

INTENTIONS AND ADVISORIAL RELATIONS

by

D.B. Mulqueen II

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

in Philosophy

at

The University of Wisconsin-Milwaukee

August 2014

ABSTRACT
INTENTIONS & ADVISORIAL RELATIONS

by

D.B. Mulqueen II

The University of Wisconsin-Milwaukee, 2014
Under the Supervision of Luca Ferrero

In this paper, I develop a novel account of the role of intentions in diachronic intentional agency that I call the *advisorial model*. §I begins by offering an analysis of advice based on agents' first and second-order reasons coupled with trust and deliberative authority. I then proceed to argue (§§I.iii-II) that the individual components of the analysis can be leveraged against executive agential states such as intentions. This provides a course-grained analysis of future-directed intentions that is refined by noting pertinent differences from advice. Along the way I detail two, which hinge on what I call the *accessibility* and *sustenance provisions*. I then use the resulting advisorial model of intentions supplemented by the provisions to develop a more robust picture of human agency that includes higher-order executive states such as *commitments* (§II.iii). The impetus for employing a comparative methodology that proceeds from simpler to more complex notions is twofold. I do so first, in order to characterize an anomalous phenomenon and second, in an effort to account for the temporal unity and self-governance that Michael Bratman claims is distinctive of human agency (§II.iv). Finally, §III

levels one specific criticism against his planning theory of action based on a pair of intuition pumps in §III.ii.

TABLE OF CONTENTS

O. Introduction.....	page 1
I. Other-Directed Advice, Intending, & Self-Advice.....	4
II. Sketching an Advisorial Model: Intention and Commitment.....	15
III. Bratman's Planning Model: An Objection.....	24
IV. Closing Remarks.....	29

0. INTRODUCTION

Imagine you are consuming copious quantities of the potent draught M and burgling local homes to feed your habit with the easy money. Suddenly—in a brief reprieve from the deranged throes of your addiction—you realize the error of your ways. You decide to stop robbing homes, check into a rehabilitation facility, and permanently quit M at some date in the relatively near future (say, for instance, in three weeks' time when a spot in the nearest rehab opens up). Then consider a second instance, where you are now a straight-laced, middle-aged accountant. Your childhood best friend has just dropped by and confessed with some remorse that he is addicted to M and robbing local homes to fund his habit. Seizing the opportunity provided by his fleeting moment of clarity, you sit your friend down and try to convince him that he should cease committing burglaries, check into a rehabilitation facility, and permanently quit drinking M. Persuaded by your reasoning, he agrees to check into rehab in three weeks when the first spot becomes available.

What differs between the two scenarios? For one thing, in the first instance you form an *intention*. Specifically, you form a complex future-directed intention¹ to quit M and stop burgling homes. In the second instance, you do not. Rather, you *advise* a second party to quit M and stop burgling homes (in effect, suggesting that party form and execute a complex future-directed intention). More interesting than these differences, however, are the striking parallels. The basic structure is the quite similar, aside from the addition of another party in the second scenario. Notice what

¹ It is standard in the current agency literature to distinguish between simple and complex intentions, and between present and future-directed intentions. I follow this useful and intuitive convention throughout.

happens if we alter the first case's wording ever-so-slightly and rearticulate your act of intention-formation in terms of a present self issuing counsel to a future self—the diminutive gaps close further.

If it seems strange to think of issuing advice to oneself, consider the fact that you do not currently have direct executive control over your future self. At best, you have direct executive control only over your present self. This is, however, precisely the same state that you will (hopefully) find yourself in when the relatively near future becomes the present.² If intending minimally involves deciding³ to act for some (perceived) reason and the action must be executed by a temporally distant self that is not currently under one's direct executive control, the analogy between intending and advising might actually run quite deep.

Precisely how deep this resemblance runs is a matter of some concern. If it turns out that the parallels are so tight that no substantive dissimilarities remain, the notion of intention will collapse into advice entirely. This would be quite an oddity. We rarely—if ever—form intentions and explicitly consider them advice to our future selves. The act of intending is hardly unfamiliar, and what it feels like to intend can be discovered via introspection. Yet on its own, this isn't particularly illuminating.

² I do not here intend to implicitly endorse either presentism or four-dimensionalism; the arguments should hold either way.

³ I use 'deciding' and 'decision' in a loose sense throughout the paper that is not meant to endorse any particular view about mental actions or the relationship between decisions and intentions generally. I am specifically interested in future-directed intentions, which are typically preceded by deliberations that are concluded by decisions.

In the ensuing pages, I want to explore the parallels between advisorial relations and intending to see exactly where they lead and what they offer. I'll pay particular attention to future-directed intentions and then to commitments. The exploration's ultimate goal is to better characterize a distinct phenomenon that Michael Bratman has described as lying at the very heart of diachronic agency.⁴ Intention occupies this crucial position, he contends, because it is the means by which agents procure temporal unity and diachronic self-governance: unique, defining features of human agency.⁵ While attempting to maintain that intending is in fact distinct from advising, I will argue that exploiting the similarities—and, of course, taking stock of any notable differences—offers the best available avenue of explicating an anomalous phenomenon. The methodology is fairly straight-forward; I begin with the simpler of the two notions (advising) and then proceed to leverage the various components of the analysis against the more complex and intricate (intention) in an attempt to move beyond bare appeals to phenomenology.

§I starts by examining the structure of advice issued by an advisor to a separate advisee, arguing that there are at least two types of other-directed advice, and that both exhibit pertinent similarities to intending (§§I.i-I.iii). §I.iv then considers and eschews the potential reduction of intention to other-directed advice or self-advice, and pins down one crucial difference. The second section more fully sketches an account of intentional agency based upon the analysis in §I. The

⁴ See his (1992, 1998, 1999a, 2000) for just a few examples.

