in Hamilton (2000).
9. As Hacker (1996, 255) notes: for instance, I cannot intend to A when I desire to A but know that it is practically impossible to A.
10. Harman, for instance, comments that 'It represents an important aspect of rationality, an aspect that promotes stability in one's plans . . . in the midst of them, one will be strongly motivated to continue so as not to waste what has gone before' (1976, 462–3).
11. The metaphor is normally used to contrast desire and belief, and is discussed by M. Smith (1994, 111–6) and Wedgwood (1995, 274–5); it may originate with J.L. Austin, as Humberstone (1992) suggests.

Though often attributed to Anscombe, for instance by Platts (1979, 256), she does not use the metaphor herself, but makes a similar point in her contrast between a mistake in the performance of someone buying from a shopping list, with a mistake in a record which a detective is keeping of the shopping purchased (Anscombe 1963, sec. 32). As Danto comments (1999, 56), the shopper may make a mistake by putting in the basket something that is not on the list – beer when the list says ‘beef’; the detective may make a mistake in writing ‘beef’ when the basket contains beer. The shopper rectifies the mistake by replacing the beer with beef; the detective by writing ‘beer’ instead of ‘beef’. A shopping-list looks more like an expression of intention than of desire, it might be argued. Bernard Mayo (as reported to me by Roger Squires) imagined a mechanic dipping the dipstick into the oil sump, noting that the liquid falls below the line, and redrawing the line at where the oil comes to! This is the difference between checking the oil and recording it. An insightful discussion of ‘direction of fit’ is found in Yang (2006).

12. See Anscombe (1963, 2–6).
13. Harman (1976, 453) continues: ‘To say how desire is relevant to reasons for intentions, but not to reasons for predictions, is to say how coherence with intrinsic desire is a factor in practical but not in theoretical reasoning’.

## Notes on contributor

Andy Hamilton teaches Philosophy at Durham University, and has published numerous articles on philosophy of mind, aesthetics and the history of philosophy. He is also the author of *Aesthetics and Music* (Continuum, 2007) and *Lee Konitz: Conversations on the Improviser’s Art* (University of Michigan Press, 2007).

## References

- Anscombe, G.E.M. 1963. *Intention*. Oxford: Blackwell.
- Danto, A. 1999. Basic actions and basic concepts. In *The body/body problem*, 56. Berkeley, CA: University of California Press.
- Guttenplan, S., ed. 1994. *Blackwell companion to the philosophy of mind*. Oxford: Blackwell.
- Hacker, P. 1996. Wittgenstein: Mind and will. Vol. 4 of *Analytical commentary on the philosophical investigations*. Oxford: Blackwell.
- Hamilton, A. 2000. The authority of avowals and the concept of belief. *European Journal of Philosophy* 7, no. 2: 20–39.
- . Forthcoming. *Memory and the body: Exploring self-consciousness*.
- Hampshire, S. 1959. *Thought and action*. London: Chatto and Windus.
- . 1963. Disposition and memory. *International Journal of Psychoanalysis* 42: 59–68.
- Harman, G. 1976. Practical reasoning. *Review of Metaphysics* 29.
- Humberstone, L. 1992. Direction of fit. *Mind* 101, no. 401: 59–83.
- Kavka, G. 1983. The toxin puzzle. *Analysis* 43: 33–6.
- Pears, D. 1980. Intentions as judgments. In *Philosophical subjects: Essays presented to P.F. Strawson*, ed. Z. Van Straaten. Oxford: Clarendon.
- Platts, M. 1979. *Ways of meaning: An introduction to the philosophy of language*. London: Routledge and Kegan Paul.
- Scott, M. 1996. Wittgenstein’s philosophy of action. *Philosophical Quarterly* 46, no. 184: 347–63.
- Smith, M. 1994. *The moral problem*. Oxford: Blackwell.
- Velleman, J. 1989. *Practical reflection*. Princeton, NJ: University Press.
- . 2000. *The possibility of practical reason*. Oxford: Clarendon Press.



- Wedgwood, R. 1995. Theories of content and theories of motivation. *European Journal of Philosophy* 3: 273–88.
- Wittgenstein, L. 1958. *Philosophical investigations*. Oxford: Blackwell.
- . 1988. *Wittgenstein's lectures on the philosophy of psychology 1946–7*, ed. P.T. Geach. Chicago, IL: Chicago University Press.
- Yang, S. 2006. Emotion and intentionality. PhD diss., University of Durham.