

Intrapersonal Ascriptions of Responsibility

by
Maureen Sie

[...] theoretical and practical ethics lie on a continuum that leads from the particular to the general, from the application of values to the analysis of moral concepts, from shared moral beliefs to justified moral principles, and vice versa. This in itself is a complex and highly ambitious undertaking that renders any sharp demarcation between the two fields obsolete. (Gerhold K. Becker in Sie 2005: xi)

Introduction

Recent research in the behavioral, cognitive, and neurosciences reveals that sincere, fully developed moral agents who are mentally healthy are nevertheless capable of grave mistakes concerning their so-called agential contribution. That is, they sometimes judge that they acted whereas it can be shown that they did not, and, vice versa, they sometimes fail to recognize that they acted whereas it can be convincingly established that they did act. More about this later.

In the wake of these findings a host of researchers have developed experiments designed to test our so-called moral agency: our capacity to judge and act for/on the basis of moral reasons. The results suggest that in this central human area too, we lack crucial agential transparency, offering reasons for our moral judgments that played no role in generating them and sticking to our moral judgments even when we are no longer capable of providing sound reasons for them.

These findings are challenging, for, first of all, the practice of giving and asking for reasons is a central feature of our moral practices.¹ Secondly,

¹ Of course there is a huge range of moral attitudes and practices in every society, many of which might be quite unaffected by insights concerning the lack of agential transparency and the causes of our moral judgments. When speaking of our “moral practices” in this general manner I refer to the basic and crucial assumptions without which the whole idea of “morality” and our nature as moral beings seem to collapse, e.g., the assumption that we are able to and often act *in accordance with* norms and values.

many philosophers believe that agential authority—whatever that might mean exactly—is a crucial ingredient of our (moral) agency. Hence, if we lack such authority, what does this tell us about our moral practices?

In this essay I argue that the lack of agential transparency is part and parcel of our ordinary moral practices. Therefore, recent developments in the sciences strengthen rather than challenge these practices: the developments in the behavioral, cognitive, and neurosciences indicate that we should consider ascribing moral responsibility to ourselves even for “doings” we do not entertain any authoritative relation to and/or not even recognize as ours.

I To Do or Not to Do

At some stage in their development, little children discover the “I did not do that on purpose” excuse: they do something, but when held responsible for it will tell us that, since they did not do it on purpose, they do not deserve punishment or blame. They elaborately try out this excuse on different occasions also, for example, when pinching their siblings or throwing a ball skillfully in the “wrong” direction (on the dinner table, on the garden flowers). Of course, their parents tell them that pinching and throwing things are not the kind of actions one can *do* accidentally.

They tell them this, even though there sure is a difference between pinching or throwing a ball *on purpose* and doing so “accidentally.” Sometimes children just get “carried away” and do perform intentional actions accidentally. One of the reasons why we do not hold children fully responsible for their behavior, even though they often act on purpose and intentionally, is that they are not *well-practiced* self-governing agents, yet.²

Of course—if only for educational reasons—parents often act as if they held their children fully responsible, voicing their moral indignation and resentment whenever their children use the excuse of “doing it by accident” on pinching or throwing occasions; or, for that matter, on other occasions that do not qualify for the use of the “I did not do that on purpose” excuse.

Part of every proper moral education is not only to learn what we should and should not do, but also on what occasions to *take* responsibility and on what occasions there is no immediate need for that. What we communicate

² Cf. Paul Benson, who argues that children are not fully developed normative evaluators yet, able to evaluate their own behavior on the basis of normative standards “[...] and to make competent critical evaluations, in light of those norms, of open courses of action.” (Benson 1987: 476).

to one another is that one cannot do certain things except “on purpose,” i.e., that such doings need a kind of “agential contribution” that excludes its classification as something one did “accidentally.”

Also, we learn that in some situations, accidentally doing something should actively be prevented by an appropriate amount of concentration and attention. A teenager who—upon your disapproving frown—blurts out at you that “surely she did not step on your toes on purpose!” does not quite grasp what is expected of her in this situation. That is, she is expected not to step on your toes, not even accidentally. She is expected to be careful *not* to step on your toes; and, if she accidentally fails to fulfill that expectation, she is expected to apologize, or, at the very least, feel awkward for having stepped on your toes.