⁵ Temporal unity is perhaps best understood as a special kind of diachronic continuity resulting from the exercise of agency that transcends mere endurance or perdurance of the agential object. Self-governance is essentially the intentional autonomy we take ourselves to have over our actions.

resulting model aspires to illuminate both the general diachronic nature of agency and the specific nature of intending. §II.ii nails down a second departure from the mold set by advice. In addition, the advisorial model will engender explanation of important higher-order executive states (§II.iii). The next section considers the leading contemporary account in the tradition descending from Donald Davidson's work in philosophy of agency—Michael Bratman's influential planning theory—as a competitor to the account offered in §II.⁶ Though I'll suggest that the advisorial model better captures deeper features of our agency, I will concede that it is impossible to decisively rule out the planning theory by appealing to ordinary exercises of intentional agency. The planning model's failure to account for a pair of non-standard cases, however, should render my account preferable to Bratman's. Finally, §IV offers some concluding remarks.

I. OTHER-DIRECTED ADVICE, INTENDING, & SELF-ADVICE

§I.i

On the most general level of description, all genuine advice is suggestion or counsel issued by an advisor to an advisee about what the advisee ought to do. Suggestion or counsel is paradigmatic of advice since: (1) advisors do not enjoy executive control over their respective advisees⁷; and (2) suggestion differentiates the counsel of advice from mere commentary, on the one hand, and stronger imperatives such as commands and orders on the other. A full and comprehensive analysis of advice must, however, move beyond the notion of suggestion and

⁶ Many others have similarly noted the importance of Bratman's work, which has—to borrow a phrase from Sergio Tenenbaum (ms, p. 2)—become “the new orthodoxy” in philosophy of agency.

⁷ Self-directed advice is an exception to this rule, and is taken up later in the section.

address the crucial role first of *reasons* and then of *trust* and *authority* in explaining the issuance, reception, and general function of advice. In both instances, distinguishing the perspective of the advisor from that of the advisee is vital.

Reasons are key to understanding advice because any advice that did not take there to be some *perceived* reason for the suggested action would be, by definition, nonsensical and unintelligible *as* advice.⁸ Thus we can likewise say of the relation between advice and reasons what Donald Davidson (1963, p. 685) says more generally of reasons and actions:

What is the relation between a reason and an action when the reason explains the action by giving the agent's reason for doing what he did? We may call such explanations *rationalizations* and say that the reason *rationalizes* the action [...] A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his

⁸ I've deliberately couched the discussion in terms of 'perceived' reasons to avoid debates about what sorts of things properly constitute authentic reasons. As Alfred Mele has noted, "Philosophers' intuitions about which items are reasons for action and which are not tend to vary in ways that are predictable given whether the philosophers are more concerned with the explanation of actions or with the evaluation of actions (or agents)" (2003, p. 83). Although I am interested in both and believe an advisorial model will enable explanations and evaluations, all that matters for the sake of the analysis at this stage is the fact that an agent *considers* there to be a reason. While in the remainder of the paper I follow Donald Davidson and others in treating both wantings and beliefs as constitutive of reasons, any such discussion should similarly be understood as referring merely to 'perceived' or 'apparent' reasons. Likewise, although I sometimes talk as though there is at least a conceptual or descriptive difference between motivating and normative reasons, no specific understanding of reasons is essential to the account of advising on offer/under consideration. We could make do equally well with accounts that take all reasons to be facts, or accounts that hold that all reasons are ultimately reducible either to motivating or to normative reasons with a little extra conceptual legwork. For Davidson's account, see "Actions, Reasons, and Causes" (1963). For a more in depth discussion of the distinction between motivating and normative reasons, see Michael Smith's *The Moral Problem* (1994, especially chapters 4 and 5) and Bernard Williams's (1981) discussion of internal and external reasons, which is closely related. In her *Kinds of Reasons* (2010) Maria Alvarez provides an account of all reasons as facts. For an account that takes all reasons to be normative, see Jonathan Dancy's *Practical Reality* (2010).

action—some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable.

Reasons rationalize advice in precisely the same manner.⁹ Namely, they reveal what the advisor thinks she sees in making her suggestion and whatever the advisee accordingly does or does not see in evaluating it. Consequently, reasons are equally important—though variant in function—from the vantage points of both advisor and advisee.

An advisor cannot issue authentic and sincere advice if she does not think there is a reason for the advisee to do what she has advised. Here reason has two senses. The advisor must believe there is a suggestion-independent reason for the advisee to act as counseled, else she would be unable to advise the action.¹⁰ Yet even in the event that she does think there is such a reason, it would nonetheless be futile

⁹ Although I do not want to become bogged down in debates regarding the nature of rationality, and hope to remain as neutral as possible on that front, it bears mentioning that I am most interested in thinking about accounts of rationality in terms of coherence, rather than responsiveness to reasons. While this is contentious it is not unmotivated and has a number of implications for the account of intentions on offer. Even a coherence theory of rationality is going to rely on an agent's responsiveness to reasons in its evaluations. But, importantly, since coherence accounts of rationality are rather flexible compared to their counterparts, the reasons under consideration need not be any particular sorts of reasons, so long as they are minimally reasons that the agent takes herself to have. The reasons being responded to hence do not have to be 'objective reasons' of which she may be unaware, permitting greater neutrality in evaluations of agents' behavior, while also allowing evaluators to hone in on what makes the most sense from the agent's perspective, by her own lights. This seems particularly relevant in an advisory model of temporally-extended agency, where one agent must play roles analogous to both advisor and advisee.

¹⁰ This might call to mind instances such as bullying where one individual 'suggests' (perhaps even using this diction) that another do something or other that there would not be a reason independent of the pseudo-suggestion to do. But these instances are really more akin to commands than advice, even if they manifest under the guise of suggestion.

and impractical of her to counsel the action if she did not additionally harbor the conviction that there was cause for the advisee to seriously consider her advice. In some instances, her confidence that she can present the independent reason(s) for the suggested action persuasively will suffice to generate her perception that the advisee ought to mull her counsel over in earnest. Yet not all advice explicitly cites the first kind of reason—nor need it—in which case the nature of relation the two bear to one another provides the reason for the advisee to take the advisor’s counsel seriously. Here advice again mirrors action more generally, insofar as, “many explanations of actions in terms of reasons that are not primary do not require mention of the primary reasons to complete the story” (Davidson, 1963, p. 688).¹¹ Hence in some cases the second sort of reason essentially amounts to an overt presentation of the first, while in others it is something entirely additional.

