



BIROn - Birkbeck Institutional Research Online

Garnett, Michael (2023) Ghost-written lives: autonomy, deference, and self-authorship. *Ethics* 133 (2), pp. 189-215. ISSN 0014-1704.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/48959/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html> or alternatively contact lib-eprints@bbk.ac.uk.

GHOST-WRITTEN LIVES: AUTONOMY, DEFERENCE, AND SELF-AUTHORSHIP*

Certain forms of practical deference seem to be incompatible with personal autonomy. I argue that such deference undermines autonomy not by compromising the governance of an authentic self, nor by constituting a failure to track objective reasons, but by constituting a particular *social relation*: one of interpersonal rule. I analyse this social relation and distinguish it from others, including ordinary relations of love and care. Finally, I argue that the particular form of interpersonal rule constituted by dispositions of practical deference in itself entails a partial, though not total, disconnection from the good of a self-authored life.

Theories of personal autonomy often struggle to make sense of internalised oppression, the phenomenon whereby people are rendered unfree or nonautonomous by internal, psychological forces in ways that help to preserve and reinforce structures of social injustice. The most recurring and widely discussed putative example of this in the recent autonomy literature has been that of the *Deferential Wife*:

She buys the clothes *he* prefers, invites the guests *he* wants to entertain, and makes love whenever *he* is in the mood. She willingly moves to a new city in order for him to have a more attractive job, counting her own friendships and geographical preferences insignificant by comparison. She loves her husband, but her conduct is not simply an expression of love. She is happy, but she does not subordinate herself as a means to happiness. She does not simply defer to her husband in certain spheres as a trade-off for

his deference in other spheres. On the contrary, she tends not to form her own interests, values, and ideals, and when she does, she counts them as less important than her husband's... she is quite glad, and proud, to serve her husband as she does.¹

Many share the intuition that people like the Deferential Wife lack full freedom or autonomy. Yet since her deference may be uncoerced, and since the internalised norms from which it springs may run extremely deep, this intuition is notoriously difficult to vindicate.

In this paper I suggest a new strategy for doing so. I provide an account of practical deference as distinct from ordinary attitudes of love and care, and argue that such deference is inherently incompatible with autonomy. Then, adapting a familiar distinction, I use this account to argue that while a person's *choice* of a deferential relationship may be perfectly autonomous, *what is thereby chosen* is necessarily a life of diminished autonomy. That is, deferential lives necessarily lack a degree of substantive or first-order autonomy, and so may be fairly subject to perfectionist criticism, even in cases in which they have been autonomously chosen and are thereby immune to procedural criticism. Finally, I show that such a loss of first-order autonomy suffices for a partial, though not total, disconnection from the good of a self-authored life. Those who willingly allow others to determine their choices in deferential ways, I suggest, live lives that are not so much *self-authored* as *ghost-written*. The result is an analysis of the Deferential Wife case that enables us to make sense of the peculiar way in which deferential choices can at the same time appear both to be, and not to be, exercises of autonomy.

The argument proceeds in the following way. Section I sets out the deep theoretical problems facing attempts to understand how people like the Deferential Wife necessarily lack

some degree of freedom or autonomy. Section II attempts to make headway with these problems via an account of practical deference, understood as a disposition to treat the desires of certain others as weighty, content-independent reasons for action. Section III shows how practical deference in this sense can be distinguished from deference to experts, and Sections IV and V show how it can be distinguished from attitudes of love and care. Section VI draws out some further implications of the proposed account of deference and rounds out its defence. With that account in hand, then, Section VII clarifies the strictly first-order nature of deference-inconsistent autonomy, and argues that deferential choices can nevertheless be fully autonomous at the second-order level. Section VIII draws out some normative implications from this, arguing that self-authorship requires both first- and second-order autonomy and that, therefore, people like the Deferential Wife necessarily fail to fully realise this traditional liberal good. Section IX concludes the argument.²

I. THE DILEMMA

If the case of the Deferential Wife is to help us to think clearly about the phenomenon of internalised oppression, it must be supplemented with some artificial assumptions. In particular, we must stipulate that the Deferential Wife does not act under external coercion: that she defers to her husband's desires neither under threat of violence, nor under threat of an economically and reputationally ruinous divorce—nor, even, because deference represents the best of an excessively bad option set. That is, we must imagine that the Deferential Wife acts, in at least one important sense of the term, *voluntarily*.

These stipulations are unrealistic: many, probably most, women who commit to unequal romantic relationships do so because these constitute the best of a bad bunch of options given the way those options have been structured by social norms.³ That is, many women confront a social world wherein certain bads, such as deference, have been made instrumentally necessary for the acquisition of important goods, such as social standing or economic stability. Such contexts of unjustly restricted choice may be taken to compromise the freedom and autonomy of the choices made within them. However, they are not best understood as modes of *internalised* oppression. Compare: a mugger may make it instrumentally necessary for a person to choose a bad (such as handing over her purse) in order to obtain a good (such as not getting stabbed), and the resulting choice may not be fully free or autonomous for this reason. Yet we would not think of the mugger's victim as here contending with her own internalised oppression; to the contrary, the problem she faces is wholly external. The same applies in the structurally parallel case of the woman who grudgingly accepts her deferential role as the only available route to material security: her limited options directly constitute an unjust restriction of her *external* freedom, not of her internal freedom.

By contrast, genuine cases of internalised oppression are those in which the relevant bad, such as deference, has also come to be accepted or desired for reasons not grounded merely in a person's awareness of their deficient options. This includes, centrally, the case of the person who actively embraces her deferential role, who regards it as a positive ideal, or as a matter of social, moral or religious duty. Such a person would, we are to assume, choose deference even in cases where it is not necessitated by the specific external constraints mentioned in the previous paragraph. Insofar as we are inclined to regard such a person as oppressed (a controversial designation, the appropriateness of which is a central topic of this paper), we must therefore locate

that oppression not in the world she currently faces and the options it provides, but in the values and ideas that she has come to have: either in their content, or in the processes by which she has come to have them. *This* is the form of oppression that I take to be picked out by the qualifier ‘internalised’.⁴ And while internal and external oppression can be difficult to separate in practice, being both interconnected and mutually supporting, it is in order to separate them theoretically, and thereby to isolate these internalised (and, to my mind, theoretically most difficult) aspects of oppression, that we must unrealistically stipulate for the purposes of this paper that the Deferential Wife acts voluntarily in this sense.

Given this, wherein might lie her lack of autonomy? Certainly not in the domain of negative freedom, seeing as we are assuming that she suffers no coercion or physical prevention.⁵ Nor in the domain of republican freedom (at least as that is now commonly understood), for the same reason: her husband may also lack the social and legal capacity to obstruct or to coerce her arbitrarily.⁶ Instead, it would seem to lie within the broad domain of ‘positive freedom’: the internal province of psychological self-government. In deferring to her husband’s desires as to what she should do, it seems, she fails to govern herself. But here we encounter a problem.

The problem takes the form of a dilemma. Accounts of psychological autonomy either analyse individual self-government in terms of governance by the self, or they don’t. If they do analyse self-government as governance by the self, then they can’t explain the Deferential Wife’s lack of autonomy, since it’s perfectly possible that her deferential attitude springs from her deepest self. If they don’t, then they bear no clear conceptual connection to the idea of self-government, and so don’t look like accounts of autonomy at all. Let me explain.⁷

Psychological self-government is standardly taken to require government by the self. On many views, that is, autonomy is thought to be primarily a matter of one's self—typically, one's real or true or authentic self, or that set of attitudes with which one identifies or stands behind, or with respect to which one is truly active—being in charge of one's behaviour. Yet while the Deferential Wife's deferential motivations may be external to her self so considered, there's no general reason to think that they must be. And were they to be internal to her self—were they to go, as it were, *all the way down*—then her deferential attitudes would be manifestations of, and not obstacles to, government by her self, and hence manifestations of her autonomy.

