

# Motivational Internalism

EDITED BY GUNNAR BJÖRNSSON,  
CAJ STRANDBERG,  
RAGNAR FRANCÉN OLINDER,  
JOHN ERIKSSON,

*and*

FREDRIK BJÖRKLUND

**OXFORD**  
UNIVERSITY PRESS

**OXFORD**

UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

Oxford New York  
Auckland Cape Town Dar es Salaam Hong Kong Karachi  
Kuala Lumpur Madrid Melbourne Mexico City Nairobi  
New Delhi Shanghai Taipei Toronto

With offices in  
Argentina Austria Brazil Chile Czech Republic France Greece  
Guatemala Hungary Italy Japan Poland Portugal Singapore  
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trademark of Oxford University Press  
in the UK and certain other countries.

Published in the United States of America by  
Oxford University Press  
198 Madison Avenue, New York, NY 10016

© Oxford University Press 2015

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer.

Library of Congress Cataloging-in-Publication Data  
Motivational internalism / edited by Gunnar Björnsson, Caj Strandberg,  
Ragnar Francén Olinder, John Eriksson, and Fredrik Björklund.  
pages cm. — (Oxford moral theory)  
Includes index.  
ISBN 978-0-19-936795-5 (hardcover : alk. paper)  
1. Ethics. 2. Moral motivation. 3. Internalism (Theory of knowledge)  
I. Björnsson, Gunnar, editor.  
BJ45.M68 2014  
170'.42—dc23  
2014018542

9 8 7 6 5 4 3 2 1  
Printed in the United States of America  
on acid-free paper

## 13

## Tempered Internalism and the Participatory Stance

KATE MANNE

Motivational internalism (also known as judgment internalism) is a position designed to capture the intuitively tight connection between moral judgments and motivations. It's common to hear it billed in casual philosophical conversations—around the water cooler, so to speak—as being the claim that moral judgments are intrinsically or necessarily motivating.<sup>1</sup> But this claim is likely stronger than it needs to be to capture the connection. For one thing, it doesn't factor in the way in which the connection between moral judgments and motivations seems strongest and most direct when it comes to moral thinking that is distinctively *first-personal*.<sup>2</sup> The connection seems likely to be looser when it comes to moral judgments regarding distant others.

To see this, suppose that an agent judges that a geographically or historically remote group of people are or were engaged in morally bad behavior.

<sup>1</sup> Note that 'moral judgment' is intended throughout this chapter to be a neutral theoretical term for our basic moral thoughts or responses. As such, it is meant to be compatible with non-cognitivist accounts of moral thought and language, rather than assumed to be belief-like in nature. Note too that I mean to be individuating judgments psychologically rather than according to their contents throughout this chapter—i.e., as the mental act/state of judging rather than that which is judged to hold.

<sup>2</sup> For another thing, it's not clear that moral judgments need to be construed as *necessarily* motivating—i.e., because of properties intrinsic to the relevant mental states. As Jon Tresan has pointed out, moral judgments themselves might be perfectly ordinary, motivationally inert beliefs (2006). Maybe they just won't count as being moral judgments unless they play a certain role in motivating the agent who makes them. So moral judgments might entail motivation, i.e., always or necessarily be accompanied by motivation, without involving or necessitating motivation by their very nature. This view would naturally make life easier for the cognitivist looking to accommodate the connection.

Why think that moral judgments like this one will entail motivations on the part of the agent to do anything in particular? There might be nothing that she *can* do to discourage this behavior, in which case it seems perfectly normal and healthy for her not to want to do so. And even if she could do, she might think that it would be totally inappropriate for her to intervene—such that she wouldn't even dream of it, or be inclined to do so at all. It's also at least unclear that the agent would be motivated to discourage such behavior from being engaged in closer to home (or anything along these lines). After all, what's judged morally bad for the goose may be deemed permissible for the gander. So the connection between moral judgments and motivations seems intuitively more remote when the subjects of the judgment are more remote as well.<sup>3</sup>

In light of this, I suggest that we should understand the core claim of motivational internalism as the weaker claim that a moral judgment made by an agent about *what she herself (currently) ought to do* will entail motivations on her part to act in accordance with this judgment—even if such motivations are subsequently overridden, or subject to *akrasia* (say).<sup>4</sup> This way of finessing the claim is simple enough, and it is certainly not unusual.<sup>5</sup> But I believe it nonetheless contains a valuable clue as to how to weaken or temper motivational internalism so as to render the position more plausible, while preserving the intuitive kernel of truth which I believe the position to have. This kernel has

<sup>3</sup> It might still be insisted that the agent must be prone to blame the people involved, or to feel disapproval for their actions, or some such. I suspect that such claims would be too strong as well, and for similar sorts of reasons. Intuitively, we seem to be able to engage in dispassionate moral reflection about the distant past, in particular. Considering this issue properly would take me too far afield. But, however it turns out, establishing a connection between moral judgments and affect would not suffice to answer my driving question about their connection with motivations proper, which I take to either be or to entail dispositions for action of some kind. And my primary interest in motivational internalism has to do with whether and how morality might be distinctively practical or action-guiding. More on this shortly.

<sup>4</sup> Many questions admittedly remain about how best to understand, develop, and extend this thesis further. For example, should moral judgments that are *pro tanto* rather than all things considered (e.g., judgments about the moral reasons which one has) be held to be subject to this claim as well? There is also the question of whether the claim should be generalized so as to encompass certain other types of normative judgments (e.g., prudential judgments). Another issue is how strong the relevant motivations should be held to be, in proportion to various measures of the strength of the judgment. That is, how disproportionately weak should these motivations be allowed to be, compatibly with a commitment to motivational internalism? Fortunately, I can afford to remain neutral on these further issues here.

<sup>5</sup> Some formulations of motivational internalism have it that judging some state of affairs to be good or bad motivates us to promote or prevent it, respectively. But explicitly first-person-centered versions of motivational internalism along the lines of the above claim seem at least as common. For a helpful discussion of this issue, which leads to a similar take on the core claim of motivational internalism, see Jeanette Kennett and Cordelia Fine, on 'in situ' moral judgments (2008).

to do with the fact that morality is meant to tell us what to do. It is supposed to *speak* to us, somehow, in our capacity as agents. And one way in which its prescriptive and practical character would plausibly show up is in our being disposed to do what we think is morally being asked of us. Otherwise, we wouldn't really be taking ourselves to be subject to a *moral* requirement (so the thought continues). It would be hearsay, not testimony, to stick with the analogy. That is, we would merely be giving our sense of what morality is said to require by true believers—among other more or less familiar possibilities.<sup>6</sup>

There clearly seems to be something to this. Still, we should be careful not to exaggerate either the universality or the individual reliability of the disposition to form the subsequent motivations. Some people appear to have views of some kind about what morality requires of them. And yet they seem less than anxious—or indeed, completely unmotivated—to do as they think they're told. Nor are such amoral characters purely the stuff of nightmares or Dostoyevsky novels. Sometimes, we simply feel too tired to do all of the things we take morality to ask of us (tiredness notwithstanding). We are lazy about recycling our newspapers; we fail to send that awkward e-mail; we allow ourselves to capitulate on some small matter of principle. Sometimes we feel guilty about these run-of-the-mill lapses; but at other times we don't. Maybe a little bit of an amoralist lurks within us all, then.