The fact that the first sort of reason does not always require direct expression highlights the prominent role that trust plays in advisorial relations.¹²

Trust is important from the perspective of the advisor in demarcating her specific

¹¹ While Davidson uses ‘primary reason’ as a term of art in a manner that I am not, the relevant point of overlap involves the implication that there are—in Davidsonian explanations—such primary reasons even when they are not mentioned, just as genuine advice implies that the advisor perceives there to be what I’ve called first-order, advice-independent reasons—even if they are not explicitly stated.

¹² Here I try to remain as neutral as possible about what trust actually is to avoid debates that would take the paper too far afield. Edward S. Hinchman’s “Advising as Inviting to Trust” (2005) provides an in depth treatment of the relation of trust to advising. The account on offer here does not follow his, though he also notes that advising, in some instances, is “just inviting someone to act on a reason he already has” (p. 355). I am indebted to his Spring 2013 epistemology seminar on self-trust and higher-order evidence for helping me to think about the general importance of trust. His “Trust and Diachronic Agency” (2003) also helped in initially thinking about advising and intention.

standing relative to a specific advisee, in a particular advisorial context. An advisor that realizes she occupies a position of trust in a given context will not need to (though she nevertheless may) cite the reasons why she believes an advisee ought to heed her suggestion; an advisor that realizes she does not occupy a privileged position of trust will accordingly need to (though she may still fail to) cite the reasons why an advisee ought to follow her counsel. In either case, the advisor can cite reasons in favor of the suggested action,¹³ and/or reasons why she merits the advisee's trust (say, for example, that she is an expert on the topic of advice). But trust is especially important from the perspective of the advisee, where it impacts reception of the counsel. When an advisee trusts an advisor, he will naturally treat her suggestion differently than when he does not. Trust seems to *add* something to advice.

§I.ii

The variety of ways in which reasons and trust impact the issuance and reception of other-directed advice (henceforth ODA) indicate that the relative strength of advice is scalar. On a coarse-grained analysis, there are at least two very different types of ODA at either extremity of the spectrum. I will refer to the two as *strong advice* and *weak advice*, respectively.

When an advisee trusts an advisor, that relation adds substantive force, of a particular variety, to her counsel. Here, however, we must tread carefully so as to avoid double-counting reasons. While the auxiliary support resembles the first-order reasons, it cannot be an additional reason of the same stripe as those directly

¹³ I am not here attempting to endorse what Mele (2007) calls an 'objective favorers' account of reasons. The favoritism referred to here may equally well be subjective.

underlying advice. Rather, trust endows counsel with *advisorial authority*, or—perhaps alternatively—presumes the counsel possesses and deserves this authority. Significant trust in an advisor thus amounts to a presupposition on the advisee’s part that there really are reasons to act as suggested. If this is sufficient to act upon, it is only because it is tantamount to a presumption that the suggestion-independent reasons are sufficient to act upon.¹⁴ Yet it cannot be a bootstrapped reason on top of those reasons.¹⁵ Instead, the type of reason that gives some counsel its additional force is closer in kind to what Luca Ferrero (following Joseph Raz) has called an *exclusionary reason*. These are second-order reasons, whereupon an “agent at the time of action is justified in acting directly out of the exclusionary decision-based reason only insofar as she is warranted in believing that, were she to consider the matter anew at that time (that is, independently of her past decision), she would come to the same conclusion” (Ferrero 2010, p. 2). Advice often generates an analog of second-order, exclusionary decision-based reasons that stem from an advisee’s conjecture that were he to consider the matter from his advisor’s more privileged position of authority, he would reach the same deliberative conviction that she has in her advice.

Counsel an advisee takes to derive from a trusted authority on a given topic—and which consequently gives rise to a second-order, exclusionary advice-based reason—accordingly falls under the marquis of strong advice. Occasionally,

¹⁴ In a similar vein, renowned boxing coach Cus D’Amato once quipped to a young pupil: “If you do it because *I said it*, you’re a fool. You gotta do it because *what I said was right* [...] all I do is remind you.”

¹⁵ For further discussion of the bootstrapping worry, see Bratman (1999a) and Ferrero (2010).

however, counsel from advisors without this authority also comes to constitute strong advice when the advice-independent reasons are presented persuasively enough to instill trust in an advisee. In these instances, the advice results in what will later be treated as exclusionary reasons because the advisee has actually considered the independent, first-order reasons himself and decided they are sufficient to act on. Typically, however, counsel that an advisee doesn't take to derive from a trusted authority falls under the heading of weak advice, and doesn't add an exclusionary reason in favor of the advised course of action. While weak advice might serve as a psychological prompt—causing the advisee to consider the suggested course of action on his own—if the first-order reasons implicitly underpinning the advice are not presented, the advice won't decisively tip an advisee's deliberations in one direction rather than another.¹⁶

Hence Finnegans' unreflective decisions to say fifteen Hail Marys in penance for missing Mass yet again *and* to water his wilting shamrocks with Guinness merely because Father Shannon has counseled him to do so are intelligible as paradigmatic instances of strong advice. Maureen's laughter, in turn, when kooky old Finnegans suggests that *she* water *her* wilting shamrocks with Guinness is no less intelligible; he is not an authority for Maureen—far from it—and his counsel constitutes weak advice. By basing the analysis on a combination of reasons, trust, and authority, we can thus account not only for the general aspects of counsel's issuance, reception,

¹⁶ Occasionally, counsel from an initially trusted advisor will come to constitute weak advice when an advisee's suspicions are aroused. This might occur when an initially authoritative advisor decides to make an overt presentation of her reasons, but fails to do so persuasively, or when an advisee learns relevant facts about an advisor that instill doubt (say, for example, that an advisor has some ulterior motive or significant bias on a given topic).

and function, but also for why it enjoys a specific strength in one context, while it falls entirely flat in another.

§I.iii

Notice that intentions are susceptible to a very similar analysis: their formation¹⁷ is explicable in terms of perceived reasons; their authority derives from or gives rise to the level of trust an agent does or does not have in her deliberations at the time of intention-formation; and later execution or lack thereof often depends on the strength of an authority-induced exclusionary reason. After comprehensive investigation of advice, the individual components can be utilized to explicate intentions in a kindred manner.