Indeed, no prominent modern theory of true selfhood is capable of systematically excluding deferential attitudes from the domain of the self. Note, for instance, that a disposition of deference need be neither unreflective nor uncritical; to the contrary, it might express settled ideas about one's proper role-based duties. In this way, the Deferential Wife might endorse and be satisfied with the first-order motivations from which she acts, and these motivations might themselves be sensitive to her judgements about her reasons.⁸ Moreover, inasmuch as her deferential values do lack a kind of reasons-responsiveness by virtue of being foundational parts of her character, they need be no less reasons-responsive—and no more alien—than any other possible set of bedrock values.⁹ Finally, deferential values may be acquired by means of the very same processes of education and socialisation by which people acquire fundamental values in general. Hence there need be nothing about the history of her deferential disposition, nor about her current attitude towards its history, that formally distinguishes it from any other disposition.¹⁰

Moreover, there's a deep reason why modern theories of selfhood are systematically incapable of excluding deferential attitudes—or, indeed, any set of attitudes characterised in terms

of their substantive content—from the domain of the self. It lies in the fact that such theories are generally concerned to avoid ‘transcendental’ or ‘overly metaphysical’ conceptions of the self, something they seek to do by grounding authentic selfhood in procedural or structural features of our contingent psychologies (such as the nature of, or relationship between, our desires, beliefs, values, or patterns of reasons-responsiveness). Yet human psychology is so plastic, and people so diverse, that there can be no plausible relevant limits on the contents of the attitudes that might successfully play these structural roles. We may easily imagine, for instance, that the Deferential Wife feels profoundly identified with her subservient disposition, and that she values it, believes it to be rationally justified, and so on; and unless we are to invoke a conception of her true self wholly decoupled from these contingent aspects of her actual psychology, we must allow that the boundaries of her self may genuinely extend to include her deferential attitudes in this way.

So the conclusion that deferential attitudes can be part of a person’s deepest self looks unavoidable. If autonomy is essentially a matter of rule by the self, then, the Deferential Wife needn’t lack autonomy. That is, no theory of autonomy that takes seriously the putative connection between autonomy and rule by the self can place substantial limits on the possible contents of one’s autonomous attitudes, because there can be no substantial content-based limits on which attitudes might genuinely count as a person’s own. Call this *the problem of the plastic self*.

Some are happy to accept this result: that, so long as the Deferential Wife acts reflectively, for coherent reasons, and so on, she may be perfectly free and autonomous. Later I’ll argue that this response captures an important part of the truth, but not all of it. First, however, I address those inclined to the view that the Deferential Wife can’t be fully free or autonomous *even if* she meets the relevant procedural requirements. For these other theorists, the fact that procedural approaches

can't account for the Deferential Wife's lack of freedom or autonomy is a decisive strike against them. In light of this, such theorists typically argue that we must abandon the analysis of autonomy as rule by the self in favour of a more substantive, content-based view.

For example, some have argued that the Deferential Wife's nonautonomy lies in the fact that her deferential attitudes are grounded in values that are objectively false, and that in acting she thereby fails to track the objective practical reasons.¹¹ Importantly, this kind of account needn't proceed via ideas about the nature of the authentic self. The thought isn't supposed to be that one can be *identified* only with values that are objectively correct. Indeed, proponents of content-based approaches have generally wished to avoid the implication that their accounts are offered as analyses of the authentic self. The idea, instead, is that possession of certain attitudes undermines one's autonomy directly: that the Deferential Wife lacks autonomy in virtue of subscribing to incorrect norms, say, even if these incorrect norms indeed form part of her true self and so do nothing to compromise its governance.

Yet this brings us to the second horn of the dilemma. The problem now is that, in rejecting the theoretical framework of the ruling self, this approach threatens to sever all connection to the underlying idea of self-rule or self-government. It invites the objection that it has merely changed the subject, and is no longer a theory of autonomy at all. That is, proponents of content-based approaches must do more than simply point to the presence of attitudes with defective content. They must also explain how such defective content amounts to a failure of self-government. Put simply, they must explain how a condition that *ex hypothesi* has no bearing on the question of whether one is governed by one's self is nevertheless crucial to the question of whether one is self-governed. Call this *the intension problem*.

To illustrate: a common idea amongst content-based accounts is that autonomy requires the ability to track objective values. Clearly, both autonomy and the ability to track objective values are good things. What is needed, however, is a reason for thinking that they are one and the same good thing. We need an explanation as to how the ability to track values constitutes the *specific* good of self-determination or self-direction or self-rule. Moreover, we need an explanation that doesn't proceed via the idea of the authentic or ruling self, on pain of returning us to the first horn of the dilemma. Thus generic appeals to the idea of 'rational self-governance' won't cut it here, for it's precisely the idea of a rational ruling self that traditionally serves as the bridge between the idea of conformity with objective rational norms and that of self-government.¹²

Others have tried to make headway by narrowing the class of autonomy-inconsistent attitudes to objectively false attitudes of some particular type. For example, Sonya Charles argues that autonomy is undermined only by false beliefs with specific *causal upshots*, namely those that (1) 'lead to subordination' and (2) 'perpetuate oppression'.¹³ Yet this doesn't solve the problem, for it remains unexplained how my having an attitude that tends towards the reduction of (my and others') self-determination *in the future* should constitute a lack of self-determination on my part *now*. If I believe that I ought to sell myself into servitude tomorrow, then I may not be self-governed for long; but we don't yet have any reason to doubt that I'm self-governed today. This requires a further argument. And it's the absence of precisely this kind of further argument—one that conceptually connects the relevant problematic attitudes with some recognisable notion of self-rule—that constitutes the intension problem.

This paper aims to resolve the dilemma by developing an analysis of the Deferential Wife case that both (1) avoids the intension problem and (2) manages to do so without invoking any

idea of the authentic self. To do this, we'll need a way of thinking about self-government that stays true to the concept but doesn't analyse it as governance by the self. It's to this task that I now turn.

II. THE NATURE OF DEFERENCE

As Marina Oshana has pointed out, at least part of what it is to be self-ruled is not to be ruled from without.¹⁴ This means that self-rule doesn't have to be construed positively, as rule by the self. It can also be construed negatively, as the absence of rule by others. From this perspective, the aim of a theory of autonomy isn't to elucidate a relevant notion of *selfhood* (the 'self of self-rule'), but a relevant notion of *interpersonal rule*. Here, then, is an alternative thought: if deferential attitudes don't compromise autonomy by undermining the rule of the authentic self, perhaps they do so by constituting a relationship of interpersonal rule. That is, perhaps deferential attitudes, despite being wholly internal to the psychology of the deferential party, nevertheless help to constitute a certain type of social relation, and it is in virtue of constituting this type of social relation that they are inherently incompatible with self-rule.¹⁵

When the Deferential Wife draws up a guest list, or buys a new dress, she takes the fact that her husband wants her to invite these particular guests, or to buy this particular dress, to be an extremely weighty reason for doing so in itself. It's not that her husband helps her to come to see reasons for acting that obtain independently of his preferences, such as that certain clothes are good value; his preferences simply *are* her reasons. And were her husband to have wanted her to invite different guests, or to buy a different dress, she would've taken this to be an equally weighty reason for doing *that*. This willingness to do whatever her husband wants just because he wants it is key to, indeed constitutive of, her deference.¹⁶

It's also a key feature of coercion. Suppose that a robber says: 'your money or your life!' This gives you a particularly weighty reason to do what the robber wants, namely to hand over your money. Moreover, this is a reason to do not only *what* the robber wants, but *whatever* the robber wants: if the robber demanded instead that you hand over your ring, or that you stand on one leg, you would have an equally weighty reason to do this. The strength of your reason doesn't depend on the content of the robber's wishes. The robber doesn't help you come to see independent reasons for performing the act in question (such as that charity is, after all, a virtue). The function of the threat is to give you content-independent reason to comply with his desire simply *qua* desire.

That's part of the essence of a coercive threat.¹⁷ Although the Deferential Wife doesn't act under coercion—she has no *external* reason to treat her husband's desires as weighty, content-independent reasons for action—she's disposed to treat them as such reasons all the same. And it's this feature of the Deferential Wife's practical identity, one shared with the victim of coercion, that I hereby contend holds the key to explaining her reduced autonomy. It does so by helping to make sense of the thought that, in deferring to her husband, the Deferential Wife subjects herself to a certain form of rule-from-without.¹⁸

We may specify this *disposition of practical deference* as follows:

B has a *disposition of practical deference* just in case *B* is disposed to take *A*'s wanting it to be the case that *p* to be an especially weighty reason in itself for bringing it about that *p*, independently of the content of *p*.

