Filosofia Unisinos
Unisinos Journal of Philosophy
17(2):144-148, may/aug 2016
Unisinos – doi: 10.4013/fsu.2016.172.08

PHILOSOPHY SOUTH

Sympathy and resentment

Alejandro Rosas[1]

ABSTRACT

I defend resentment as a legitimate and necessary moral attitude by neutralizing an objection that points to its hostile and morally repugnant character. The argument proceeds by embedding resentment in a view of morality as a social and communicative practice, supported by a common knowledge of apparently inborn moral expectations. In virtue of these natural expectations, every person is pre-institutionally entitled to goodwill and to the pleasure that arises from showing and receiving goodwill from others. Resentment arises naturally when these expectations are violated: it is a reaction to an offence, leading to a dynamic exchange that aims to restore a broken moral relationship between persons. The offender participates in this communicative exchange by undergoing a form of punishment that is self-imposed and expressed through feelings of guilt and contrition.

Keywords: blame, compatibilism, goodwill, moral reparation, reactive attitudes, sympathy.

Introduction: Reactive attitudes and moral reparation

Compatibilism is the view that an agent can act freely and be held responsible even if her actions are determined by natural (including psychological) causes. One difficulty with this view is that the social practice of blaming and holding responsible requires that the agent could have done otherwise; but natural causation seems to preclude this. A standard compatibilist response is that the agent could have done otherwise if she had considered other reasons; and she can be blamed for not considering them. Against the obvious regress in the argument (e.g., determinism precludes that the agent could have considered those other reasons), the compatibilist must quickly remark that blaming and holding responsible ultimately function as ways of influencing the character of the blamed person, such that in similar circumstances in the *future* the agent will act differently (Nowell-Smith, 1948; Smart, 1961). The practices of moral evaluation, of holding someone responsible, of blaming and even punishing make sense because they instantiate forward-looking social-psychological processes, which ultimately operate to adjust characters to moral expectations.

Strawson (2008 [1974]) offers a broadly phenomenological argument to the effect that the future-oriented tenor of these practices, as interpreted in classical compatibilism, distorts morality if it is not complemented by the backward-looking dimension of reactive and self-reactive attitudes. He presents the argument as part of a refurbished compatibilist alternative. Before Strawson's paper, compatibilism reduced blaming and holding responsible to impersonal practices aiming at social control in a purely forward-looking way, without the reactive attitudes

[1] Philosophy Department, National University of Colombia, Sede Bogotá. Kra. 30 # 45-03, Ciudad Universitaria, Bogotá D.C., Colombia.
E-mail: arosasl@unal.edu.co

(Nowell-Smith, 1948; Smart, 1961). The function of such practices was to deter future wrongdoing and to promote the general welfare of society. Thoughts of desert as a justification of punishment were to be eradicated, and not only from folk-views regarding responsibility and blame; they should also disappear from our legal practices, which should be solely organized around deterrence (Greene and Cohen, 2004). Strawson was not happy with this view. He observed that it completely ignored vital psychological dimensions of doing and suffering wrong. I believe that he was right, for something crucial about morality falls out of the consequentialist forward-looking reading. Strawson located the missing elements in the backward-looking feature of reactive and self-reactive attitudes like resentment, indignation, gratitude, guilt, and remorse. In the following, I argue that the importance of reactive and self-reactive attitudes is that they combine both the backward and the forward-looking dimensions of moral practice in a communicative dynamics that is lost in the narrow consequentialist attitude.

