



Oxford Studies in Agency and Responsibility Volume 1

David Shoemaker

Print publication date: 2013

Print ISBN-13: 9780199694853

Published to Oxford Scholarship Online: September 2013

DOI: 10.1093/acprof:oso/9780199694853.001.0001

Responsibility, Naturalism, and “The Morality System” *

Paul Russell

DOI:10.1093/acprof:oso/9780199694853.003.0008

Abstract and Keywords

Even those who follow the general strategy of P. F. Strawson’s enormously influential “Freedom and Resentment” accept that his strong naturalist program needs to be substantially modified, if not rejected. An important effort to revise the Strawsonian program has been provided by R. Jay Wallace. This chapter argues that Wallace’s narrow construal of reactive attitudes, as they are involved in holding an agent responsible, comes at too high a cost. Related to this point, it is also argued that Wallace’s narrow conception of responsibility is a product of his effort to construct his account within the confines of the morality system and that this way of construing responsibility turns on series of unnecessary and misleading oppositions. A more plausible middle path, it is maintained, can be found between Strawson’s excessively strong naturalist program and Wallace’s narrow and restrictive view of responsibility.

Keywords: Responsibility, naturalism, the morality system, moral skepticism, genealogy, P. F. Strawson, R. Jay Wallace, Bernard Williams

Theory typically uses the assumption that we probably have too many ethical ideas ... Our major problem now is actually that we have not too many but too few, and we need to cherish as many as we can.

Bernard Williams, *Ethics and the Limits of Philosophy*

Lying at the heart of P. F. Strawson’s core strategy in “Freedom and Resentment” is his effort to direct his naturalist claims and observations against not only the philosophical extravagance of a general skepticism about moral responsibility but also against all nonskeptical attempts to provide responsibility with some form of external rational justification (Strawson 1962: 23). According to Strawson, efforts of this kind are not only misguided and unconvincing in themselves, when they fail they encourage a general skepticism about moral responsibility. The alternative strategy that Strawson pursues is one that places the foundations of responsibility on our natural, universal emotional propensities and dispositions relating to moral sentiments or reactive attitudes. This naturalistic turn invites us to focus our attention on familiar facts about human moral psychology, and to drop our focus on the analysis of the concept of “freedom” as a way of dealing with the threat of determinism. Beyond these general features, however, the details of Strawson’s strategy become both complex and layered. As a consequence of this, interpretations and assessments of his arguments differ greatly. There **(p.185)** is, nevertheless, a general consensus among both followers and critics alike that there are significant strands in Strawson’s specific naturalistic arguments that are implausible and unconvincing and that some “retreat” from the original strong naturalist program that he advanced is required.

In this paper I take up two closely related issues arising out of this overall problem. First, I want to consider if the right way to amend and modify the naturalistic program is to adopt a “narrow” construal of our moral reactive attitudes along the lines proposed by R. Jay Wallace, one of Strawson’s most prominent followers on this subject. The narrower approach, as I will explain, involves a substantial retreat from Strawson’s original naturalistic program and has significant implications for Strawson’s core claim that our commitment to responsibility requires no external rational justification. The second issue that will be addressed is whether or not we should interpret moral responsibility entirely within the confines of what Bernard Williams has described as “the morality system.” (Williams 1985: ch.1). The narrow construal of responsibility requires that we understand moral responsibility within the conceptual resources provided by the morality system, making notions of obligation, wrongness, and blame essential to the analysis of moral responsibility. There is, therefore, an intimate connection between these two issues. With respect to both these issues I will argue that the narrow approach, while it has legitimate criticisms to make of Strawson’s original strategy, nevertheless takes us in the wrong direction and involves an unacceptable distortion and truncation of moral responsibility.

1. Two Modes Of Naturalism

Strawson’s naturalistic account of moral responsibility insists that a proper understanding of this matter must begin with a description of what is involved in *holding* a person responsible. The key to his analysis is the role that reactive

attitudes play in this sphere, where this is understood in terms of our natural emotional responses to the attitudes and intentions that we manifest to each other. Strawson’s naturalistic strategy, as based on this general observation, has two aspects or dimensions that need to be carefully distinguished—although this is not done in his own presentation. There is, in the first place, a strong form of naturalistic argument that involves the claim that even if we had some theoretical reason to abandon entirely or altogether suspend our reactive attitudes (e.g. as required by skeptical arguments based on considerations of determinism), it would be psychologically impossible for us to do this (Strawson 1962: 9–12,18). A systematic repudiation of all reactive attitudes of this kind, he argues, **(p.186)** would result in “a thoroughgoing objectivity of attitude” to others, with the resulting loss of all our interpersonal personal relations (1962: 12–3). While the objective attitude may be appropriate when we are dealing with the abnormal and incapacitated, and may even on occasion be available to us when dealing with normal adults, we cannot “do this for long, or altogether” (1962: 9–10). Armed with these claims, Strawson suggests an easy way to deal with the skeptic. No matter what theoretical arguments may be presented, our human nature is such that we will continue to feel or entertain *tokens* of reactive attitude. According to this view, skeptical arguments will inevitably fail to dislodge or wholly eradicate these attitudes and responses. Let us call this strong line of reply token-naturalism, since it turns on the claim that our tokens of reactive attitudes cannot be systematically eliminated or abandoned whatever (philosophical) arguments may be advanced against them.

Most philosophers have found Strawson’s token-naturalism too strong and unconvincing. There are two basic objections to be made against it.¹ First, the psychological claim that it makes is doubtful in point of fact. It is not at all obvious that we are constitutionally incapable of entirely ceasing to entertain tokens of reactive attitude should we be persuaded that they are systematically unjustified. Moreover, even if the psychological claims were true, this does nothing to remove or discredit the objections put forward by the skeptic, since these claims do not address the justificatory issue that concerns us. It would, in fact, be disturbing to discover that we will naturally continue to entertain tokens of reactive attitude in face of well-founded grounds for discrediting these emotional responses to others.

