

Accountability and the Thoughts in Reactive Attitudes

Abstract: As object-directed emotions, reactive attitudes can be appropriate in the sense of *fitting*, where an emotion is fitting in virtue of accurately representing its target. I use this idea to argue for a theory of moral accountability: an agent S is accountable for an action A if and only if A expresses S's quality of will and S has the capacity to recognize and respond to moral reasons. For the sake of argument, I assume that a reactive attitude is fitting if and only if its constituent thoughts are true, and I argue for the above theory by determining thoughts partly constituting resentment and gratitude. Although others have argued that the capacity to recognize and respond to moral reasons is necessary for accountability, the argument here is significantly better in two respects. First, it does not rely on intermediary ethical principles, supplementary arguments, or assumptions about the nature of reactive attitudes specifically. Instead, it simply assumes that reactive attitudes, like all emotions, have cognitive content. Second, the argument here is more powerful because it brings to light the quality of will condition and has the resources to flesh out the capacity to recognize and respond to moral reasons.

Reactive attitudes – e.g., resentment, indignation, and gratitude – are emotional reactions to people that hold them accountable for their actions. As object-directed emotions (as opposed to mere feelings or moods), reactive attitudes have representational content. In this paper, I argue that the representational content of reactive attitudes entails that a person is morally accountable for an action if and only if the action expresses her quality of will and she has the capacity to recognize and respond to moral reasons.

I start by assuming, as is common in the literature, that an agent is accountable for an action if and only if a reactive attitude is appropriate toward her on account of the action. In other words, an agent is blameworthy for an action if and only if a negative reactive attitude is appropriate toward her on account of it, and an agent is praiseworthy for an

action if and only if a positive reactive attitude is appropriate toward her on account of it.¹ (I remain neutral on why this biconditional holds and hence do not assume that the appropriateness of reactive attitudes towards an agent explains her being accountable.²)

To use the biconditional, we need an understanding of what it means for reactive attitudes to be “appropriate.” Different understandings have been proposed. For example, R. Jay Wallace (1994) understands “appropriateness” as “moral fairness,” whereas Gideon Rosen (2015) understands it as “accuracy.” Reactive attitudes can be assessed in terms of accuracy because, as object-directed emotions, they have representational content, making them accurate if and only if their representational content is true. In this paper, I assume that we should understand “appropriateness” as “accuracy.” Specifically, I assume that a reactive attitude is appropriate if and only if it accurately presents its target – i.e., if and only if it is *fitting*.

I adopt this understanding of “appropriateness” because it is plausible and, I will show, leads to powerful results. It is plausible because reactive attitudes are emotions with representational content, and all such emotions can be appropriate in the sense of fitting (D’Arms and Jacobson 2000). In fact, I think that this understanding is better than that of “moral fairness,” since it is difficult to see how, say, unexpressed resentment can be morally fair or unfair, given that unexpressed resentment does not affect others. Yet I need not pursue this point. Even if there are multiple plausible ways to understand the appropriateness of reactive attitudes, the one that I adopt is one of them. Further, I will demonstrate that the results of understanding “appropriateness” as “accuracy” or

¹ Thus, to use the biconditional, I set aside actions for which an accountable agent might be neither blameworthy nor praiseworthy, such as merely doing one’s duty. In adopting the biconditional, I do not assume that reactive attitudes are the only way to hold agents accountable. Further, I mean “blameworthy” and “praiseworthy” in the accountability sense, since agents can be blameworthy and praiseworthy in the attributability sense of responsibility too. I discuss attributability below.

² Strawson (1962/2003) famously argues that being an accountable agent is a matter of being an appropriate target of reactive attitudes, and many have agreed, including Wallace (1994) and Rosen (2015). My argument does not depend upon this idea. Even if being accountable explains the appropriateness of reactive attitudes, we can recognize that an agent is not accountable for an action by recognizing that she is not the appropriate target of reactive attitudes for it.

“fittingness” is striking. By substituting this understanding into the biconditional, I develop a simpler and more powerful argument than other arguments in the literature for the idea that accountability requires the capacity to recognize and respond to moral reasons. The argument that I develop is simpler in that it does not rely on intermediary ethical principles, supplementary arguments, or assumptions about the nature of reactive attitudes specifically, as the other arguments do. The argument that I develop is more powerful in that it yields a more detailed understanding of the capacity to recognize and respond to moral reasons than the other arguments do, and it reveals the further quality of will condition.

This paper proceeds as follows. In Section 1, I explicate the fittingness of emotions. For the sake of argument, I adopt a plausible cognitivist understanding of fittingness: an emotion is fitting if and only if its constitutive thoughts are all true. However, for my argument, all that matters is that emotions have representational content, and so any view of fittingness would work. In Section 2, I show how we can apply this understanding of fittingness to reactive attitudes specifically in order to illuminate moral accountability. Then, in Section 3, I examine everyday cases of resentment and gratitude in order to bring out the thoughts constitutive of those reactive attitudes, and I use those thoughts to argue for the above-mentioned view of accountability. Finally, in Section 4, I discuss the advantages of this argument over other arguments in the literature with respect to deriving the requirements of accountability.

1. Appropriateness as Fittingness

To understand appropriateness as fittingness, notice that emotions present their targets as having certain evaluative features. Envy, for example, presents its object as having features that make it enviable, such as having something good that the one envying lacks. Whether a particular emotion is fitting depends upon whether it accurately presents

its target – i.e., on whether the emotion’s evaluative presentation is correct. If the emotion accurately presents its target, the emotion is fitting; otherwise, it is unfitting. Hence, if Susan envies Tom because he is an excellent violinist but she is a better one, her envy is unfitting. Her envy presents Tom inaccurately as having something good – being an excellent violinist – that she lacks.³ Because the fittingness of an emotion is a matter of the emotion’s accuracy or correctness, fittingness is distinct from moral rightness (D’Arms and Jacobson 2000). An emotion may be fitting because it accurately presents its object but may nevertheless be morally wrong to feel and vice versa. Thus it may be fitting for Tom to envy Susan because she is a better violinist, even if envy is morally wrong.

Because emotions can be accurate or inaccurate, they are at least partially constituted by states that represent the world. (They are also plausibly partially constituted by conative states, such as desires.) Cognitivist theories of emotions hold that these representational states are propositional attitudes, such as beliefs or judgments.⁴ Perceptual theories hold that these representational states are evaluative perceptions, akin to visual perceptions.⁵ The argument below can be run using any of these theories, since what matters for the argument is just that an emotion has representational content.