So how are we to make sense of the seemingly tight connection between first-person moral judgments and motivations to act, without going too far in the opposite direction? The rough idea behind my proposal about how to temper motivational internalism is this: sometimes, when an agent makes a first-person moral judgment about what she herself (currently) ought to do, she is telling herself what to do but responding only passively, almost as if she is overhearing a direction being issued to someone else. And she may also feel like a spectator to her subsequent behavior. All told, she fails to relate to herself as an active *participant* in moral practice. At one level, the agent acknowledges that the moral judgment she makes properly applies to her, and it is at least open to us to take her judgment at face value. So she might say or think to herself “I really ought to  $\phi$ ,” and this statement may be quite sincere and not even half-hearted, exactly. Rather, or better, it is bloodless—or, alternatively, mindless. For, her thinking might be more perspicuously represented as involving two somewhat disconnected thoughts, namely: “In such circumstances, one

<sup>6</sup> The most famous view along these lines was of course defended by R. M. Hare, when he talked about moral terms being used in their inverted commas senses (1952, 124–125, 164–173). Alternatively, we might be giving our sense of what would be in the extension of the best candidate replacement or cleaned-up version of the messy, folk concept. Another possibility is that this is a tentative or half-hearted moral judgment which is accompanied by similarly or indeed disproportionately weak motivations, which are hence easily averted, overridden, and overlooked.

really ought to do such-and-such” and (limply, ineffectually) “I seem to be in these circumstances, and to be that one.” The agent’s sense of *agency* has somehow gone missing. She may have views about what to do, but it’s as if she forgets herself.

So that is the basic idea behind my proposal in this chapter. The view that I develop and call tempered internalism says that a moral judgment made by an agent about what she herself (currently) ought to do will entail motivations on her part to act in accordance with this judgment, *provided that she takes the practical stance toward herself which is fitting for such judgments*—that I call the ‘participatory stance,’ following P. F. Strawson (1962/2008). The normative notion of ‘fittingness’ plays an important role here. It would not be very impressive to show that a currently applicable—or, as we might also say, ‘actionable’—first-personal moral judgment made from the participatory stance will inevitably be accompanied by the relevant motivations.<sup>7</sup> For, as we will see, such motivations are plausibly an integral component (although they are by no means exhaustive) of this stance. Rather, the idea here is that these judgments *ought* to be made from the participatory stance; they are otherwise defective. And while few would disagree that *agents* ought to make such moral judgments in this mode (since they ought to view themselves as participants or insiders within the moral community), it will still be controversial to say that such a judgment would be defective *qua* judgment insofar as it is not. This is what I’ll be going on to argue. I’ll then argue that the psychological component which may be missing from such judgments is plausibly none other than the agent’s sense of herself as an active participant within the relevant practice—from which motivations to comply with the norms held to govern her conduct therein would follow.

Tempered internalism is, of course, not the only way of trying to soften the core claim of motivational internalism so as to deal with apparent counterexamples, while preserving at least something of the spirit of the original position. In particular, some theorists have moved to restrict motivational internalism so as to only apply to moral judgments made by agents who are psychologically normal, or practically rational, or akin to the majority of people in their communities when it comes to such judgments (see, e.g., Simon Blackburn [1998], Michael Smith [1994], and Jon Tresan [2009], respectively, for views along these lines). But I think there are advantages to tempering

<sup>7</sup> Since the phrase “actionable first-person moral judgment” doesn’t exactly roll off the tongue, I will often drop the ‘actionable’ qualifier in the remainder of what follows. But bear in mind throughout that the relevant class of judgments is not intended to include judgments about past or future obligations, only current or ongoing ones, which the agent is (perhaps in addition to this) actually in a position to act upon.

motivational internalism by making an exception for certain putatively defective moral *judgments* in the first instance, insofar as they are made from the wrong stance, rather than employing the more indirect strategy of making an exception for certain judgments based on the defective or abnormal global condition of the *agent* who makes them. For one thing, the latter way of going plays into the externalist's suspicion that it is not the fact that a moral judgment is being made that entails and explains the presence of the relevant motivations. Rather, these motivations might be held to derive from an independent *desire* to adhere to moral standards that will be harbored by all or most of the agents who meet the relevant condition. For another thing, I think that my proposal may be better placed than at least some of the aforementioned ones to accommodate the fact that lacking moral motivation can be a local phenomenon. We all have our moments, as I suggested in opening. Or, more cautiously, we who are less than saintly will often have more or less specific moral weaknesses, in the sense of being unmotivated to do as we say. Similarly, some moral judgments that would usually motivate us to act can fail to be motivating under certain conditions—for example, when we are 'tapped out' or 'beside ourselves,' so to speak. Tempered internalism is well-placed to account for such possibilities. For it seems plausible to think that we may temporarily take leave of the participatory stance toward ourselves when it comes to some actions if we are tired, or dejected, or emotionally exhausted (say).<sup>8</sup> The participatory stance can in fact be exited in many different ways and for many different reasons, as we will be seeing. But the idea is nonetheless intended to provide a kind of unifying story at the level of moral psychology about the judgments that constitute genuine exceptions to the internalist rule—whatever the overall nature of the agents they belong to.

So that is the backstory; here is the plan. In section 13.1, I'll introduce the Strawsonian idea of the participatory and the spectator stances, in connection with particular first-person moral judgments. Then, in section 13.2, I'll be in a position to tell a story about when and why such judgments may float free of motivation, insofar as we may fail to relate to ourselves in the participatory mode for the purposes of the action which the judgment is about. In section 13.3, I will compare my approach to such judgments with that of an externalist who takes an agent's first-personal moral judgments to have no motivating force independently of her desire to be moral or to be a good person. I will argue that an externalist of this sort will have trouble explaining why there is an intuitive disanalogy between moral judgments and other normative judgments that clearly do not motivate, absent some independent, accompanying

<sup>8</sup> This seems more plausible than the idea that we pop in and out of seemingly more global states like psychological and social abnormality, in particular.

desire to adhere to the relevant norms or standards. For the former, but not the latter, will tend to sound *hollow*—as if they are missing some psychological component which such judgments ought to have or be accompanied by. I'll then suggest that what they are missing is the agent's sense of herself as an active participant in moral practice, as the tempered internalist position I develop would have it.