In this schema, *B* is an individual agent, and *A* is either some specified individual or corporate agent, or a representative member of some specified class of such agents.¹⁹

It may seem, however, as though such dispositions are held by a lot of perfectly autonomous people. For example, deferring to another person's expertise need not compromise autonomy, and yet this would seem to involve treating another's directives as weighty reasons for action. Similarly, love and care need not compromise autonomy, and yet this would also seem to require treating another's desires as weighty reasons for action. However, I now show that neither a willingness to take advice from experts nor an attitude of love or care need involve a disposition of practical deference in the specified sense. That is, dispositions of practical deference necessarily constitute a particular form of social relation—one of external rule—that these other dispositions need not.

III. DEFERENCE TO EXPERTS

When my doctor tells me to do something, I generally do it. It may look like what I'm doing here is taking the doctor's desire that I undergo a particular course of treatment to be an especially weighty and content-independent reason for undergoing that treatment (inasmuch as I'll follow whatever course of treatment my doctor recommends). But this isn't actually what I'm doing. It isn't my doctor's *desire* that I take to constitute my reason for following her prescription, but her *judgement about my medical interests*. What my doctor happens to desire is neither here nor there; what matters is what my doctor judges to be best for me.

Since doctors generally want what's best for their patients, this can be a subtle distinction. But it can be clarified by imagining cases in which the two come apart. Suppose that my doctor

strongly recommends that I take medicine *X*, and is then called away from her computer. Peering at the screen, I'm horrified to learn in an open email that she has a long-standing vendetta against me and wishes me to take *X* because it will do me harm—as opposed to *Y*, which she actually judges best for someone with my condition. If I still wish to defer to my doctor's expertise in this situation, that will mean taking medicine *Y*: that is, doing not what she wants, but what she judges to be in my medical interests.

Since what constitutes a person's expertise is her judgment, and not her will, this just is what it is to defer to a person's expertise. Hence even in more ordinary cases, in which my doctor wants what's best for me, it's my doctor's judgement about what's best for me, and not her wanting it, that matters. Suppose my doctor tells me that although she judges it best for me to take *X*, she very much wants me to take *Y* instead. If I treat this—her mere desire—as an especially weighty reason for taking *Y*, then I defer not to her expertise, but to her will, and in that case I *am* in a genuine position of deference to her.

A disposition to defer to expert opinion is therefore distinct from the disposition of practical deference that I claim characterises both the Deferential Wife and the victim of coercion. Of course, this doesn't mean that a disposition to defer to the judgement of another can't sometimes undermine autonomy too. My claim is that a disposition of practical deference is sufficient for diminished autonomy, not that it's necessary. There are many ways of falling short of autonomy. Moreover, it's likely in practice that one who defers to the will of another in an autonomy-undermining way will also defer to their judgements in an autonomy-undermining way. Nevertheless, telling us which forms of *epistemic* deference threaten autonomy and which don't is part of the job of a theory of testimony, and I don't offer such a theory here. As described above,

the Deferential Wife doesn't primarily lack autonomy by virtue of a defective epistemic relationship with her husband's judgements, but by virtue of a defective practical relationship with his will, and it's this that I'm trying to understand.

IV. TO LOVE, CHERISH, AND OBEY

What is it to love and care for another person? At least in part, it's to be motivated to prioritise that person's needs and to help to realise her projects. From here it's a short step to the thought that to love or care for someone is to take that person's desires to be especially weighty, content-independent reasons for action. But to care for someone is not, in fact, to be disposed to act so as to *satisfy their desires*. Rather, it's to be disposed to act so as to *promote their interests*.²⁰

Imagine a woman who finds the company of her elderly father very difficult, but who nevertheless decides to invite him to move into her home. Consider two versions of this case. In the first, her father informs her that he wishes to move in, and she moves him in simply because he desires it. This is a pure case of filial obedience. In the second, she moves her father in because she judges it to be in his best interests to do so, and she's willing to prioritise his interests over her own. This is a case of filial love. The difference is that in this second case her reason for action is grounded not in her father's desires *per se* but rather in his best interests.

Of course, people generally want what's in their interests. Nevertheless, there's an important difference between treating a person's wants as reason-giving and treating a person's interests as reason-giving. To draw this out, suppose that the father misjudges his own interests: although he desires to live with his daughter, his physical needs are in fact so complex that he'd be better off in a good care home. Were his daughter to know this, then the act of care would be to

disobey his wishes and to refuse to invite him to move in, whereas the act of obedience would be to do as he wishes despite knowing that it's not what's best for him.

This last point is further complicated by the fact that, plausibly, the satisfaction of one's desires is an important component of one's well-being. This means that people's desires not only often coincide with their interests, but also help to constitute those interests. Even so, however, one's objective interests aren't fully reducible to the indiscriminate satisfaction of one's actual desires. This is true on any plausible theory of interests: even on preference-based accounts, for instance, preferences are invariably subject to some manner of idealisation or restriction.²¹ So there remains a stable distinction between being motivated to satisfy someone's desires *qua* desires, and being motivated to satisfy someone's desires *qua* components of their interests.

To illustrate: I might give special deliberative weight to the fact that my toddler wants to wear her red shoes, on the grounds that I care about her interests and that it's in her interests to exercise some control over what shoes she wears. But I give very little weight to the fact that she wants to wear her red shoes as a fact *independent* of her interests. Were I to do so—that is, were I to treat my toddler's desires as especially weighty reasons for action regardless of what I judge to be in her interests—then I would indeed be in a position of deference and subservience to my toddler, as opposed to one of love and care.

So, to return to the woman and her elderly father, it's true that any calculation as to what's in his best interests must give significant weight to his desires. For instance, it might pain him greatly for his request to move in with his daughter to be rebuffed, and she'd need to factor this into her judgement as to which option would serve him best. In some cases, his desires may well be what tips the scales in one direction or another. However, that she gives weight to his desires

in this way doesn't mean that she acts from a disposition of deference. In the case of deference, her father's desires short-circuit this kind of deliberation, directly generating weighty reasons disconnected from any broader thinking about his interests.

None of this is to deny that deference and love can coexist; love takes many forms. The obedient daughter who accedes directly to her father's desire to move in with her may also love him very much. Indeed, if she believes that he'd be better off in a care home, she might be deeply troubled about the consequences for him of his poor decision, and might go to great lengths to try to mitigate its effects—even though actively disobeying his wishes never crosses her mind.

V. AUTONOMY AND THE POLITICS OF CARE

I now address an objection to the preceding section's analysis. Consider the *Doting Wife*. Unlike the Deferential Wife, the Doting Wife doesn't treat her husband's desires as strongly reason-giving in themselves. Instead, she prioritises his interests far above her own. This means that she sometimes defies his will in the service of his good, such as (say) by refusing to enable his smoking habit. Nevertheless, she mostly behaves similarly to the Deferential Wife, buying the clothes that it benefits him that she wear, uprooting herself for the sake of his career, and so on—all without *him* giving any special priority to *her* interests.²² According to the analysis developed so far, the Doting Wife isn't in a position of obedience and deference to her husband, but one of (perhaps extreme) love and devotion. Nevertheless, the Doting Wife may seem just as problematic, morally and politically, as the Deferential Wife.

However, I now argue that what troubles us about the Doting Wife isn't best understood as a lack of autonomy. Here, in summary form, is why. If we believe that the Doting Wife lacks

autonomy in virtue of her deep concern for her husband's interests, then we risk having to accept that dispositions of love and care are incompatible with autonomy *per se*. But if dispositions of love and care are incompatible with autonomy *per se*, then autonomy is revealed as an uncaring ideal of social isolation. To block this result, we'll need to identify some other distinctive feature of the Doting Wife case that isn't shared by less problematic cases of love and care. But there is no such feature. Moreover, the Doting Wife's problems are better and more satisfyingly understood in terms of other forms of injustice. I now expand on each of these points.

First, consider two further cases. One is the Doting Couple, a doting wife married to a symmetrically doting husband: while she's out buying clothes to please him, he's out buying clothes to please her; when she declares herself ready to uproot herself and move across the country for the sake of his career, he declares himself ready to turn down the job for the sake of her career; and so on. This version of the case seems morally and politically untroubling. But if the Doting Wife lacks autonomy in the non-reciprocal case, then she surely lacks it in this one too.²³

The other case is that of a devoted parent of a baby. The parent prioritises the baby's interests far above her own and acts predominantly in the service of the baby's good. This also seems morally and politically untroubling. Again, however, it's difficult to see how, if the Doting Wife lacks autonomy, the parent doesn't: the parent's concern certainly isn't any less intense than the Doting Wife's, and in neither case is the care reciprocated. (We might think also about cases involving love for an elderly relative with dementia, or for a sibling with severe cognitive disabilities.) There's no obvious reason why the baby's lack of moral agency or mental capacity should, in itself, render the parent any more self-determined with respect to her own attitudes of love and care.