For the sake of argument, I adopt a cognitivist theory of emotions on which emotions are partly constituted by thoughts.⁶ When those thoughts are true, the emotion is fitting; otherwise, it is unfitting. Thus I assume the following view:

The True-Thoughts View: an emotion is fitting if and only if its constitutive thoughts are true.⁷

³ In cases of misleading evidence, an emotion’s evaluative presentation may be unfitting but yet justified. Although there is a sense in which an emotion is appropriate in virtue of being justified, I set aside this sense. I thank Nathan Ballantyne for raising this issue.

⁴ For examples of cognitivist theories, see Solomon (1976) and Nussbaum (2001).

⁵ For a perceptual theory, see Prinz (2004). For views close to the perceptual theory, see Roberts (2003) and Oddie (2005).

⁶ For this view, see Greenspan (1988) and Rosen (2015).

⁷ Rosen (2015) adopts the True-Thoughts View, but he calls it the “Alethic View” and takes it to be an alternative to fittingness views. I diverge from his terminology to highlight both that the True-Thoughts

How should we understand thoughts? I take thoughts to be belief-like propositional attitudes, representing how things seem to the agent. They may be beliefs, but they need not be, as agents may withhold their assent from these seemings. For example, someone who is afraid of flying may think that he is in danger while on the airplane, in that it seems to him that he is in danger, even if he does not believe that he is. Or a woman raised in a teetotal home may feel guilty indulging in a drink because she has the thought that she is acting wrongly – i.e., it seems to her that she is acting wrongly – even if she does not believe it. We could imagine her saying, “I just keep thinking that I shouldn’t be drinking this martini, even though I know it is fine.”

Before proceeding, note that, for the purposes of my argument, I can safely ignore the conative states that also plausibly partially constitute reactive attitudes. There are two reasons for this. First, suppose that a reactive attitude is not “appropriate” unless it is both fitting and its conative states are a certain way. For example, suppose that disproportionately strong reactive attitudes – e.g., strong feelings of resentment for a minor infraction – are inappropriate even if they are fitting. Even so, we can simply assume that the conative states are as they should be in the examples below, as doing so will not affect determining the conditions that make agents accountable for their actions. Second, if disproportionately strong reactive attitudes are unfitting, that could only be so in virtue of their having inaccurate representational content. Yet, as we will see, we need not determine all of the representational content of reactive attitudes in order to find the conditions that make agents accountable, and we will plausibly find all of those conditions without examining the possibility that the conative states of reactive attitudes possess representational content.

View is a view of fittingness on my understanding and that my argument would go through equally well with any view of fittingness.

2. From Thoughts in Reactive Attitudes to Accountability Conditions

Because I understand appropriateness as fittingness, reactive attitudes are appropriate if and only if they accurately present their targets. How do reactive attitudes present their targets? Reactive attitudes present their targets as accountable – as blameworthy in the case of negative reactive attitudes and praiseworthy in the case of positive reactive attitudes. This is why reactive attitudes are fitting responses toward an agent if and only if he is accountable.

On the True-Thoughts View, for reactive attitudes to present a target as accountable, the thoughts partly constituting them must represent the target as accountable. This could happen in either of two ways. First, reactive attitudes may be partly constituted by the thought that the target is blameworthy or praiseworthy. This thought would be compatible with any view of accountability and so would not illuminate the nature of accountability. Second, reactive attitudes may be partly constituted by thoughts that represent all of the “accountability conditions” – i.e., the conditions that make an agent blameworthy or praiseworthy. As we will see from examples, it is reasonable to conclude that the second is the case, and hence we can determine the accountability conditions from the thoughts partly constituting reactive attitudes.

Because reactive attitudes are partly constituted by thoughts representing all of the accountability conditions, reactive attitudes are specially suited to illuminate the accountability conditions. By way of contrast, consider trying to find the accountability conditions using the belief that an agent is accountable. (After all, the following biconditional also holds: an agent is accountable for an action if and only if the belief that he is accountable for the action is true.) We cannot determine the accountability conditions from this belief alone, since it simply represents the agent as accountable. Suppose then that we discover that the belief that an agent is accountable rises and falls entirely on the basis of two further thoughts, thoughts X and Y. Hence people take the truth of the

accountability belief to turn on the truth of thoughts X and Y. Could we conclude that thoughts X and Y represent accountability conditions?⁸ We could not. Even without being skeptical about the reliability of our belief-forming practices, it could be, say, that those thoughts represent useful heuristics for determining whether agents are accountable. The problem is this: establishing a causal relationship between independently-existing representational states – in this case, thoughts X and Y causing the belief that an agent is accountable – does not itself establish a relationship between the representational contents of those states. We therefore cannot conclude that thoughts X and Y represent accountability conditions.

We avoid this problem by working with reactive attitudes. Unlike the belief that an agent is accountable, reactive attitudes do not simply represent an agent as accountable and then rise and fall in response to independently-existing thoughts. Instead, as discussed above, reactive attitudes represent an agent as accountable in virtue of being *partially constituted* by thoughts representing all of the accountability conditions. This means that, through those thoughts, the accountability conditions are built into the representational content of reactive attitudes. As a result, reactive attitudes are distinctively suited to elucidate the accountability conditions.⁹

Importantly, not all of the thoughts partially constituting reactive attitudes represent accountability conditions. The thoughts partly constituting reactive attitudes are about the

⁸ I thank an anonymous referee for this example and for asking me to address what is gained by working with reactive attitudes to find the accountability conditions.

⁹ Note that reactive attitudes are partially constituted by thoughts representing the accountability conditions no matter which side of our original biconditional explains the other. (Recall our original biconditional: an agent is accountable for an action if and only if reactive attitudes are appropriate toward her on account of it.) Hence reactive attitudes are specially suited to illuminate the accountability conditions even if being accountable explains the appropriateness of reactive attitudes. You might worry that, if being accountable explains the appropriateness of reactive attitudes, then the thoughts partially constituting reactive attitudes could misidentify the accountability conditions. We can set this skeptical worry aside. Even if being accountable explains the appropriateness of reactive attitudes, our practice of holding people accountable – of which reactive attitudes are an integral part – is the best evidence that we have of the nature of accountability. Plus, at the very least, the representational content of reactive attitudes tells us what we take the correct accountability conditions to be. The onus would then be on the skeptic to demonstrate that our practice has gotten it wrong.

agent, his action, and the person experiencing the reactive attitude. This is because the appropriateness of X holding Y accountable for action A depends upon X, Y, and A. Yet only the thoughts about Y, the target of the reactive attitude, can represent accountability conditions. After all, it may be inappropriate for me to resent someone because my resentment would be partly constituted by some false thought about me – e.g., the thought that that I was wronged by his action, when in fact only someone else was. Yet this only shows that it is inappropriate for *me* to resent him. It may still be appropriate for the other person to resent him. Further, it may be inappropriate for me to resent someone because my resentment would be partly constituted by some false thought about the action – e.g., the thought that he acted wrongly, but gratitude toward him may be appropriate. Hence, to determine the accountability conditions, we need only determine the thoughts about the agent.