### 13.1. The Participatory Stance

P. F. Strawson's celebrated paper, "Freedom and Resentment" (1962/2008) introduced an important distinction between the interpersonal and the objective stances that we can take to other people, from the second-person point of view. This is the distinction that everyone remembers. However, Strawson also suggests that we can take an analogous pair of stances toward *ourselves*—that is, we can relate to ourselves in the participatory and the non-participatory modes. I'll call this self-directed, non-participatory stance the 'spectator stance,' so as to have a nice label for it.<sup>9</sup>

How should we think about the interpersonal and the objective stances? And how should we understand the first-personal analogues of these second-personal attitudes?<sup>10</sup> These are the questions to which I'll now turn.

Here is how Strawson explains the distinction between the interpersonal and the objective stances, in a rich and suggestive passage:

What I want to contrast is the attitude (or range of attitudes) of involvement or participation in a human relationship, on the one hand, and what might be called the objective attitude (or range of attitudes) to another human being, on the other. . . . To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of senses, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided, though this gerundive is not peculiar to cases of objectivity of attitude. The objective attitude may

<sup>9</sup> As we'll see, Strawson doesn't distinguish the two sets of labels, and often intermingles the terminology (e.g., referring to what I call the second-personal 'interpersonal stance' as the 'participant reactive attitude'). But it will be helpful for my purposes to keep each of these pairs of labels distinct.

<sup>10</sup> I call them 'attitudes' (as well as 'modes') for convenience, and in keeping with Strawson, in a sense that is not supposed to prejudge the issue of whether they have a partly conative or affective nature. I discuss what stances might be in terms of our mental life below.

be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in interpersonal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. If your attitude towards someone is wholly objective, then though you may light him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him. (1962/2008, 10)

The interpersonal and the objective stances thus involve several different, albeit inter-related elements that contrast with one another. They are a “network of human attitudes,” tending in opposite directions. They involve *seeing* someone in one way rather than another—as a fitting interlocutor, say, versus something else or someone lesser. And they involve *treating* someone, or at least being disposed to treat him, in one way rather than another—by being prepared to interact with him, rather than keeping one’s distance, literally or psychologically. These attitudes also constrain and enable the relevant emotional and relational possibilities in various important ways. For example, one can love someone to whom one does not relate interpersonally in some ways but not others, Strawson tells us. And, crucially, such a person is outside “the game of giving and asking for reasons,” to use Wilfrid Sellars’s phrase. Strawson held, most famously, that when one takes the objective stance to someone, it is then inapt (or does he mean literally impossible?) to resent or to blame him.<sup>11</sup> These ‘reactive attitudes’ are restricted in their proper targets to people who we are relating to in the interpersonal mode—which, among other things, involves seeing him as someone who similarly experiences himself as a free agent in practice.<sup>12</sup>

<sup>11</sup> One reason for holding that it is merely inapt is that there are plenty of cases that look (from the outside) as if the agent is relating to someone in the objective mode, and also blaming her for everything. And I assume I am not alone in sometimes feeling a wave of irrational anger when a person accidentally treads on my toes or snores loudly through the night (say). What seems hard to pull off is maintaining (let alone defending) both of these attitudes from the inside, i.e., in full cognizance of both of them. So if we still feel angry at the person, we may start to build a case against them as oblivious or negligent or otherwise irresponsible. In other words, we contrive to find reasons for maintaining the stance toward them from which our lingering resentment is still eligible to be apt.

<sup>12</sup> In trying to disentangle these various different aspects of the second-personal stances, I don’t mean to be trying to explain them away, which I think would be somewhat contrary to the spirit of Strawson’s views. Rather, I take a stance to be an illuminating holistic explanatory

If relating to another is partly a matter of recognizing that he is an “I” to himself, as this might suggest, then our relationship with ourselves becomes a natural further subject. And if we can relate to others as insiders or outsiders to the moral community at large, then perhaps we can relate to ourselves in a similar manner.<sup>13</sup> So, having reminded you of the best-known pair of attitudes now, we can consider the first-personal analogues of the interpersonal and the objective stances. Here’s how Strawson introduces these ‘self-reactive’ modes of thinking:

... the picture is not complete unless we consider also the correlates of these attitudes on the part of those on whom the demands are made, on the part of the agents. Just as there are personal and vicarious reactive attitudes associated with demands on others for oneself and demands on others for others, so there are self-reactive attitudes associated with demands on oneself for others. And here we have to mention such phenomena as feeling bound or obliged (the ‘sense of obligation’); feeling compunction; feeling guilty or remorseful or at least responsible; and the more complicated phenomenon of shame. (1962/2008, 16)

Unfortunately, Strawson does not say much more about the specific contours of these first-personal attitudes. But we can fill in the picture by drawing on the parallel with the second-personal stances, as I’ve moved to understand them. Taking the spectator stance toward oneself would then naturally involve seeing oneself as *passive* rather than active, and treating oneself as *exempt* rather than implicated. And the participatory stance will involve the contrasting set of tendencies, in terms of self-conception and self-governance.<sup>14</sup> It would

category within moral psychology, which is meant to crystallize a certain characteristic, overall ‘take’ one can have on another—or, as I’ll now argue, oneself.

<sup>13</sup> And, as Strawson points out, we can also view others’ treatment of others as happening within our moral community—or not, as the case may be. These attitudes are thus the third-personal analogues of the second-personal stances. This pair of stances evidently bears on my opening suggestion that moral judgments about the behavior of distant others do not necessarily have a behavioral or even affective profile associated with them (see n. 4). Such a profile would plausibly only obtain when one’s perspective on the action is more like the stance of an engaged or involved bystander, who is ready (if not necessarily able) to step in on behalf of the wronged or injured party.

<sup>14</sup> For the purposes of the discussion, I am going to adopt the (no doubt overly simplistic) assumption that the participatory and the spectator stances are mutually exclusive and exhaustive. This is why I take the stances to involve merely contradictory, rather than opposite or contrary, tendencies. But we should not forget the existence of subversives (i.e., apparent immoralists), who are motivated to undermine or disrupt the relevant social practice.

thus naturally be bound up with a ‘sense of obligation,’ and a tendency to take responsibility, as Strawson claims in the above passage. It will also involve a tendency to feel genuine guilt and shame for admitted improper behavior—along with a readiness to submit to proper punishment unresentfully, Strawson adds (1962/2008, 23). On the flipside, we might think that the participatory stance opens up the possibility of a higher grade of self-respect or self-love as well. My supplementary suggestions here are offered merely by way of (I hope) plausible suggestions as to how the framework might be filled in. We will have to engage in detailed moral psychological work in order to say more about how the various components of these stances would naturally hang together. I will turn to one such question in a moment—namely, the sorts of motivations involved in the participatory versus the spectator stances.

But first, we need to get clearer on the relationship between our stance toward a person and the moral judgments that we might hope to level in their direction. We adopt a stance with respect to a person—be it another or ourselves—with respect to particular (although not necessarily isolated) pieces of behavior which they might exhibit.<sup>15</sup> But it is only when we view them as a person and as an agent that we can credit these bits of behavior *to* them as being their actions—that is, as expressive of their agency and for which they will have to answer.