Strawson was not invoking the common retributivist view. In fact, a common brand of retributivism serves as an ally to the narrow consequentialism that Strawson was criticising. The alliance between retributivism and consequentialism (Goodenough, 2004) focuses restrictively on violent punishments, namely inflicting one of the following on offenders: physical pain and/or loss of property, limb, freedom, or life. The difference between them is that retributivism inflexibly prescribes such punishments, while consequentialism is tempted to withdraw them if they do not guarantee deterrence. But these punishments severely narrow the idea of moral reparation (Bennett, 2002; Holroyd, 2010). Violent punishment, though it may give some satisfaction to victims, does not really repair or restore anything. I want to distinguish it categorically from a different kind of punishment that could be labeled restorative, in reference to the concept of restorative justice (Johnston, 1999). But I am not fixed on this label. Perhaps a neutral anthropological one, like 'ancestral' punishment, would be better. One example of this type of punishment is a group-level withdrawal of goodwill directed at the offender. This phenomenon is known in anthropology as ostracism. It is the preferred form of punishment among hunters and gatherers and similar egalitarian societies (Boehm, 1999). It is severe, because it confronts the offender with the loss of every form of communication and hence of the cooperative relationships previously entertained with members of her group. Another example is the self-imposed punishment usually expressed in feelings of guilt and contrition. These two kinds of punishment must have been common among our ancestors (Trivers, 1971). In contrast, evidence suggests that violent punishment, typical of retributivism and consequentialism, is virtually inexistent among hunters and gatherers (Boehm, 1999).

In this paper I am solely concerned with self-imposed punishment expressed in feelings of guilt and contrition. Its importance lies in the fact that it combines the backward- and forward-looking dimensions: it looks back at the wrong done and acknowledges it, but it also looks forward to the restoration of the moral link with the victim, only possible in virtue of its backward-looking feature. Linking resentment and other hostile reactive attitudes to this self-imposed punishment, whereby the offender restores the moral link to the victim and the community, provides a reply to the moral criticisms leveled against the "hostility" of such reactive attitudes (Pereboom, 2008; Kelly, 2008; Nussbaum, 2015). The clue lies in interpreting resentment as including a demand directed at the offender to undergo the self-imposed pain that arises from acknowledging one's wrong actions. It is part of a communicative exchange between victim and offender (Nozick, 1981; Bennett, 2002; McGeer, 2012). Reactive and self-reactive attitudes are usually the vehicles of this exchange: resentment in the offended party, followed by guilt and contrition in the offender. I should here insert a caveat: I use "resentment" to refer to feelings like indignation and moral anger, which I believe was Strawson's meaning. I do not use it to refer to the slow-burning emotion that perhaps would be more properly captured with terms like "rancour" and "grudge".

Typical of the narrow consequentialist understanding of freedom and responsibility is the view that the kind of moral reparation that happens in a communicative-emotional exchange between victim and offender is a pure delusion, because once a wrong is done, it cannot be undone. But discarding this exchange amounts to shedding reactive and self-reactive attitudes—which are its vehicles. Strawson believes that this is equivalent to dumping morality altogether. It would reduce morality to deterrence, while reducing deterrence, in turn, to violent punishment, i.e., the infliction of externally imposed pain or some kind of material loss on the offender. Moral motives would thus be reduced to a selfish deliberation or calculation, by means of which would-be offenders weigh the probability of a rather heavy loss as a consequence of their misdeeds, against the selfish benefits intended through those very misdeeds. Without moral reparation, moral demands are nothing but the manipulation of others through threats (and sometimes incentives), whereby addressees only need the capacity to reason in an egoistic way. No role is left for the communicative process of moral reparation that begins with resentment in the offended party, followed by the offender's acknowledgment to having done wrong. It seems that a consequentialist view of punishment (and its common retributivist ally) "treats everyone as outside the moral community", as Nozick once said (Nozick, 1981, p. 372).

A psychological ritual of restoration is built into our reactivity and self-reactivity to moral wrongdoing (Walker, 2006). In moral reparation, the parties undergo a ritual, as it were, through reactive and self-reactive attitudes: they look back at the wrong done and aim forward at restoring the moral link broken by the offence. Resentment proclaims that a wrong has been committed. The offender may respond in sympathy, by sharing the *evaluative* perspective of the offended party, who condemns the action. In the self-reactive attitude known as contrition, the offender feels pain at having wronged the offended party. The pain should be felt on ac-

count of the wrong done, not just in the form of a 'contagion' that comes from observing the victim's pain. Strawson does not go into such detail, but he does mention the complementarity between reactive and self-reactive attitudes: the readiness of the offended party to punish corresponds to a readiness in the offender to accept punishment (Strawson, 2008 [1974], p. 23-24). In what follows, I shall explain resentment's role in moral repair. What needs repair is a state of mutual goodwill: a communicative and public intersubjective state that serves as the foundation of morality, akin to mutual sympathy in Smith's *A Theory of Moral Sentiments*.