This is largely because future-directed intentions also *function* much like advice, prompting agents to act in accordance with their reasons at the appropriate times. Having arrived at the time of action execution, agents often treat their earlier deliberations and resulting intentions to φ as a sufficient basis for φ -ing, sans further deliberation. Ferrero has noted that a “distinctive feature of our agency is the ability to bind our future conduct by making future-directed decisions” and that, in many cases, “[d]ecisions appear to have *rational authority* over the agent’s future conduct. When the time of action comes, the agent is normally guided by no other rational consideration but her past decision” (2010, p. 1). Here intention operates like strong advice, having created a decisive exclusionary reason to φ . But while it

¹⁷ Mele (2009), among others, talks of ‘intention acquisition’ in lieu of intention formation, which has significant implications. While I prefer intention formation, especially when the focus is future-directed or complex intentions, I do not mean to imply that we never acquire intentions in a less active manner than ‘formation’ might imply, especially when they are present-directed and/or simple.

might be tempting to think that—at the very most—only one of the advising relations could purport to parallel intending in any noteworthy way, intention actually mirrors both. Agents do not always follow through with their intentions. Upon approaching the point of action execution, agents also frequently re-open deliberation, treating their earlier intentions as though they *do not* offer decisive exclusionary reasons to φ . In such cases, intention looks much more like weak advice.

§I.iv

Precisely how the parallels between intention and advice are construed, however, is quite important. If the similarities run so deep that there are no distinguishing features, then the notion of intention will collapse entirely into that of advice. Indeed, at this juncture it would be natural to wonder why we shouldn't think that intention isn't merely a particular form of advice. There is at least this difference between intention and ODA: when one intends, the agent forming and executing the intention are one and the same individual (though temporally extended), and importantly aware of it. But the agent's cognizance of this fact cannot by itself preclude a radical reduction of intention to advice if it turns out that intending just happens to be a specific—albeit accidental—self-directed instance of an advisorial relation.

Perhaps, then, we ought to treat intention as a form of self-advice. These two cannot be identical, however, since it is possible to advise oneself without meaning to heed the counseled course of action, i.e. without forming a cognate intention to act in the fashion one considers advisable. *Akrasia* often illustrates this when

accompanied by an internal monologue wherein one recognizes that by one's own lights, one ought not φ , yet φ 's regardless. While the same is true of ODA—one can receive counsel from another without forming a cognate intention—it is noteworthy here in virtue of the proposed reduction of intention to self-advice. As long as there is one example of self-advice that is not accompanied by an intention, the reduction to self-advice is blocked. Furthermore, while the agent that forms a future-directed intention is continuous with the agent that executes the action, the two are temporally separated. When one advises oneself, by necessity the advisee and advisor are contemporaneously identical at the moment counsel is issued. This contrasts sharply with ODA, where the advisor is distinct from the advisee both synchronically and diachronically. So, while present-directed intentions may be more akin to self-advice, future-directed intentions plainly share more common ground with ODA. Yet—based on the foregoing—intention clearly cannot be reduced to either. Though a simple subtraction of advising from intention would appear to deliver the distinguishing features of intention, pinning down the nature of the divergences is much trickier.

Nonetheless, I suspect the crucial difference is closely related to the very special 'non-observational' awareness of our own actions and intentions that Elizabeth Anscombe pointed to in *Intention* (2000). While Anscombe thought the character of this awareness properly constituted a unique kind of knowledge, Paul Katsafanas's (2006) recent forays into neo-Nietzschean allegations of systemic self-ignorance may be cause for some measure of reticence in going so far. All that matters here, however, is the fact that there is a clear qualitative disparity resting on

what I'll call the *accessibility provision*: namely, that each of us *feels* like we have a very particular sort of access to our respective histories, faculties, wantings, beliefs, shortcomings, intentions, and actions. Though this access is not unlimited and we are far from infallible, normally-equipped cognitive agents typically have the sense that they occupy rather privileged positions with regard to their own intellectual networks. The character of the disparity highlighted by the accessibility provision thus constitutes a key discrepancy. While the *structural* parallels between intending and advising are quite tight, there is a substantive difference in intention's phenomenology, and consequently its authority's *mode of transfer*. The analogy breaks down in its failure to capture this divergence, but it does not preclude the parallels with advisorial relations from providing a host of details about the psychological framework of intention when properly coupled with the accessibility provision. Nor does it prevent them from helping us to glean temporal unity or self-governance from the model. In fact, the added provision will aid in securing these.

The purpose of continuing to pursue the analogy between intending and advising then is threefold. First, we stand to gain structural insights into an anomalous phenomenon. Second, by wedding those structural insights to the accessibility provision, we can better understand the origin, force, and function of intention. Finally, by stressing any further divergences, we can continue to isolate the unique features of intention and build a robust theory of intentional agency.

II. SKETCHING AN ADVISORIAL MODEL: INTENTION & COMMITMENT

§II.i

Recall that if it seems odd to think of issuing advice to oneself, you do not ever contemporaneously have direct executive control over your future self. At best, you have direct executive control only over your present self. Restated, the worry seems to be that, “[a] successful theory of future-directed decisions must account for the distinctive rational guidance of decisions and show how they can be effective without being manipulative” and that, “it is unclear whether a *future-directed* selection can be truly effective in the mode of rational guidance. The fact that a present-directed selection is usually (and immediately) effective does not guarantee that selections can have a similar power over an extended temporal interval” (Ferrero, 2010, p. 1; p. 3). One goal of pursuing the correlation with advisorial relations is to thereby mollify such worries. An intention aims at some future action whose completion is not currently—but later will be—under one’s executive control. Thus an intention to φ is advisorial insofar as it mimics counsel to be moved to action by reasons present at its formation, but at a time and to an agent in the future.