Hence, when we make the second-person moral judgment “*You* ought to  $\phi$ ,” we typically relate to the other interpersonally for the purposes of  $\phi$ -ing. That is, we take her to be actively responsible for  $\phi$ -ing, and are prepared to remonstrate with her, or at least to muster up resentment, if she fails to follow through. Similarly, when we make the first-person moral judgment, “*I* ought to  $\phi$ ,” we typically relate to ourselves as agents for the purposes of  $\phi$ -ing. That is, we take ourselves to be actively responsible for  $\phi$ -ing, and are prepared to admonish ourselves, or at least to feel compunction, if we fail to follow through.

However, such attitudes can be suspended, and the corresponding judgments withdrawn, under certain conditions. We will suspend the interpersonal attitude to another if she seems to be ‘not in her right mind,’ to be ‘not herself,’ or if she is only a child.<sup>16</sup> Whatever the reason for the suspense, we may then go into ‘management mode,’ treating the person in question as a problem to

<sup>15</sup> I cast things in terms of particular pieces of behavior because there seems no principled limit as to how isolated the exceptions we need to make for people in moral matters can be. But of course we will often adopt the same stance toward another person—or ourselves—with respect to whole regions within moral practice.

<sup>16</sup> Strawson also talks about a person who is deemed to be “an idiot or a moral idiot” (1962/2008, 13)—an unfortunately dated choice of words, which should still give us pause here, since it brings up the crucial question of how the objective stance might be squared with recognizing the full humanity of the less than fully rational.

be handled, rather than someone to be reasoned with (or alternatively, put up with, as someone who is entitled to be doing her own thing). And perhaps we also suspend the participatory attitude toward ourselves—along with the corresponding moral judgment—if we realize that we are currently indisposed in such ways. We may then go into ‘self-management mode,’ psychological deficits permitting. Or we may hand ourselves over to the custody of another.<sup>17</sup>

These are the sorts of cases that are by far the most salient and hence widely discussed. However, and crucially for my purposes, the interpersonal mode can be suspended for other reasons too. Strawson thought that we sometimes resist the “strains of involvement” with another person because we’re emotionally exhausted—or just plain lazy (see 1962/2008, 10, 13, 18<sup>18</sup>). The moral judgment “You ought to  $\phi$ ” may then remain intact, but the usual enforcement mechanisms will tend to fall by the wayside. We will be more or less unperturbed if the person fails to do what we take to be required of her—or, as we might then put it, what is generally required of someone in this sort of position. We cannot work up the energy to care. We do not want to deal with it. So we turn a blind eye; we choose to let it go. I would similarly suggest that we sometimes resist the “strains of involvement” even in our own affairs—because, among other possibilities, we’re overwhelmed with other responsibilities, or again, simply lazy. The moral judgment “I ought to  $\phi$ ” may then remain intact, but the usual self-enforcement mechanisms will be similarly lacking. We will be more or less unperturbed by our own failure to do what we take to be required of us—or, more congenially, what is required of someone in the position in which we seem to find ourselves. We let ourselves go; we cut ourselves some slack; or, we give up the ghost. Even though it is our business, it is treated as if it is not—as if it is the behavior of some distant third party, say, whose conduct we have little interest in or, alternatively, no control over. The usual attendant motivations thus fail to gain traction.

The usual motivations being what? It seems safe to say that, in order to relate to myself as a participant or agent for the purposes of some action, I must be

<sup>17</sup> There are of course epistemic difficulties involved in doing so; we have to be sufficiently ‘with it’ to recognize that we are sufficiently ‘out of it’ to merit such a reprieve. And it’s also doubtful that we should or would typically cut ourselves as much slack as we cut other people who appear to have lost it. It would be interesting to think more about whether and why this is so, in connection with Bernard Williams’s notion of ‘agent regret.’ But cases of this kind are not my primary concern here.

<sup>18</sup> Strawson: “The objective attitude is not only something we naturally tend to fall into in cases . . . where participant attitudes are partially or wholly inhibited by abnormalities or by immaturity. It is also something which is available as a resource in other cases too . . . [W]e can sometimes look with something like the same eye on the behavior of the normal and the mature. We have this resource and can sometimes use it: as a refuge, say, from the strains of involvement . . .” (1962/2008, 10).

motivated to abide by the normative standards which I take to apply to such actions, and which I take to have implications about whether or not I should be performing them.<sup>19</sup> For, if the participatory stance did not involve that much, then it is hard to see why it would be centrally connected with the capacity for the characteristic self-reactive attitudes that are integral to this stance. In what sense could one be said to feel genuine guilt or shame for violating a norm which one would not be moved to abide by, were one to find oneself now placed in the same circumstances again? And what sense could we make of the idea that a person so unmotivated nonetheless has a genuine sense of responsibility regarding that action?

I see little reason to suppose that things will be any different when it comes to moral matters, specifically. In which case, if someone views herself as a participant or agent for the purposes of some action, then she will be motivated to act in accordance with her moral judgments about whether or not she should be performing it. From the opposing spectator stance, however, all such bets are off. That is, while the agent *might* be motivated to perform the action for other reasons, or under another description, the entailment will no longer hold simply in virtue of her having made this judgment. This moral judgment will no longer come with such motivational guarantees or backing.

The idea of the participatory stance is now tolerably clear, I take it. And, like the proverbial happy family, this proactive, ‘can-do’ attitude may be more or less alike whenever it occurs, whereas the spectator stance appears on the face of it to come in a variety of forms and to serve a variety of purposes (as is also plausibly the case when it comes to the objective attitude which we take to other people). In the next section, I will canvass some of the major possibilities here, in connection with first-personal moral judgments that may float free of motivation.

## 13.2. Disengagement in Action

According to the story so far, when an agent makes a currently applicable or ‘actionable’ first-person moral judgment that floats free of motivation, then she must be taking the spectator stance toward herself with regard to whatever actions this judgment is supposed to regulate. Her judgment may be *about* her various moral responsibilities, but she does not relate *to* herself as morally

<sup>19</sup> A complication is that actions are of course performed under various descriptions. So, strictly speaking, we should add that the description of the action for the purposes of which I take myself to be an agent, and the description of the action which figures in the normative judgment, are a tolerably close ‘match.’ But for the sake of simplicity, I leave this qualification implicit.

responsible for what she does in so making it. There is subsequently a mismatch between her practical point of view and the grammatical point of view evinced in the *content* of her judgment. And this, I'll argue in the next section, constitutes a defect in her judgment, rather than merely being evidence of an attitude problem generally.<sup>20</sup>

But first, how is the flight from the participatory stance meant to go in practice? An example will serve to illustrate some of the psychological mechanisms at work in the mundane sorts of cases with which we're all too familiar. Imagine someone who is prepared to agree (at least when he is pushed) that eating meat is wrong, that he ought to give it up. Still, even as he thinks this, he has no intention whatsoever of giving up "his meat" (as people sometimes put it). Indeed, he refuses to so much as even feel guilty about eating the ham sandwich currently halfway to his mouth. How might he pull off this (seeming) psychological trick? Lots of ways come to mind (and, to be candid, have come to *my* mind in fending off similar qualms). He might be somewhat in denial about what he is doing—framing it as having a Croque Monsieur, or simply, lunch. He doesn't think too hard about the contents of his sandwich, lest he lose his appetite. Or he might go for distraction (enter the ham sandwich). He might make a variety of excuses for himself, thereby refusing to apply his moral judgments about what ought to be done generally, or all else being equal, to his current situation. He really needs the protein, say (unlike everyone else). So *ceteris*—and he—ain't *paribus*. Or, alternatively, he happens to be especially fond of ham. He just can't help himself, he says (as he helps himself to another sandwich). And if 'ought' implies 'can' . . . well, you can see where this is going.