Next notice that an accountability condition C must be represented by some thought in *each* reactive attitude, since all reactive attitudes hold agents accountable. Yet this does not mean that C is represented by a thought common to all reactive attitudes. Indeed, we will see that, for each accountability condition C, C is entailed by some thought in each reactive attitude, although the thought that entails C is different for negative and positive reactive attitudes.

For my purposes, we can focus on determining the thoughts about the agent that partly constitute resentment and gratitude, the standard negative and positive reactive attitudes. I primarily support these thoughts by appealing to cases, showing how these thoughts are manifest in an agent's resentment or gratitude. (I will also mention how these thoughts are manifest in other reactive attitudes too.)

This brings me to an important methodological point. In arguing that a certain thought partly constitutes resentment or gratitude, I draw on intuitions about what agents are thinking in having reactive attitudes and about what thoughts extinguish reactive attitudes, but that does not mean that any proposed thought is irrefutable. If you can

intuitively have an emotion without having a thought of a certain kind, or if you think that an emotion is fitting in spite of a proposed thought being false, then that is good reason to think that such a thought does not partly constitute the emotion. Further, as Rosen (2015) points out, thoughts partly constituting resentment and indignation must be naïve in the sense that any agent capable of having that emotion must be capable of having that thought.

3. The Main Argument

In this section, I determine the thoughts about the agent that partly constitute resentment and gratitude and use them to argue for the following view of accountability: an agent S is accountable for an action A if and only if 1) A expresses S's quality of will and 2) S has the capacity to recognize and respond to moral reasons.

3.1. The Quality of Will Thought

I claim that the following thought partly constitutes resentment: in doing A, S expressed insufficient good will toward me. To see that this is plausible, consider an example. Suppose that Fred has promised to pick you up from the airport. He fails to show, and you resent him for it. “How thoughtless of him,” you think. “If he cared more about me, he would have remembered.” As you are brainstorming ways to tell him off, a mutual friend calls to say that Fred just lost his job and is completely distraught. Your resentment naturally disappears (if you are rational). Why? It is not because breaking his promise was permissible. Rather, the news negates your thought that Fred failed to pick you up from thoughtlessness or lack of caring – i.e., from insufficient good will – toward you. This example illustrates that resentment is partly constituted by the thought, under some guise, that the agent acted from insufficient good will.

I say “under some guise” because, as this example illustrates, the thought “in doing A, S expressed insufficient good will” may be put in different ways. Here it is put in terms of thoughtlessness or lack of caring. It may also be put in terms of cruelty, lack of respect, etc. The important point is that this thought, in some form, partly constitutes resentment, since we cannot resent someone without thinking some thought that amounts to the idea that he acted with insufficient good will.

This thought plausibly partly constitutes other negative reactive attitudes as well. When you feel indignant toward an agent, you think that he expressed insufficient good will toward someone else. If you find out that he actually acted from sufficient good will, your indignation disappears (if you are rational). Because negative reactive attitudes are plausibly partly constituted by the thought that the agent acted from insufficient good will, acting from insufficient good will is a condition of blameworthiness.

Of course, it is not a condition of praiseworthiness. Yet positive reactive attitudes are also partly constituted by a thought about the agent’s quality of will. Consider gratitude. Gratitude toward someone for an action, I claim, is partly constituted by this thought: in doing A, S expressed more than sufficient good will.¹⁰ “What a great guy to go out of his way for me,” you might think. If you discover that the benefit to you was accidental, or that he had hoped to harm you, your gratitude disappears (if you are rational) because this thought is false.

The above thoughts partly constituting resentment and gratitude show this: that reactive attitudes are only appropriate toward agents when their actions express their quality of will. This result is in line with Strawson’s (1962/2003) idea that reactive

¹⁰ You may object that gratitude is sometimes appropriate when an agent expresses mere sufficient good will, such as when the stakes are high or doing the right thing is costly. In these cases, I think that doing the right thing likely requires more than sufficient good will. Yet, I need not decide this issue. Even if gratitude is sometimes appropriate toward someone who expresses mere sufficient good will, the conclusion that I draw still stands.

attitudes are responses to agents' quality of will as manifested in their actions. Hence agents are only accountable for actions that express their quality of will.

On my view, this result brings us to the idea of responsibility as *attributability*. Although I cannot fully argue for the connection to attributability here, let me briefly say a few things to motivate it. To start, consider an example to illustrate attributability and how it differs from accountability. Imagine that your coworker Alice takes pleasure in demeaning others, and one day, you are the victim of her caustic remarks. You fume to a fellow coworker who responds: "I'm so sorry. Alice is cruel. But you shouldn't hold it against her. She was so mistreated as a child that it would have taken a miracle for her to turn out differently." Here, I claim, your coworker supposes that Alice is morally responsible for her caustic remarks in one sense but not another. In evaluating Alice as cruel for her remarks, he takes her to be *attributionally responsible* for them. An agent is attributionally responsible for those actions that express her moral identity, making it appropriate to appraise her morally for them.¹¹ In evaluating Alice as cruel, your coworker takes Alice's remarks to express her moral identity and so takes her to be attributionally responsible for them. Yet your coworker also claims that you should not hold Alice's caustic remarks against her due to her terrible upbringing and hence that it is inappropriate to resent her for them. In saying this, he asserts that she is not *accountable* for her remarks. Your coworker may be wrong, but he is not inconsistent to claim that Alice is blameworthy for her remarks in the attributability sense but not the accountability sense.