Since there are no relevant differences between the Doting Wife case and these two further cases, we must treat them equivalently. So in order to sustain the claim that the Doting Wife lacks autonomy, we have to accept that the Doting Couple and the devoted parent lack autonomy, too, and we're thereby driven towards the conclusion that dispositions of love and care are inconsistent with autonomy *per se*. Yet we have good reason to resist a conception of autonomy that necessarily excludes love and devotion. On such a conception, the truly autonomous person would be one who loved and cared for no one. To be sure, such a view could be tempered by insistence on the fact that autonomy isn't the only good, and that it can be sensibly traded for other things of value. Yet it's hard to see the normative *point* of a conception of autonomy that treats even healthy loving relationships as necessarily involving the extensive sacrifice of something of serious moral importance. Such a conception certainly seems less fruitful than one that's capable of prying apart ordinary loving attitudes from more obviously problematic ones. For the Deferential Wife's problem isn't fundamentally one of social entanglement, or of commitment, or of love. It's not that she cares about another's good, but that she renders herself subservient to another's will.

Why, then, do we find the case of the Doting Wife so troubling? As we've just had occasion to note, the case wouldn't trouble us so much if the husband reciprocated his wife's feelings, or if he lacked moral capacity. This is instructive, since although the husband's degree of reciprocation and his moral capacity seem conceptually irrelevant to his wife's degree of psychological self-determination, they're both clearly relevant to the justice of their marriage. And this suggests that the Doting Wife's fundamental problem isn't heteronomy, but injustice.

More specifically, note that while the Doting Wife prioritises her husband's interests far above her own, her husband asymmetrically prioritises his own interests at the expense of hers.

This means that his interests are double-counted, being looked after by both parties in the relationship, whereas hers are zero-counted, being looked after by neither. This renders the relationship unequal. Moreover, insofar as the husband acquiesces in this inequality and makes no attempt to redress it, he takes unfair advantage of his wife's devotion. This renders the marriage exploitative. What's more, in many societies, including our own, women are typically socialised to care for others in such ways to a greater extent than are men. Where non-reciprocal relationships of this kind are widespread, such double-counting of men's interests is a general feature of society as a whole and so an instance of broader social injustice.²⁴

It's therefore important to distinguish the moral and political problems associated with insufficiently reciprocated love from those associated with dispositions of deference. The distinction is subtle and often overlooked; indeed, Hill Jr.'s description of the Deferential Wife is ambiguous between them, and this is a large part of what makes the case so difficult to analyse. Yet the distinction is as follows. On the one hand, a person might be disposed to prioritise another person's *interests*. Such a person is in a position of love and care, and there's no reason to think that her autonomy is thereby compromised. However, if that love is unreciprocated, her situation may well be one involving objectionable inequality, exploitation and social injustice. On the other hand, a person might be disposed so as to do whatever another person merely wants. *In addition* to the problems often associated with unreciprocated love, then, such a person is in a psychological position of obedience, subservience or deference, and this constitutes at least one clear sense in which her autonomy is necessarily diminished.

VI. HIDDEN DEFERENCE AND PSEUDO-DEFERENCE

Two final points are needed in order to round out this analysis of autonomy-compromising practical deference. The first is that deference can sometimes be *hidden*. For example, a person may be very adept at citing non-deferential reasons for actions that they nevertheless perform on deferential grounds. Such a person won't tell you that they're acting in the service of another's mere desires; instead, they may convincingly list a number of considerations that justify the course of action independently of those desires. (In some cases, these considerations may have been helpfully supplied by the very person to whose desires they defer, under the guise of rational persuasion.) Yet if they aren't actually moved by these independent reasons, and they're instead moved by the other's mere desires, then they act deferentially nonetheless.

This means that it can be very difficult to tell, in isolated cases, whether a person is acting deferentially—even, or perhaps especially, when that person is oneself, given the rich opportunities for self-deception here. Over time, however, one may come to understand which considerations a person's deliberations are really tracking. The crucial question is whether the person's choices are properly *responsive* to her judgements concerning the full range of reasons, or whether, in the end, they're responsive only to her beliefs about the other's preferences. In short: a person who always seems to find reasons to do whatever the other happens to desire, and who is relatively unresponsive to critical engagement with respect to these reasons, is likely to be in a position of practical deference.

All of this helps to explain why dispositions of deference so often seem to correlate with failures of interpersonal practical justification (for instance, of the type analysed by Andrea Westlund).²⁵ Those who are disposed to act deferentially may come under pressure to hide this fact, both to others and to themselves, and so may tend to supply insincere or mistaken

justifications for their behaviour—attempted justifications that are evasive, parroted, or otherwise fail to stand up to critical scrutiny. But it isn't in these failures of interpersonal rational justification themselves that a deferential person's lack of autonomy ultimately lies; after all, some deferential people are perfectly clear-eyed and unconflicted with respect to their deference. It's in the substantive relationship of interpersonal rule that their lack of autonomy is located.

The second point is the converse: that what appear to be cases of deference are sometimes cases of *pseudo-deference*. That is, a person may give the appearance of tracking another's mere desires when they're actually tracking something else. Imagine a woman who, since her youth, has aspired to a specific form of *haute-bourgeois*, conservative lifestyle. Her aspiration is to live in a grand house, to wear Cartier and Chanel, to host dinners and parties and charity fundraisers, and to have a traditional marriage to a man who shares her values and wants his wife to behave in precisely these ways. She marries such a man, and on the surface seems to act deferentially: he wants her to dress a particular way, and she does; he wants her to entertain certain guests, and she does. But in fact she doesn't do these things because he wants her to; to the contrary, she was careful to marry a man whose desires fit with the life she already had planned, and thus the explanatory connection runs the other way. Were her husband one day to declare himself fed up with their staid existence, and to propose that they abandon everything in favour of a new life in a beach hut in Bermuda, she wouldn't obediently follow him there. Instead, she'd find herself a new husband, one who could be better trusted to play his allotted role.

Deference isn't always easy to spot. Nevertheless, I submit that the account of practical deference developed here, one which seeks to understand it in terms of a particular practical relationship that a person can have with another's mere desires, a relationship that's also partly

characteristic of coercion, is capable of capturing its distinctive character. In particular, it provides a principled explanation for how attitudes of deference differ from more ordinary attitudes of love and care, thereby avoiding the unwelcome conclusion that substantive autonomy is ultimately an asocial ideal of isolation from human relationships in general.

VII. CHOOSING DEFERENCE

We began with a dilemma facing attempts to understand the Deferential Wife's diminished autonomy. On the one hand, no theory that analyses autonomy primarily in terms of rule by the self can explain it, since the plasticity of human selves requires all such accounts to be purely procedural. On the other hand, and for the same reason, any substantive account must analyse autonomy in terms other than mere rule by an authentic self, and this renders it mysterious how these alternative substantive conditions on autonomy might be conditions on *self-rule*. I proposed that the way out of this problem is to demonstrate that the Deferential Wife's deferential attitudes constitute a form of interpersonal rule, since this looks like the only route to an analysis of her lack of autonomy that (1) is grounded in an intuitive notion of self-rule and (2) doesn't essentially depend upon ideas about the authentic self.

I've assumed throughout that the Deferential Wife is relevantly unconstrained, and that no external thing forces her to relate to her husband in a deferential manner. However, this may seem to undermine the claim that there's an important sense in which she's ruled from without: if, *ex hypothesi*, her deference is self-chosen, and if her form of life is a genuine expression of her deferential values, then perhaps she's autonomous after all. If we're truly to vindicate the intuition that the Deferential Wife lacks autonomy, that is, it may not be enough to show that she lacks a

degree of self-rule *within* the terms of her deferential marital relationship; we may need to show, instead or as well, that she lacks self-rule with respect to her overarching *choice* of a deferential marital relationship.