The origin of resentment in three steps

Strawson mentions that reactive attitudes arise from mutual expectations of goodwill (Strawson, 2008 [1974], p. 5). Reactive attitudes are part of our nature, and the same applies to the mutual expectations of goodwill that support them. These natural attitudes and expectations justify pre-institutional attributions of entitlements and desert (Scanlon, 2013). I therefore interpret these expectations as containing a foundational moral principle, which gives resentment and other reactive attitudes their moral content. I cannot say whether Strawson would have agreed to this interpretation or not, but I will let it stand, because it allows me to build a bridge to Adam Smith's theory of morality based on sympathy; a natural ally, I believe, of Strawson's views. We can explain resentment—and in a sense justify it—by telling a plausible story about how it arises from mutual expectations of goodwill. We are apparently born with these expectations, for infants also seem to have them, as shown by empirical research (Bloom, 2013; Baillargeon *et al.*, 2015). The story develops in three steps. (i) The first thing is to realize that expectations of goodwill, insofar as they are both shared and public, involve higher-order intentionality and are, therefore, communicative. (ii) They communicate a basic normative stance, namely, that everyone participating in this shared space is entitled to goodwill and to the pleasure arising from showing and receiving goodwill. (iii) A violation of these expectations for selfish reasons denies moral status to the victim; it causes resentment, which is a combination of pain, disapproval, and a demand to re-establish mutual goodwill.

The three steps unpacked

Communication of mutual expectations

Expectations of goodwill are mutual and also shared. The concept of shared expectations is similar to the concept of shared intentions (Tomasello, 2009). It implies common knowledge and higher-order intentionality, i.e., agents have common knowledge about the fact that they have such expectations with regard to one another. This would usually be expressed by saying that agents know that they know that they know (recursively) that they have these expectations with regard to one another. I do not believe that common knowledge requires an infinite recursion of higher-order intentional states, but the point is unsubstantial here. In any case, common knowledge gives mutual expectations the quality of being communicative, or rather, of being the basis for a communicative dynamics, because common knowledge places the content of the expectation in a shared, public domain, as if it had been publicly proclaimed. Common knowledge of mutual expectations of goodwill is constitutive of the *public* sphere and of the entitlement to goodwill of anyone who participates in this public sphere.

The basic norm: Everyone is entitled to goodwill

What is the content of the communication embedded in mutual expectations of goodwill? It will help to compare goodwill to sympathy in Adam Smith's moral theory. In Smith, sympathy has both cognitive and evaluative components. I use it here to refer primarily to the evaluative element. Through sympathy we confer value on the wellbeing of others, similar to the value that each claims, at least implicitly, for her own person. This statement describes what sympathy does to our motivational system. But once we acquire this motive, and in virtue of common knowledge, we enter a public sphere in which agents expect mutual respect from all participants. The propositional content communicated is a generalized demand for mutual goodwill and its corresponding generalized endorsement. When we confer, in sympathy, value on the wellbeing of others, we acknowledge the entitlement we both have to equal consideration. We give each other this entitlement naturally, precisely in mutual goodwill or mutual sympathy. We can withdraw it, normally in reaction to an offence, for an offence implicitly declares a prior withdrawal of sympathy by the offender. We give each other mutual sympathy by nature, I assume with Smith, perhaps through an adaptive social instinct mediated by memory and self-awareness (Darwin, 1981 [1871], Chapter 3).