While Strawson’s token-naturalism may be judged too strong, there is a weaker form of naturalism that is more convincing and has attracted greater interest. What is crucial to naturalism, on this account, is that we are all *liable* or *disposed* to reactive attitudes. It is from within the framework of this weaker form of naturalism that Strawson develops his general rationale of excuses and exemptions. While we may all have a natural liability to reactive attitudes, particular tokens can be discredited by reference to excuses and exemptions. Excuses operate by way of showing that the agent’s conduct was consistent with the underlying general demand for “some degree of goodwill or regard” (1962:

7). In cases of this kind, involving an accident, ignorance and other such factors, the conduct in question does not manifest any ill-will or malicious intent, even if some injury has **(p.187)** occurred. With respect to exemptions, however, we are asked to view the individual as an altogether inappropriate target of our reactive attitudes on the general ground that we are dealing with someone for whom we cannot make the moral demand due to an abnormality or incapacity of some relevant kind. Strawson employs this two-level account of excuses and exemptions to show that the thesis of determinism fails to engage any of these recognized considerations and cannot, therefore, constitute a basis for systematically discrediting all our reactive attitudes associated with moral responsibility.

This weaker naturalist approach may be described as a form of *type-naturalism*, where this is understood in terms of our natural disposition to reactive attitudes (just as we have a natural liability to love, fear, joy and other basic emotions). What is crucial, however, is that unlike token-naturalism, type-naturalism offers no easy way of discrediting the skeptic. On the contrary, it is essential, on this approach, that a plausible account of excusing and exempting conditions is provided consistent with compatibilist commitments. At this level, concerning our natural liability to reactive attitudes, we must still engage in arguments that counter the skeptical challenge. At the same time, type-naturalism does insist on the “internal” nature of these replies to the skeptic (1962: 23). While it is possible that our token reactive attitudes could be systematically discredited from within, there remains no need or possibility of providing an external, rational justification of a more general kind for our (natural) propensity to these emotions (any more than we need to do this for our similar liability to love, fear, etc.). Although justificatory issues remain with us, and cannot be evaded by way of token-naturalist claims, these justificatory requirements do not take the form of a demand for general or external rational justifications.

2. Reactive Attitudes And Narrow Responsibility

Having established a distinction between token- and type-naturalism, let us now consider Wallace’s amended account of the Strawsonian project, as developed on the basis of his narrow construal of the reactive attitudes. Wallace’s compatibilist account weaves together two distinct strands of thought. The first is a broadly Strawsonian description of holding people responsible, interpreted in terms of our reactive attitudes (Wallace 1994: 8-12). The other strand is his Kantian theory of reflective self-control or moral agency (1994: 12-15). Taken together, these two strands constitute **(p.188)** what Wallace calls his “normative interpretation” of responsibility, which maintains that the correct way to understand what it is to be a morally responsible agent is by way of describing those conditions under which it is *fair* to hold an agent responsible (1994: 5,15, 64). Our stance of holding a person responsible must itself be understood in terms of the mutual dependence between expectations and reactive attitudes (1994: 20-5). To hold someone to an expectation, or a demand

of some kind, is to be susceptible to reactive attitudes when the expectation is violated (1994: 21). We are susceptible to a reactive attitude if we either feel this emotion or believe that it would be appropriate to feel it in these circumstances (1994: 23, 62). While these moves are generally consistent with Strawson’s original approach, Wallace aims to substantially modify and amend this approach by providing a narrower and more fine-grained account of moral reactive attitudes.

According to Wallace, Strawson’s account of reactive attitudes is too “inclusive” and needs to be refined into several different categories relating the various emotions we are concerned with (1994: 25–40). On Wallace’s analysis we need to draw two overlapping distinctions (1994: 33). The first is between moral and nonmoral reactive attitudes. Both forms of reactive attitudes depend on their reciprocal relationship with a system of expectations. In the case of moral reactive attitudes the relevant set of expectations are justified with reference to moral reasons, and the expectations they justify are obligations (1994: 35–6). However, not all expectations are backed by moral reasons, as we find in the case of etiquette, where a breach may generate resentment even though no distinctive moral claim has been violated (1994: 36–7). According to Wallace, the central cases of reactive attitudes are the emotions of resentment, indignation, and guilt (1994: 29–30), which are all negative in character. This is a feature of our reactive attitudes, he suggests, that explains their “special connection to the negative responses of blame and moral sanction” (1994: 62, 71).

Wallace draws out several significant points from this narrow account of the reactive attitudes. The first is that reactive attitudes are not coextensive with the emotions we feel towards people with whom we have interpersonal relationships, since reactive attitudes must be identified with reference to their “constitutive connections with expectations” (1994: 31). There are many interpersonal emotions we may experience—such as attachment, friendship, sympathy, love, and so on—that cannot be counted as reactive attitudes since they do not have any relevant connection with expectations. It follows from this that we must reject Strawson’s claim that a life without reactive attitudes would commit us to an (impossible) universal adoption of the “objective attitude.” People may continue to entertain various other forms of interpersonal emotions and **(p.189)** relations even in the *complete absence* of reactive attitudes narrowly conceived. Moreover, against Strawson, Wallace argues that even if some form of interpersonal relations are an inescapable feature of human life, it does not follow that the reactive attitudes are “similarly inevitable” (1994: 31). This point leads Wallace to his second distinction relating to his account of the reactive attitudes.