I now argue against this line of thought. As we have stipulated her to be, the Deferential Wife is, I suggest, *autonomously nonautonomous*. That is, the autonomy of her overall choice of life needn't dispel the nonautonomy of the life that she thereby chooses. Hence we may recognise that she lacks a meaningful form of autonomy in her interpersonal relations while also allowing that, in a broader sense, she enjoys a valuable form of self-determination. Indeed, I argue that doing so is key to making sense of the normative complexity of the case.

This idea of a person's being autonomously nonautonomous is one that has been invoked before, and may be illustrated by means of the following example. Imagine a person who's always dreamt of serving her country in the military; she freely enlists during a time of war and is sent to a combat zone. As a soldier subject to military discipline, and now unable to leave her post (deserters are shot), her autonomy is strictly curtailed. Yet this curtailment of autonomy is an integral part of the military life that she's always wished to live, and in *this* sense she's successful in living a self-directed and autonomous life. She has autonomously chosen a life of reduced autonomy. Adopting Clare Chambers' terms, we may say that she enjoys *second-order autonomy*, but lacks *first-order autonomy*.²⁶

Unlike the soldier, of course, the Deferential Wife isn't subject to coercive threats, and so it may not be immediately clear in what way this distinction applies to her case. To see how it does, we first need the idea of a *deferential choice*. This may be characterised as a choice made from a disposition of practical deference, in the following way:

B makes a *deferential choice* to bring it about that *p* just in case *B* takes some *A*'s wanting it to be the case that *p* to be an especially weighty reason in itself for so choosing, and *B*'s choice is motivated by this reason.

A choice may be more or less deferential depending upon (1) the degree of reason-giving weight afforded to *A*'s desires and (2) the motivational effectiveness of these reasons. Moreover, not every choice made by a person with a disposition of deference need be a deferential choice. We may imagine, for instance, that many of the Deferential Wife's day-to-day choices are non-deferential, either because they concern matters that her husband doesn't care about, or because her deference doesn't extend to certain domains. Nevertheless, many of her choices are likely to be deferential in this sense.

In choosing deferentially, the Deferential Wife substitutes her husband's will for her own and thereby allows herself to be ruled from without. This, I've argued, suffices for a loss of first-order autonomy. Were she *forced* to choose deferentially, then she'd lack second-order autonomy too. As noted, this is true of many real-world deferential wives, for whom defiance may foreseeably lead to physical, emotional, or financial retaliation on the part of their husbands, condemnation and isolation from their communities, and formal legal or religious censure. However, we've been focusing on the theoretically most puzzling case, that in which no external thing forces the Deferential Wife to choose deferentially. In this case, the Deferential Wife *could*—in at least one perfectly ordinary sense of 'could' (such as the standard compatibilist sense)—choose to bracket her husband's desires, consider the balance of other reasons, and act solely on

the basis of the latter. The deferential and autonomy-undermining nature of her choices is therefore something over which she exercises a form of voluntary control. It's in this sense that she lacks first-order autonomy but retains second-order autonomy: like the soldier, her loss of autonomy proceeds from an exercise of her own capacity for autonomous self-direction.²⁷

Nevertheless, many have opted for the alternative strategy of arguing that the Deferential Wife lacks autonomy at the second-order level. This, I believe, is for two main underlying reasons. One is pessimism about the prospects of locating her loss of autonomy at the first-order level, due to the difficulties involved in distinguishing relations of deference from relations of love and care—a pessimism I hope to have already shown is unwarranted. The other is implicit acceptance of the idea that losses of first-order autonomy cease to be morally interesting when they are sanctioned by exercises of second-order autonomy. The thought is that if the Deferential Wife has freely chosen to trade some of her first-order autonomy for other goods, such as community or belonging or pleasingness to God, then that's her prerogative; her nonautonomy shouldn't matter, morally or politically, if it's what she autonomously wants. So if we want to show that the Deferential Wife lacks a meaningful form of autonomy, we have to show that she lacks second-order autonomy.

As I argue in the next section, however, this presumption is false: second-order autonomy isn't the crucial issue. Moreover, this is fortunate, since the claim that the Deferential Wife (as imagined) must lack second-order autonomy is extremely difficult to vindicate. In the remainder of this section I briefly explain why this is.

It's undeniable that the values and ideas on the basis of which the Deferential Wife makes her choice of life are likely open to serious criticism. Feminist theorists have described in detail

the ways in which choices such as this typically proceed from underlying values and motivations shaped by patriarchal forces grounded in unjust relations of power.²⁸ Even if externally unconstrained, then, the Deferential Wife's choice of a life of deference is likely to be based on a distorted normative viewpoint, one involving ideas that are objectively false (for instance, concerning the relative moral standing of men and women), as well as reflecting a lack of self-respect or self-worth.

However, any attempt to move from these facts about the false, patriarchal basis of the choice to its nonautonomy must contend with the dilemma described in Section 1. The problem is this. On the one hand, there's no obvious reason to think that governing oneself *badly*—for instance, in accordance with objectively false norms—amounts to a failure to govern oneself *simpliciter*. (It's clearly possible for polities to govern themselves poorly; why should things be different for individuals?) At the very least, some argument is needed to take us from the idea that one's choices are grounded in objectively false judgements to the idea that one lacks self-rule. On the other hand, one can't solve this problem by reverting to the familiar doctrine that self-government requires government by a true or authentic self, since it may be that the relevant patriarchal norms *flow through* and *partly constitute* the Deferential Wife's true self, such that she authentically desires her life of deference. (These are, respectively, the intension problem and the problem of the plastic self.)

I've argued that this dilemma can be resolved at the first-order level by focusing on the substantive relationship of interpersonal rule in which the Deferential Wife places herself. The prospects for escaping the dilemma by a similar route at the second-order level are, however, unpromising. In choosing to embrace a deferential marriage, the Deferential Wife needn't be ruled

or dominated by anyone: we've already stipulated that her choices are uncoerced. And I see no way to make sense of the only other possibility here, this being that she's subject to the rule of those responsible for the *transmission* of the false, patriarchal values on which her choice is based. For those values may have been transmitted in the same way, and by the same people, as other values that we don't regard as autonomy-undermining. That is, there's no reason to regard her as *any more* other-ruled in virtue of her acquisition of patriarchal values than in virtue of the general fact that she's a socially embedded human being.²⁹

So it's very difficult to avoid the conclusion that the Deferential Wife, as we have imagined her, is autonomous at the second-order level. This is the truth in the proceduralist view, mentioned earlier, that even choices of deference may be fully autonomous. However, this conclusion isn't as significant as it seems. As I now argue, her lack of first-order autonomy has important moral and political implications that aren't simply washed away by her possession of second-order autonomy. Thus we needn't tie ourselves in theoretical knots trying to avoid the unavoidable conclusion that she may be perfectly autonomous at the second-order level.

VIII. GHOST-WRITTEN LIVES

The basic concept of autonomy is that of self-government. This idea is one that plays a number of different moral-theoretic roles for us. One is the *deontic* role of marking out which choices qualify as legitimate exercises of our normative power to waive rights: it is in this sense, for instance, that we say that a token of consent is morally valid only if it has been autonomously given. Another is the *axiological* or *eudaimonic* role of marking out which choices contribute to a person's enjoyment of the traditional liberal good of a self-authored life.³⁰ So far, I have been considering

practical deference just in relation to the root idea of autonomy as self-government. In this final section, I focus more specifically on the idea of autonomous choice as it functions in the second of these roles, i.e. as part of an account of self-authorship.³¹

Part of what it is for a life to go well is for it to be self-directed, lived on one's own terms, and shaped by one's own agency. Clearly, however, not all choices have the power to confer these properties. For example, choices made under coercion tend to make one's life *other*-directed; people subject to extensive coercion are cut off from the good of self-authorship. I've argued that voluntary deference defeats first-order autonomy but preserves second-order autonomy. Where does this leave a person with respect to the good of self-authorship?

Suppose that a young man has long dreamt of being an architect, but comes from a strict family with proud traditions and a farm that has been in the family for many generations. When his father tells him to abandon his dreams and to take over the family farm, he complies on the basis of a belief that it's proper to defer to the wishes of his parents. Yet he never lets go of his dream, and despite never judging that he made the wrong decision, he often thinks about the life that he would have had as an architect. Our question is: to what extent can he be said to have been the author of his own life?