The agent might also deflect or prevaricate, perhaps by focusing on ways in which we are *all* bound up in morally bad practices, or otherwise complicit. He feels powerless, he says, to do everything he ought to, morally, to live up to all of his many and various responsibilities. So why even try? Or perhaps he will maintain that this sandwich is actually something of a one-off, and not something he would have gone out of his way to have, had the choice been left up to him. But his mother packed his lunch, you see. He might additionally allege, as ambivalent meat-eaters do not infrequently, that he only really eats meat when it is served *to* him as a guest in other people's houses. There seems to be a surprising amount of meat being foisted on us by our families and friends.

<sup>20</sup> Indeed, such a mismatch may not even count as being evidence of a problem. Taking up 'non-matching' practical perspectives can seemingly be neutral or even positively praiseworthy. For example, taking the participatory stance rather than the spectator stance from one's position in the grandstands seems like a plausible way of envisaging empathetic or vicarious participation—which hardly seems a problem.

Depending on how sheepish he seems to be, and how loudly he doth protest, we might begin to wonder if he *is* in fact somewhat motivated to put down the sandwich. He might begin to seem at least a bit guilty, his protestations to the contrary notwithstanding. But, in any case, it's a bit of a moot point in the end. For, his guilt doesn't last. And nor does his sandwich.

The tempered internalist's approach to the amoralist thus clearly differs—and is meant to be an improvement—on having to insist (as per internalism unmodified) that an unmotivated person's apparent moral judgment must be insincere, or conceptually aberrant, or mean something different in his mouth from the equivalent-looking judgment made by a motivated maker. (Of course, the tempered internalist will acknowledge that such possibilities *may* obtain; the point is that they need not.) The moral judgment in question need not be insincere or even tentative, so much as disengaged. And the aberration in question need not show up at the level of the *content* of the judgment “I ought not to eat meat;” rather, it may have to do with the agent's way of relating to *himself* when it comes to its content.<sup>21</sup>— Specifically, he stands at a distance from the actions which the judgment would prescribe or proscribe. In eating his ham sandwich, he is a case in point. He avoids his own gaze; he is a passive consumer; he chows down mindlessly, thanks to psychological mechanisms like the ones noted above (among other possibilities). And what is common to these various mechanisms, by the lights of tempered internalism, is that they help the agent to distance and detach himself psychologically from the sense of being morally responsible for what he is putting in his mouth—making for a moral judgment that therefore may not last very long, and may not be accompanied by motivation even if it does. But on this picture, there is no need to deny that such a judgment may be present and the motivation absent. For, this absence can in turn be explained by the absence of something else—namely, the agent's sense of himself as such, that is, as active and acting. And, given this, it isn't too surprising that his first-personal moral judgments would be lacking the normal motivational oomph of the “I.”

As well as more or less isolated moral weaknesses, such as might be exhibited by the recalcitrant carnivore above, there are more global kinds of disengagement and disaffection to consider. An agent might be morally disengaged, specifically—being something like the prototypical amoralist, who has no sense of involvement or engagement in more or less any moral matter. In particular, she may not see herself as responsible for her actions as described in terms of their impact on other people (see n. 20). Psychopaths and sociopaths are the real-life characters who are the most obvious candidates for fitting this

<sup>21</sup> So if anything deserves inverted commas here, which I tend to doubt, it is the “I.” Compare the notion of the royal ‘we.’

general profile—assuming that they even make the relevant judgments to begin with.<sup>22</sup> But an agent might also have difficulty regarding herself as an agent for *any* purpose, in having a marked tendency toward ‘dissociation,’ to use a term which is more common in psychology than philosophy.<sup>23</sup> As Marlene Steinberg and Maxine Schnall describe the phenomenon, such dissociation will often involve a sense of “detachment from yourself or looking at yourself as an outsider would,” and/or feeling “detached from your emotions, like a robot or an automaton” (2001, 31). People in such dissociative states might be described as not really being present, or not being fully *there*. The agent isn’t quite herself, or quite at one with herself—she is somewhere else, or perhaps nowhere whatsoever. She may be doing various things (at least in some sense of the word ‘doing’) but feeling passive all the while, as if the things which she is doing are merely happening *to* her or *via* her body. She may watch himself act as if watching a stranger.

Albert Camus’s *The Outsider* (also translated as *The Stranger*) is the classic literary illustration of the more extreme end of this kind of phenomenon (which shades into the everyday, as I’ll shortly argue). But for people like Camus’s anti-hero Meursault, their behavior is less a matter of failing to do something or, similarly, doing nothing. It is more of a matter of sinking to the level of being a “nothing doing”—a phrase that crops up more or less randomly, but appropriately enough, at the beginning of one translation<sup>24</sup> (1942,

<sup>22</sup> But, according to Kennett and Fine, psychopaths tend to “. . . make exceptions for themselves. Their hypothetical or third-personal moral judgments are inconsistent with their moral judgments about their own actual situations and actions, and indeed we question whether they ever spontaneously engage in moral assessments of their own projected actions” (2008, 176). They also worry that Adina Roskies’s putative counterexample to motivational internalism, patient EVR, may not “recognize when he is in a situation that falls under a third-personal hypothetical moral judgment[.] Can he translate the knowledge into the first person?” they ask (2008, 183). So their take on the empirical evidence regarding psychopathy and acquired sociopathy is evidently quite congenial to the views I am developing here.

<sup>23</sup> For some representative contemporary treatments, see, e.g., Steinberg and Schnall (2001), Simeon and Abugel (2008), and Dell and O’Neill, eds. (2009). Note that dissociation is sometimes construed somewhat more broadly than I intend here, so to include distancing and detachment from one’s immediate environment, with ‘depersonalization’ figuring as a more specialized concept that is closer to my meaning. (And the following quotation in the text is actually part of Steinberg and Schnall’s characterization of depersonalization.) Also note that multiple personality disorder, always a controversial diagnosis, was replaced in the DSM-IV by what is now known as ‘dissociative identity disorder.’ Some theorists have argued that the problem here initially was not too many identities, but rather too few—or, perhaps better, too flimsy a grip on the self. This may have allowed some therapists to unwittingly play a role in the construction of ‘multiples.’ See Ian Hacking for the classic, and enduringly controversial, discussion of this issue (1995, Chapter 1).