Accidentally doing something, lack of intention or intentionality, and even lack of purposeful bodily movements—e.g., when one is literally thrown off one’s feet against another by some abrupt movement of the bus—do not necessarily apply as a disclaimer; i.e., as a condition or circumstance that mitigates our personal responsibility.

In our everyday moral practice it depends upon the situation whether we believe talk in terms of “mere doings,” “actions,” or “blameworthy actions” is appropriate. One set of bodily movements might even qualify for different descriptions. E.g., although one is usually blamed for getting into fights and expected to take responsibility for staying out of fights, one can at the same time be praised for the bodily movements made *in* the fight. One does not necessarily need to excuse oneself for a specifically well placed punch during the fight; on the contrary, after making peace with one’s opponent, one might even receive her admiration and praise for well placed punches.

To make clear this thoroughly situational aspect of our moral practices, I refer to our moral practices in terms of “normative expectations,” i.e., the specific “shoulds” and “should nots” that regulate and constitute our shared practices.³ We expect our children not to pinch and to be careful when playing with their ball. And “expect” is meant here in a specifically *normative* sense, though there are some predictive aspects to our expectations as well.⁴ Likewise, we expect a teenager not to step on our toes even in a crowded bus, we expect her to *attend to* not stepping on our toes accidentally. (Although we might of course excuse or exempt the teen if we believe that being in one’s teens is like being in a—hormonally induced—*enduring*

³ Cf. Jay Wallace’s reactive account of moral responsibility. Wallace, R. J. (1994).

⁴ For the relation between normative and predictive aspects of our expectations, cf. Sie, M. (2005), Chapter 3.

“normative expectations transgressing mood.”) Finally, we expect one another to avoid getting into fights; but, once in a fight, we might also expect everyone participating “to stand up and fight as well as one can.” It belongs to the dynamics of the fight, its intrinsic aim, its purpose, and its rationality, to defend and fight as well as possible.

II The Normative Disagreement View on Responsibility

In our day-to-day dealing we tend to fulfill these normative expectations effortlessly and most of the time without preceding deliberation, specific attention, and perhaps sometimes even without much *conscious* control. All of us do such things as, for example, running to be in time for appointments, answering questions truthfully, ignoring/obeying traffic-regulations (even) when in a hurry, and so on. If we had time to stop and reflect on what should be done on each and every particular occasion, we might sometimes judge that we should have done otherwise than we in fact did.⁵ Consider how difficult it is (and the ample practice it takes most of us) to resist answering questions of, e.g., telephone enquirers. This is so not because we need time to decide or convince ourselves that these enquirers are annoying, but because we are so naturally inclined to provide an answer as a response to a question. Or consider what happens when we are in a rush, how easy it is to ignore or forget things (our children desperately in need of a hug) even when—with hindsight—the appointment that sets us off in a hurry turns out to be not worth a hurry to begin with.

As I argue elsewhere, many “ordinary” everyday wrongdoings arise from the fact that our practices are regulated by different—sometimes even conflicting—normative expectations and that we are inclined to fulfill these normative expectations often in an automatic manner, without much forethought (Sie 2005, Chapter 4; Sie 2006). E.g., I might ignore a red traffic light at a dangerous spot in a reckless attempt to be in time for an important meeting at work. Had I taken the time to reflect and weigh the conflicting sets of normative expectations—“one should obey traffic regulations especially if one risks the safety of others” and “one should be in time for important meetings”—I might have made the “right” decision and stopped.

⁵ Which is not to claim that we need believe that we “could have done otherwise than we actually did,” a phrase that usually refers to robust libertarian free will. Nowadays there are also compatibilist conceptions available of the thought that “we should have done otherwise,” e.g., as expressing a negative moral evaluation of what we did (Frankfurt 1969, 1971: 5-20, 1973/88, and 1987).

Ascribing⁶ responsibility to one another by default for these kinds of ordinary wrong actions brings to the surface the normative expectations we share with one another and agree about, but also those we disagree on. In the foregoing example, I realize that I did wrong and by making apologies and excuses⁷ disclose that I agree with those blaming me for not having stopped at the traffic light. However, in slightly different circumstances I might disagree that I should have stopped at the traffic light and learn that other people evaluate “adherence to traffic-regulations” in quite a different way than I do.