§II.ii

The phenomenon of diachronically extended intentional action is broadly divisible into three stages, all of which the advisorial model must adequately account for in order to succeed. The first stage is comprised of the time immediately preceding and including the formation of an intention (t_1). Stage two encompasses the time between intention formation and intended execution of the intention,

during which the agent holds the intention (t_2). The third and final stage, intention execution, involves actually carrying out the intention and φ -ing (t_3).¹⁸

Much like advice, the onset of an intention's first stage is typically ushered into existence by an agent's perceived reasons. Hence the impetus for forming any intention is traceable to the agent's desires and beliefs about (a) what she wants and/or ought to do, (b) her current circumstances, and (c) her projections about prospective future states of affairs. If her perceived reasons are sufficiently moving, the agent will acquire a corresponding intention to φ . The initial formation's basis in the force of the agent's perceived reasons thus closely resembles an advisor's counsel to φ at the appropriate (later) time. Passage from the first to the second stage is accordingly marked by the closing of deliberation (at least for the present interval), and settling on a course of action for the future.

The second stage stretches from the moment of settlement to immediately prior to execution. Intending during t_2 is markedly similar to the state of an advisee that has received and endorsed counsel. This, after all, should come as no surprise since an advisee that accepts an advisor's suggestion will need to form an intention in order to effectively implement her counsel. The course of events at t_2 can then unfurl in several ways, all of which we can account for by harkening back to the accessibility-supplemented analysis of advisorial relations. The agent may never consider her intention again, proceeding directly to t_3 and attempting execution at the appropriate time in an almost ballistic fashion. Here endeavored intention execution is ballistic insofar as the agent's intention formation functions much like a

¹⁸ φ -ing can denote either a specific action, or some sequence of actions.

trigger after the pulling of which events simply unfold along the intended trajectory. The presence of a stalwart exclusionary reason analogous to those generated in cases of strong advice goes a long way towards explicating these instances. Post intention formation, the agent treats her earlier deliberative self as an authority whose decision is sufficient to act upon.

Like advice, however, the strength of intention is scalar. Rather than proceeding ballistically to t_3 , when the agent does not *implicitly* accept her earlier counsel by unreflectively continuing to project her future φ -ing she may also more reflectively continue to endorse her prior deliberative reasoning—that is, consciously reflect on, agree with, and feel moved by the reasons cited in her deliberations at t_1 . This neither necessitates nor rules out her rehashing those deliberations. Active reflection on her part might just take the form of her thinking that her earlier self was right, whereby she manifests trust in the authority of her earlier self and treats the intention similarly to strong advice. But reflection is also compatible with her reviewing her previous deliberations, reconsidering those deliberations, and perhaps even re-deliberating.¹⁹ Which of these occurs will correspondingly reveal whether she treats her intention in the same fashion as strong or weak advice. Even without any relevant changes, the agent may choose to entirely drop her intention, therein regarding her earlier self as if she were offering very weak advice. Whether or not she implicitly accepts her earlier self's decision whenever she is not actively reflecting on the happenings of t_1 (and accordingly continues to intend) can be discovered via basic inquiry: if asked by herself or

¹⁹ One might reconsider one's deliberations as such (e.g. on occasions when one feels some doubt about earlier deliberations), without actually re-deliberating.

another if she continues to intend to φ and the answer comes out in the affirmative, she accepts the analog of her earlier self's advice. Continued intending at t_2 then, is parallel to maintaining that one's earlier self is still presently issuing appropriate counsel about what to do later.²⁰

The third stage (when reached) offers the same options as the second, but with one important codicil, perhaps best illustrated by Michael Thompson's (2008, p. 91) house-building example: *intention sustains action*. Call this the *sustenance provision*. If the agent actually endeavors to φ at t_3 , *ceteris paribus* her forward momentum will be guided largely by her intention. This must be the case since—as Mele has noted in the face of claims to the contrary—intentions are importantly efficacious executive states such that, “whenever human beings perform an overt intentional action, at least one of the following plays a causal role in its production: some intention of theirs, the acquisition or persistence of some intention of theirs, or the physical correlate of one or more of the preceding items” and that, “there is powerful empirical support for the claim that the fact that an agent has consciously decided to *A* sometimes has a place in the causal explanation of *A*-ing” (2009, p. 145). While this is most dramatic in long, drawn-out sequences of action (such as house-building), it is also evident in instances of appreciably less complex future-

²⁰ Intention modification and execution postponement are interesting phenomena that have not been dealt with. After reflecting and reviewing, reconsidering, or re-deliberating, an agent might decide to modify her intention or postpone its execution to some later time. Here the agent has not fully disregarded her earlier self's 'advice'. This is evident based on her retention of the intention to φ . But in deciding not to φ at the time settled upon at t_1 , she is diverging from the analog of advice issued by her earlier self. The significance of this divergence will naturally vary depending on the specific intention.

directed intention (such as driving to the store after work to pick up ice-cream). As a result, in appealing to intention we simultaneously seek to account for the unity of an agent's actions. This highlights another possible divergence from the model provided strictly by advice: It is not clear that counsel itself is essential to sustaining action once that counsel has been accepted. Firstly, since counsel may be endowed with or impart an exclusionary reason that aids in the sustenance of action, and second, because acceptance entails the formation of a cognate intention to φ that can later do the requisite work.

The fact that the available possibilities during progression from t_1 through t_3 run a gambit from ballistic execution of an earlier intention to full re-deliberation is a benefit of the model. As Bratman points out, agents have limited intellectual resources and deliberative stakes are often high, necessitating action without reconsideration, (1999a; 2000; 2012). It is often simply impractical given the pressures of circumstance and associated costs to do anything other than act on the basis of earlier deliberations. But it is also reasonable when such pragmatic considerations do not hold (or, minimally, are not sufficiently moving in an agent's eyes) to reflect on earlier decisions and reconsider, potentially even re-opening deliberation prior to action. By allowing for reflection while accommodating concerns for mental economy, the model thus seeks to strike a balance that avoids hyper-intellectualizing intention and action.

To recapitulate, then, something will count as a future-directed intention according to the advisorial account if and only if:

- i). it is intelligible based minimally on an agent's (perhaps only perceived) reasons,
- ii). it is describable using the structure and taxonomy outlined in §§I-II and,
- iii). both the accessibility and sustenance provisions hold.