Now return to the idea that agents are only accountable for actions that express their quality of will. An action that expresses an agent's quality of will expresses her moral identity, making it appropriate to appraise her morally for it. For example, if an agent's action expresses ill will, we can appropriately appraise her negatively – e.g., as selfish or

¹¹ Here I follow Gary Watson (1996/2004), who first distinguished attributability and accountability. Watson claims that attributable actions are those actions that express an agent's practical identity. I focus more narrowly on an agent's moral identity in discussing moral responsibility.

unkind (at least on that occasion) – for it. The converse also holds: an action that expresses an agent’s moral identity expresses her quality of will. Hence actions that express an agent’s quality of will are those for which she is attributionally responsible. Notice that I have assumed that “out of character” actions that express an agent’s quality of will on a particular occasion still express the agent’s moral identity. I cannot argue for that idea here. (I argue for it elsewhere.¹²) If that is right, then attributability is necessary for accountability. Yet either way, using the True-Thoughts View, we see that agents are only accountable for actions that express their quality of will.

3.2. The Capacity Thought

At least one more thought about the agent constitutes resentment. This is because young children and the mentally ill can act wrongly from insufficient good will, but they are not usually blameworthy. I propose this thought: “S could have done better,” understood to mean that S could have acted with a better quality of will. Note that this thought is a *normative* thought, since it says that S could have done *better* and not merely that S could have acted differently. Note also that this thought is sufficiently naïve, as even young children have the concept of being able to act better.

In this section, I argue for this proposal by appealing to cases. I then show that, for it to be true that S could have done better, S must have the capacity to recognize and respond to moral reasons (understood *de re* – i.e., S must have the capacity to recognize and respond to actual moral reasons). Having this capacity is an accountability condition, I argue, because it is also entailed by a thought partly constitutive of gratitude and, plausibly, the other reactive attitudes. It is worth noting in advance that, although this section contains my main argument for the idea that accountability requires the capacity to recognize and respond to moral reasons, I support and expand upon this argument in the

¹² See Strabbing (2016a), p. 291 and Strabbing (2016b), p. 740.

following sections by responding to objections and by showing how the True-Thoughts View gives us the resources to flesh out that capacity.

Here is the first case. Nora resents her cold and critical mother for not expressing love to her as a child. One day, while talking to an old family friend, Nora learns that her mother’s parents were also cold and critical, rarely showing affection to their children. Further, her mother was young when she had Nora and felt overwhelmed with the responsibility. After hearing this, Nora feels compassion for her mother rather than resentment. She thinks to herself: “although my mom made a lot of mistakes, she did the best that she could.”

I contend that Nora moves from resentment to compassion by thinking that her mother did the best that she could. It becomes sad to Nora that her mother, due to her upbringing, could not show more affection. This supports the idea that resentment is partly constituted by the thought that the agent could have done better. Nora no longer resents her mother because she no longer thinks that her mother could have done better.

In thinking that her mother could not have done better, Nora means that her mother could not have acted with a better quality of will. After all, the information that she receives does not reveal that her mother had a different quality of will, that she was actually kind and loving. Rather, it is evidence that her mother could not have been other than cold and critical and so could not have done better than to treat Nora coldly and critically. Hence, in thinking that her mother could not have done better, we should take Nora to mean that her mother, given her difficult past, could not have acted with a better quality of will toward her.

Consider now criminals with terrible upbringings. In such cases, we often think: “no wonder he acts as he does!” This often cuts into our resentment if we are the victims or into our indignation if someone else is. Why? It is not simply because we view the criminal as a victim of others’ mistreatment. After all, we can imagine someone resenting the criminal and saying: “I feel badly that he had such an awful childhood. But other people

with awful childhoods don't act like that.” Such a statement is meant to provide evidence that the criminal could have done better in spite of his terrible upbringing. Thus, when our resentment toward the criminal is mitigated, I claim that it is because we think that he could not have done better due to that upbringing.

Again we should understand the thought “he could not have done better” to mean that he could not have acted from sufficient good will. After all, we can have that thought while recognizing that the criminal might have performed a better action if, say, someone had bribed him. Hence, when we say that he could not have done better due to his upbringing, we mean that he could not have acted with moral concern for others due to his upbringing.

These cases support my proposal that resentment is partly constituted by the thought that the agent could have done better. Does this thought explain why young children and the mentally ill are (generally) not accountable? It does. We often say that we should not blame young children when they act wrongly because they do not know better, even when they act from ill will. When we say this, we mean that they *could not* have known better, and as a result, they could not have done better in the sense that they could not have acted from sufficient good will, due to their lack of moral knowledge. This plausibly explains why they are not accountable. As for the mentally ill, it is reasonable to think that, when they act from insufficient good will, their condition makes it the case that they could not have acted from sufficient good will.

I have argued that resentment is partly constituted by the thought that the agent could have done better, understood to mean that he could have acted from sufficient good will. An agent who acts from sufficient good will is, I claim, one who recognizes and responds to moral reasons *de re*.¹³ In other words, he recognizes and responds to the actual moral reasons, whether or not he conceptualizes them as such. This is because good will

¹³ Arpaly (2003, p. 79) also makes this point.

is that which we expect of one another as members of the moral community, and we expect one another to regulate our behavior according to moral considerations *de re*. Hence, if an agent could have acted from sufficient good will, then he could have recognized and responded to moral reasons *de re*. The upshot is this: the thought “he could have done better” plausibly entails that, to be a fitting target of resentment, an agent must have the capacity to recognize and respond to moral reasons *de re*.¹⁴

As for other negative reactive attitudes, it is reasonable to suppose that the thought “he could have done better” partly constitutes indignation and that the thought “I could have done better” partly constitutes guilt. In support of the latter, we sometimes seek to counter another’s guilt by telling him that he did the best that he could. Thus the capacity to recognize and respond to moral reasons (*de re*) is necessary for blameworthiness.

What about praiseworthiness? First return to the thought partly constituting gratitude discussed in the previous section: “in doing A, S expressed more than sufficient good will.” This thought itself entails that the agent has the capacity to recognize and respond to moral reasons *de re*. After all, as I claimed above, an agent who acts from good will is an agent who recognizes and responds to moral reasons *de re*. Thus, an agent who acts from more than sufficient good will (or even just good will) is an agent who has and exercises the capacity to recognize and respond to moral reasons *de re*. Hence that capacity is necessary for praiseworthiness as well as blameworthiness and is therefore an accountability condition.