This, in a nutshell, is the basis of what I have called “the normative disagreement view” on responsibility: we can justify our practices of moral responsibility

1. regardless of metaphysical complications concerning the questions
 - a. whether free will is a necessary prerequisite for responsible agency and
 - b. whether free will and determinism are compatible;
2. provided that we restrict ourselves to everyday moral wrongdoings and exclude unintelligible and extraordinary “moral monster”-like crimes (Sie 2005, Chapter 4; Sie 2006).

This view takes as its starting point a picture of human agency crucially lacking what we could call “moral transparency.” Because we so often act in an automatic and unaware fashion when fulfilling all kinds of unexplicated normative expectations, default ascriptions of moral responsibility are indispensable for clarifying in what ways we desire to deal with one another and/or be dealt with. In view of recent research findings from the behavioral, cognitive, and neurosciences, this starting point is fortunate: these findings indicate that deliberative awareness and conscious control are all but pervasive aspects of our everyday interactions with one another. However, the lack of transparency disclosed by these developments seems to go a bit further, revealing a fundamental lack of transparency that stands in the way of any practice aimed at clarification. Let me explain.

⁶ I use the phrase “ascribing responsibility” instead of, e.g., “holding responsible” to make it clear that I am here specifically concerned with the moral attitude and/or sentiments that *express/disclose* our idea that the agent is responsible for the action (including behavior and mere bodily movements). Cf., e.g., the negative moral sentiments (of blame and resentment) and the positive ones (of praise and moral admiration) made influential in the discussion on the legitimacy of our practices of responsibility by Peter F. Strawson (Strawson 1962: 59-81).

⁷ Often the mix of excuses and apologies forms a slightly incoherent whole, because, on the one hand, we want to explain how it could happen that we did something we ourselves also judge to be wrong (excuses), while on the other hand we want to take responsibility and humbly acknowledge that we did wrong (apologies).

III The Adaptive Unconscious⁸

The overall paradigm arising from the behavioral, cognitive, and neurosciences is the paradigm of the Adaptive Unconscious, so named in honor of Kihlstrom's "cognitive unconscious" (Kihlstrom 1987) later recast as the "new unconscious" (Wilson 2002). Unlike the Freudian unconscious, this new unconscious is not a collection of highly complex interacting drives and conflicts, blindly seeking gratification "without regard to constraints of reality and society" (Uleman 2005: 5). Rather it is an efficient, though also highly complex, interacting "module"⁹ that enables us to cope with the world in a much more adequate way than when acting solely on the basis of our conscious information-processing and decision-making procedures. Hence, the name *adaptive* unconscious.

According to the adaptive-unconscious paradigm, our interactions with the world take place primarily at an "unconscious" level, in an "automatic" manner (Bargh and Chartrand 1999).¹⁰ To be sure, the concepts "unconscious" and "automatic" call for extensive further clarification (Levy and Bayne FC), as do the concepts of awareness, consciousness, and attention. Determining the exact meaning of these concepts and their importance in our understanding of such crucial notions as free will and agency is not an easy matter; it is the heart of many contemporary discussions in the philosophy of mind (Pockett, Banks, et al. 2006). However, what is important for the purposes of this paper does not depend on the *exact* meaning of those concepts. Rather, what counts is the fact that at least part of what we do takes place at a level unavailable to prior conscious deliberative processes and decision-making procedures, and the fact that we are all but infallible in judging whether we did or did not do a particular thing and for what reasons we did it. Not only do we effortlessly make up reasons for bodily movements that are actually triggered by causes of which we are unaware; we are not even always aware of doing so and, vice versa, we seem quite capable of performing intentional bodily movements without being aware of it too (Wegner and Wheatley 1999; Wegner 2002).

⁸ Earlier stages of the same thoughts are expressed in "Delegated and Delegating Moral Agency: Some Observations on Everyday Reason-Talk," a paper under review for the volume *Understanding Other Minds and Moral Agency*, to be edited by Karsten R. Steuber.

⁹ The term "module" is not (necessarily) to be understood *literally* as an empirically identified "part of our brain." How exactly our brains are made up is subject to great controversy.

¹⁰ Cf., e.g., the extensive work conducted on the somatic marker hypothesis (Saver and Damasio 1991; Damasio 1994; Phelps 2005; Bechara, Tranel, et al. 1993) and on what Wegner calls "conscious will" (Bechara, Tranel, et al. 1993; Wegner and Wheatley 1999; Wegner 2002).

