

Mental Overpopulation and Mental Action: Protecting Intentions from Mental Birth Control

ANDREI A. BUCKAREFF
Franklin and Marshall College
Lancaster, PA 17604-3003
USA

Many, I suspect most, philosophers of action afford intentions a central role in theorizing about action and its explanation. Furthermore, current orthodoxy in the philosophy of action has it that intentions play a causal role with respect to the etiology and explanation of action. But action theory is not without its heretics. Some philosophers have challenged the orthodox view. In this paper I will examine and critique one such challenge. I will consider David-Hillel Ruben's case against the need for intentions to play a causal role in the etiology and explanation of *mental* actions.¹ *Contra* Ruben, I will defend the orthodox view that intentions play an indispensable causal and explanatory role with respect to mental actions.

1 David-Hillel Ruben, 'Mental Overpopulation and the Problem of Action,' *Journal of Philosophical Research* 20 (1995) 511-24; and *Action and its Explanation* (New York: Oxford University Press 2003)

I Ruben on Mental Overpopulation and the Causal Theory of Action

Ruben's critique of a causal and explanatory role for intentions with respect to mental actions issues in part from his declaration of war on 'mental overpopulation' in the philosophy of mind and action. One target theory of action he claims overpopulates the mental is the causal theory of action (CTA). According to the CTA, some behavior *A* of an agent *S* is an action if and only if there are some non-actional mental events or states that proximately cause *S*'s *A*-ing and constitute *S*'s reasons for *A*-ing.² Regarding the implications of the CTA for mental overpopulation Ruben writes that

Assuming that we engage in a great deal of genuine activity, the demands of the CTA will ... require of agents an implausibly rich mental life, over full of reasons, beliefs, desires, and intentions. The CTA must inflate the mental, as a precondition for reducing action. Its slogan might be: no (action) reduction without (mental) inflation.³

The CTA 'overintellectualizes action,' according to Ruben. So 'the friendliest thing one can do in the case of action is to prune (but not, of course, to eliminate) the mind's contents.'⁴ So he sets himself to practicing what he calls a form of 'mental birth control.'⁵

Focusing on mental action, against theories that overpopulate our mental lives, Ruben offers a theory of mental action which allows for mental actions to be 'spontaneous actions.' For Ruben, a spontaneous action, either overt or mental, is one that 'has no *rationalising* mental

2 This is a rough CTA schema offered by Jing Zhu and I in our paper 'Causalisms Reconsidered,' *Dialogue* 43 (2004) 147-55, at 147. Alfred R. Mele helpfully articulates the explanatory role of reasons in his book, *Motivation and Agency* (New York: Oxford University Press 2003). He writes, 'Causalists traditionally appeal, in part, to such goal-representing states as desires and intentions (or their neural realizers) in their explanations of human actions, and they take acceptable teleological explanations of human actions to be causal explanations' (38).

3 Ruben, 'Doing Without Happenings: Three Theories of Action' in *Contemporary Action Theory, Vol. 1: Individual Action*, G. Homstrom-Hintikka and R. Tuomela, eds. (Dordrecht: Kluwer 1997), 267-86, at 268; cf. his similar remarks in *Action and its Explanation*, 98.

4 Ruben, *Action and its Explanation*, 98. See also 'Mental Overpopulation and the Problem of Action,' 512, and 'Doing Without Happenings,' 268.

5 *Action and its Explanation*, 98

cause or explanation' such as a belief-desire complex or an intention.⁶ When an agent acts intentionally, according to Ruben, 'it does not follow that the act was preceded by his having an intention to act.'⁷ With respect to mental actions, Ruben writes that, 'Mental life is simply not rich enough to insure that the requisite rationalising items will always be there.'⁸ This does not mean that actions are uncaused, lacking any motivation. It is just that it is not a necessary condition for some behavior to be actional that it has a rationalizing mental cause or explanation, according to Ruben.

I will not take up all the issues raised by Ruben in his case against theories of action that are guilty of 'overpopulating' our minds, or the problems I have with his theory of action more generally.⁹ Much of what

6 'Mental Overpopulation and the Problem of Action,' 514. See also *Action and its Explanation*, 147-8. It is worth noting that in *Action and Its Explanation*, 104-10, Ruben argues that intentions cannot be *rationalizers* of action. I will not take up Ruben's arguments for this separate claim here. I take it that the executive role of intentions and their function in guiding and coordinating the activity of agents would be part of the story one would have to tell in making sense of intentions as rationalizers of action. For the time being, I am solely concerned with defending the necessary role of intentions in the etiology and explanation of intentional action, particularly mental action. So I will take it for granted that intentions can play a role as rationalizers of action. I hope, however, to take up this matter in detail in the near future. For a defense of the claim that intentions can be rationalizers of action, see J. David Velleman, *Practical Reflection* (Princeton: Princeton University Press 1989), especially chapters 4 and 7. I will go on record as not agreeing entirely with Velleman's strategy, given that he endorses a doxastic theory of intention. For more on the role of intentions in practical reasoning (and so as potential rationalizers of action) see Michael Bratman, 'Intention and Means-End Reasoning,' *The Philosophical Review* 90 (1981) 252-65 and *Intention, Plans, and Practical Reason* (Cambridge, MA: Harvard University Press 1987).