§II.iii

As an added windfall, treating advisorial intentions as building blocks also engenders explanation of other essential higher-order agential states. Though I am going to focus on what I'll call *commitments*, these are just one such example.²¹

Commitments are easily subsumed under the model as another instance roughly analogous to an agent advising her later self. There are two salient differences, however, between basic future-directed intentions and commitments. The first difference lies in the strength of the earlier deliberator's authority. Commitment implies something much stronger than basic intention. If the former analogy with advice is any guide, then there is a question first about the origin of commitment's strength, and then about the strength of the resulting mandates as such.

Commitment owes its strength largely to the second major difference, which involves generality and applicability. This is essentially a difference of scope. While basic future-directed intentions aim at specific actions or courses of action, commitments are a general, umbrella-like phenomenon.²² They issue counsel stemming from deeper concerns, pertinent to a broader array of instances.

²¹ Though it is common in the literature to describe many different agential states as 'committing', I am attempting to lay out a more technical sense of commitment here in order to pick out a specific executive state.

²² One possible way cashing this out is by appealing to the type/token distinction. While basic future-directed intentions aim at specific tokens, commitments aim at types. Thanks to Andrei Buckareff for pointing this out.

Accordingly, the reasons underlying commitments must be perceived to be stronger and less specific in application than those of intentions. The resulting state will also, therefore, typically be much stronger. Consider, for example, an agent's intention to be kind to her crotchety next-door neighbor when she sees him tomorrow morning, as compared to her commitment to be generally kind to others. Differentiating commitments from basic future-directed intention therefore involves examining the respective state's strength, breadth and depth of base, and scope of application. As such, *ordinarily*, commitments are accorded precedence over basic future-directed intentions. This does not mean commitments are not subject to reflection in the same manner of review, reconsideration, and re-deliberation. But it does mean that they will be more resistant to—and perhaps even reinforced by—the pressures of reflection in virtue of the deeper plane from which they issue and the more prolific range of instances to which they apply.

Thus construed, higher-order commitments bear some resemblance to Bratman's *policies* and Richard Holton's *resolutions*.²³ In spite of the superficial continuities among the three, however, there are several relevant differences. One might form a resolution to quit smoking on Friday, and make it a policy to couple dark shoes with dark slacks whenever possible. Yet one could not make it a policy to

²³ See, for instance, Bratman's (1999b; 2000) discussions of policies and Richard Holton's (2009) discussion of resolutions. While they have particular accounts of the notions that I am not specifically outlining here, the advisorial model being sketched is also equipped to explain them in the same fashion that it does intentions and commitments. Though Bratman holds that policies are just a specific breed of intention, Holton argues that resolutions must be a distinct state. The advising model necessitates neither view of commitments, though it would appear to be compatible with either. Whether or not commitment can be reduced to intention is, I take it, an open question that is irrelevant to the present purposes.

quit smoking on Friday (given both the act's particularity and required strength of state). Nor could one form a resolution to couple dark shoes with dark slacks whenever possible (given both the generality of state that is needed and relative lack of strength required). Commitment, on the other hand, is an umbrella-like state that is broad in scope *and* exhibits substantial strength. So if one were to plot out the respective terrain of intentions, resolutions, policies, and commitments, the resulting topography would track a map something like the following:

Strength Scope	-	+
-	Intention	Resolution
+	Policy	Commitment

Intention is the weakest in strength and narrowest in scope, while commitment is strongest and broadest. Resolution mirrors commitment's strength, but is—like intention—narrower in scope. Opposite resolution, policy is far-reaching—like commitment—yet also weaker, resembling intention. As I have depicted them, intention and commitment thus bookend resolutions and policies.²⁴

§II.iv

Utilizing the resources outlined above, it is possible to account for the characteristically human temporal unity (or special kind of diachronic continuity resulting from exercises of agency that transcend mere endurance or perdurance of the agential object) and self-governance (or intentional autonomy over our actions)

²⁴ Thanks to Luca Ferrero for pointing out that as described, the notions fit into a matrix like the one above.

that we desire. The coupling of lower-order and higher-order executive states (e.g. intentions and commitments, respectively) is essential. Like advice, both stem from an agent's perceived reasons, and their efficacy as executive states in turn depends on the trust an agent exhibits in her deliberative authority. The model meets practical demands by allowing agents to act out of prior decisions sans further deliberation. Yet it also permits reflection. Reflection meshes nicely with the accessibility provision (§I.iv), which helps describe an important phenomenological facet of human agency, the source of deliberative authority, and its resulting unique mode of manifestation. This was initially useful in distinguishing intention from advising, and of further use in more thoroughly characterizing intention. The accessibility provision helps separate commitments from intentions in a like manner. Naturally, the strength of a commitment is often bolstered by access to a past history of reflection. But access and reflection do more than just this. Together, they also begin to explain how a specific intention might be derived from a more general commitment, and how agents might arrange their respective commitments and intentions into hierarchies that link their actions over time and govern them. The picture would still be incomplete, however, without a second stipulation. The sustenance provision—based on agents' perceived reasons, trust in their deliberative authority, and reflective access to their intellectual networks—explains how complex, temporally-extended, autonomous and genuine action (in stark contrast with mere behavior such as muscle spasms) crucially relies on executive states such as intention and commitment for guidance and subsistence. The

conjunction of the model's many elements thus provides the requisite fodder for an account of temporal unity and self-governance.

III. BRATMAN'S PLANNING MODEL: AN OBJECTION

§III.i

According to Bratman, at the most general level, "Intentions are plan states" (2010, p. 9). We must treat intentions as plans, Bratman holds, because:

- 1) [A]ppeal to such planning structures allows us to articulate basic features of intention and decision, features that allow us to distinguish such phenomena both from belief and from ordinary desire (1999b, p. 5).
- 2) We do not simply act from moment to moment. Instead, we settle on complex—and, typically, partial and hierarchically structured—future-directed plans of action, and these play basic roles in support of the organization and coordination of our activities over time. (2007, p. 26).