For my purposes, then, it does not matter whether gratitude is partly constituted by a thought analogous to the thought “he could have done better” in resentment. If gratitude contains no analogous thought, this would support the idea, advocated by Wolf (1990) and

¹⁴ You may wonder whether being moved by a moral consideration *de dicto*, such as being moved to perform an action by the fact that it is morally obligatory, may sometimes express good will. I think that it may, so long as this happens because of an underlying recognition of a moral consideration *de re*. For example, you may be moved to perform an action by the fact that it is obligatory because you recognize (perhaps unconsciously) that, say, in performing obligatory actions, you treat others well or respectfully. I thank Justin Coates for raising this issue.

Nelkin (2011), that blameworthiness and praiseworthiness are asymmetrical in that blameworthiness requires the ability to have performed a different action, whereas praiseworthiness does not. Yet the following analogous thought may partly constitute gratitude: “he could have done less well.”¹⁵ In support of this, consider how people describe heroic acts. Imagine that a child is walking a pier in winter when he slips and falls into the icy water below. Bystanders are screaming and pointing when a woman hurtles herself over the edge of the pier, grabs the child, and swims him to shore. Later, we can imagine the child’s parents expressing gratitude like this: “She could have stayed on the pier and waited for the rescue team, but she knew that the water was too cold. So she risked her life to save our son. She is a hero.” Part of the parents’ gratitude toward the rescuer is the idea that she could have stayed on the pier. She could have done less well, and it is a credit to her that she did not do less well. I take it that the thought “she could have done less well” means that she could have failed to act from more than sufficient good will. This could happen in either of two ways: she could have failed to rescue the child or she could have rescued the child from worse motives, such as rescuing the child just to achieve fame on the morning talk-show circuit.

If the thought “she could have done less well” partly constitutes gratitude, it too entails that the agent has the capacity to recognize and respond to moral reasons. After all, as just mentioned, this thought means that the agent could have failed to act from more than sufficient good will, which implies that she has and exercised the capacity to recognize and respond to moral reasons de re.

3.3. Fleshing Out the Capacity to Recognize and Respond to Moral Reasons

Using thoughts in resentment and gratitude, I just argued that the capacity to recognize and respond to moral reasons is necessary for accountability. I will now show

¹⁵ I thank John Mackay for this suggestion.

that those thoughts provide the resources to flesh out that capacity. Doing so will also allow me to respond to some potential objections.

Consider again Nora’s mother and the criminal with a terrible upbringing. You might worry that they could have done better because it is not impossible for either to recognize and respond to moral reasons. After all, some people with similar pasts are loving parents and law-abiding citizens.

I agree that some people with similar pasts do better. Yet this does not entail that Nora’s mother and the criminal have the capacity to recognize and respond to moral reasons. Such backgrounds are evidence, not proof, of incapacity, and the same background factors may lead some but not others to develop the incapacity.

Further, when we say that someone could not have done better, we need not suppose that there was zero probability of their doing better. After all, the capacity to recognize and respond to moral reasons comes in degrees, and at a certain point, someone has the capacity to such a small degree that it makes sense to say that they do not have the capacity. I think that we should understand the degree to which someone has the capacity in terms of difficulty, such that if it is sufficiently difficult for someone to recognize and respond to moral reasons, then he lacks the capacity to recognize and respond to moral reasons. Consider an analogy. It makes sense to say that I lack the capacity to make a putt beyond five feet, since it would be incredibly difficult for me to make such a putt. Similarly, given their pasts, it would be very difficult for Nora’s mother and the criminal to recognize and respond to moral reasons, since they were given substantial misleading information as children about how to behave. Hence they lack the capacity to recognize and respond to moral reasons in that it is very difficult for them to do so.

Importantly, we can see this point about difficulty directly from a thought in resentment, since the thought “she could have done better” is sometimes expressed as a matter of difficulty. After all, we often think, “it is not that hard to do x,” while resenting someone for failing to do x, and if we think that it actually would be hard for that person

to do x , that would mitigate our resentment. For example, imagine that Alena’s sister Greta is very self-focused. One day Alena has exciting news, but she never gets a word in edgewise as she is regaled with stories from Greta’s life. Alena resentfully hangs up the phone. “This is ridiculous,” she thinks. “How hard is it to ask someone how they are doing?” This thought expresses that Greta could have done better by pointing out that it is easy to ask others how they are doing. Hence, by looking at other ways to think “she could have done better,” the True-Thoughts View has the resources to support the idea that the capacity to recognize and respond to moral reasons comes in degrees, depending upon how difficult it is for someone to recognize and respond to those reasons.

The True-Thoughts View also supports the idea that we should interpret the capacity to recognize and respond to moral reasons as a specific capacity – i.e., as a capacity to recognize and respond to the moral reasons that bear on an agent’s action in his particular situation. This contrasts with R. Jay Wallace’s (1994) claim that accountability just requires the general capacity to recognize and respond to moral reasons, analogously to the general ability to speak a language. To see that the True-Thoughts View supports the specific capacity view, consider a person with a “moral blind spot,” like an otherwise loving and morally-upstanding father who opposes his daughter’s dream to become an engineer because, due to a patriarchal upbringing, he considers it to be a “man’s profession.” His daughter may think that it is inappropriate to resent him for this, and when pressed to explain why, we can imagine her saying that it would be very difficult for him to recognize that he should support her dream, given his upbringing. Such an explanation is reasonable, even though her father has the general capacity to recognize and respond to moral reasons.

These two points about how to interpret the capacity to recognize and respond to moral reasons are crucial for making sense of mundane cases of resentment within personal relationships. First notice that, in support of the idea that the thought “he could have done better” partly constitutes resentment, a target of resentment commonly responds with the

protest, “I did the best that I could!” For example, if you resent your partner for not defending you when his family mentions yet again that being a philosopher is not a real job, your partner may respond: “I did the best that I could! It was such a tense situation; I didn’t want it to escalate.” This response means to undercut your resentment, I claim, by countering your thought that your partner could have done better.

Next, notice that, in saying that he did the best that he could, your partner does not claim that he acted well. Rather, he claims that it was very difficult, given the tense situation, to determine what to do and to act accordingly. The idea is that, to some significant degree in the particular situation, he lacked the capacity to recognize and respond to the reasons to defend you to his family in the sense that it was very difficult for him to recognize and respond to those reasons.¹⁶ The point about difficulty is reinforced by the fact that you could justify your resentment as follows: “What do you mean, you did the best that you could? Just say that you don’t want to hear it anymore. It’s not that hard.”