We also need to draw a distinction, Wallace says, between moral reactive attitudes and other kinds of moral sentiment. Not all moral sentiments take the form of moral reactive attitudes, with some identifiable tie to expectations and

obligations. Among the examples of nonreactive moral sentiments Wallace cites are shame, gratitude, and admiration; all of which involve different kinds of “modalities of moral value” (1994: 37–8). This distinction allows us to acknowledge that there are other cultures with forms of ethical life that do not have any commitment to reactive attitudes but may have other kinds of moral sentiment, as we find in shame cultures (1994: 31, 37–40). Evidently, then, on Wallace’s narrow interpretation of reactive attitudes, Strawson’s naturalism is excessive not only at the token level, but also at the type level, since it is entirely conceivable that there are cultures where members are not subject or liable to reactive attitudes at all. For Wallace, abandoning naturalism at both the token and type levels is a price well worth paying, as it is the only way to avoid an overly inclusive account of reactive attitudes and a false dichotomy between the reactive attitudes and the objective attitude. The narrow construal, Wallace maintains, is more faithful to the relevant psychological and sociological facts and also permits us to identify more accurately the justificatory issues that arise with respect to issues of moral responsibility.

Wallace is well aware that his narrow construal of the reactive attitudes commits him to an interpretation of moral responsibility understood entirely in terms of the conceptual resources of “the morality system.”² The morality system is understood as a particular form of ethical life, associated with our modern, Western, Christian culture. Its central normative concepts are obligation and blame, along with related notions of wrongness and voluntariness. These are all the same key elements that feature in Wallace’s narrow construal of moral responsibility. The narrow account renders moral responsibility, so interpreted, as a local and contingent cultural achievement, involving a legalistic, neo-Kantian view of morality. Understanding moral responsibility in narrow terms presents its adherents with their own set of difficulties. Some of these difficulties are anticipated in Wallace’s discussion.

(p.190) 3. The Costs Of Going Narrow

One of the most obvious costs of analyzing moral responsibility in terms of a narrow interpretation of our reactive attitudes is that it commits us to an “asymmetrical” account with respect to “worthy and unworthy actions” (Wallace 1994: 71). Since our moral reactive attitudes are, on the narrow interpretation, aroused only when expectations that we endorse are violated, it follows that any moral sentiments we experience that are positive responses to other “modalities of moral value” are not strictly reactive attitudes—and so no part of moral responsibility. While we may feel gratitude or admiration for a morally worthy act, the moral emotions involved are not to be accounted for in terms of the specific structure of beliefs about the violation of moral obligations (1994: 37–8). Wallace is unapologetic about this asymmetrical feature of the narrow view on the ground that it accounts for the “special connection” that exists between holding people responsible and our retributive practices involving blame and punishment (1994: 71). Beyond this, Wallace also claims that holding a person

responsible for a worthy action “does not seem presumptively connected to any positive emotion in particular” (1994: 71).

Wallace’s defense of the asymmetrical features of the narrow account of moral responsibility is unconvincing for several reasons. First, while we may agree that there is a close connection between moral reactive attitudes and our retributive practices, this does not imply that we need to interpret our reactive attitudes as *exclusively* negative in character, and as *always* connected with blame and moral sanctions. If we allow that there are reactive attitudes of a positive kind, based on beliefs concerning worthy actions and admirable character traits, then these too may be connected with positive retributive dispositions such as praise and rewards. Second, it is not obvious that our emotional resources with respect to our responses to morally worthy actions and traits are any more impoverished or limited than in the case of our negative reactive attitudes. As Wallace’s own observations suggest, we not only have “thin” ethical concepts such as approval and praise (correlates to disapproval and blame), we also have many “thicker” concepts, including gratitude and admiration. Finally, and most importantly, any asymmetrical account of the kind advanced by the narrow view inevitably truncates and distorts our experience of moral life and the various ways in which our reactive attitudes are grounded and directed. A one-sided view that is exclusively concerned with negative reactive attitudes, focusing entirely on their connection with blame and retribution, offers us an impoverished and unbalanced interpretation of responsibility and fails to properly accommodate the constructive role of reactive attitudes in endorsing and supporting morally worthy or admirable actions and traits.

(p.191) Another objection to the narrow interpretation, which Wallace also anticipates and tries to fend off, is that it presents an account of moral responsibility that has a “local” bias toward (our own) modern, Western Christianized culture—i.e. toward “the morality system.” It follows from the narrow account that the ancients Greeks and shame cultures, among others, have practices that are at best “analogous” to ours, based as they are on different moral beliefs with distinct patterns of moral response (1994: 65–6) Considered in terms of the narrow interpretation, these alien forms of ethical life do not share *our* understanding of moral responsibility, not simply in the sense that they have a *different* understanding but rather that they have *no commitment* to moral responsibility (since their ethical responses cannot be understood in terms of the narrow account of reactive attitudes). This view of things renders moral reactive attitudes and moral responsibility as both local and contingent, and thereby places a conceptual barrier between ourselves (modern, Western, etc.) and alien forms of ethical life that are removed from us in historical time and geographical space. From one point of view this narrow account *oversimplifies* our own (modern) attitudes and practices, which are not perfectly or purely represented by “the morality system” and evidently involve dimensions of holding people responsible that cannot be compressed into the

narrow and rigid framework provided. From another point of view, it denies us the *critical* apparatus and resources to question and challenge the way we (moderns) have (locally) arranged and structured our own attitudes and practices relating to responsibility. When we are confronted with other cultures and forms of ethical life outside “the morality system” they must, according to the narrow account, be set aside as—by definition—no longer possessing *any* conception of responsibility. Even if it is granted that we moderns are straightforwardly committed to “the morality system” and its narrow construal of responsibility, this still puts unnecessary and excessive (conceptual) distance between ourselves and these alternative forms of ethical life. More specifically, it erects a conceptual barrier to any genuine critical exchange and confrontation on the subject or question of responsibility itself—since there is, on this account, no shared or common ethical life with respect to the attitudes and practices that are actually constitutive of moral responsibility.³