7 *Action and its Explanation*, 107. Ruben adds that 'Nor does it follow, on my view, that there was no such intention.' Of course, I would add that those who defend a central role for intentions in the etiology and explanation of action who do not accept the Simple View of intentional action would not insist that an intention to *act* would always precede an intentional action (all that is necessary is that an action was within the motivational potential of an intention). In fact, I am not certain that all defenders of the Simple View would insist on such an intention. See n. 16 below for more on the Simple View of intentional action.

8 'Mental Overpopulation and the Problem of Action,' 513

9 For a fuller picture of Ruben's theory of action beyond what he says about mental action, see *Action and its Explanation*, especially 174-84. See also 'Doing Without Happenings: Three Theories of Action'; and 'The Active and the Passive,' *Proceedings of the Aristotelian Society, Supplementary Volume* 71 (1997) 229-46. It is worth noting in passing that Ruben dispenses with any attempts at identifying some

I will ignore about some of Ruben's work on mental action has been taken up elsewhere by Alfred Mele and, to my estimate, effectively addressed.¹⁰ But I will also address some issues taken up by Mele.¹¹ However, I will pursue a different line of argumentation against Ruben than Mele.

II Trying and Mental Action

Before moving to consider Ruben's case against mental overpopulation in theories of action, I should lay out my cards and offer a useful criterion for distinguishing between actional and non-actional mental behavior. I propose applying what Joelle Proust calls 'the try test' as a means of determining whether some behavior that occurs is actional. She suggests that, 'To determine whether some piece of behavior qualifies as an action ... just consider whether it can meaningfully be tried.'¹² Proust notes that such a test is not foolproof. But it seems to be at least one way to test whether some behavior is actional or mere behavior.

The features of a full-blown account of trying go beyond the scope of this paper. But, following Frederick Adams, I suggest that we take mental trying to be a composite event that consists of some relevant mental state or event or some cluster of mental states and events M_1 causing some relevant state or event M_2 for a reason.¹³ Among the

property or properties that distinguish actions from non-actions. He claims in *Action and its Explanation* that 'there is no feature, F , other than the feature itself of being an action (or a feature which presupposes that anything that has it is an action). What makes an action what it is, is nothing other than its being an action' (184).

10 What I will ignore that Mele takes up is Ruben's claim that desires alone cannot provide rationalizing explanations of actions. See Alfred R. Mele, 'Agency and Mental Action,' *Philosophical Perspectives* 11 (1997) 231-49, at 238-40.

11 In *ibid.*, 235-8.

12 'A Plea for Mental Acts,' *Synthese* 129 (2001) 105-28, at 108

13 'Cognitive Trying,' in *Contemporary Action Theory*, Vol. 1, G. Holmstrom-Hintikka and R. Tuomelo, eds., 287-314; see 289-91. For similar views, see Fred Adams and Alfred R. Mele, 'The Intention/Volition Debate,' *Canadian Journal of Philosophy* 22 (1992) 323-38. Different deflationary accounts of the will are also offered by David Armstrong, *The Nature of Mind* (Ithaca: Cornell University Press 1980), 68-88; and Timothy Cleveland, 'Trying Without Willing,' *Australasian Journal of Philosophy* 70 (1992) 324-42. Compare the volitionist accounts endorsed by Carl Ginet, 'Trying to Act,' in *Freedom and Determinism*, J.K. Campbell, M. O'Rourke, and D. Shier, eds. (Cambridge, MA: MIT Press 2004), 89-102; Jennifer Hornsby, *Actions* (London:

components of a trying would be a proximal intention that would causally initiate and sustain the action in the case of a successful trying. Furthermore, *trying on such an account is not a distinct mental action, but every action — both overt and mental — is itself a successful trying*. An event would fail to be an action if it is not a trying.¹⁴ Finally, on such an account, an intention is a component of a trying, as opposed to the trying being a distinct product of an intention.¹⁵ A trying would commence once an agent has acquired a proximal intention to act — i.e., an intention to act immediately. Such a theory of trying is consistent with a version of the CTA and overcomes some of the difficulties that have plagued earlier versions of the CTA that have merely afforded intentions and other mental states and events a ballistic role in causally initiating action.