One final point: the thought “he could have done better” shows that, to be blameworthy, an agent must be able to recognize and respond to moral reasons that support his doing better. After all, an agent who acts wrongly may act on a moral reason that is outweighed by other moral reasons. In such a case, the agent will only be blameworthy if he could have recognized and responded to moral reasons in favor of acting better. It is beside the point that he can recognize and respond to a moral reason in favor of his wrong action. (Note that I am not saying that the agent must have the capacity to recognize and respond to moral reasons in favor of the *right* action. If he could have acted less wrongly by responding to moral reasons in favor of some better but still wrong action, he can still be blameworthy for his action.)

¹⁶ Even if the reasons here are grounded in personal relationships rather than morality, such cases would still show that the thought “he could have done better” partly constitutes resentment, and this thought would still support the idea that, to be accountable, an agent must have the capacity to recognize and respond to certain reasons: moral reasons for moral accountability and reasons deriving from personal relationships for accountability within those relationships.

3.4. Distinguishing Amongst Cases

Given the cases considered so far, you may wonder whether I have misidentified the condition that exempts agents from accountability. I have claimed that Nora's mother and the criminal with the terrible upbringing are not accountable because they lack the capacity to recognize and respond to moral reasons; however, you might think that they are not accountable because their wrong values and false beliefs about morality are *reasonable* given the substantial misleading evidence that they were given.¹⁷ Similarly, you might think that young children are not accountable for their actions because, prior to receiving a moral education, it is reasonable for them to have mistaken values and beliefs about what they should do.

I agree that the reasonableness of these agents' beliefs and values matters to their lack of accountability, but it is not the fundamental explanation. Instead, these agents are not accountable because, due to the reasonableness of their faulty values and false beliefs, it is very difficult for them to recognize and respond to the relevant moral reasons, and so they lack the capacity to recognize and respond to those moral reasons (at least to a substantial degree). For example, suppose that, given Nora's mother's evidence, it would be unreasonable for her to value being affectionate to her children. As a result, it would be extremely difficult for her to recognize the reasons for being affectionate to her children, and so she would lack the capacity to recognize those reasons (at least to a substantial degree). Hence the above agents are not accountable because they lack the capacity to recognize and respond to moral reasons. It is just that the reasonableness of their beliefs and values explains why they lack this capacity.

You may wonder why we should accept this explanation instead of simply appealing to the reasonableness of their beliefs and values. I see two reasons in favor of this explanation. First, it fits the phenomenology of reactive attitudes. As I have argued,

¹⁷ I thank Sarah Buss for pressing this worry.

reactive attitudes are partly constituted by a thought about the agent's capacity – e.g., resentment is partly constituted by the thought that the agent could have done better. This thought is not about the reasonableness of the agent's beliefs and values, although certainly, as just discussed, the reasonableness of those beliefs and values may lead us to conclude that he could not have done better.

Second, this explanation provides a unified and plausible account of why agents are not accountable across a range of cases. This is because misleading evidence is just one factor that makes it difficult for agents to recognize and respond to moral reasons, but there are other such factors. For example, certain mental illnesses make it difficult for agents to properly assess good evidence about what is valuable or about what they should do, thus making it difficult for them to recognize and respond to moral reasons. This difficulty, I claim, explains why agents with these mental illnesses are not accountable. Or consider strong addictions, in which agents find it extremely difficult to respond to certain moral reasons, even if they recognize them. On my view, this difficulty explains why these agents are not (or are not fully) accountable for actions resulting from their addictions. Thus, the idea that accountability requires the capacity to recognize and respond to moral reasons provides a unified and plausible explanation for why agents are not accountable across a range of cases. Importantly, this unified explanation accounts for the appeal of the other explanations, such as the reasonableness of an agent's false beliefs given her misleading evidence, by saying that these other explanations *tell us why* an agent has difficulty recognizing and responding to moral reasons.

3.5. Fundamental and Derivative Accountability

You might wonder how the argument in this paper fits with the distinction between fundamental and derivative accountability. An agent is *derivatively accountable* for an action if and only if she is accountable for it only in virtue of being accountable for another

action that leads to it. An agent is *fundamentally accountable* for an action for which she is non-derivatively accountable.¹⁸ This distinction is important in assessing the claim that accountability requires the capacity to recognize and respond to moral reasons, since this is only so for fundamental accountability. An agent who lacks this capacity may be derivatively accountable for an action. For example, suppose that Rob gets so drunk that he purposely smashes your car window. While so drunk, assume that he lacks the capacity to recognize and respond to moral reasons; however, he knows that he behaves terribly while drunk but got drunk anyway. It is reasonable to think that, in this case, he is fundamentally accountable for getting drunk and derivatively accountable for smashing your window.¹⁹ Now notice that it seems appropriate to resent people for actions for which they are derivatively accountable – e.g., it seems appropriate to resent Rob for smashing your window. Does the appropriateness of resentment in cases of derivative accountability fit with the idea that resentment is partially constituted by the thought that the agent could have done better?

It does. Consider Rob again. With respect to getting drunk, he could have done better morally: he could have stayed sober (since he knows that he acts morally badly while drunk). With respect to smashing your car window, he could have done better morally: he could have avoided smashing it *by staying sober*. Notice that Rob’s derivative accountability is clear in how he could have done better. Thus, the thought “he could have done better” covers cases in which an agent is blameworthy for an action because he culpably lacks the capacity to recognize and respond to moral reasons. Hence, as we would expect, the same thought partly constitutes resentment for both derivative and fundamental blameworthiness.

¹⁸ Rosen (2004) makes this distinction, using the terms “original responsibility” and “derivative responsibility.”

¹⁹ What makes him derivatively accountable for smashing your car window? Presumably it matters that he could foresee that he would act badly while drunk, but I need not determine here what makes agents derivatively accountable. All that matters here is having a clear case of derivative accountability.

3.6. Frankfurt Cases

You might worry that Frankfurt cases (Frankfurt 1969/2003) count against my proposal. In a Frankfurt case, it seems that the agent could not have done better, but we think that resentment toward him is appropriate. Imagine that Smith is deliberating about whether to paralyze Jones. Black wants to ensure that he does, and so he secretly attaches a device to Smith's brain that works as follows. If Smith is about to decide to paralyze Jones, the device does nothing; if he is about to decide not to paralyze Jones, the device causes him to paralyze Jones. Now imagine that Smith is about to decide to paralyze Jones, and so the device does nothing. Smith then paralyzes Jones. It is clearly appropriate for Jones to resent Smith, but it seems that Smith could not have done better, as he would have paralyzed Jones either way.