The comparison with sympathy suggests that feelings of a special kind are at the very root of morality. As Smith explains, pleasure arises from sympathy and mutual sympathy. Pleasure may be correlated with future benefit, but it does not arise instrumentally at the thought of future benefit; it arises immediately when experiencing mutual sympathy (Smith, 1982, *TMS*, I.i.1.1; I.i.2.1; I.i.2.2; I.ii.4.1; III.2.6). When we are targets of sympathy, we feel—and think of ourselves as—approved of and loved. Mutual goodwill, like mutual sympathy, is the reciprocal giving of approval and love.

A strong objection against this view immediately comes to mind. It is common knowledge that humans can hardly be said to love strangers, solely on account of their being human. The sympathy we bring towards strangers is meager and

the quality of being human deserves either love or hatred depending on what other qualities a person displays. There is no love of mankind as such, Hume said (Hume, 1978 *THN*, 481). But Hume also said that if a European encountered a Chinese person on the moon, they would feel a sparkle of mutual love, just on account of both being human (Hume, 1978, *THN*, 482). So perhaps there is a love of *humankind* after all.

In fact, the love in question is not directed at *humankind* as such. It is a type of love that we normally give to strangers, provided they are entitled beings, beings with whom we can sympathize (whether human, intelligent or sentient) and that respond, in some degree, to our sympathy. This type of love explains the nature of justice. Justice is the minimum of mutual goodwill that we naturally give entitled beings. Even when we pay our taxes for the sole reason that it is just to do so, we are expressing our goodwill towards the collective that will benefit from the public goods that tax revenues help produce. The minimum of goodwill involves the duty not to harm others, i.e., to respect their life, their property, reputation, and agreements reached with them, unless they themselves show ill will first. It does not include the virtue of benevolence, or the duty to help, and for this reason we may fail to realize that a kind of love is required to see others as being entitled to the minimum amount of goodwill, namely justice.

Resentment demands the restoration of mutual goodwill

Violating expectations of goodwill means failing to display the minimum level, which is justice. Unjust acts, namely those that attack, for selfish reasons, the life, property, or reputation of another person clearly violate expectations of goodwill (even infants have moral expectations, see section "The origin of resentment in three steps"). They treat their target as if she has no moral status and no claim to goodwill. But since, *per hypothesis*, selfishness motivates this treatment, either the offender *publicly* denies any moral commitments, which would attract a severe form of group-punishment, e.g. ostracism, or she violates this expectation only towards a specific target and hopes her authorship will remain unknown. But even if only the target is aware of the act's authorship, the publicity condition holds, because the wrong thus enters the public domain that both target and offender occupy. The target can respond in several ways, but she basically demands re-establishment of goodwill from the offender. Resentment communicates this demand, but it also comprises a feeling of pain: being denied sympathy for selfish reasons causes pain in the target, just as receiving sympathy is accompanied by pleasure in both the subject and the target of sympathy.

This concludes my attempt to explain how resentment, in both its communicative and emotional aspects, arises from expectations of goodwill. This view of its origins will obviously tell us something about its nature, described in the following two sections.

The nature of resentment

Resentment is a part of the communicative interaction in which we are spontaneously engaged (Nozick, 1981; Bennett, 2002; McGeer, 2012, 2013) by virtue of shared and mutual expectations of goodwill. I remind the reader that I use "resentment" to refer to the moral indignation felt by the target of injustice immediately after suffering an offence. As such, it has propositional content. On one level, the recipient of an offence morally condemns the action. This moral condemnation is justified because we spontaneously demand and expect mutual goodwill from all those who we can assume to share this expectation and have common knowledge of the sharing. Shared expectations generate a shared norm about everyone's entitlement to goodwill. When somebody violates this norm, resentment arises. The victim disapproves of the offender and her action, publicly accuses her of violating the basic moral norm (whereby common knowledge in both the offender and the victim suffices for publicity), and finally demands repair and restoration. Disapproval, public accusation, and demand for repair are three connected propositional attitudes that resentment implicitly expresses.