Given that trying has intending as a component, I take it that if all mental actions are attempts then they are intentional actions under some description. Suppose we endorse either the Simple View of intentional action, which holds that intentionally *A*-ing implies that an agent has an intention to *A*, or the Single Phenomenon View, which takes the intention in intentionally *A*-ing not to have to be an intention to *A*.¹⁶ If all mental

Routledge and Kegan Paul 1980), ch. 3; Hugh J. McCann, *The Works of Agency: On Human Action, Will, and Freedom* (Ithaca: Cornell University Press 1998), ch. 5; Brian O'Shaughnessy, 'Trying (As the Mental "Pineal Gland")', in *The Philosophy of Action*, A. Mele, ed. (New York: Oxford University Press 1997), 53-74, at 56 and *The Will*, 2 vols. (Cambridge: Cambridge University Press 1980).

- 14 Adams and Mele write in 'The Intention/Volition Debate' that 'Roughly, trying to *A* is an event or process that has *A*-ing as a goal and is initiated and (normally) sustained by a pertinent intention. Successful tryings to *A*, rather than causing *A*-ings, are *A*-ings' (326).
- 15 Thanks are due to Randolph Clarke for pointing out this distinction between a component versus product view of the role of an intention in trying.
- 16 For my purposes here, I am remaining non-committal with respect to the two views. It is worth noting that on the Single Phenomenon View, *S*'s intention may include *A*-ing in what Michael Bratman calls its 'motivational potential' although *S* does not, strictly speaking, intend to *A*. See his paper, 'Two Faces of Intention' in *The Philosophy of Action*, A. Mele, ed., 178-203, at 198. For defenses of versions of the Single Phenomenon View, see the article by Myles Brand, 'Intention and Intentional Action,' in *Contemporary Action Theory, Vol 1*, G. Holmstrom-Hintikka and R. Tuomela, eds., 197-217; and Alfred R. Mele, *Springs of Action: Understanding Intentional Behavior* (New York: Oxford University Press 1992), ch. 8. For defenses of the Simple View, see Frederick Adams, 'Intention and Intentional Action: The Simple View,' *Mind and Language* 1 (1986) 281-301; David K. Chan, 'A Not-So-Simple View of Intentional Action,' *Pacific Philosophical Quarterly* 80 (1999) 1-16; Jorge Garcia, 'The Intentional and the Intended,' *Erkenntnis* 33 (1990) 191-209; and Hugh McCann, *The Works of Agency*, ch. 10.

actions are intentional actions under a description, and the CTA is true, then a proximal intention is at least part of the cause and explanation of any mental action. More accurately, given that it is a component of the trying, the proximal intention is a cause of a proper part of the trying that is identical to the action — being the proximal cause of some actional behavior that is a component of the trying. Thus it plays an indispensable role in the etiology of the trying, although the trying is not a distinct product of the proximal intention. However, if Ruben is right, then mental action need not be intentional under any description and the account of trying I have recommended cannot be applied to test whether some instance of mental behavior is a mental action. But Ruben is not right, as will become apparent in the next section.

III Ruben's FORGETFUL THINKER

In his paper 'Mental Overpopulation and the Problem of Action,' Ruben asks his readers to consider a scenario that involves an agent he christens the 'FORGETFUL THINKER.'¹⁷ This person has been unsuccessfully trying to solve a puzzle or perform some other mental task such as recall a name. This bothers him. He resolves not to think about the puzzle or try to recall the name any longer. After a bit, the FORGETFUL THINKER's mind is 'diverted, wanders, and he forgets, momentarily at least, his firm resolve.'¹⁸ Ruben writes, 'Before he is aware of it, once again he is reflecting on that puzzle or trying to recall the name. And when he realises that he has been thinking about what he resolved not to think about yet again, he is likely to chastise himself for forgetting his resolve.'¹⁹

In such a scenario, according to Ruben, the FORGETFUL THINKER forms an intention not to reflect on the puzzle or search his memory for the name again and stops performing whatever mental action he was performing. But, after a time, he winds up reflecting or attempting to recall the name. Ruben claims that only two descriptions of what is going on in such a scenario are possible:

17 On pages 520-1. A version of the case of the FORGETFUL THINKER appears in *Action and its Explanation*, 149-50. But the version in 'Mental Overpopulation' is more fully developed and thus a better target.

18 'Mental Overpopulation and the Problem of Action,' 520.