(p.192) Finally, there is a further objection to the narrow construal of moral responsibility, which raises difficulties that Wallace does not anticipate or directly address. Granted that moral responsibility narrowly interpreted in terms of the concepts provided by the morality system is both local and contingent, it follows that we must reject Strawson’s original claims concerning our *type*-naturalist commitments to the reactive attitudes. If this is the case, then a fundamental plank of Strawson’s original naturalistic strategy must be abandoned: namely, the claim that we do not need and cannot provide any external rational justification for moral responsibility. If we accept the narrow construal, then there is no natural, universal liability to reactive attitudes and they do not serve as a natural foundation for all recognizably human forms of ethical life. Since the framework of moral responsibility is erected around culturally local forms of moral emotion, confrontations with other cultures and forms of ethical life will place us in the position of needing external rational justification for the entire framework of moral responsibility so conceived. Internal justifications, provided in terms of a rationale of excuses and exemptions, will not serve this purpose even if it is successful in fending off the (internal) skeptical challenge based on worries about determinism. To this extent, the skeptical threat remains with us not just at the token level but also at the type level. Nor is it an option to retreat back to other “analogous” forms of moral emotion since, if the narrow account is correct, this will not secure or preserve responsibility properly understood. The crucial point remains that if we embrace the narrow construal then, contrary to Strawson’s core original view, we are faced with the task of providing the whole edifice of moral responsibility with an external rational justification.

4. Type-Naturalism And Broad Responsibility

Whatever difficulties may be found for the narrow construal of moral responsibility it is important to begin with a full appreciation of Wallace’s critique of the original Strawsonian strategy, much of which is justified. There

are four particular features of Wallace’s critique that should be endorsed as clearly justified.

- (1) Token-naturalism is, as we have noted, psychologically implausible and fails, in any case, to discredit the justificatory issues advanced by **(p. 193)** the skeptical challenge (e.g. as based on worries about the implications of determinism).
- (2) We do require a more fine-grained and less inclusive account of the reactive attitudes. In particular, it is essential that we exclude interpersonal emotions that lack any relevant cognitive element containing ethical content, as in the case of emotions such as love, sympathy, friendly feeling, and so on.
- (3) It is also essential that any plausible naturalistic strategy is one that is historically or genealogically sensitive, displaying an awareness of the considerable variation in human ethical life and the range of moral emotions this may involve. In particular, we must avoid any crude form of naturalism that projects *our* (local) sentiments onto (alien) others.
- (4) Finally, it is entirely correct to argue, as Wallace does, that we need a theory of moral capacity to provide the basis for an account of exemptions, in order to answer the (internal) skeptical challenge that determinism would somehow render us all morally incapacitated and thus inappropriate targets of reactive attitudes. This last issue is, however, not itself part of the revised Strawsonian analysis of *holding* people responsible and is not, therefore, a matter for our present concern.⁴

Our concerns rest with the issues arising out of the first three items on the list above. In the case of all three of these items, I will argue, Wallace’s narrow view construal of responsibility involves a series of unnecessary and misleading oppositions. It is possible for us to avoid the weaknesses identified in Strawson’s original strategy without collapsing into the excessively narrow view of moral responsibility that reflects the meager resources of “the morality system.”

We may begin by noting that we can readily reject token-naturalism without rejecting type-naturalism. If we adopt this approach then, it is true, we will be denied the sort of easy way with skepticism that token-naturalism encourages. Moreover, as already explained, the type-naturalist approach, building on our natural liability to reactive attitudes, still leaves us needing a theory of excuses and exemptions that is consistent with compatibilist commitments, if we are to defeat the skeptical challenge. Although the skeptical effort to discredit all tokens of reactive attitudes is one that we must take seriously, if we accept type-naturalism, we do not **(p.194)** need any external rational justification for reactive attitudes (i.e. justification at the foundational level). Whether this further claim is acceptable or not will depend on whether we accept the narrow construal of our reactive attitudes. Clearly if we go narrow, then type-naturalism and the associated claim regarding the dispensability of external rational

justifications must be dropped. The resolution of this issue depends, therefore, on our interpretation of reactive attitudes.

As I have argued, although we do need a narrower account of our reactive attitudes, we need to make sure we do not go too narrow, as otherwise we will generate some of the difficulties that have already been noted. Wallace’s narrow view places considerable and appropriate emphasis on the “propositional content” involved in the beliefs that serve to delineate our reactive attitudes (1994: 11,19,74).⁵ The narrow view would restrict the contents in question to the limited range provided by the conceptual resources of “the morality system.” It is these limits that result in the problems of asymmetry and localism, as we have described them. We need to find, therefore, a *middle path* that avoids the inclusiveness of Strawson, on one side, and the excessively narrow approach of Wallace on the other. To put ourselves back on the right track we may turn again to Bernard Williams’s critique of “the morality system.”

The narrow view, as we have seen, presents ethical considerations in highly restricted terms, specifically with reference to obligation and blame, which is appropriate when obligations are violated. Williams identifies these features as central to the morality system (1985: ch. 10). This tendency to reduce and simplify is also manifest in ethical theory, a philosophical project which is itself intimately linked to the assumptions and prejudices of the morality system (1985: ch. 1). One aim of ethical theory is to provide an account of morality that will provide an exact boundary between ethical and nonethical considerations. This is done primarily by reducing the diversity of ethical (and nonethical) considerations, with a view to identifying a narrow and strict range of ethical considerations that may serve as moral reasons available to all rational agents—the universal constituency. The most notable features of moral theory are its simplicity, reductionism, and systemization of our ethical concepts and claims. Williams’s critique of the morality system involves challenging and rejecting these assumptions. In the first place, while our ethical considerations certainly include obligations, under some **(p.195)** interpretation, they extend well beyond this. The scope of the ethical relates more broadly to “the demands, needs, claims, desires, and, generally the lives of other people, and it is helpful to preserve this conception in what we are prepared to call an ethical consideration” (1985: 12; see also 1985: 153, where Williams mentions our need to share a social world in relation to these various ethical considerations). What is required, from this perspective, is an account of ethical considerations that also includes forward-looking concerns relating to welfarism and utilitarianism, as well as ethical considerations that relate to our ideals and self-conceptions that mark out actions that fail the standards and boundaries that we may set for ourselves (e.g. considerations of what we regard as demeaning, base, dishonorable, etc.). When we interpret ethical considerations in this *broader* manner we find that these interests are plural and varied in their nature and secure no sharp boundary between ethical and nonethical considerations.