Different aspects of the case pull in different directions. On the one hand, there looks to be a clear sense in which he isn't living the life that he wants: by allowing his father to determine his choice of career, he fails to determine himself. On this basis, one might argue that first-order autonomy is what matters for self-authorship. This would be broadly in line with the position taken, for instance, by Oshana, according to whom autonomy turns entirely on first-order matters (such as one's degree of financial self-sufficiency, rationality and self-respect—though not practical

deference as I've analysed it here).³² Yet first-order autonomy can't be the whole story. For were this man to lack second-order autonomy *as well*—for instance, were he forced to take over the farm at gunpoint—then his life would be other-directed in a further sense. The fact that he freely chooses to defer to his father's wishes means that he retains an important kind of control over his life, and we need an approach that captures this.

On the other hand, then, one might point to the broader sense in which this man *is* living the life that he wants, inasmuch as it's he and no one else who chooses to defer to his father's wishes. One might argue on this basis that what matters for self-authorship is second-order autonomy alone. This verdict would be broadly in line with the position taken, for example, by Ben Colburn, according to whom autonomy turns primarily on second-order matters, specifically (in this case) the fact that the thwarted architect continues to endorse his deferential choice.³³ Yet second-order autonomy can't be the whole story either, and the metaphor of self-authorship is helpful in explaining why. Suppose that you're writing a novel, but that instead of deciding the plot points for yourself you choose to structure it in accordance with someone else's story frame. It's your decision to do this, and in that sense you maintain ultimate control over the finished work. Nevertheless, it's the case that you're not as fully the author of the novel as you would have been had you devised the plot for yourself. Authorship is in part a first-order matter: it's not enough that one ghost-write a story dictated by someone else. One must write one's own story. Yet in allowing his father to determine the course of his life, the thwarted architect does precisely this: he lives a life that isn't so much self-authored as ghost-written.

The truth is that a fully self-authored life must manifest autonomy at both the first-order *and* the second-order levels. This helps to explain the seemingly paradoxical sense that, when the

Deferential Wife packs her bags to follow her husband across the country, she both is and isn't determining the course of her own life. She isn't determining it inasmuch as her choices are subject to the will of another. But she is determining it inasmuch as she freely allows her choices to be so subject. This makes her more autonomous than a person who acts under coercion, but less autonomous than one who acts for non-deferential reasons.

An important upshot of this analysis is that dispositions of deference are *inherently counter-autonomous*. Once one forms a desire to be subject to the will of another, one is condemned to nonautonomy either way: both if one succeeds, and if one fails. If one succeeds, one's life is directed by others and so is not self-authored; if one fails, one is forced to forge one's own path against one's will, and so falls short of self-authorship insofar as one fails to live the other-directed life that one wants. This means that a political morality committed to the protection and promotion of individual autonomy (in the 'self-authoring' sense) must also be broadly committed to the dismantling and elimination of dispositions of deference.

To flesh this out: suppose that one holds a broadly eudaimonist view of the good, according to which a central part of what it is for a human life to go well is for it to be autonomous in the sense of being partly self-authored. A moral upshot of the forgoing analysis, then, is that to have a disposition of deference is to fail to live a fully flourishing human life, and so to be *pro tanto* worse off. (The qualifier '*pro tanto*' is important: one need not deny that some deferential lives may be substantially better, from the point of view of human flourishing, than some non-deferential lives, *on balance*.) And a further corollary of this is that to intentionally or negligently bring it about that someone else has a disposition of deference is to commit a *prima facie* wrong.

From this normative perspective, teaching one's daughters to defer to the wishes of men is *pro tanto* harmful and *prima facie* wrongful in exactly this way. So is teaching the children of colonial subjects to defer to the wishes of their colonial oppressors, or servants to defer to the wishes of their masters, or any other form of what we might call *subordination training*. It's harmful inasmuch as, if successful and not undone, it makes it impossible for them to lead self-authored lives, by putting them in an ongoing position of actual or frustrated subservience to others. And it's wrongful inasmuch as it's the intentional or negligent bringing about of a harm. Of course, difficult moral questions lurk here concerning the culpability of parents and carers, themselves typically subject to oppressive internalised norms, in perpetuating those norms in others. But insofar as it's a duty of parents to raise their children in a way that enables them to lead flourishing lives, and insofar as dispositions of practical deference impede the flourishing of human lives, then intentionally raising one's children so as to manifest these dispositions, or negligently failing to push back against the social forces that tend to inculcate them in one's children, is *prima facie* wrongful.

Dispositions of practical deference may arise from social processes that defy analysis in terms of individual culpability, too. Yet even where there is no wrongdoer, a harm remains. A social environment that encourages dispositions of practical deference is a harmful social environment, and this generates (defeasible) reasons to work to change it. *How* we do so, however, will depend upon the further details of our political morality—in particular, on whether state action in pursuit of the ideal of individual self-authorship is permitted or forbidden—and I do not pursue that question here.

IX. CONCLUSION

In admitting that the Deferential Wife may autonomously choose her life of deference, we don't automatically cancel the moral and political implications of her first-order nonautonomy. We don't, therefore, need to insist that she lacks second-order autonomy in order to do justice to these moral and political concerns. This allows us to say that while her choice of a deferential marriage may be criticisable on many grounds, it needn't be criticisable on grounds of nonautonomy. And this entails that her choice of a deferential marriage may demand the moral respect generally commanded by the autonomous choices of others.

In choosing deference, however, the Deferential Wife chooses, *inter alia*, a life of diminished first-order autonomy. In treating her husband's mere desires as weighty, content-independent reasons for action, she makes herself subject to his will in a manner partly characteristic of a victim of coercion. In so doing, she accepts a form of rule from without and hence a loss of autonomy. This loss of first-order autonomy then prevents her from truly partaking in the good of a self-authored life: the decisions she makes from *within* her marriage are not the fully autonomous decisions of a self-author, but those of one leading a life that's partially ghost-written.

This analysis avoids many of the familiar pitfalls of the recent debate. Instead of requiring us to view her deferential attitude as an alien force external to her true self, it allows us to do justice to her self-conception as a person fully identified with her deferential role. And it doesn't require us to insist on conditions of autonomy that, despite their purported extensional adequacy, lack any obvious intuitive connection to the idea of self-rule. Instead, by locating her loss of autonomy in

the inherently heteronomous content of her deferential disposition, it connects that loss to the familiar idea that to be self-ruled one must be relevantly free of the rule of others.

* Thank you to the editors of *Ethics*, an anonymous reviewer for *Ethics* and Ben Colburn for especially helpful comments on a previous draft. Thank you also to the following people for generously providing feedback on the paper at various stages of its development: Robyn Bluhm, Kirstin Borgerson, Danielle Bromwich, Sarah Fine, Amanda Greene, Ami Harbin, Jennifer Hornsby, Andrew Huddleston, Fleur Jongepier, Charlotte Knowles, Jonathan Knutzen, Matthew Kramer, Hallvard Lillehammer, Douglas MacKay, Joseph Millum, Massimo Renzo, Andreas Schmidt, Liam Shields, Daniel Statman, Natalie Stoljar, Marika Warren, Han van Wietmarschen, Andrea Wilson and Chiara Zucchelli. Thank you also to participants at the following meetings and conferences for helpful discussions of prior drafts: a 2018 meeting of the Birkbeck Philosophy Work in Progress Seminar; the 2019 UK Analytic Legal & Political Philosophy Conference; the 2019 New Scholars in Bioethics Symposium; a 2019 meeting of the Oxford University Philosophy Society; a 2020 Kings College London Feminist Philosophy Reading Group; a 2020 meeting of the Institute of Philosophy PPE Seminar; and the 2020 Erasmus University Rotterdam workshop on ‘Domination and Wrongful Interference’.

¹ Thomas E. Hill Jr, ‘Servility and Self-Respect’, *The Monist* 57, no. 1 (1973): 87–104. See also Marilyn Friedman, ‘Moral Integrity and the Deferential Wife’, *Philosophical Studies* 47, no. 1 (1985): 141–50; Marcia Baron, ‘Servility, Critical Deference and the Deferential Wife’,

Philosophical Studies 48, no. 3 (1985): 393–400; Andrea C. Westlund, ‘Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?’, *Philosophical Review* 112, no. 4 (2003): 483–523; Marina Oshana, *Personal Autonomy in Society* (Aldershot, Hants, England; Burlington, VT: Ashgate Publishing Limited, 2006), 57–60; Sonya Charles, ‘How Should Feminist Autonomy Theorists Respond to the Problem of Internalized Oppression?’, *Social Theory and Practice* 36, no. 3 (2010): 409–28; Jules Holroyd, ‘Substantively Constrained Choice and Deference’, *Journal of Moral Philosophy*, 2010, 21; Suzy Killmister, ‘Autonomy and the Problem of Socialisation’, *Social Theory and Practice* 39, no. 1 (2013).