19 *Ibid.*, 520

Either (a) he conjectures or searches his memory, without every [*sic.*] intending to do so; *or* (b) before he conjectures or searches yet again, he does intend to do so, and so either (b)(i) he has two simultaneous but contrary intentions, to think about the puzzle or the name *and* not to think about it, or (b)(ii) he changes his intention from an intention not to think about it, to an intention to think about it.²⁰

Ruben thinks that (a) is the correct description and that both horns of (b) are wrong. I think this is a mistake. And I believe that *if* Ruben is right about (a) being the right description, the FORGETFUL THINKER's thinking is non-actional. Additionally, Ruben's assessment of (b) is mistaken. If (b) is the correct description, then it makes sense to take an intention to be the cause and explanation of the mental behavior of the FORGETFUL THINKER. But (b) can also be described as involving a non-actional event if an intention is not afforded the proper causal and explanatory role.

I will start with (b). But I should first note why I am starting here and even taking up the task of showing that Ruben's assessment of (b) is mistaken since I claim that there are problems with accepting (a) as the correct description of what occurs in the case of the FORGETFUL THINKER.²¹ Considering both horns of (b) is important for making the strongest possible case against Ruben. Even if I am right about description (a), it is still important to address the alternative explanations of what occurs in the case of the FORGETFUL THINKER. This is so if for no other reason than to preempt any defensive moves that Ruben and those sympathetic to his views may offer.

So suppose the FORGETFUL THINKER intends not to think about the puzzle or the name he forgot. Regarding the first horn, (b)(i), which takes the FORGETFUL THINKER to have two contrary intentions, Ruben claims that the problem would merely be that the agent is irrational. Either course of action, Ruben notes, 'would be in breach of exactly one current intention.'²² He suggests that nothing about such a scenario would explain why the FORGETFUL THINKER feels worse after he commences the mental behavior he intended not to engage in. Neither

20 *Ibid.*, 521. John Bishop suggested a third interpretation of what occurs in the case of the FORGETFUL THINKER is that it is a case of *akrasia*. Of course, such an interpretation may also threaten the CTA given that it would be an instance of *akrasia* where the *akratic* action is not caused by any relevant intention. There would be a failure to act in a way that conforms to an agent's intention. Such an interpretation is challenging. But I will not consider it here.

21 Both John Bishop and an anonymous referee for this journal advised that I explain why discussing (b)-scenarios is not otiose.

22 'Mental Overpopulation and the Problem of Action,' 521

intention seems to be privileged, according to Ruben, 'since either course of action (or forbearance) would be in breach of exactly one current intention.'²³ But suppose that the intention not to think about the puzzle or the name is a higher-order intention (I offer a characterization of what such a state is like below). Ruben expresses general skepticism about the 'rather overblown' attribution of such mental states to agents. Ruben seems to think that the onus is on his opponent to supply a higher-order intention that could serve as a candidate cause of the FORGETFUL THINKER's 'thinking about what he has resolved not to think about.'²⁴ The problem is that those who want to posit such metamental states to explain scenarios like that of the FORGETFUL THINKER need to explain why we should expect that agents will always have such intentions right when they are needed (e.g., to account for cases like the FORGETFUL THINKER). Such a move strikes Ruben as *ad hoc* and 'motivated only by the need to save various theories from counterexample.'²⁵

Ruben too quickly writes off the possibility of a conflict between a higher-order intention and a first-order intention. The FORGETFUL THINKER's earlier intention to stop thinking was clearly a metamental intention that was formed because he reflected about his object level mental behavior and acquired the higher-order desire to cease thinking about what he was thinking about. The attitude takes as its object a mental state or event and is acquired as a consequence of metamental reflection — i.e., reflection about one's mental life. The effect of the acquisition of the desire would be the formation of a proximal intention to cease thinking about what he was thinking about.²⁶ After the cessation of thinking, the agent acquires a more general intention not to perform mental actions of a certain type — viz., thinking about what he intended to stop thinking about. The intention represents a commitment he has to avoid thinking about the puzzle or name. The desire and intentions were expressions of the agent's control that were responsive to feedback from

23 *Ibid.*, 521

24 *Ibid.*, 522

25 *Ibid.*, 522

26 The capacity for such metamental control by agents is discussed in the psychological literature on metacognition. For interesting discussions, see Cesare Cornoldi, 'The Impact of Metacognitive Reflection on Cognitive Control,' in *Metacognition and Cognitive Neuropsychology: Monitoring and Control Processes*, G. Mazzone and T. Nelson, eds. (Mahwah, NJ: Lawrence Erlbaum 1998), 139-59; and Diego Fernandez-Duque, Jodie A. Baird, and Michael I. Posner, 'Executive Attention and Metacognitive Regulation,' *Consciousness and Cognition* 9 (2000) 288-307.

his object level mental behavior. Both states are metamental states and, therefore, higher-order. When the FORGETFUL THINKER commences thinking again, assuming that he acquired a proximal intention at some point to think about the puzzle or the name, the acquisition of his intention is first-order, occurring at the object level. As soon as the intention causally kicks in, and if the agent is trying at that point to think, then he is performing an action. But he may still have the higher-order distal intention not to think. But he has stopped trying not to think about the puzzle or the name long before he starts to think. His distal intention only figures in his plans to keep himself from thinking. But he has failed to stop himself from thinking. This may be where his frustration and regret come from. He wants and intends not to do what he is doing, but his intention at that point is either causally inefficacious, or his first-order intention to resume thinking is a stronger motivational force, vitiating any motivational efficacy of the agent's higher-order intention.