Vagueness, conflict, and diversity—contrary to the demands of “theory” and the prejudices of “the morality system”—are of the essence of human ethical life.

Our own ethical reactive attitudes must be understood in these broader and vaguer terms. One of the implications that Williams draws from this is that we should be skeptical of the effort to understand reactive attitudes in the reductive, thin language of binary judgments; approval and disapproval, guilt and innocence, and so on (1985: 37, 177, 192). If we reconfigure our ethical reactive attitudes in terms of a broader construal of the ethical considerations that ground them and serve as their propositional content, then we may acquire a very different understanding of the scope and content of moral responsibility, as based upon these emotions. Our ethical qualities are manifest in the “deliberative priority” and “importance” that we give to ethical considerations as expressed in our conduct and character. So interpreted, ethical reactive attitudes may be construed as *reactive ethical value*, where this is understood as emotional responses to the weight and value given by an agent or person to ethical considerations widely conceived (i.e. in terms of our human needs, interests, welfare, claims, and the requirements of social cooperation). Ethical reactive attitudes involve coming to see a person in a certain ethical “light” based on these lower-order evaluations of their ethical qualities. Clearly we have varied and diverse ethical norms and standards that serve as the relevant basis for evaluating an agent’s ethical qualities understood in these terms. These evaluations of agents based on their ethical qualities serve to generate or arouse a myriad of ethical reactive attitudes which may be either “positive” or “negative” in nature. As Williams argues, our ethical and emotional language here is not at all “thin” (e.g. praise and blame etc.) but is “thick” and varied, involving notions such as “being creepy,” or a “cad,” and so on—all of which are responses loaded with ethical significance. **(p.196)** Different cultures and different forms of ethical life will not only have different lower-order ethical norms, they also deploy a different or variable set of ethical reactive attitudes (reflecting their variable propositional content).

The significance of this criticism of “the morality system,” along with the style of ethical theory associated with it, for our understanding of ethical reactive attitudes should be clear. The revised account is broad enough to accommodate positive ethical reactive attitudes (e.g. gratitude, admiration, etc.) as well as “alien” reactive attitudes (e.g. shame) all under the umbrella of those ethical considerations that serve to ground or justify them. This avoids the costs of going too narrow, by way of relying on the limited and restricted resources of the morality system. The broader construal is, nevertheless, controlled and focused enough to exclude elements that do not relate at all to ethical considerations and ethical qualities (e.g. friendly feeling, sympathy, romantic love, etc. do not count as ethical reactive attitudes because they are not reactive to *ethical* qualities as such). On this analysis, we should not be surprised or disappointed to find that there is no sharp or clear boundary between ethical

reactive attitudes and other emotional responses to qualities and features of those with whom we are dealing. No such sharp boundary should be expected if we want an accurate understanding of the nature of ethical life and the way in which it “bleeds” into human life in general.⁶

Taking this broader approach to ethical reactive attitudes has other significant advantages as well. It avoids, for example, the “legalism” of the narrow account, which turns moral responsibility into a model of legal responsibility—eliminating the more nuanced and complex set of responses we have outside legal contexts. We also avoid the failings of what Williams refers to as “progressivism,” the assumption that we moderns alone have access to a full and complete concept of moral responsibility and are “better off” than those who lack our own understanding.⁷ It is Williams’s view that not only should we be open to the possibility that we might learn from the ancients, this is in fact our situation. Learning from the ancients is possible—and desirable—precisely because we *share* a concept of moral responsibility with them, however differently we may interpret various key elements associated with it (1993: 55).

For our present purposes, our concern is not to present a worked-out alternative to the narrow model of reactive attitudes—not the least because, **(p.197)** for reasons given, this may itself be a problematic ambition driven by the aims of “ethical theory.” What is important, however, is to insist on finding some middle ground that can accommodate ethical reactive attitudes broadly conceived without expanding this set to include interpersonal emotions that have no relevant ethical content (i.e. which do not involve our emotional reactions to a person’s ethical qualities). Wallace’s own observations suggest that this can readily be done, since he allows “analogous” forms of responsibility and also speaks of “responsibility for worthy acts” (Wallace 1994: 38–40, 64–6, 71). To see, in a particular case, how this middle ground between an excessively narrow and overly inclusive view may be found let us consider *shame*. Shame may be based on standards and norms that have no ethical content, as in the case of concern about one’s physical appearance (e.g. my frail constitution) or economic status (e.g. my family’s poverty). In other cases, however, the relevant standards and norms may move into the territory of our ethical qualities and characteristics, such as feeling shame about being lazy or being vulgar. Whether a response is an ethical reactive attitude or not will depend on the *nature* of the quality or consideration it is a *reaction to*. There are, moreover, clear cases of ethical shame that cannot be analyzed or understood in terms of the apparatus of obligation and doing wrong. We may, for example, feel ashamed of failing to live up to our own ethical ideals and standards, even when we are well aware that we have not failed to comply with any obligations and cannot be blamed for our conduct. An example of this is provided in Joseph Conrad’s *Lord Jim*, where Jim is ashamed of himself because he fails to act heroically and is, therefore, disappointed in himself in these ethical terms but not in terms of any recognizable requirements of “the morality system.”⁸ The general point here is

that there is a wide range of ethical reactive attitudes lying outside the theoretical schema of the narrow interpretation that, nevertheless, do not collapse into an overly “inclusive” set of interpersonal emotions lacking any relevant ethical content. Some cases of shame will be cases of ethical reactive attitudes and others will not. What will settle this issue will be the specific content and target of what we are ashamed of. Moreover, the fact that there are no sharp or precise boundaries to draw here is a failing only if we assume the prejudices of the morality system and the forms of “theorizing” associated with it.