Resentment would be pointless if it never achieved the goal of restoring the intersubjective state of mutual goodwill. The point of resentment is to elicit in the offender a receptive attitude to its demands. If the offender responds positively, she generates agreement with the evaluative perspective of the victim. Through contrition and guilt, the offender shares the evaluative perspective of the victim, approves her indignation, and feels pain on account of the pain inflicted on her. Contrition also entails a self-evaluation: harming the victim was wrong. In resentment, we want the offender to undergo this response, enact these feelings, and make these evaluations. In going through this process, offenders show sympathy—*post hoc*—towards their victims; and in this way they contribute to restoring the state of mutual goodwill that was denied through the offence.

A desire for the offender to suffer

When offenders respond through contrition and feelings of guilt, they feel pain. These feelings and this pain are good news for the victim, who desires that the offender undergo them. But, can it be moral to want someone else to suffer? (Nozick, 1981, p. 363ff; Mackie, 1985, p. 206ff.) Maybe it can, if the suffering that one desires for another consists in the painful feelings they must undergo to repair the broken moral link (see Bennett, 2002, p. 151-153; Holroyd, 2010). In the context of mutual sympathy, which is by nature communicative, pain is a necessary part of the communicative dynamics that begin with an offence. An offence causes pain to the victim, expressed in resentment. If the offender shows contrition, she takes on some pain in response to the pain she inflicted. All this must be public, in the sense that both agent and patient, at least, should have common knowledge of this dynamic. Here is a type of pain that the victim can morally desire that the offender undergo, for

this suffering is a necessary element in the process that restores a broken moral link. This clarification answers the objection pointing to the immorality of resentment.

There is one last issue on which I wish to comment. Couldn't we achieve the same goal of reparation by substituting resentment with dispassionate messages or reminders, as in "detached blaming" (Pickard, 2013)? This kind of blaming also demands goodwill, but omits the reactive feeling of resentment. This sounds plausible. I believe, however, that the desirability of "detached blaming" appeals mainly to those for whom the "strains of involvement" (Strawson, 2008 [1974], p. 10) weigh heavy, so that they seek refuge in a procedure that can be followed without feelings. For them, moral interventions like blaming and punishing are a duty, with no feelings attached. There is also the peril of relapsing into purely forward-looking consequentialism. I would therefore favor a slightly different interpretation of "detachment": it should not imply a total lack of feeling, but rather self-command over "unsocial" emotions (Smith, 1982, *TMS*, I. ii. 3), such as the exaggerations of resentment or anger that would lead to revenge and to inflicting violent punishment on the offender. A total lack of feeling towards offenders would probably lack effectiveness in morally shaping the emotions and characters of those who need some adjustment to meet moral expectations. Moreover, if blaming should be purged of reactive feelings, should we not say the same of contrition? But try to picture a detached version of contrition: what would contrition be without psychological self-reactive pain? Anyhow, the purpose of this paper is not to discuss the moral appropriateness of these detached alternatives. My purpose was merely to vindicate resentment as a morally legitimate feeling.

References

BAILLARGEON, R.; SCOTT, R.M.; HE, Z.; SLOANE, S.; SETOH, P.; JIN, K.; WU, D.; BIAN, L. 2015. Psychological and Sociomoral Reasoning in Infancy. *In:* M. MIKULINCER; P.R. SHAVER (eds.), E. BORGIDA; J.A. BARGH (assoc. eds.), *APA Handbook of Personality and Social Psychology: Vol. 1. Attitudes and Social Cognition*. Washington, American Psychological Association, p. 79-150. https://doi.org/10.1037/14341-003

BENNETT, C. 2002. The Varieties of Retributive Experience. *Philosophical Quarterly,* **52**(207):145-163. https://doi.org/10.1111/1467-9213.00259

BLOOM, P. 2013. *Just Babies: The Origins of Good and Evil*. New York, Crown, 273 p.

BOEHM, C. 1999. *Hierarchy in the Forest: The Evolution of Egalitarian Behaviour*. Cambridge, Harvard University Press, 292 p.