Assuming that the foregoing is plausible, is the FORGETFUL THINKER performing both mental and metamental actions? No, he is not. Assuming that he is aware of both intentions, if he begins to perform the mental action of thinking about the puzzle or name he forgot, he is best described as trying to think while having a causally inert metamental intention not to think — he is certainly practically irrational because of the failure of his intentions to match up. But he cannot be trying to do both. That is psychologically impossible. I should say, his *consciously* trying to do both, knowing what he is trying to do in each case, is psychologically impossible.

But there is another description of what could happen in such a case if the metamental attempt fails and some behavior occurs at the object level. Suppose that when THE FORGETFUL THINKER formed the higher-order intention not to think about the puzzle or the name he did not actually try not to think about it by taking effective measures, like distracting himself. If that were the case, then he may best be described as having fallen prey at the object-level to the 'ironic processes of mental control' that are much-discussed in the literature on ironic process theory in social psychology.

Experiments performed by one of the leading researchers on ironic process theory, Daniel M. Wegner, and his colleagues have shown that unwanted thoughts will often arise in instances when agents attempt to control their thoughts — particularly when trying to suppress a thought.²⁷ The reason such processes of mental control are 'ironic' is that

27 Wegner discusses the findings in Daniel M. Wegner, David J. Schneider, Samuel R. Carter III, and Teri L. White, 'Paradoxical Effects of Thought Suppression,' *Journal*

it is not only difficult for agents to suppress a thought, but the attempt to suppress a thought makes agents 'especially inclined to become absorbed with the thought later on.'²⁸ In fact, Wegner notes, according to ironic process theory 'attempts to influence mental states require monitoring processes that are sensitive to the failure of the attempts and ... these processes act subtly yet consistently in a direction precisely opposite the intended control.'²⁹

In one experiment performed by Wegner and his colleagues, subjects were asked not to think about a white bear. Each subject was placed in a room at a table with a microphone and a bell. Each was asked to spend five minutes saying everything that came to mind into the microphone. At the end of five minutes, the experimenter came in, and asked each person to continue. However, each subject was told not to think about a white bear. If they thought about a white bear, they were to ring the bell and continue talking. Wegner writes that, 'On the average, people in the predicament rang the bell more than 6 times in the next five minutes and mentioned white bear [*sic.*] out loud several times as well.'³⁰ Some of the subjects were asked immediately after their session to continue for another five minute period, only they were asked to *think* about a white bear. Wegner notes that, 'These people became unusually preoccupied with white bear thoughts, ringing the bell 16 times and mentioning a white bear 14 times.... The people thinking about a white bear tended even to show an *acceleration* of white bear thoughts over time.'³¹ While we often fall prey to the ironic processes of mental control, we are not without means of exercising mental control. Wegner catalogues various means of exercising relatively effective mental control. These typically involve distracting oneself, among other things, and not merely resolving not to think a thought or just suppressing a thought.³²

I take it that, in the case of the FORGETFUL THINKER, if the agent's attempt to stop thinking about the puzzle or the name he forgot fails —

of Personality and Social Psychology 53 (1987) 5-13; Daniel M. Wegner, 'Ironic Processes of Mental Control,' *Psychological Review* 101 (1994) 34-52; and Daniel M. Wegner, *White Bears and Other Unwanted Thoughts: Suppression, Obsession, and the Psychology of Mental Control* (New York: Guilford 1994).

28 Wegner, *White Bears and Other Unwanted Thoughts*, 4

29 'Ironic Processes of Mental Control,' 34

30 *White Bears and Other Unwanted Thoughts*, 2

31 *Ibid.*, 4

32 See *ibid.*, chs. 2-6

due to a failure to take the proper measures to bring it about that he stops thinking about it — he will fall prey to the ironic processes of mental control like the subjects in Wegner's experiments. The thinker desires not to think about the puzzle or the name and tries to stop thinking about it, only the attempt fails. And if thinking about the puzzle or the name comes about in the way the white bear thoughts occurred in the minds of the people in the experiments, such thinking is hardly worth calling *actional*.³³ It is mere behavior. It is not action. This may be the best description of what happens in at least one sort of case where the FORGETFUL THINKER'S first-order mental behavior is non-actional.³⁴ I believe Ruben would agree that in such an instance what occurs is non-actional.