We have noted that it is essential that the naturalist approach to moral responsibility should be sensitive to historical and cultural variations with regard to our understanding of moral responsibility and the specific and **(p. 198)** various forms which our ethical reactive attitudes may take. It should be pointed out that Strawson is himself alive to these concerns. Speaking of our increased “historical and anthropological awareness of the great variety of forms which these human attitudes may take at different times and in different cultures” Strawson says:

This makes one rightly chary of claiming as essential features of the concept of morality in general, forms of these attitudes which may have a local and temporary prominence. No doubt to some extent my own descriptions of human attitudes have reflected local and temporary features of our own culture. But an awareness of variety of forms should not prevent us from acknowledging also that in the absence of *any* forms of these attitudes it is doubtful whether *we* should have anything that we could find intelligible as a system of human relationships, as human society. (1962: 24–5. Strawson’s emphasis)

It may be argued, along the lines of Wallace’s criticisms of Strawson’s claims about the objective attitude, that Strawson’s remarks in this passage run together two distinct issues. One claim is that we could not recognize a society as truly human without any ethical reactive attitudes. Clearly this need not be the case, so long as we do not overly expand the class of ethical reactive attitudes to include all interpersonal emotions. Nevertheless, the general point that Strawson is primarily concerned to make in this context still stands: namely, without *some form* of ethical reactive attitudes we could not recognize or find intelligible a system of human relationships that would qualify as a *human ethical life*. In other words, an ethical life devoid of all forms of ethical reactive attitudes is not recognizable or intelligible *to us* as a form of *human ethical life*.

It should be clear, in light of Strawson’s observations, that we do not need to choose between naturalism and genealogy, where this is understood in terms of sensitivity to historical and cultural variation and diversity. On the broad construal, ethical norms and the ethical considerations to which they give

weight, may vary greatly from one culture and historical period to another. With these variations we will also find variations in the particular forms of ethical reactive attitudes that are adopted and endorsed. One form these ethical reactive attitudes may take is the narrow form encouraged by the morality system—which makes obligation and blame its central features. While this form of ethical reactive attitude may be local and contingent, it does not follow that ethical reactive attitudes *broadly construed* are local and contingent, unless we take the local form to be the sole legitimate representative form of moral responsibility. Since the broad construal neither interprets ethical reactive attitudes nor moral responsibility in these restrictive terms, it is able to acknowledge that the ancient Greeks, among others, have ethical reactive attitudes that are recognizably **(p.199)** continuous with our own conception of moral responsibility. This can be done without in any way denying that there are significant differences between their culture and our own with regard to the attitudes and practices involved (e.g. with respect to issues of voluntariness and intention). The irony about this situation is that it is the narrow construal, which rejects (type) naturalism, which cannot accommodate genealogical sensitivity to cultural and historical variation. Given that the narrow construal insists that moral responsibility be understood in terms of the concepts of the morality system, and that this interpretation alone constitutes genuine or real moral responsibility, it is compelled to exclude all other understandings (i.e. as based on a broader construal of ethical reactive attitudes) as falling outside the parameters of moral responsibility. It is, therefore, the narrow construal that is insensitive to genealogical variation and diversity as it arises within the framework of moral responsibility. For reasons that have been explained, this is not simply a verbal issue, as it involves and encourages a truncated and distorted understanding of moral responsibility and the way in which it is naturally rooted in human ethical life.

It has been shown that “asymmetry” and “localism” are unnecessary and unacceptable costs of the narrow approach. What, then, about the further issue relating to type-naturalism and external rational justifications? If we adopt the broad construal then no external rational justification of our liability to ethical reactive attitudes is required, where ambitions of this kind involve what Strawson describes as the tendency to “over-intellectualize the facts” concerning the natural foundations of moral responsibility (Strawson 1962: 23). In contrast with this, the narrow construal needs to provide some relevant external rational justification since, *per hypothesis*, it is a local (modern, Western) achievement. Clearly the broad view can accept that the *local forms* of ethical reactive attitudes (e.g. as based on the requirements of the morality system) may come and go. Considered from this perspective, the morality system and its associated narrow interpretation of moral responsibility may be judged vulnerable to extinction for several related reasons. First, as already argued, it suggests a truncated and distorted account of our own ethical concerns (i.e.

even from a modern, Western perspective). Second, it generates (unnecessary) problems of asymmetry and localism as we have described them. Third, the narrow account is also especially vulnerable to internal skeptical collapse due to worries about determinism. (Although Wallace believes these internal skeptical objections can be defeated, not all those who endorse the morality system believe this can be done, much less done on the basis of compatibilist commitments.) It follows from all this that there is a real prospect of this *local* form of ethical reactive attitudes collapsing under both internal and external skeptical pressure. What does not follow from this, **(p.200)** however, on the broad construal, is the total collapse of moral responsibility in any recognizable form. We may still have available other forms of ethical reactive attitudes that are not similarly vulnerable in any of these dimensions. Clearly it would be a mistake, therefore, to present the collapse of the local form of moral responsibility associated with the morality system as putting us in the predicament of having to adopt some form of utilitarian, forward-looking approach that has no place for ethical reactive attitudes *of any kind*. Alternative forms of reactive attitudes remain available and viable within the structure of ethical life that is still recognizably *human* and intelligible to us.⁹

The question we must now turn to is what is the relationship between skepticism and type-naturalism as understood on the broad approach? Skepticism may take the form of aiming to discredit *local* understandings of our ethical reactive attitudes. This does not, as has been argued, show that all forms of ethical reactive attitude are thereby discredited or that moral responsibility, as such, cannot be vindicated. It remains true, nevertheless, that whatever local forms our ethical reactive attitudes may take, they (all) remain vulnerable, in principle, to internal, global skeptical challenge at the *token* level. That is to say, our commitment at the type level, even on a broad construal, does not secure any general immunity from potential global skepticism with respect to *all* tokens of our ethical reactive attitudes.¹⁰ The crucial point that needs to be emphasized, however, is that even in these circumstances, the skeptic remains committed to ethical reactive attitudes at the *type* level (unless, of course, the capacity to feel and experience ethical reactive attitudes is itself damaged). In other words, although the skeptic may systematically disengage from all *tokens* of ethical reactive attitude from “the inside,” she cannot abandon the propensity or liability to these attitudes. That is a project that, from one point of view, would require radical intervention with her own nature (e.g. by way of genetic engineering or medical surgery) **(p.201)** and, from another point of view, would place her outside the recognizable human ethical community.¹¹