² This paper follows previous discussions in considering the phenomenon of internalised deference mainly as it functions as a mode of gender oppression. This is not, however, to deny the role that internalised deference may also play with respect to other axes of oppression. On the correlation between deference and social class, for instance, see Robert E. Freeland and Jesse Hoey, ‘The Structure of Deference: Modeling Occupational Status Using Affect Control Theory’, *American Sociological Review* 83, no. 2 (2018).

³ See, e.g., Marilyn Friedman, *Autonomy, Gender, Politics* (Oxford: Oxford University Press, 2003), chap. 6; Clare Chambers, *Sex, Culture and Justice: The Limits of Choice* (University Park, PA: The Pennsylvania State University Press, 2008).

⁴ Natalie Stoljar argues for a broader understanding of the term, one which includes a person’s *awareness* of their external oppression and of the reactive behaviours that it demands of them (‘“Living Constantly at Tiptoe Stance”: Social Scripts, Psychological Freedom, and Autonomy’, in *Personal Autonomy and Social Oppression*, ed. Marina Oshana (New York: Routledge, 2015), 105–23). By contrast, I here seek to maintain a distinction between the *internal aspects of*

oppression, which may include the broader psychological phenomena discussed by Stoljar, and *internalised* oppression proper, which requires some manner of psychological endorsement of oppressive norms on the part of the oppressed.

⁵ Ian Carter, *A Measure of Freedom* (Oxford: Oxford University Press, 1999); Matthew H. Kramer, *The Quality of Freedom* (Oxford: Oxford University Press, 2003).

⁶ Philip Pettit, *Republicanism: A Theory of Freedom and Government* (Oxford: Clarendon Press, 1997); Philip Pettit, *On the People's Terms: A Republican Theory and Model of Democracy* (Cambridge; New York: Cambridge University Press, 2012). On the account I develop below, internalised deference is incompatible with autonomy inasmuch as it constitutes a form of self-subjection to the arbitrary will of another. While this is clearly a republican analysis in some sense, it is not, and cannot be, a mere application of republican ideas as developed by contemporary neo-republican theorists like Pettit. This is because such theorists treat republican liberty primarily as an *external* matter related to the capacity of others to interfere with us arbitrarily, whereas my analysis is an attempt to understand dispositions of deference insofar as they have been deeply internalised. (See also footnote 15.)

⁷ This way of posing the fundamental problem, in terms of approaches that do and do not take seriously the basic idea that self-rule is rule by the self, is intended to offer an underlying conceptual *diagnosis* of the particular strengths and weaknesses of procedural and strong substantive analyses of autonomy as originally mapped out by Catriona Mackenzie and Natalie Stoljar in their 'Introduction: Autonomy Refigured', in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. Catriona Mackenzie and Natalie Stoljar (Oxford: Oxford University Press, 2000), 3–31.

⁸ See Harry G. Frankfurt, 'Freedom of Will and the Concept of the Person', *Journal of Philosophy* 68, no. 1 (1971): 5–20; Harry G. Frankfurt, 'On the Faintest Passion', in *Necessity, Volition and Love*, ed. Harry G. Frankfurt (Cambridge: Cambridge University Press, 1999); Gary Watson, 'Free Agency', *Journal of Philosophy* 72 (1975): 205–20; Diana T. Meyers, *Self, Society, and Personal Choice* (Columbia University Press, 1991); J. David Velleman, 'What Happens When Someone Acts?', *Mind* 101 (1992): 461–81; Friedman, *Autonomy, Gender, Politics*; Richard Moran, 'Frankfurt on Identification: Ambiguities of Activity in Mental Life', in *Contours of Agency: Essays on Themes from Harry Frankfurt*, ed. S. Buss and L. Overton (Massachusetts Institute of Technology Press, 2002).

⁹ Put differently, we *all* must sometimes act from reasons that we take to be basic. See Westlund, 'Selflessness and Responsibility for Self'; Killmister, 'Autonomy and the Problem of Socialisation'.

¹⁰ See Alfred R. Mele, *Autonomous Agents: From Self-Control to Autonomy* (New York and Oxford: Oxford University Press, 1995); John Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge: Cambridge University Press, 2009). For further critical discussion along these lines, see Paul Benson, 'Autonomy and Oppressive Socialization', *Social Theory and Practice*, 1991, 385–408; Natalie Stoljar, 'Autonomy and the Feminist Intuition', in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. Catriona Mackenzie and Natalie Stoljar (Oxford University Press, 2000); Nomy Arpaly, *Unprincipled Virtue: An Enquiry into Moral Agency* (New York: Oxford University Press, 2003), 126–27; Ann Levey, 'Liberalism, Adaptive Preferences, and Gender Equality', *Hypatia* 20, no. 4 (2005): 127–43.

¹¹ Benson, ‘Autonomy and Oppressive Socialization’; Stoljar, ‘Autonomy and the Feminist Intuition’; Charles, ‘How Should Feminist Autonomy Theorists Respond to the Problem of Internalized Oppression?’. See also Susan Wolf, *Freedom Within Reason* (Oxford University Press, 1993).

¹² Nor, for similar reasons, will a by-the-book Kantian analysis do the trick. Perhaps there’s an important sense in which the Deferential Wife lacks autonomy in virtue of failing to act from principles of pure practical reason. But, if so, this is a kind of heteronomy that she shares with many others. There’s a *distinctive* loss of autonomy suffered by the Deferential Wife, in virtue of her internalised disposition of deference, that needn’t be suffered by (say) an egoist or a utilitarian, and it’s this that we’re trying to understand. See also Holroyd, ‘Substantively Constrained Choice and Deference’.

¹³ ‘How Should Feminist Autonomy Theorists Respond to the Problem of Internalized Oppression?’, 419.

¹⁴ *Personal Autonomy in Society*. See also Michael Garnett, ‘The Autonomous Life: A Pure Social View’, *Australasian Journal of Philosophy* 92, no. 1 (2014): 143–58; Danielle M. Wenner, ‘Nondomination and the Limits of Relational Autonomy’, *International Journal of Feminist Approaches to Bioethics* 13, no. 2 (2020): 28–48.

¹⁵ Such an approach departs from Oshana, who follows Pettit in treating substantive independence as an *external* condition of social immunity to arbitrary interference by others (*Personal Autonomy in Society*, 86–90). As previously discussed (above, pp. 4–6), however, while autonomy may well have such external conditions, and while real-world deferential wives may well fail to meet them,

these are not the conditions we need in order to understand the inherently autonomy-defeating nature of *internalised* oppression. That is, while I concur with Oshana that the deferential wife lacks autonomy ‘due to her personal relations with others’ (59.), it is in those personal relations as constituted by *her* attitudes and dispositions, and not as constituted by the *others*’ powers and behaviours, that the genuinely internalised aspects of her oppression lie.

¹⁶ Friedman, ‘Moral Integrity and the Deferential Wife’; Westlund, ‘Selflessness and Responsibility for Self’.

¹⁷ Michael Garnett, ‘Coercion: The Wrong and the Bad’, *Ethics* 128, no. 3 (2018): 545–73.

¹⁸ Note that, on this analysis, a victim of coercion may be understood as someone who has been *deliberately forced* by another to adopt a specific and limited disposition of deference (the ‘especially weighty reason’ to comply with the other’s wishes, in the schema presented immediately below, being supplied by the coercer’s threat). Purely internalised deference, then, is the internal aspect of coercion without its external aspect. This may also help to explain why, in practice, internalised deference is so often a residual product of prior coercion: even when the external compulsion stops, its internal corollary may still remain.

¹⁹ Three further brief points about this schema may help forestall misunderstandings. First, note that *A*’s desires must constitute *especially* weighty reasons for *B*. Some moral theories require us to treat the preferences of everyone as at least weakly reason-giving; this kind of general moral concern is insufficient for a disposition of practical deference, since if all desires are equally reason-giving then no particular person’s desires are especially weighty. Second, *A*’s desires are to be understood as providing *B* with tenseless reasons. The fact that *A* will want it to be the case

that *p* tomorrow may give *B* reason to act today. Finally, where *A* has conflicting desires, *B* must be disposed so as to prioritise those desires that are temporally closer over those that are temporally more distant, and those that are motivationally stronger over those that are motivationally weaker (with the former consideration taking priority over the latter).