One more explanation of what happens if (b)(i) is true is worth considering.³⁵ Perhaps the case of the FORGETFUL THINKER is a case of forgetting an intention. In such a case, the agent formed an intention not to think. After some time, he forgot about his intention and acquired an intention to think about the puzzle or name. In this case, only one intention is present in the consciousness of the agent at the time he commences the action of thinking. In such a case, the agent is not obviously irrational because he is only aware of one intention. He may later remember his earlier resolve and intention not to think and become frustrated with himself. But he did not execute his later intention being aware of his earlier intention.

What about (b)(ii) — viz., the agent changes his intention from one not to think about the puzzle or name, to an intention to so think? Ruben claims that the FORGETFUL THINKER could not have 'dropped his intention not to think about the name or puzzle ... for otherwise he would have no reason for chastising himself for not remembering the intention

33 In conversation, Rebekah Rice noted that it seems that if the agent's thinking is like Wegner's subjects, then the activity would lack a goal; but it seems that the FORGETFUL THINKER's thinking is aimed at some end — viz., remembering a name or solving a puzzle. So the FORGETFUL THINKER'S mental behavior is different from that of Wegner's subjects. While this may be right, the etiology of the mental behavior of both the subjects and the FORGETFUL THINKER is similar in ways that suggest that both fall prey to the ironic processes of mental control. So in spite of the apparent purposefulness of the FORGETFUL THINKER's mental behavior, its etiology seems sufficiently similar to that of Wegner's subjects to warrant the comparison.

34 Of course, as Hugh McCann pointed out to me, it seems unlikely that the agent's thinking would be unintentional and non-actional if the agent continued to think about the puzzle or name for very long.

35 This was suggested to me by Hugh McCann.

he has at the crucial moment.³⁶ The FORGETFUL THINKER would not have done anything he intended not to do if he dropped his earlier intention, according to Ruben. Ruben writes, 'At most, he could be annoyed with himself for altering his intentions, but not for reconsidering the puzzle or trying to recall the name, and that simply does not seem [to be] what is going on.'³⁷ Ruben contends that the agent must still have the intention not to think. The FORGETFUL THINKER cannot be annoyed with himself if he no longer has the intention. His resolve not to think, even if forgotten, cannot be dropped if we are to make sense of the agent's frustration and regret.

There are two problems with Ruben's reply. First, it seems overly restrictive because he suggests that one cannot feel regret or be annoyed with oneself for what one is doing when there has been a change in one's intentions and plans. But perhaps in the case of the FORGETFUL THINKER he has adopted a personal policy not to think about these matters anymore, but still changes his intention. He can be annoyed with himself because he acquired a proximal intention and commenced acting in a way that runs contrary to a personal policy he had — viz., not to think about the puzzle or name any longer. The policy was represented in a policy-based intention he had. But he no longer has an intention that sustains his past resolve.

Second, Ruben's reply seems psychologically naïve because he seems to suppose that we cannot shift from being spectators to actors with respect to certain processes in our mental lives. But clearly one can be overcome with thoughts, being totally passive and a spectator, and then shift midstream to being an active agent with respect to some mental behavior. In the case of thinking about the puzzle or the name, the resolve of the FORGETFUL THINKER may not be strong enough. After a time he begins to think about what he earlier formed the proximal intention not to think about and still resolves not to think about. But he becomes overcome with the thoughts, or just caves in. But once he caves in, assuming that he acquires the proximal intention to think and attempts to think about it some more, he is now acting. Before that point he was not acting. And bringing the two points I have made together, the agent may change his intention once he begins to act, but he is still acting contrary to a policy he developed when he formed his previous intention and tried to stop thinking. He can be upset with himself for failing to do what he had intended earlier to do and acting contrary to a

36 'Mental Overpopulation and the Problem of Action,' 521

37 *Ibid.*, 521

policy he formed. He is not merely disappointed and annoyed with himself for changing his intention. Because of his earlier intention, he can be accurately described as annoyed with himself for thinking about the puzzle or the name he forgot.

This brings me to description (a). Recall that on this description of what happens, the FORGETFUL THINKER thinks about the puzzle or the name he forgot after having intended earlier not to do so. The FORGETFUL THINKER thinks about the puzzle or the name again without ever forming an intention to do so. This is Ruben's preferred description of what is going on. If the FORGETFUL THINKER's thinking is actional, then no intention causes and explains his thinking. Moreover, even if it is motivated somehow by some type of pro-attitude, one cannot provide a rationalizing explanation in terms of the pro-attitude, according to Ruben, because it is a spontaneous action. What is most important for my purposes is that if no intention causes and explains the FORGETFUL THINKER's thinking, then one can perform a mental action without it being intentional under some description and so a trying.