<sup>24</sup> It is used to translate “ce n’était pas possible,” which is how a character responds to Meursault’s request to turn off one of several lamps glaring at his mother’s wake—since, given the way these lights work, “c’était tout ou rien.” And the word *rien* echoes again some 80 times

7). Here is how the narrator Meursault (who notably lacks a first name) comes to commit a murder with no ostensible motive. He is blinded by the sun to the reality of his actions:

The heat was beginning to scorch my cheeks; beads of sweat were gathering in my eyebrows. It was just the same sort of heat as at my mother's funeral, and I had the same disagreeable sensations—especially in my forehead, where all the veins seemed to be bursting through the skin. I couldn't stand it any longer, and took another step forward. I knew it was a fool thing to do; I wouldn't get out of the sun by moving on a yard or so. But I took that step, just one step, forward. . . .

Beneath a veil of brine and tears my eyes were blinded; I was conscious only of the cymbals of the sun clashing on my skull, and, less distinctly, of the keen blade of light flashing up from the knife, scarring my eyelashes, and gouging into my eyeballs.

Then everything began to reel before my eyes, a fiery gust came from the sea, while the sky cracked in two, from end to end, and a great sheet of flame poured down through the rift. Every nerve in my body was a steel spring, and my grip closed on the revolver. The trigger gave, and the smooth underbelly of the butt jogged my palm. And so, with that crisp, whipcrack sound, it all began. (1942, 38–39)

He sinks four more bullets into the dead body, “on which they leave no visible trace.” He does not see his shots land so much as he hears them—as “loud, fateful raps on the door of my undoing” (1942, 39). Such was the “undoing” of a stranger to himself. The sun and the sea (his namesake) absorb him. But long before that, he was a dead man walking.

No doubt losing oneself as totally and seemingly irrevocably as this is rare. I am not suggesting that people walk around in full-blown dissociative fugue states very often. But I think it is common to mentally withdraw to some degree or other in the course of ordinary life. Indeed, some theorists have argued that mild dissociation is a normal and perhaps even near-universal human experience—a common way of dealing with stress, even boredom (see, e.g., Steinberg and Schnall 2001, 20–21). The range of ways we have of talking about disengagement in action tends to bear this out. We talk about acting absent-mindedly, mindlessly, passively, numbly, and vacantly. We may not have

throughout the original text. At the trial, for example, the Prosecutor opines on the state of Meursault's soul: “He said he'd studied it closely—and had found a blank, “literally nothing, gentlemen of the jury” (1942, 63). His lawyer demurs: “I have found something there” (1942, 65). This, and he, does not convince.

our heads in the proverbial game. We talk about freezing and “zoning out,” and getting “locked in” to some pattern of behavior. We may report feeling trapped, or stuck, or caught up in something. We drift around aimlessly. We talk about feeling detached, disconnected, alienated, and estranged from ourselves, even though we are doing things we (in some sense) choose to do. We talk about taking leave of our senses, and of being beside ourselves. We talk about feeling like robots, and zombies, and the feeling of sleepwalking through waking life. This would naturally extend to the moral life as well, I take it. Hence the common Australian moral rebuke, “You better wake up to yourself, sunshine.”

But I do not mean to be suggesting, and I am little inclined to believe, that we cannot be blamed for what we get up to when we’re not quite ourselves.<sup>25</sup> For one thing, I think that not facing up to what we’re doing, or what our actions are designed to achieve deep down, is often at least somewhat under our control. We are not so much unable as we refuse to see what we are up to, or what we are doing in spite of ourselves (as we say). There is also the phenomenon of voluntary intoxication—and, by extension, dissociation. Consider, for example, the time-honored strategy of drinking to excess in the face of temptation. One thing is then said to lead to the other. Which, even if true, is all too predictable and none too impressive. People can certainly be blamed for letting themselves go in these ways, even if their subsequent offense was committed whilst blotto.

It is also common to check out, I think, when we act with a sense of *delegated* agency—when we see ourselves as acting on behalf of someone or something else. One of the best examples of this comes from the famous Milgram experiments, wherein the subjects became grimly locked into their assigned role of doing the indefensible within a laboratory setting (1974/2009). The majority of people (around two-thirds in the basic experimental condition) went all the way to administer apparently agonizing electric shocks to an apparently helpless victim<sup>26</sup> (1974/2009, 35). How did this happen? What were they thinking?

By and large, it wasn’t sympathy that seemed to be missing in those who capitulated to the experimenter’s prodding (“the experiment *requires* that

<sup>25</sup> Although it does seem to me that blaming someone like Meursault would be inapt, if we are prepared to take him at his word with regards to the following: “I have never been able really to regret anything in all my life. I’ve always been far too much absorbed in the present moment, or the immediate future, to think back” (1942, 65). But this is obviously a large question, and I can’t do justice to it here.

<sup>26</sup> This apparent victim being, of course, a confederate of the experimenter who played the part of the ‘learner,’ in a study presented to the participants as an investigation of the effects of punishment on memory. The experiment began with a rigged draw, such that the participant appeared to be randomly assigned to the role of the ‘teacher’ rather than that of the ‘learner.’ This was also intended to ensure that they would meet the learner in person—an affable, middle-aged accountant type. Most people reported finding him likable (1974/2009, 16).

you continue”; “please, go on”) (1974/2009, 21). Rather, it was efficacy. Most people visibly oozed stress—variously sweating, smoking, crying, chanting, and groaning. (1974/2009, Chapter 5). But they seemed to have lost the sense of themselves as responsible for the human consequences of their actions. They no longer saw themselves as actively choosing—and therefore capable of refusing—to do what they were doing to another human being.<sup>27</sup> In addition to this, many people felt responsible for something else entirely—namely, living up to their end of what was clearly a very bad bargain.<sup>28</sup> One person kept muttering to himself: “It’s *got* to go on. It’s *got* to go on”<sup>29</sup> (1974/2009, 9). Milgram saw him and other subjects as laboring under a “sense of obligation”—a highly misplaced “conception of his duties as a subject” (1974/2009, 6).

But this sense of obligation appeared to be a temporary addendum to, rather than actually supplanting, the person’s customary moral sense. Many of those who continued, who went all the way, remained clearly opposed to what they were doing, even as they were doing it—as opposed to losing touch with their standing moral judgments in the heat of the moment. Here is Milgram himself making much this point:

Many of the subjects, at the level of stated opinion, [felt] quite as strongly as any of us about the moral requirement of refraining from action against a helpless victim. They, too, in general terms know what ought to be done and can state their values when the occasion arises. . . . [But] many people were unable to realize their values in action and found themselves continuing in the experiment even though they disagreed with what they were doing. (1974/2009, 6)

<sup>27</sup> Milgram calls the state that people tended to get into the ‘agentic state,’ somewhat confusingly given the present context. But the idea is actually very much in keeping with the notion of the spectator stance as I am understanding it. A person in the agentic state “divests himself of responsibility . . . [and] sees himself not as a person acting in a morally accountable way but rather as the agent of external authority” (1974/2009, 8). He “no longer views himself as acting out of his own purposes but rather comes to see himself as an agent for executing the wishes of another person.” Milgram contrasts this attitude with an attitude of autonomy, wherein a person “sees himself as acting on his own” (1974/2009, 133).