Rather than counting against my proposal, I think that Frankfurt cases help to determine the truth conditions of "S could have done better." We may have wrongly taken the "can" in this thought to be modal: an agent can do better if and only if he does better in nearby possible worlds. Frankfurt cases show that this is not so. When we move to the nearest possible worlds in which Smith is about to decide not to paralyze Jones, he still paralyzes Jones. Fortunately, there is another option: the "can" of structural capacity. Smith's capacity to do better – his capacity to recognize and respond to moral reasons – is grounded in structural features of his. Because Black's device remains inactive, it does not affect Smith's structural features, and so Smith can still do better.

To see that this response is reasonable, consider an analogy. A glass is fragile because it has certain structural features. Because of those features, if a glass is not suitably protected, it breaks when struck. However, if in nearby possible worlds the glass is packed in Styrofoam, the fact that it does not break when struck does not mean that it is not fragile. Similarly, although Smith cannot perform a different action because of Black's device, he still has the capacity to do better when the device is inactive, since that capacity is grounded

in his structural features. In support of this, remember that doing *better* does not entail having the ability to perform a different action. Smith could have done better by not paralyzing Jones on his own, so that Black's device would have caused him to do it. Hence Jones can appropriately resent Smith because he can appropriately think that Smith could have done better.

3.7. Other Thoughts?

I have argued that the following two thoughts about the agent partly constitute resentment:

- i) In doing A, S expressed insufficient good will (toward me).
- ii) S could have done better.

Is resentment partly constituted by other thoughts representing accountability conditions? We can reasonably conclude no. First, we have accounted for the standard conditions that exempt agents from accountability. For example, thought i) is false for infants and those under hypnosis or brain stimulation because their actions do not express their quality of will, either because they lack a quality of will (infants) or are disconnected from it (hypnosis or brain stimulation). Thought ii) is false for young children, the mentally ill, and those with terrible upbringings – agents whose actions express their quality of will but who are nevertheless exempt. Since these thoughts cover the standard exemptions, we require no other thought in resentment to account for them. Further, I doubt that we can glean a novel exemption from another thought in resentment, as no other thought about the agent is readily apparent. Since the thoughts in resentment are ones that anyone capable of resentment can have, we would expect such a thought to be readily apparent if there was one. I therefore conclude that we have found all of the thoughts in resentment representing accountability conditions and so have found all of the accountability conditions. Yet

importantly, if there is another accountability condition, the argument of this paper shows how one should argue for it: by arguing that it is entailed by some thought in each reactive attitude.

I have just argued that we have plausibly found all of the accountability conditions, since we have accounted for the standard circumstances that exempt agents from accountability. This means that, even if there are other plausible ways to understand “appropriateness,” it is reasonable to think that pursuing those options would not yield new accountability conditions. This is further supported by the fact (discussed above) that reactive attitudes present their targets as accountable, just as envy presents its target as enviable, and so we can plausibly conclude that any accountability condition can be derived from a reactive attitude’s evaluative presentation – i.e., from its fittingness conditions.

4. The Advantages of This Argument over Others

In this section, I highlight the advantages of the argument presented here over other arguments in the literature for the idea that accountability requires the capacity to recognize and respond to moral reasons.

Start with R. Jay Wallace’s (1994) argument. As mentioned above, Wallace claims that we should understand the appropriateness of reactive attitudes as moral fairness, and I have my doubts about that approach. But even setting aside those doubts, we should still prefer the argument in this paper. That is because of the next step in Wallace’s argument. To move from the idea of appropriateness as moral fairness to the idea that accountability requires the capacity to recognize and respond to moral reasons, Wallace appeals to the following moral principle of fairness: it is unreasonable to demand that someone do something – in a way that potentially exposes him to the harms of sanctions – if he lacks the capacity to recognize and respond to the reasons that support the demand. The main issue with this step in the argument is that unexpressed reactive attitudes are not plausibly

sanctions, and so the ethical principle of fairness to which Wallace appeals does not entail that it is morally unfair to have unexpressed reactive attitudes towards those who lack the capacity to recognize and respond to moral reasons. Yet unexpressed reactive attitudes hold people accountable. Hence Wallace's argument does not support the idea that the capacity to recognize and respond to moral reasons is required for accountability.

Perhaps Wallace could respond by claiming that the connection between reactive attitudes and sanctions is tighter than I am allowing, such that even unexpressed reactive attitudes potentially expose their targets to the harms of sanctions. Yet even if we grant this point about reactive attitudes, we still have good reason to prefer the argument in this paper because it is simpler and more powerful. It is simpler because, unlike Wallace's argument, it does not appeal to an intermediary ethical principle or an assumption about the nature of reactive attitudes specifically in order to arrive at the idea that accountability requires the capacity to recognize and respond to moral reasons. It is more powerful because, unlike Wallace's argument, it gives us the tools to flesh out the capacity to recognize and respond to moral reasons and brings to light the quality of will condition of accountability.

Next consider Gary Watson's (1987/2004) argument that responsibility requires moral understanding. According to Watson, reactive attitudes are forms of moral address²⁰ – e.g., negative reactive attitudes address to their targets “a demand for reasonable regard.”²¹ Watson then points out that, for reactive attitudes to be intelligible as forms of moral address, their targets must have moral understanding.

Even if Watson's idea is correct, his argument does not tell us how to spell out moral understanding. It is natural to interpret moral understanding as the ability to recognize moral reasons, but additional argument is required to get to that conclusion. After all, a target of resentment may understand that, according to the one resenting, he has

²⁰ Watson (1987/2004), p. 231.

²¹ Watson (1987/2004), p. 229.

violated the demand for reasonable regard, and yet he may lack the ability to understand what reasonable regard amounts to in particular cases. Further, Watson’s argument does not entail that the ability to *respond* to moral reasons is necessary for responsibility. The argument in this paper, in contrast, is powerful enough to flesh out moral understanding, and it entails that accountability also requires the capacity to respond to moral reasons. Further, the argument in this paper is simpler because it does not rely upon intermediary principles about reactive attitudes as forms of moral address but only upon the idea that reactive attitudes, like all emotions, have fittingness conditions.