We have already noted that whereas the narrow construal would place the ancient Greeks (and other shame cultures) outside the framework and fabric of moral responsibility, the broad construal does not. What, then, about the skeptic? In contrast with individuals who are engaged participants in forms of ethical life involving ethical reactive attitudes, the skeptic has systematically

disengaged from all such participation or involvement. Disengagement of this kind requires (internal) doubts about the justification for *any* proposed tokens of ethical reactive attitude. So described, the skeptic cannot evade the challenge of providing some account of the excusing and exempting considerations that apply universally in such a manner that all *tokens* of ethical reactive attitude are discredited. If the skeptic fails or refuses to provide any such rationale for her (disengaged) stance then her skeptical stance has not been vindicated or justified. Contrast the skeptic with another distinct character, who we may call the “Vulcan.”¹² Vulcans are understood to be entirely rational but incapable of human emotion. As such, Vulcans may rationally understand (human) ethical norms but are incapable of feeling or entertaining ethical reactive attitudes (or similar moral emotions with an attitudinal aspect). Vulcans have, in other words, no type-naturalist commitments with regard to ethical reactive attitudes. It is, for this reason, a mistake to assimilate the skeptic to a Vulcan, as plainly the skeptic is not a Vulcan. The Vulcan faces no skeptical problem with respect to ethical reactive attitudes. They have no token commitment because they have no type commitment to this range of (ethical) emotion. For the (human) skeptic, however, the skeptical challenge is *real* because their type propensities require something to be said with respect to disengaging all tokens of these reactions to the ethical qualities of others in their community. In this way, both the skeptic and the anti-skeptic, in contrast with the Vulcan, can accept Strawson’s type-naturalism and dispense with the search for external rational justifications. What divides them is the issue of whether or not a theory of excuses and exemptions can handle relevant internal skeptical worries (e.g. as based on the implications of determinism).

(p.202) 5. Against The Narrow Construal Of Moral Responsibility

It has been argued that the narrow construal of reactive attitudes and its associated account of moral responsibility has unacceptable costs. While it is true that there are significant failings in Strawson’s original naturalistic project that need to be addressed and corrected (e.g. we should reject token-naturalism), we should retain the core feature of type-naturalism. In order to do this we need to provide a *broader* account of ethical reactive attitudes that extends beyond the constraints and limits of “the morality system” and its conceptual structures. It is only by taking this route that the difficulties we have described relating to “asymmetry” and “localism,” as well as the fruitless and misguided search for external rational justifications, can be avoided. We may summarize the significance of these observations in the following points.

(1) The narrow construal of moral responsibility, as developed on the basis of the morality system, both distorts and truncates our understanding of human ethical life as it relates to moral responsibility. In particular, it makes it impossible to accommodate both positive ethical reactive attitudes and alien ethical reactive attitudes as they may arise from outside our (modern, Western) culture. Even

our own *local* understanding of moral responsibility is not fully or adequately captured by this narrow construal.

(2) It is the broad construal, along with its commitment to type-naturalism, which is able to accommodate genealogical sensitivity to historical and cultural variation in relation to our understanding of moral responsibility. The narrow construal excludes all alternative forms that do not fall into the constraints imposed by “the morality system” as mere analogues or prototypes of moral responsibility. As such, the narrow construal constitutes a form of conceptual imperialism with regard to (real, true) moral responsibility and also commits us to an implausible “progressivism” concerning our own (modern, Western) views. In contrast with this, the broad approach recognizes the variation in modes and forms of ethical reactive attitude within a wider understanding and appreciation of the emotional fabric of moral responsibility.

(3) Type-naturalism, as understood on the broad construal, provides no easy way of dealing with a potential internal skeptical challenge (i.e. in contrast with the aims of token-naturalism). Even allowing for our natural liability to ethical reactive attitudes, on a broad construal, we must still formulate some relevant schema of excuses and exemptions. From this perspective it is always conceivable that a systematic or global skepticism (**p.203**) could be generated from “the inside” (i.e. extending to all our token ethical reactive attitudes). This possibility does not, however, license a search for external rational justifications, since our liability or propensity to such emotions is natural and not rationally grounded. The skeptic remains committed to ethical reactive attitudes at this level, even if she has entirely abandoned or disengaged any commitment to tokens of these attitudes (in light of internal skeptical pressures of some kind).

(4) Much of the motivation behind Wallace’s narrow construal of the reactive attitudes is to find a satisfactory compatibilist account of moral responsibility consistent with the core requirements and constraints of the morality system. From this perspective the internal skeptical challenge is especially acute, since it is targeted on the notions of wrongness, blame, desert, and retribution that are central to moral responsibility as the morality system interprets it. It is evident, however, that the broad construal of ethical reactive attitudes, along the lines that has been sketched, significantly *deflates* these (internal) skeptical pressures. The reason for this is that a broader and more liberal conception of ethical reactive attitudes does not place such heavy weight or emphasis on the very elements of the morality system that have proved especially vulnerable to skeptical criticism (i.e. desert, blame, etc., along with their apparent dependence on ultimate or absolute agency). Even if—contrary to what Wallace argues—it proves impossible to vindicate this local interpretation of moral responsibility, as understood on the narrow construal, it does not follow, given a broad interpretation of ethical reactive attitudes, that global skepticism results. All that follows from the success of the skeptical challenge, so described, is that

the *local* understanding of moral responsibility encouraged by the morality system cannot survive critical reflection.¹³

(p.204) In sum, we may contrast the relative strengths and weaknesses of the broad and narrow accounts in these terms. The narrow construal not only generates a partial and incomplete account of moral responsibility, it also leaves the entire edifice of moral responsibility, so understood, vulnerable to both internal and external skeptical threat. The broad construal not only avoids the significant difficulties that the narrow construal encounters (e.g. asymmetry), it provides for the complexity, variation and nuance that we find in this sphere. Moreover, the broad construal, by moving away from the rigidities and (peculiar) demands of the morality system, deflates the internal skeptical threat and eliminates all worries relating to the misguided ambition of providing a satisfactory external rational justification. These are fundamental points relating to moral responsibility and the defects of the morality system that the discussions of both Strawson and Williams converge on.