<sup>28</sup> Their having been paid four dollars, plus fifty cents carfare, for their troubles (1974/2009, 15).

<sup>29</sup> Interestingly, however, no one kept going after hearing the fourth and final prompt which the experimenter was instructed to use, in a last-ditch attempt to keep them from leaving: “You have no other choice, you *must* go on” (1974/2009, 21). Subjects tended to react angrily to being told this and were then able to storm out. One person exclaimed indignantly: “I *do* have a choice. Why don’t I have a choice? I came here of my own free will” (1974/2009, 51). Indeed.

Some people were clearly in some sense motivated to stop, but could not bring themselves to do so. They subsequently “protested even as they obeyed” (1974/2009, 10). But not everyone in the experiment had even that much of a fight in them. Some people seemed to instantly withdraw and sink into themselves, despite their seeming sensitivity to the victim’s apparent suffering. One subject’s conduct was described by Milgram as follows: “When the victim’s first protests are heard, he turns toward the experimenter, looks sadly at him, then continues . . . throughout, a sad, dejected expression shows on his face.” He stayed at his post until the bitter end. Afterward this subject (a middle-aged Black man from the South) attributed most of the responsibility for what happened to the experimenter. “I merely went on. Because I was following orders. . . . I was told to go on,” he kept repeating hopelessly (1974/2009, 49–50). Milgram’s elegant little book describing his experiments is filled with depressing anecdotes about individual subjects’ behavior therein. But to me, this is easily one of the saddest.<sup>30</sup>

### 13.3. Tempered Internalism

If we take the phenomenon of disengagement in action seriously, then where does this leave us, when it comes to the key reasons for espousing motivational internalism? Some theorists who are attracted to the position may be animated by some optimism that people’s moral sensibility deeply shapes the way they live; that they have a moral compass and that they use it in practice. If this is why we were originally attracted to internalism, then it would evidently be something of a pyrrhic victory to shore up the connection between

<sup>30</sup> And we shouldn’t assume that all of those who managed to extricate themselves from the situation had recovered their sense of themselves as active participants in the experiment (as per the now favored terminology in experimental psychology). What disposed people to leave in droves is instructive. It seemed that being led by example made defiance come naturally. Seeing a fellow participant walk out enabled around 90 percent of subjects to then follow suit (1974/2009, 119). Being in the same room as the learner, or having to actually press his hand onto the electric shock plate, were also strong predictors of non-compliance (1974/2009, 35). This is heartening, of course, at least up to a point. But did people then refuse to comply because their sense of moral responsibility was quickened? Did these people really come to be more morally proactive? Or did they instead come to view themselves as subject to the demands of another? Namely, the learner, whose every cry of pain might seem to register protest. If so, then the following possibility might occur to us: perhaps many of the people who walked out hadn’t actually recovered their sense of moral agency. Rather, they just managed to glom onto to yet another set of directions implicitly issued by someone else. One person who refused to go on said as much explicitly: “If he doesn’t want to continue, I’m taking my orders from him” (1974/2009, 48). Is this an act of moral conscience or not, we might wonder.

self-directed moral judgments and motivations by effectively pointing to the possibility of inward misdirection. What is the point of having a moral compass if you are often in the dark about your current location?

The point of tempering motivational internalism in the way I am recommending is decidedly not this, then. Rather, tempered internalism is intended to help accommodate certain intuitive features of moral judgments that plausibly derive from the broadly prescriptive or practical nature of their subject matter. One such intuition which I find powerful is that something is going *wrong* when someone makes a moral judgment to the effect “I really ought (not) to  $\psi$ ,” but seems entirely unmotivated (not) to do what he takes to be instances of  $\psi$ -ing. The motivational internalist can plausibly explain what is going on here. Namely, the agent is purporting to make a moral judgment when, in reality, he is not. He is essentially mistaken about what he is up to.

Still, is not as if the motivational internalist’s standard explanation here is the only game in town (even if it is granted that this explanation is intuitively plausible, which is certainly not beyond dispute). In particular, are we so convinced that what has gone wrong in such cases is that a judgment hasn’t actually been made, as opposed to its having been made, but *badly*? Our intuitions are generally not so fine-grained as to rule this out. Compare: does a supposed belief in some empirical proposition that is unresponsive to counter-evidence fail to qualify as a belief at all, or does it merely violate *norms* which govern such beliefs? Similarly, is someone who doesn’t seem to see anything good about some possible state of affairs failing to desire it, or are they merely harboring a *defective* desire? The weaker position is not ruled out, on the face of it. And moral judgments might similarly be subject to norms which require such products of first-person moral thinking to have a motivational upshot. Why this might be so is a question I’ll come back to.

But first, we need to consider an alternative possibility. Maybe what has gone wrong here is that the person is unconscientious, unprincipled, wanton. That is, something may have gone wrong at the level of her *character* rather than the level of her judgment. There is nothing the matter with the judgment itself. This is the basic thrust of Sigrún Svavarsdóttir’s rival hypothesis, when she suggests that motivational internalists may be overgeneralizing from the perfectly legitimate intuition that moral judgments will be motivating for morally committed agents—that is, people who have an ingrained *desire* to be moral or to be a good person (1999, 185–187). So, when such an agent judges that morality requires her to do such-and-such, a standard belief-desire pair will mobilize her for action. Moreover, we would naturally expect that morally committed agents will tend to be more invested in making moral judgments, not to mention liable to engage in moral thinking in the first place—which will naturally sometimes lead to their making such judgments. But that doesn’t

mean that the morally uncommitted amongst us *can't* make moral judgments if and when they choose to. Svavarsdóttir's well-known character of Patrick, described in observational terms, seems to be a case in point here (1999).

Svavarsdóttir's hypothesis has the virtue of parsimony (in not constraining the nature of moral judgments unnecessarily) and seems viable on the face of it. But I'm not convinced that it can explain everything that needs to be explained here. For it seems to me that, when characters like Patrick advance a moral judgment to the effect that they are morally required to do something, but they are no more inclined to do it than they would be in any case, then there is something that strikes us as 'off' about these judgments themselves. Specifically, such judgments will tend to strike us as 'hollow.' (Think back, for example, to the moral judgment made by the recalcitrant carnivore above.) They may still be genuine moral judgments, best taken at face value. And, absent a general theory of moral thought and language—and perhaps even thought and language quite generally—it is hard to see what would justify an insistence that they are not.

I think that Svavarsdóttir is right about all of this. But I still have the sense that these judgments are not all that they should be as judgments. They seem to be missing psychological components or connections that such judgments ought to have, their being the specific kind of judgments that they are.