The lack of awareness and the automaticity of our bodily movements is not in itself a problem, neither practically nor philosophically speaking. We know from everyday experience that it is extremely efficient to “delegate” our actions so that they take place outside of our immediate attention—e.g. driving our car to work while listening to the newest audio book on our iPod. As long as we can trust that we ourselves—somehow—determine the general direction of where we are heading, such delegating possibilities enhance rather than undermine our agency.¹¹ Moreover, they do not necessarily seem to lessen our responsibility for the things done. What actions can, and what actions can not be delegated, and in what way, is structured by our moral practices: e.g., I can daydream while standing in a crowded bus, but only if I am used to standing firm even on bumpy roads; I can listen to my iPod while driving, but only on the condition that I am able to snap back if my full awareness is needed to react to some unforeseen happening on the road. If we are, respectively, clumsy or absent-minded persons, daydreaming in a crowded bus and listening to an iPod while driving are “morally inadvisable,” to say the least. In this sense, the lessons from our youth, “not to pinch” and “not to throw balls accidentally,” continue in our adult lives.

However, unlike this unconscious automaticity familiar to us from everyday experience and regulated by our moral practices, the adaptive-unconscious paradigm suggests that there also exists something we could call “undelegated” agency: finding ourselves doing things and making up reasons *only as we go along*. (Bargh and Chartrand 1999; Wilson 2002; Fine 2006). There are people who argue that more often than we might suppose we are not directly aware of what drives our actions but infer reasons on the basis of *a priori* causal theories, confabulating them if we cannot find reasonable explanations (Nisbett and Wilson 1977). Moreover, it is argued that so many causal factors escape consciousness that confabulation seems the rule rather than the exception (Wilson 2002; Fine 2006). Even deliberative actions like our moral judgments—embedded in a widely shared and sophisticated justificatory framework—seem based on intuitions that are not, or only partially, accessible to introspection.

On the basis of such experiments and views, some people have argued that the reasons we come up with to justify these judgments are *mere* post-hoc rationalizations that played no role in their generation (Haidt 2001).¹² Of course it remains to be seen how exactly we should interpret and evaluate these theories. What they establish, though, is that our reason-

¹¹ On this cf. Fine 2006: Chapter 6; Haidt 2006; Waal, Macedo, et al. 2006.

¹² He soon modified these claims; cf. Greene and Haidt 2002: 517-523.

giving practices—insofar as they are taken to disclose the reasons for which we act—are not to be taken at face value, at least not without further argument. That is, we need to accommodate the fact that at least sometimes the reasons we offer do not track the true causes of our behavior. As a consequence, we need to accommodate the fact that we lack agential transparency in a much more fundamental way than can be neutralized by our current practices of responsibility ascriptions.

As put forward in the previous section, the idea that we can accommodate a lack of moral transparency in the case of interpersonal ascriptions of responsibility, relies on its function to explicate and illuminate the normative expectations that regulate our shared practices. However, how can we speak of illuminating and explicating normative expectations if we also lack this more fundamental agential transparency? Does this not undermine the very practice of giving and asking for reasons that our interpersonal ascriptions of responsibility invite?

IV Interpersonal and Intrapersonal Ascriptions of Responsibility

Practical philosophers tend to agree that it would be “patently absurd” to doubt the common assumption that there is a trustworthy—though, of course, defeasible—relation between our rational deliberations and our actions (Smith 1994: 132). Provided that we are not too strict regarding the application of the concept “reason,” it is safe to say that we act in accordance with reasons all the time: We stop because the traffic light is red, I get angry because you insulted me, you take the item from the shop because you think it is a free sample. That is not to claim that in all these cases we are aware of what moves us in the way it does. If someone stops for a red traffic light because she believes it important to obey traffic regulations, her hitting the brakes need not involve any *thought* whatsoever. To say that we “acted for reasons” does not exclude that we acted in an automatic and unaware fashion.

As long as we can understand the bodily movements as somehow standing in “a proper connection” to the agent’s desires, values, evaluations, and/or judgments, these movements qualify as an action. A contrast is usually drawn here with bodily movements that merely happen to the agent, such as sneezes or tics. As put forward in the first section of this paper, I believe that it depends upon the situation whether we believe talk in terms of “mere doings,” “actions,” or “blameworthy actions” to be appropriate. Part of our education and learning to be a self-governing (moral) agents is mastering the appropriate application of excuses and exemptions concerning

control and the lack of control. I therefore do not believe that there is an easy distinction to be made between mere bodily movements on the one hand and actions on the other. Someone suffering from a ticcing disease such as Tourette's syndrome is able to withhold tics, e.g., to operate; hence, we cannot simply say that in all circumstances ticcing and/or sneezes are not things we do. The details of what constitute "a proper connection" between our desires, values, evaluations, and/or judgments and our actions are to be worked out in practice, by sheer *practicing*.