Finally, consider Gideon Rosen’s (2015) argument. As mentioned above, Rosen also understands the appropriateness of reactive attitudes in terms of the truth of their constituent thoughts. Yet he proposes the following “retributive thought” as partly constituting resentment: “S deserves to suffer for what he has done.” To explain why this thought is false for young children and seriously impaired adults, Rosen appeals to Wallace’s above-mentioned principle of fairness: it is unfair to sanction someone for violating a moral rule if, through no fault of his own, he lacked the capacity to recognize and respond to the moral reasons for complying with it.

We should prefer the argument in this paper to Rosen’s argument for three reasons. First, I have already shown that the thought “he could have done better” plausibly partly constitutes resentment across a range of cases, including the cases with which Rosen is concerned.

Second, as Rosen admits, his proposal has a counterintuitive implication. We usually think that children and seriously impaired adults do not deserve sanctions because they are not blameworthy. Rosen’s view says the reverse: that young children and seriously impaired adults are not blameworthy because they do not deserve sanctions. My proposed thought avoids this counterintuitive implication.

Third, Rosen relies on Wallace’s intermediary ethical principle to show that the capacity to recognize and respond to moral reasons is necessary for accountability. We

should not accept this indirect route when a direct route is available, especially not when we understand the appropriateness of reactive attitudes in terms of the truth of their constituent thoughts, as Rosen does. Here is why: reactive attitudes present their targets as accountable, and we should therefore expect to derive the accountability conditions just from the constitutive thoughts of reactive attitudes, as I have done, rather than by combining a constitutive thought with an independent principle.

In sum, the argument in this paper is simpler and more powerful than the other arguments in the literature that move from the nature of reactive attitudes to the idea that accountability requires the capacity to recognize and respond to moral reasons. It is simpler because it does not rely on intermediary ethical principles, supplementary arguments, or assumptions about reactive attitudes specifically. Instead the argument in this paper relies only on the idea that reactive attitudes, like all emotions, have representational content. Further, by just relying on the representational content of reactive attitudes, the argument here takes appropriately seriously the idea that blame represents its target as blameworthy. Finally, the argument in this paper is more powerful than the other arguments because, as we have seen, it brings out the quality of will condition and has the resources to flesh out the capacity to recognize and respond to moral reasons.

5. Conclusion

By understanding the appropriateness of reactive attitudes as fittingness, I argued for the following view of accountability:

S is accountable for action A if and only if:

- 1) A expresses S's quality of will.
- 2) S has the capacity to recognize and respond to moral reasons.

What matters for my argument is that reactive attitudes have representational content, and so my argument would work with any view of fittingness. For the sake of argument, I

assumed the True-Thoughts View: an emotion is fitting if and only if its constituent thoughts are true.

I argued for the above view by showing that Conditions 1) and 2) are each entailed by some thought constitutive of resentment, gratitude, and plausibly the other reactive attitudes. Condition 1) is entailed by the thought partly constituting resentment that the agent expressed insufficient good will in acting, and the thought partly constituting gratitude that the agent expressed more than sufficient good will in acting. I then briefly connected Condition 1) to the idea of the agent being attributionally responsible for the action. Condition 2) is entailed by the thought partly constituting resentment that the agent could have done better, and the above thought partly constituting gratitude. Gratitude may also be partly constituted by the thought that the agent could have done less well, which also entails Condition 2). Since Conditions 1) and 2) account for all of the circumstances that we intuitively think exempt agents from accountability, they are plausibly jointly sufficient, as well as individually necessary, for accountability. Finally, I demonstrated that the argument in this paper is both simpler and more powerful than other arguments in the literature that use reactive attitudes to argue that accountability requires the capacity to recognize and respond to moral reasons.

Acknowledgments I thank Nathan Ballantyne, Lara Buchak, Sarah Buss, Gideon Rosen, Michael Smith, and anonymous referees for very helpful comments on previous drafts of this paper. I also thank the participants of the following workshops for valuable comments and questions: the Moral Responsibility Workshop at Rutgers University (2017), the Speculative Ethics Forum at St. John’s University (2015), and the 2013–14 New York City Early Career Ethics Workshop.

References

- Arpaly, Nomy. (2003) *Unprincipled Virtue*, Oxford University Press.
- D'Arms, Justin and Daniel Jacobson. (2000) "The Moralistic Fallacy: On the 'Appropriateness' of Emotions," *Philosophy and Phenomenological Research*, 61 (1): 65-90.
- Frankfurt, Harry. "Alternate Possibilities and Moral Responsibility," in Gary Watson (ed.) *Free Will*, Oxford University Press, 2003. (Original Work Published 1969).
- Greenspan, Patricia. (1988) *Emotions and Reasons*, New York: Routledge.
- Nelkin, Dana. (2011) *Making Sense of Freedom and Responsibility*, Oxford University Press.
- Nussbaum, Martha. (2001) *Upheavals of Thought: The Intelligence of Emotion*, Cambridge University Press.
- Oddie, Graham. (2005) *Value, Reality, and Desire*, Oxford University Press.
- Prinz, Jesse. (2004) *Gut Reactions: A Perceptual Theory of Emotions*, Oxford University Press.
- Roberts, Robert C. (2003) *Emotions*, Cambridge University Press.
- Rosen, Gideon. (2004) "Skepticism about Moral Responsibility," *Philosophical Perspectives* 18 (1): 295-313.
- Rosen, Gideon. (2015) "The Alethic Conception of Moral Responsibility," in Randolph Clarke, Michael McKenna, and Angela M. Smith (eds) *The Nature of Moral Responsibility: New Essays*, Oxford University Press.
- Solomon, Robert C. (1976) *The Passions*, Garden City, NY: Anchor Books.
- Strabbing, Jada Twedt. (2016a) "Attributability, Weakness of Will, and the Importance of Just Having the Capacity," *Philosophical Studies* 173 (2): 289–307.
- Strabbing, Jada Twedt. (2016b) "Responsibility and Judgment," *Philosophy and Phenomenological Research* 92 (3): 736–760.
- Strawson. P.F. (1962/2003) "Freedom and Resentment," in Gary Watson (ed.) *Free Will*, Oxford University Press.
- Wallace, R. Jay. (1994) *Responsibility and the Moral Sentiments*, Harvard University Press.

Watson, Gary. (1987/2004) “Responsibility and the Limits of Evil,” in his *Agency and Answerability*, Oxford University Press.

Watson, Gary. (1996/2004) “Two Faces of Responsibility,” in his *Agency and Answerability*, Oxford University Press.

Wolf, Susan. (1990) *Freedom within Reason*, Oxford University Press.