In order to bring this out, it will be helpful to consider a case that clearly does have the structure which Svavarsdóttir suggests may be at issue in the moral case—that is, where the judgment about the standard, and the desire to conform to the standard, are genuinely independent. So suppose we agree (for the sake of argument) that philosophical writing ought to be grammatical and clear. So, when a talented student of yours is falling short of the mark, writing opaque prose and flouting proper syntax, you point these problems out to him in the hope and expectation that he will work on improving them. But it turns out that the student fully agrees that, from the point of view of grammaticality and clarity, he ought to change his writing in this or that way. But he does not *want* to write philosophy papers that are grammatical and clear, he says; he believes good philosophical writing is more a matter of expressing yourself creatively and pushing the syntactic envelope. In this case, I do not think that the student's judgment to the effect that, grammatically speaking, he ought to do such-and-such will tend to sound *hollow*. His judgment itself might well seem perfectly in order. The problem would then rather seem to consist in the student's failure (at least by our lights) to sufficiently *value* grammaticality and clarity—hence his not being moved to comply with normative judgments like this one, which he has no problem making.

The evidence provided by this disanalogy is of course defeasible, but it is nonetheless suggestive. Namely, what it suggests is that the hollow ring of

first-person moral judgments unaccompanied by motivation is not mere ‘transference’ (as it were) from our belief that people ought to value and to want to conform to moral standards. The intuition that a judgment is hollow seems to point to something defective in the judgment itself.

And what might this defect be? That is, what is supposed to be missing here, at the level of moral psychology, from actionable first-personal moral judgments that float free of motivation? I think the answer is not far to seek, in light of the foregoing discussion. I suggest that what’s missing is plausibly *the agent herself*—or at least her sense of herself as an active participant in practice. For, as I argued in section 13.1, if she were taking the participatory stance toward herself with respect to the relevant actions, then she would be motivated to act in accordance with her judgments about whether or not to perform them. But she doesn’t, so she isn’t. Moreover, if this motivation was present, then this judgment would *not* sound hollow (or at least, not for this reason). And this hollowness is supposed to evince and hence entail a certain defect. So, insofar as the agent takes the participatory stance rather than the spectator stance toward herself with respect to the actions her moral judgment is about, then this judgment will not suffer from this particular psychological defect. It is on this basis that I propose that this stance constitutes the appropriate ‘fit’ for moral judgments of this kind.<sup>31</sup> And all told, this yields the tempered internalist claim in full now: namely, that a moral judgment made by an agent about what she herself (currently) ought to do will entail motivations on her part to act in accordance with this judgment, provided that she takes the participatory stance toward herself with respect to these actions—which has now also been argued to be the *fitting* stance from which to make such judgments.

I think that what is left of motivational internalism, thus tempered, is plausible. And I also would suggest that it captures something important. Our moral demands on ourselves are *supposed* to be undertaken with a sense of ourselves as active and responsible for complying. And the corresponding judgments will then be accompanied by motivation—which is just as it (or rather, they) ought to be.

There admittedly remains a question about *why* this norm might hold. That is, why would the participatory stance be the appropriate or fitting stance from which to make these judgments? But it’s not too hard to see what a natural answer to this question might look like, at least in broad outline. Consider the observation with which we began: namely, that morality is supposed to tell us what to do. It is supposed to speak to us, somehow. Hence, such actionable, first-person

<sup>31</sup> At least, setting aside the (I think, dim) possibility that judgments which are made in the participatory mode suffer from further, and worse, defects than judgments made in the spectator mode—in which case, taking the participatory stance would not constitute an overall improvement in the resulting moral judgment. But I see no general reason to think that this would be the case.

moral judgments can plausibly be construed as exercises in telling *ourselves* what to do—often, presumably, on behalf of someone else. But what is the point in telling ourselves to do something if the subject of the address has effectively gone walkabout? What purpose will be served by taking another’s perspective, or identifying with her plight, if we then step out on ourselves as would-be moral respondents? We are calling on ourselves, but there is nobody home. Although, on reflection, it might seem that there may sometimes be perverse incentives for the most conscientious amongst us to arrange to be out. We wouldn’t allow ourselves to succumb to certain temptations in good conscience. Even if we are sorely tempted, there is no way that this is happening on our watch. But sometimes we seem to manage or contrive to get around this. We look away; it happens; and then it just happened. And what’s done is done. So what can you do? <sup>32</sup>

## References

- Blackburn, S. 1998. *Ruling Passions*. Oxford: Clarendon Press.
- Camus, A. 1942. *The Stranger*. Trans. S. Gilbert. New York: Vintage Books.
- Dell, P. F., and J. A. O’Neill, eds. 2009. *Dissociation and the Dissociative Disorders: DSM-V and Beyond*. New York: Routledge.
- Hacking, I. 1995. *Rewriting the Soul: Multiple Personality and the Science of Memory*. Princeton, NJ: Princeton University Press.
- Hare, R. M. 1952. *The Language of Morals*. Oxford: Oxford University Press.
- Kennett, J., and C. Fine. 2008. “Internalism and the Evidence from Psychopaths and ‘Acquired Sociopaths.’” In *Moral Psychology*, Vol. 3, ed. W. Sinnott-Armstrong, 173–190. Cambridge, MA: The MIT Press.
- Milgram, S. 1974. *Obedience to Authority*. Reprinted 2009. New York: Harper Perennials.
- Roskies, A. 2003. “Are Ethical Judgments Intrinsically Motivational? Lessons from ‘Acquired Sociopathy.’” *Philosophical Psychology* 16: 51–66.
- Simeon, D., and J. Abugel. 2008. *Feeling Unreal: Depersonalization Disorder and the Loss of Self*. New York: Oxford University Press.
- Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell.
- Steinberg, M., and M. Schnall. 2001. *The Stranger in the Mirror: Dissociation—the Hidden Epidemic*. New York: Harper Perennial.
- Strawson, P. F. 1962. “Freedom and Resentment.” *Proceedings of the British Academy* 48: 1–25, reprinted in *Freedom and Resentment and Other Essays*, 2nd ed., 2008, 1–28. London: Routledge.
- Svavarsdóttir, S. 1999. “Moral Cognitivism and Motivation.” *The Philosophical Review* 108: 161–219.
- Tresan, J. 2009. “The Challenge of Communal Internalism.” *The Journal of Value Inquiry* 43: 179–199.
- Tresan, J. 2006. “De Dicto Internalist Cognitivism.” *Noûs* 40, no. 1: 143–165.

<sup>32</sup> Thanks to Christopher Lewis, Guy Fletcher, Kevin Vallier, David Plunkett, Daniel Manne, and two anonymous reviewers for helpful comments on earlier drafts of this paper. Thanks also to audience members at the Moral Motivation: Evidence and Relevance (MMER) workshop in Gothenburg in May 2012, where this paper was first presented, and to the MMER organizers.