I have tried to show that an advisorial model can likewise delineate the notion of intention, take stock of our resource limitations, provide the necessary structures for cross-temporal coordination, and hierarchically structure our activities by employing counsel-like notions of intention and commitment. While I also believe that the advisorial model does a better job capturing deeper, less superficial features of our agency, it is nevertheless impossible to decisively rule out Bratman's planning theory of intention merely by appealing to standard exercises of diachronic agency; the planning theory is equipped with conceptual tools that allow it to address them. The planning theory's failure to account for two non-standard cases (§III.ii), however, ought to tip the scales in the advisorial model's favor if it can succeed where Bratman's planning theory does not.

Bratman contends that stability is one of intention's core features. Stability both exists as a psychological feature of our executive state of intending and helps give rise to distinctively diachronic norms of practical rationality.²⁵ Specifically, stability is linked to norms of non-reconsideration. Bratman maintains that, "plans [intentions] have, in this sense, a certain stability: there is, normally, rational pressure not to reconsider and/or abandon a prior plan" and that, "plans [intentions] introduce characteristic forms of practical reasoning" which impart their own norms to intention (2007, pp. 26-27).

Intentional agency is important to us, in turn, due largely to our concerns for temporal unity and self-governance. The need for temporal unity and self-governance combines with the existing psychological feature of stability to give rise to the norm of non-reconsideration. Together, the stability of intention and its accompanying norms are thereby supposed to *actually* deliver these to us (Bratman, 1992, 1998, 2000). So Bratman's nuanced planning theory includes the following four claims, which are successively embedded in one another much like Russian nesting dolls.

1. Intentional agency has the feature of diachronic stability.
2. We care about intentional agency (partially) for the sake of temporal unity and diachronic self-governance.
3. 1 and 2 combine and give rise to distinctive norms of diachronic rationality (one of which is non-reconsideration).
4. Stability and a norm against the reconsideration of basic future-directed intentions do, in fact, provide us with the desired temporal unity and self-governance.

If either stability is not a psychological feature of intention, or stability and a norm of non-reconsideration are not necessary for securing temporal unity and

²⁵ See his (1992; 1998; 2000; 2010; 2012), for example.

diachronic self-governance, then the planning theory has a serious flaw. Though I've already shown that we can glean temporal unity and self-governance from the advisorial account, more detailed examination of Bratman's theory is merited for the sake of comparison.

§III.ii

The first order of business is addressing whether or not diachronic stability is in fact a core psychological trait of intentions, since the normative link to non-reconsideration vanishes if it is not. Dismissing Bratman on this front will not be easy. Some measure of diachronic stability must exist in order for our executive states to be at all efficacious. But we might ask precisely how much stability is required specifically of intention. Bratman takes it that intentions are stable to quite a considerable (albeit somewhat vaguely specified) degree. Intentions enjoy their "characteristic stability" since "some form of stability is needed for intentions and plans to play their cross-temporal organizing roles" and because they are "involved in the forms of cross-temporal integrity and self-government that we normally value" (2007, p. 10). The last point is particularly important; on Bratman's model, the defining characteristic of intentions is the manner in which they serve as plans for the future (1999a, 2010). We care about stability and non-reconsideration for the sake of unity and self-governance—not the other way around.

Consider the following not so far-fetched cases.

1. *Completion*: Colleen frequently forms intentions at t_1 to φ at t_3 , specifically regarding to the completion of school assignments due at t_D . But, as it turns out, when or right before t_3 arrives, Colleen frequently fails to execute her intentions to φ . It isn't true of Colleen that she *never* completes her work. She almost uniformly completes her homework *after* t_3 , at some later time t_L , nearer its due-date. Despite her best efforts, she is not able to do so until a deadline looms directly overhead.

When such a deadline is creeping over the nearby horizon, however, Colleen does effectively complete her assignments. She is generally dedicated to finishing her work by t_D , and to certain standards that she is able to meet. It is only when she (rather frequently) intends to do her work ahead of time at t_3 (and perhaps for good reasons) that she is not able to execute her intentions. Such intentions do not ever seem—in Colleen’s case—particularly stable. Furthermore, she is willing to reconsider any basic future-directed intentions to get her work done early, for almost any reason. Even the slightest temptation causes Colleen to re-deliberate prior to the last possible moment in which she might complete her work.²⁶

2. *Spontaneity*: Seamus is what some people might call flighty. He re-deliberates, reconsiders, revises, abandons, and acquires new future-directed intentions in a manner consistent with the blowing of the wind. In fact, he attributes a particularly high value to living in this fashion. Seamus places a very high premium on spontaneity. He takes it that his spontaneity is a defining feature of his character that structures his life. And so after forming the intention at t_1 to φ at t_3 , he frequently reopens deliberation whenever something else comes up in between. He forms the intention on Monday to exercise on Wednesday, but re-deliberates and goes to the beach, to lunch, or to the movies instead almost every time the opportunity affords itself. Even in rare instances where Seamus does decide to follow through and go to the gym, he minimally reopens deliberation first. Thus his intentions are, in addition to being far from stable, hardly resistant to reconsideration.

While Colleen frequently fails to execute her more basic intentions, she nonetheless behaves in a fashion consistent with her over-arching dedication. This dedication effectively directs her behavior, producing tangible, consistent results visible across her academic record. But her more basic future-directed intentions pertaining to her assigned schoolwork are rarely stable. If she feels any normative pressure not to reconsider such intentions, this does not prevent her from doing so prior to the time period immediately preceding her assignments’ due-dates.

Seamus’s behavior is erratic at best. But this is precisely the point. His dedication to spontaneity entails that he is willing to reopen deliberation about future-directed

²⁶ I have deliberately avoided calling this a case of procrastination, since ‘procrastination’ has increasingly become a philosophical term of art, with some maintaining that no genuine intention is actually formed in instances of procrastination.

intentions whenever a seemingly enticing possibility presents itself. When we examine his behavior over time, it looks like his choices and resulting actions are both governed and unified by this dedication in a manner that renders them intelligible.