²⁰ Although this isn't the only way of understanding dispositions of love and care, it's the one that poses the greatest challenge to attempts to distinguish them from dispositions of deference. On other approaches the distinction is more readily apparent. For example, love might alternatively be understood as *aspiration towards a union of wills*: in making significant life decisions, I give serious deliberative weight to my partner's preferences, and our love and joint commitment seems partly to consist in our being disposed to treat one another's preferences as reason-giving. In this case, it's not my partner's mere preferences, but rather her reasons for so holding them, to which I accord deliberative weight. Similarly, Kyla Ebels-Duggan argues that love requires that one treat one's partner's desires as *evidence* of the existence of good reasons for pursuing their objects; so long as one treats such evidence as genuinely defeasible, this needn't amount to deference either. (See also §VI, on the topic of hidden deference.)

²¹ See, e.g., the comprehensive survey of views given by Chris Heathwood in 'Which Desires Are Relevant to Well-Being?', *Nous* 53, no. 3 (2019): n. 2, as well as the positive view developed in that paper.

²² As with the Deferential Wife, I assume for the sake of the discussion that the Doting Wife's devotion is fully voluntary, in at least the minimal sense of not having been adopted in response to coercive pressure or an excessively restricted option set. It should be emphasised again that although this assumption is dialectically necessary, it's not particularly realistic.

²³ Note, by contrast, that it's impossible for practical deference to be reciprocated. If what *A* most desires is whatever will best satisfy *B*'s desires, and if what *B* most desires is whatever will best satisfy *A*'s desires, and if *A* and *B* have relevant knowledge of each other's desires, then they're paralysed in a circular regress and neither can ever form a first-order desire. (Mark Fisher calls this the 'deadly embrace' (*Personal Love* (London: Duckworth, 1990), 30).) Thus whereas love may be mutual, deference is *inherently hierarchical*.

²⁴ See Meyers, *Self, Society, and Personal Choice*, pt. 3; Sandra Lee Bartky, *Femininity and Domination: Studies in the Phenomenology of Oppression* (London: Routledge, 1990), chap. 7; Jean Hampton, 'Selflessness and the Loss of Self', *Social Philosophy and Policy* 10, no. 1 (1993): 135–65; Friedman, *Autonomy, Gender, Politics*, chap. 6; Levey, 'Liberalism, Adaptive Preferences, and Gender Equality'.

²⁵ Westlund, 'Selflessness and Responsibility for Self'.

²⁶ Clare Chambers, 'Are Breast Implants Better than Female Genital Mutilation? Autonomy, Gender Equality and Nussbaum's Political Liberalism', *Critical Review of International Social and Political Philosophy* 7, no. 3 (2004): 1–33. Note that, as I deploy these terms here, the distinction between first-order and second-order autonomy does not directly correspond either to Meyers' distinction between episodic and programmatic autonomy (*Self, Society, and Personal Choice*, 48) or to Dworkin's distinction between local and global autonomy (*The Theory and Practice of Autonomy* (Cambridge University Press, 1988), 13). As I explain immediately below, a *one-off* deferential choice is still, on my view, nonautonomous at the first-order level but potentially autonomous at the second-order level, thereby manifesting autonomous nonautonomy

on a purely local or episodic basis. Conflict between first- and second-order autonomy need not, therefore, entail conflict between episodic and programmatic autonomy.

²⁷ On Chambers' account of choice under oppression (*Sex, Culture and Justice.*), oppressive social norms structure options such that certain people must sacrifice certain goods in order to reach other, more desired goals. For example, a woman might have two autonomous goals, material security and friendship, which are rendered mutually incompatible when social norms require her to follow her husband across the country or else face divorce and eventual financial ruin. On Chambers' account, this woman's choice to follow her husband manifests second-order autonomy inasmuch as it constitutes the successful pursuit of an autonomous second-order goal (material security), but a loss of first-order autonomy inasmuch as it is the failure to realise an autonomous first-order goal (friendship). I agree with Chambers' analysis of such cases. However, we are currently considering a version of the Deferential Wife case in which it is not a lack of acceptable alternatives that drives her to deference, but her own (socially conditioned) desire to live a certain kind of deferential life. Here, Chambers' account will likely deem her to be fully autonomous at *both* the first- *and* the second-order levels, since she faces no obstacles to the realisation of her first-order goal of a deferential marriage. Chambers' account is therefore unable to capture any sense in which the choice of a deferential life—whatever the reasons behind it—is, *inter alia*, the choice of a life of reduced self-determination. By contrast, the account of practical deference developed here provides a way of understanding how a disposition of deference can be thought to constitute a substantive relationship of interpersonal rule in itself, and so to undermine first-order autonomy even in the absence of external constraint. This allows the move that Chambers makes

in the case of external compulsion (i.e., that second-order autonomy is preserved at the expense of first-order autonomy) to be replicated in cases in which deference has been genuinely internalised.

²⁸ See e.g. Simone de Beauvoir, *The Second Sex*, trans. H. M. Parshley (London: Vintage Books, 1997); Bartky, *Femininity and Domination*; Paul Benson, 'Free Agency and Self-Worth', *The Journal of Philosophy* 91, no. 12 (1994); Catriona McKinnon, 'Self-Respect and the Stepford Wives', *Proceedings of the Aristotelian Society* 97, no. 3 (1997): 325–30; Stoljar, 'Autonomy and the Feminist Intuition'.

²⁹ See also Serene J. Khader, *Adaptive Preferences and Women's Empowerment* (Oxford: Oxford University Press, 2011), chap. 2. Khader argues, similarly, that preferences may be fully autonomous even when they are incompatible with, and have been formed under conditions unconducive to, basic flourishing. For Khader, therefore, people like the Deferential Wife needn't lack autonomy in any meaningful sense; rather, the choice of a deferential life is potentially criticisable on perfectionist grounds (180–85). However, Khader never tells us exactly what it is about deference that might render it incompatible with basic flourishing. Part of the aim of the present paper has been to do precisely this: as I have argued, practical deference is necessarily incompatible with a certain form of substantive, first-order self-government—or 'autonomy'—that is nevertheless distinct from the *second-order* autonomy that Khader targets in her discussion.

³⁰ John Stuart Mill, *On Liberty* (John W. Parker and son, 1859); Wilhelm von Humboldt, *The Limits of State Action* (Cambridge University Press, 1969); Joseph Raz, *The Morality of Freedom*, 149 (Oxford University Press, 1986); Ben Colburn, *Autonomy and Liberalism* (New York and London: Routledge, 2010).

³¹ As regards autonomy in its deontic role, I briefly note the following. Autonomy in this sense is solely a first-order matter: if I'm high on hallucinogenic drugs, then I can't validly consent to anything, regardless of whether I freely chose to take them in the first place. It might be thought obvious that, if the Deferential Wife lacks autonomy, it isn't in this sense of being incapable of entering into contracts or consenting to different forms of treatment (indeed, it may be thought highly disrespectful, and aggravating of her oppression, to suggest this). Yet matters are not so simple, at least in the more extreme cases. Imagine, for instance, that the Deferential Wife were to visit a Dutch-style euthanasia clinic, and to state in her interview that she's pursuing this option primarily because it's what her husband wants; we might have grave doubts about whether her choice were fully autonomous in the relevant sense. Or suppose that she were to respond to a request to participate in a complex medical study by saying 'ask my husband'; this would certainly raise a question about her ability to validly consent to the study. (Thank you to Joseph Millum and to Ben Colburn for helpful discussions around these cases.) I do not pursue these issues here.

³² See, in particular, her discussions of the 'Taliban Woman' and the 'Would-be Surrendered Woman': *Personal Autonomy in Society*, 60–68. These cases are mirror images of one another: in the former, a woman lacks substantive control over her life and endorses this state of affairs; in the latter, a woman possesses substantive control over her life but repudiates this state of affairs. Oshana deems the first woman nonautonomous and the second woman autonomous, thereby taking autonomy to require first-order but not second-order control over one's life.

³³ *Autonomy and Liberalism*. Raz, perhaps wisely, avoids this matter altogether, simply observing that this question 'raises tricky issues' (*The Morality of Freedom*, 374, n.1).