DARWIN, C. 1981 [1871]. *The Descent of Man, and Selection in Relation to Sex*. Princeton, Princeton University Press, 475 p. https://doi.org/10.5962/bhl.title.70891

GOODENOUGH, O.R. 2004. Responsibility and Punishment: Whose Mind? A Response. *Philosophical Transactions of the Royal Society B*, **359**:1805-1809. https://doi.org/10.1098/rstb.2004.1548

GREENE, J.; COHEN, P. 2004. For the Law, Neuroscience Changes Nothing and Everything. *Philosophical Transactions of the Royal Society London B*, **359**:1775-1785. https://doi.org/10.1098/rstb.2004.1546

HOLROYD, J. 2010. The Retributive Emotions: Passions and Pains of Punishment. *Philosophical Papers,* **39**(3):343-371. https://doi.org/10.1080/05568641.2010.538914

HUME, D. 1978 [1739]. *A Treatise of Human Nature*. 2nd ed., Oxford, Oxford University Press, 743 p.

JOHNSTON, G. 1999. Restorative Justice, Shame and Forgiveness. *Liverpool Law Review,* **21**(2/3):197-216. https://doi.org/10.1023/A:1005629612949

KELLY, E. 2008. Doing without Desert. *In*: M. MCKENNA; P. RUSSELL (eds.), *Free Will and Reactive Attitudes*. Surrey, Ashgate Publishing, p. 273-295.

MACKIE, J. 1985. Morality and the Retributive Emotions. *In:* J. MACKIE; P. MACKIE (eds.), *Persons and Values II*. Oxford, Oxford University Press, p. 206-219.

MCGEER, V. 2012. Co-reactive Attitudes and the Making of Moral Community. *In:* R. LANGDON; C. MACKENZIE (eds.), *Emotions, Imagination, and Moral Reasoning*. New York, Psychology Press, p. 299-326.

MCGEER, V. 2013. Civilizing Blame. *In:* D.J. COATES; N.A. TOGNAZZINI (eds.), *Blame: Its Nature and Norms*. Oxford, Oxford University Press, p. 162-188.

NOWELL-SMITH, H.P. 1948. Freewill and Responsibility. *Mind,* **57**:45-61. https://doi.org/10.1093/mind/LVII.225.45

NOZICK, R. 1981. *Philosophical Explanations*. Cambridge, Harvard University Press, 764 p.

NUSSBAUM, M. 2015. Transitional Anger. *Journal of the American Philosophical Association*, **1**(1):41-56. https://doi.org/10.1017/apa.2014.19

PEREBOOM, D. 2008. Revising the Reactive Attitudes. *In*: M. MCKENNA; P. RUSSELL (eds.), *Free Will and Reactive Attitudes*. Surrey, Ashgate Publishing, p. 219-234.

PICKARD, H. 2013. Irrational Blame. *Analysis,* **73**(4):613-626. https://doi.org/10.1093/analys/ant075

SCANLON, T. 2013. Giving Desert its Due. *Philosophical Explorations,* **16**(2):101-116. https://doi.org/10.1080/13869795.2013.787437

SMART, J.J.C. 1961. Free Will, Praise, and Blame. *Mind,* **LXX(**279):291-306. https://doi.org/10.1093/mind/LXX.279.291

SMITH, A. 1982. *A Theory of Moral Sentiments*. Indianapolis, Liberty Fund, 412 p.

STRAWSON, P.F. 2008 [1974]. *Freedom and Resentment and Other Essays*. New York, Routledge, 235 p.

TOMASELLO, M. 2009. *Why We Cooperate*. Cambridge, The MIT Press, 206 p.

TRIVERS, R. 1971. The Evolution of Reciprocal Altruism. *Quarterly Review of Biology,* **46**(1):35-57. https://doi.org/10.1086/406755

WALKER, M.U. 2006. *Moral Repair. Reconstructing Moral Relations after Wrongdoing*. New York, Cambridge University Press, 250 p. https://doi.org/10.1017/CBO9780511618024

Submitted on March 31, 2016
Accepted on June 17, 2016