Bratman's analysis forces him to charge both Colleen and Seamus with irrationality, since "stability of intentions is part and parcel of [...] overall skill at practical reasoning. To know how to reason is in part to know *when* to reason (or not reason) about what" (1999a, p. 66). Despite the charge's intuitive pull, it is worth noting that—minimally—Colleen and Seamus are subjectively and instrumentally rational *on some level* insofar as they successfully take means to acquiring their desired ends.²⁷ Bratman might accordingly avail himself of his two-tiered theory of rationality (1999a, p. 64), whereupon both would appear irrational at the more basic first-tier, whilst possibly rational on the higher second-order. This possibility, however, raises more troublesome questions about the existence of a new third-tier irrationality in light of the inconsistency between the first and second.²⁸

Much more significantly, as these cases illustrate, Bratman's account is wrong on two fronts: stability is often not a psychological feature of many intentions and a norm against the reconsideration of at least some intentions—even whole classes of intentions—is not essential for obtaining temporal unity and self-

²⁷ While they may be objectively irrational, this would be an odd consequence if they are unified and self-governing. Furthermore, objective rationality is expressly *not* Bratman's interest (1999b, p. 9).

²⁸ Bratman mentions a similar line of objections raised against Harry Frankfurt in the introduction to his (1999b) and in (1999b, Chap. 10).

governance. Instead of intentional stability and non-reconsideration, the extent to which Colleen and Seamus are unified and self-governing appears to be the result of their higher-order states of dedication, which look like paradigmatic instances of commitment: a psychological mechanism probed earlier in §§II.iii-II.iv for its ability to provide temporal unity and self-governance.

In response, Bratman might likewise appeal to his notion of policy, arguing that it is capable of handling such cases. This seems *prima facie* plausible, but is problematic for two reasons: First, given the central role Bratman allots stability and non-reconsideration, he would effectively be giving up much of his account of future-directed intentions if he opted to jettison the pair. Second, the fact that he holds that policies are just a specific breed or “a certain kind of intention” makes it impossible for him to sacrifice the stability and non-reconsideration of basic intentions while hoping to retain them in policies (1999b, p. 7). Hence Bratman is blocked from taking this tack.

IV. CLOSING REMARKS: A BRIEF RECAPITULATION

The driving force of the foregoing is the conviction that intentions, and intentional action, are anomalous. I began by trying to draw out a parallel between intending and advising, and proceeded to provide an analysis of the simpler phenomenon (advising) first. In broad strokes, that analysis was based on agents’ perceived reasons, the trust advisees do or do not exhibit in the deliberative authority of their advisors, and the second-order advice-based exclusionary reasons that factor into their relations. I next applied the characterization of advice to intention, arguing along the way that the accessibility and sustenance provisions

distinguish intention from advising. The resulting advisorial model of diachronic agency was shown to similarly accommodate higher-order executive states such as commitment. I then argued that the full picture and the conjunction of all its elements explained the possibility of temporal unity and agential self-governance. Finally, I conceded that Michael Bratman's planning theory is able to similarly account for typical exercises of diachronic agency, but contended that it is ill-equipped to handle at least two non-standard instances, rendering the advisorial model preferable.²⁹

²⁹ I owe an enormous debt to Luca Ferrero for numerous meetings, discussions, and rounds of comments on earlier drafts, without which many of the ideas presented herein would look nothing like they do. I also owe a huge debt to Andrei Buckareff for detailed comments on an earlier draft.

References

- Alvarez, M. (2010). *Kinds of reasons: an essay in the philosophy of action*. Oxford: Oxford University Press.
- Anscombe, G. E. (2000). *Intention* (2nd ed.). Cambridge, Mass.: Harvard University Press.
- Bratman, M. E. (1992). Planning and the stability of intention. *Minds and Machines*, 2(1), 1-16.
- Bratman, M. E. (1998). Toxin, temptation, and the stability of intention. *Essays for Gregory Kavka* (pp. 58-90). Cambridge, U.K.: Cambridge University Press.
- Bratman, M. E. (1999a). *Intention, plans, and practical reason*. Stanford, Calif: CSLI Publications.
- Bratman, M. E. (1999b). *Faces of intention: selected essays on intention and agency*. Cambridge, U.K.: Cambridge University Press.
- Bratman, M. E. (2000). Reflection, planning, and temporally extended agency. *Philosophical Review*, 109(1), 35-61.
- Bratman, M. E. (2007). *Structures of agency: essays*. Oxford: Oxford University Press.
- Bratman, M. E. (2010) Agency, time, and sociality. *Proceedings of the American Philosophical Association*, 84(2), 7-26.

- Bratman, M. E. (2012). Time, rationality, and self-governance. *Philosophical Issues*, 22(1), 73-88.
- Bratman, M. E. (2013). The fecundity of planning agency. *Oxford studies in agency and responsibility* (pp. 47-69). Oxford: Oxford University Press.
- Dancy, J. (2000). *Practical reality*. Oxford: Oxford University Press.
- Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, 60(24), 685-700.
- Ferrero, L. (2010). Decisions, diachronic autonomy & the division of deliberative labor. *Philosophers' Imprint*, 10(2), 1-23.
- Hinchman, E. S. (2003). Trust and Diachronic Agency. *Nous*, 37(1), 25-51.
- Hinchman, E. S. (2005). Advising as inviting to trust. *Canadian Journal of Philosophy*, 35(3), 355-386.
- Holton, R. (2009). *Willing, wanting, waiting*. Oxford: Oxford University Press.
- Katsafanas, P. (2006). Nietzsche on agency and self-ignorance. 2006.
- Mele, A. R. (2003). *Motivation and agency*. Oxford: Oxford University Press.
- Mele, A. R. (2007). Reasonology and false beliefs. *Philosophical Papers*, 36(1), 91-118.
- Mele, A. R. (2009). *Effective intentions: the power of conscious will*. Oxford: Oxford University Press.

Smith, M. (1994). *The moral problem*. Oxford: Blackwell Publishing.

Tenenbaum, S. (ms). Reconsidering intentions.

Thompson, M. (2008). *Life and action elementary structures of practice and practical thought*. Cambridge, Mass.: Harvard University Press.

Williams, B. (1981). Internal and external reasons. In B, Williams (Ed.), *Moral Luck*

(pp.

101-113). Oxford: Oxford University Press.