Depending upon how the FORGETFUL THINKER's thinking is caused and sustained it may or may not qualify as actional. I do not believe that the thinking in description (a) which involves no intention to think is actional. The agent in such circumstances is overcome with the thought. Such an agent is like the subjects in Wegner's experiments who begin to think about white bears when told to try not to think about them. Only in this case the FORGETFUL THINKER's attempt is a response to an inner command, but similarly it fails to be successful due to an unreliable mental control strategy. But even if the etiology of the mental behavior is different from the subjects in Wegner's experiments (perhaps because the thinking is somehow motivated but not rationalized by a pro-attitude) there is nothing actional about the relevant mental behavior that occurs. It may be behavior of a sort and not totally passive in the way that a mental state such as a belief or desire is. Something is occurring, and it passes before one's mind like a film. So also, thinking about the puzzle or the name comes over the agent, but the agent is a mere spectator with respect to the thinking — *so qua* agent, he is passive, although non-actional mental behavior is occurring within him and he is conscious of it. But if the FORGETFUL THINKER commences *trying* to think about the puzzle or the name the thinking is actional. But this requires that one form a proximal intention to think about it. As indicated earlier, this may occur even after the thinking has begun. The status of the process changes from non-actional to actional because of the shift in the agent's attitude toward the process and the effort that is exercised in attempting to think. But what Ruben describes, on its own, without trying, hardly seems like an action. It may be much more like an unwanted thought. It could be a case of a failure of mental control like the

cases discussed by Wegner and others working on ironic process theory. If the thinking is what the agent wants, there is still a failure of mental control because of the intention not to think, and the conative state that may motivate the thinking is not itself a reason for acting. If you wish, call what happens to the FORGETFUL THINKER if description (a) is correct mental behavior, but it is not actional behavior.

Of course, if mental behavior such as we get on description (a) is actional, then it seems that quite a bit of the mental behavior many would regard as non-actional is, in fact, actional. A host of thoughts, imaginings, beliefs, and other mental processes and states will be actions. Such a consequence is unacceptable and ought to be reason enough to reject Ruben's claims about mental action.

But suppose that the FORGETFUL THINKER's behavior is actional. It is characteristic of intentional action that the intention is displayed in activity by the selective direction of attention when acting — in this case, in consciously trying to solve a puzzle or remember a name.³⁸ Hugh McCann notes that 'we cannot actively concentrate on content that is already before the mind [in thought] without intending to do exactly that'³⁹ If the FORGETFUL THINKER is actively attending to what he is doing in thinking, then it seems that he means to be doing just that. That is, he intends to think. So his thinking is intentional. So unless Ruben wants to make the thinking occur beneath the radar of consciousness, it seems that there is no way that the FORGETFUL THINKER can think very long without his thinking being intentional. It would be an odd result, indeed, if such mental behavior could continue for long, with an agent having directed his attention towards the activity, meaning to do so, without the activity being intentional and, therefore, in some way caused and explained by the agent's intention.

IV Conclusion

Intentions are safe from Ruben's attempt at mental population control. Ruben's own theory of mental action avoids mental overpopulation but at a significant cost. I have shown that if we accept Ruben's favored description of the FORGETFUL THINKER scenario, we do not have a case of mental action. Ruben writes off the alternative explanations much too quickly.

38 Hugh McCann brought this feature of intention to my attention.

39 In McCann, *The Works of Agency*, 142

It is worth noting in closing that Ruben has also argued that intentions are not necessary causes and rationalizers of overt actions.⁴⁰ If intentions play an essential role as causes and rationalizers of mental action, then *a fortiori* we should expect similar results with respect to overt actions. I do not have the space here to argue at greater length in defense of this claim. However, it seems that *mutatis mutandis* the same reasons for thinking that intentions play an indispensable causal and explanatory role with respect to mental actions should lead us to conclude similarly about action *simpliciter*, and, therefore, overt action.

V Postscript: Higher-Order Intentions

I should note a difference between what I have proposed here and a proposal made by John Bishop that Ruben criticizes.⁴¹ If there is no difference, then in the case of conflicting intentions I fall prey to Ruben's arguments against Bishop.

Bishop claims, 'It is constitutive of being an intentional agent that one has certain standing higher-order intentions classifiable generally as *intentions to act in accordance with the canons of practical rationality*.'⁴² The higher-order intention plus an agent's practical judgment that it is best to perform some action cause the occurrence of the acquisition of an intention which has a causal role in the etiology of an action. In the case of deciding, the mental action being caused and explained would be the active formation of an intention by deciding (Bishop uses the example of forming an intention to wiggle one's ears). I will ignore the details of Bishop's account here because it is the general nature of the content of the higher-order intention that Ruben picks out as doing little work for those who endorse the CTA.