References

Bibliography references:

Doris, John (2002). *Lack of Character: Personality and Moral Behavior*. (Cambridge & New York: Cambridge University Press).

Russell, Paul (1992). “Strawson’s Way of Naturalizing Responsibility.” *Ethics* 102. 2.

— (2004). “Responsibility and the Condition of Moral Sense.” *Philosophical Topics*, 32, 1; 2, 287–305.

— (2011). “Moral Sense and the Foundations of Responsibility.” In *Free Will*, 2nd edn., ed. Robert Kane (New York: Oxford University Press).

Strawson, P. F. (1962). “Freedom and Resentment,” reprinted in P. F. Strawson, *Freedom and Resentment and other essays*. (London, New York and Oxford: Methuen. 1974).

Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*. (Cambridge, MA: Harvard University Press).

Williams, Bernard (1985). *Ethics and the Limits of Philosophy*. (London: Fontana).

— (1993). *Shame and Necessity*. (Berkeley, CA: University of California Press.).

Notes:

(*) I am grateful to David Shoemaker and other participants at the New Orleans Workshop on Agency and Responsibility (November 2011) for their very helpful comments and criticism. Versions of this paper were also presented at a Workshop on Moral Agency, Deliberative Awareness, and Conscious Control (Erasmus University, Rotterdam, October 2010) and at the University of Arizona (February 2012). Thanks also to those who were present on those occasions for their comments and suggestions, and especially to Michael McKenna, Maureen Sie, and Andras Sziget.

⁽¹⁾ For a more detailed development of this analysis of Strawson’s naturalism see Russell (1992).

⁽²⁾ Wallace 1994: 39–40, 64–66; and Williams 1985: esp. ch. 10.

⁽³⁾ Clearly confrontation between ethical cultures that are removed from each other in historical time or geographical space may be, as Williams observes, either “notional” or “real” (Williams 1985: ch. 9). Be this as it may, historical sensitivity about the contingency of our own “local” commitments naturally puts pressure on reflective confidence in our own attitudes and practices. To this extent our *awareness* of other modes and forms of ethical life, less attached to the rigidities of “the morality system,” may bring us to question whether our confidence in “our *modern* concept of responsibility” is altogether well-founded.

⁽⁴⁾ I have argued elsewhere that there is a more intimate relationship between our capacity for holding agents responsible and our capacity for reflective self-control than (Kantian) theories such as Wallace’s acknowledge. See, in particular, Russell 2004 and Russell 2011.

⁽⁵⁾ Wallace claims that Strawson’s account of reactive attitudes does not manage to clearly connect them with any propositional content (1994: 39). This charge seems unfair to Strawson since he is careful to ground reactive attitudes in our beliefs about the attitudes and intentions of other human beings and “the very great importance” that we attach to them (Strawson 1962: 5).

⁽⁶⁾ This is, of course, a recurrent theme throughout Williams’s writings.

⁽⁷⁾ Williams 1985: 32 n. 2; and, more generally, Williams 1993: esp. ch. 1. Williams’s view is, of course, the opposite of this, since he holds that we would be better off without the morality system (1985: 174), just as we don’t need ethical theory and should abandon its aims (1985: 17, 74).

⁽⁸⁾ For an illuminating discussion of this example see Doris 2002: 160–4.

⁽⁹⁾ Wallace refers to the utilitarian approach as “the economy of threats” model (1994: 54–61). It is crucial to his critique of this model that it lacks “depth,” where depth is provided by the “attitudinal” features of blame and retribution

(1994: 56, 75). On a broader construal, however, “depth” can be found in other forms of ethical reactive attitude, such as shame and anger—a point that Wallace comes close to endorsing in some passages. See, e.g. 1994: 89.

(¹⁰) This may well be regarded as highly unlikely or even incredible—but it is not inconceivable. Imagine, for example, the spread of some terrible disease or genetic mutation that affected us all by damaging our most basic and universal moral capacities in such a manner that exemptions applied universally (however broadly interpreted).

(¹¹) An individual who lacks any *type* commitment to ethical reactive attitudes would not be recognizably human, not because she is a systematic skeptic with respect to these attitudes but because the skeptical issue *does not arise for her* with respect to these attitudes, since she is constitutionally incapable of experiencing or entertaining such attitudes.

(¹²) Vulcans are aliens from the planet Vulcan, as described in *Star Trek*. The character of Mr Spock was half Vulcan and half human. Wallace refers to this example in a related context at Wallace 1994: 78 n.41. See also Russell 2011: esp. 212–14.

(¹³) It is true, of course, that many skeptics about moral responsibility are concerned to discredit the local conceptions of moral responsibility associated with the morality system. However, for reasons that have been discussed, skepticism of this kind does not in itself constitute global skepticism—since it does not discredit, and may not even aim to discredit, alternative forms of ethical reactive attitudes. Having said this, it is important to note that many skeptical projects of this kind either explicitly or tacitly endorse the narrow construal and its assumption that alternative accounts of ethical reactive attitudes somehow fail the standard of *real or genuine* forms of moral responsibility. When this assumption is made, the critique of our local conception of moral responsibility framed in terms of the requirements of the morality system is (mistakenly) *inflated* into a form of *global* skepticism about moral responsibility. Suffice it to say that much of the contemporary free-will debate, along with its associated worries about the skeptical threat, proceeds on this assumption of the narrow construal and the morality system.