What the adaptive-unconscious paradigm makes clear is that we do not know ourselves as well as we might have thought; that the desires, values,¹³ evaluations, and/or judgments moving us into action might not necessarily be the desires, values, evaluations, and judgments we recognize as ours and would want to authorize to speak on our behalf when brought to our full awareness. However, that does not automatically mean that we can reject responsibility for them. Perhaps with practice, these states can be brought to our awareness and/or within our control. Hence, even if it is the case that we are not infallible with regard to our judgment that we performed some action, it is sensible *and* advisable to ascribe responsibility to ourselves—within reason—whenever it looks as if our bodily movements (or lack thereof) constitute an action.¹⁴ The adaptive-unconscious research, rather than undermining the legitimacy of our ascriptions of responsibility, actually strengthens them. To reject responsibility for all bodily movement (or lack thereof) that does not fit one's own ideas about oneself—one's picture of oneself—is especially misperceived in light of the adaptive-unconscious paradigm.

Let me conclude by providing an example of the need for intrapersonal ascriptions of responsibility. One of the things that becomes increasingly clear from the adaptive-unconscious literature is that prejudices can exist at two levels; as Timothy Wilson puts it, at "an explicit level (people's conscious beliefs and feelings about other groups) and an implicit level (people's automatic evaluation of other groups of which they might not be aware)" (Wilson 2002: 133). Although at an explicit level we might resist

¹³ Under certain conceptions of what it means to "value" something, it might not be possible to endorse the idea of values of which we are not aware but that are nevertheless efficacious in our behavior. I do not think such an account of "values" would be very convincing since I believe it perfectly natural to say of someone, e.g., that she thinks she values (our) friendship, but actually does not.

¹⁴ In some circumstances or situations, the lack of bodily movements is a full-blown action, e.g., when we do not respond to someone who just greeted us.

and oppose prejudices, it can be shown that we are thoroughly influenced by them, virtually all the time.¹⁵ Perhaps, therefore, when a particular action of ours can be described as prejudiced (or is actually described as such by others), even though we do not endorse sexist or racist views, we could look at what the results are of nevertheless ascribing responsibility to ourselves.¹⁶ In time and with practice, these intrapersonal ascriptions of responsibility will perhaps reveal unsuspected opportunities to, somehow, change our ways and, consequently, the normative expectations that regulate our shared practices in this respect.

Conclusion

In this paper I have raised the question whether current research in the behavioral, cognitive, and neurosciences that can be subsumed under the heading of the “adaptive-unconscious paradigm” threaten to undermine our everyday ascriptions of responsibility to one another and, more specifically, to ourselves. I first sketched a general framework of assumptions concerning our everyday practices of responsibility. This takes as its starting point the idea that our ascriptions of responsibility to one another primarily serve the function of illuminating, establishing, and consolidating the normative expectations that should regulate our shared practices with one another. I have extended this view to include the normative expectations that regulate the range of bodily movements, actions, and behavior for which we should take responsibility. Next, I have addressed research on the adaptive unconscious and investigated whether within the bounds set by this research we could still make sense of the normative-disagreement view on responsibility. The reason for raising this question is that the adaptive-unconscious paradigm denies that we are agentially transparent. Agential transparency seems required to make sense of our reason-giving practices. According to the normative disagreement view, these reason-giving practices are crucial to the function our responsibility ascriptions serve. I have argued that the adaptive-unconscious paradigm, rather than undermining the legitimacy of ascribing responsibility to oneself, in fact broadens the range of bodily movements for which we should ascribe responsibility to ourselves.

¹⁵ For a helpful discussion see Wilson 2002: Chapter 9, and Fine 2006: Chapters 8 and 6.

¹⁶ “Ascribing responsibility to ourselves,” as I understand it, is different from the more familiar and mundane idea of “taking responsibility.” It crucially includes a hypothetical element (perhaps I did do this or that) that is not required when one takes responsibility for something.