Ruben offers the following reply to Bishop's use of higher-order intentions.

Bishop offers no example of a higher-order intention, as far as I can see, which would rationalize an agent's forming an intention to wiggle his ears.... His higher-order mental states do not provide premises for a 'valid' practical syllogism that has something about the mental action to be rationalized as its conclusion. There

40 Ruben, *Action and Its Explanation*, 104-10

41 John Bishop, 'Naturalising Mental Action,' in *Contemporary Action Theory*, Vol. 1, G. Holmstrom-Hintikka and R. Tuomela, eds., 251-66, at 259-64. Ruben's discussion of Bishop is in *Action and Its Explanation*, 150-2.

42 *Ibid.*, 260

appears to be some sort of conceptual ‘gap’ between all the higher-order intentions in the premises that he mentions and the mental action ... to be rationalized in the conclusion, with no indications of the beliefs, desires, intentions, or whatever that would be required to bridge that conceptual gap between those (alleged) premises and conclusion.⁴³

There are further problems that Ruben points out for Bishop’s account. Much of it is not relevant for showing how my proposal does not fall prey to Ruben’s critique of Bishop and other higher-order accounts.

What Bishop has proposed is an account of higher-order intentions that are less like the higher-order intentions I discuss in responding to Ruben and more like Michael Bratman’s *self-governing policies*.⁴⁴ ‘These are policies concerning what weight, if any, is to be given to desired ends in motivationally effective deliberation.’⁴⁵ There are differences between Bishop’s higher-order intentions and Bratman’s self-governing policies, but each represents a goal an agent has to bring about a desired end when acting, viz., to act in accordance with the canons of practical rationality or to act to bring about certain more desirable ends as the upshot of practical reasoning. In either case, the goal is less specific than the goal represented in a proximal intention. While I find the views endorsed by both Bishop and Bratman appealing, what I have suggested in the case of conflicting intentions is different. In the case of (b)(i) I have argued that an agent may form a *proximal* higher-order intention to stop thinking. In the case of the FORGETFUL THINKER the intention represents the cessation of thinking *right now*. It is higher-order because it is metamental. The agent reflects on his mental activity and forms an intention to cease thinking. Whether or not this can reflect a self-governing policy or a more general distal intention to perform mental actions in accordance with the canons of practical rationality is of interest. What I have argued, though, is that in the case of (b)(i) of the FORGETFUL THINKER this is not what happens when an agent forms a higher-order intention. What the agent forms is a proximal intention to stop thinking now. This intention does give way to a more general distal intention that may reflect a self-governing policy an agent has for herself. However, the formation of the proximal intention is what does the work of stopping the first-order mental activity.

43 *Action and Its Explanation*, 151

44 See Bratman’s ‘Valuing and the Will,’ *Philosophical Perspectives* 14 (2000) 249-65, especially 258.

45 *Ibid.*, 258

Admittedly, in the case of (b)(ii), I have suggested that it is possible that an agent may have a standing policy that he has violated, experiencing regret and frustration for violating it. This was not a case of a conflict of intentions such as I offered with respect to (b)(i). And in that case the agent is irrational, failing to take into account the policies the agent has which are represented in his distal intentions.

Second, suppose the higher-order intention is a general, distal intention to only perform the act of thinking under certain circumstances. It seems that the FORGETFUL THINKER could still reason from such an intention, and the intention would be part of a rationalizing explanation of why he thinks when he does (just as the formation of the proximal intention would have a similar explanatory role as the consequence of some practical reasoning). So I am not sure that Ruben has shown that even a more general intention cannot play a role in rationalizing and explaining an agent's acting as she does.

More can be said about both the nature of rationalizing explanations and the role of mental causes like the formation of intentions in providing such explanations of action that Ruben takes up.⁴⁶ To say more will take me well beyond the scope of this paper. Suffice it to say that the account I give that affords a role to higher-order intentions in causing and explaining mental actions does not fall prey to the difficulties allegedly faced by Bishop's account which Ruben discusses.⁴⁷

Received July 2004

Revised July 2005

Revised March 2006

46 Particularly in *Action and Its Explanation*, 91-110.

47 I am grateful to Jesus Aguilar, John Bishop, Randolph Clarke, Zac Cogley, Alfred Mele, David-Hillel Ruben, Rebekah Rice, Jing Zhu, two anonymous referees for this journal, and especially Hugh J. McCann for their comments on earlier drafts of this paper. While their comments have helped me to avoid many errors, the mistakes that remain are entirely my own.