References

- Bargh, John A. and Tanya L. Chartrand, "The Unbearable Automaticity of Being," *American Psychologist* (July 54, 1999): 462-479.
- Bechara, Antoine, Daniel Tranel, et al., "Failure to respond automatically in anticipation of future outcomes following damage to human pre-frontal cortex," *Society for Neuroscience* 19 (1993): 791.
- Benson, Paul, "Freedom and Value," *The Journal of Philosophy* lxxxiv/9 (1987): 465-486.
- Damasio, Antonio R., *Descartes' error : emotion, reason, and the human brain*, G.P. Putnam, New York, 1994.
- Fine, Cordelia, *A Mind of Its Own: How Your Brain Distorts and Deceives*, New York: W.W. Norton & Co, 2006.
- Frankfurt, Harry G. [1969], "Alternative Possibilities and Moral Responsibility," *Journal of Philosophy* LXVI/23 (1969): 829-839.
- [1971], "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (January 14, 1971): 5-20.
- [1973/88], "Coercion and Moral Responsibility," originally in Ted Honderich, ed., *Essays on Freedom and Action*, London: Routledge & Kegan Paul, 1973; repr. in: Frankfurt, H. G., *The Importance of What we Care About*, New York: Cambridge University Press, 1988, 26-46.
- [1987] "Identification and Wholeheartedness," in Ferdinand D. Shoeman, ed., *Responsibility, Character, and the Emotions*, New York: Cambridge University Press, 1987, 27-45.
- Greene, Joshua and Jonathan Haidt, "How (and Where) Does Moral Judgment Work?" *Trends in Cognitive Sciences* 6/12 (2002): 517-523.
- Haidt, Jonathan [2001], "The emotional dog and its rational tail: A social intuitionist approach to moral judgment," *Psychological Review* 108 (2001): 814-834
- [2006], *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom*, New York: Basic Books, 2006.
- Kihlstrom, John F., "The Cognitive Unconscious," *Science* 237 (1987): 1445-1452.
- Levy, Neil and Tim Bayne, "Doing without Deliberation: Automatism, Automaticity, and Moral Accountability," *International Review of Psychiatry* (forthcoming).

- Nisbett, Richard E. and Timothy D. Wilson, "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review* 84 (1977): 231-259.
- Phelps, Elizabeth A., "The Interaction of Emotion and Cognition: the Relation between the Human Amygdala and Cognitive Awareness," in: Hassin, R. et al., eds., *The new unconscious*, Oxford etc.: Oxford University Press, 2005.
- Pockett, Susan, William P. Banks, et al., *Does consciousness cause behavior?* Cambridge, MA: MIT Press, 2006.
- Saver, Jeffrey L. and Damasio, Antonio R., "Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage," *Neuropsychologia*, 29 (1991): 1241-1249.
- Sie, Maureen M.S.K., *Justifying Blame: Why Free Will Matters and Why it Does Not*, Amsterdam-New York: Rodopi, 2005.
- , "Ordinary Wrongdoing and Responsibility Worth Wanting," *European Journal of Analytic Philosophy* 1/2 (2006): 67-82.
- Smith, Michael, *The Moral Problem*, Oxford, UK and Cambridge, MA: Blackwell, 1994.
- Strawson, Peter F., "Freedom and Resentment," in G. Watson, ed., *Free Will*, Oxford: Oxford University Press, 1962: 59-81.
- Uleman, James S., "Becoming Aware of the New Unconscious," in R. Hassin et al., eds., *The New Unconscious*, Oxford etc., Oxford University Press, 2005.
- Waal, Frans, Stephen Macedo, et al., *Primates and Philosophers: How Morality Evolved*, Princeton, NJ: Princeton University Press, 2006.
- Wallace, R. Jay, *Responsibility and the Moral Sentiments*, Cambridge, MA: Harvard University Press 1994.
- Wegner, Daniel M., *The illusion of conscious will*, Cambridge, MA: MIT Press, 2002.
- Wegner, Daniel M. and Thalia Wheatley, "Apparent Mental Causation: Sources of the Experience of Will," *American Psychologist* 54 (1999): 480-491.
- Wilson, Timothy D., *Strangers to Ourselves: Discovering the Adaptive Unconscious*, Cambridge, MA: Belknap Press of Harvard University Press, 2002.