

*Draft Version. January the 1st, 2010  
Footnotes still to be completed  
and main text thoroughly revised*

# **THE LOSS OF CONFIDENCE IN THE WORLD.**

**An Essay in Morality and Self-Knowledge**

by

Josep E. Corbí

(University of Valencia)

## TABLE OF CONTENTS

### PRESENTATION

1. A Certain Kind of Human Suffering	6
2. The Moralist View	7
3. A Cultural Stereotype	10
4. Thought Experiments and Character	13
5. Torture and the Loss of Confidence in the World	16
6. The Real and the Imaginary	18
7. The Divided Conception of the Self	20
7.1 The Limits of Moral Projectivism	21
7.2 The Role of Principles	23
7.3 Guilt as an Immoral Emotion	24
8. The Moral Question and Self-Knowledge as an Achievement	27

### CHAPTER 1

#### THOUGHTS EXPERIMENTS, JUSTICE, AND CHARACTER

1. Presentation	30
2. John Rawls: the Original Position	33
2.1 Reasoning Behind the Veil of Ignorance	33
2.2 Reflective Equilibrium and the Original Position as a Device	34
2.3 Hypothetical vs. Actual C-Judgments	36
3. Primo Levi's Shame	38
3.1 Stereotyped Expectations	40
3.2 Was Levi's Shame Rational?	42
3.3 The Modality of 'Could'?	45
3.4 Who Can Judge the <i>Sonderkommandos</i> ?	48
4. To Sum up: Thought Experiments in Moral Deliberation	54

**CHAPTER 2**  
**THE LOSS OF CONFIDENCE IN THE WORLD**

1. Presentation	58
2. The Three Poles of Torture	61
2.1 Third Agents and the Loss of Confidence in the World	61
2.2 Interrogation and Third Agents	63
2.3 The Human World	66
3. The Nature of Our Expectations	66
3.1 Egocentric vs. Impartial Expectations	67
3.2 Non-statistical Expectations	70
4. The Loss of Confidence in the World as an Irrational Reaction	73
5. The Real of Reasons and the Faustian Ideal	77
5.1 The Faustian Ideal	78
5.2 Distance: Knowing That vs. Being Sensitive To	81
6. The Moral Reversal of Time	88
7. To Close: A Necessary Illusion	91

**CHAPTER 3**  
**THE REAL AND THE IMAGINARY IN THE SOLDIER'S EXPERIENCE**

1. An Initial Approximation	93
2. The Departure	96
3. The Battlefield	102
3.1 The Soldier as a Victim	102
3.2 The Soldier as an Executioner	105
4. The Homecoming	107

**CHAPTER 4**  
**PRINCIPLES, GUILT, AND MORAL DELIBERATION**

**Part I: Principles and Character**

1. How Should One Respond?	111
----------------------------	-----

2. Two Ambitious Generalist Projects	114
3. Weak Generalism	116
4. 'This Shouldn't Happen!'	120
5. The Relevance of Coming Across. Agents with a Character	122
 <b>Part II: Character and Guilt</b>	
6. Guilt and Morality	128
7. Principles and Guilt	130
8. Inner Figures and the Global Attack	133
8.1 Richard Wollheim's Approach	135
8.2 Inner Figures and the Human World	140
8.3 Third Agents and the Global Effect of Guilt	143
8.4 The Voice of Victims	144
9. Inner Figures and Practical Deliberation	148
10. Conclusion	153

## CHAPTER 5 A CASE FOR MORAL REALISM

### Part I: Normativity and Unmasking Explanations

1. The Bipartite Image	157
2. Moral Projectivism	159
3. A Narrative Discipline	162
4. Response-Dependent Properties	165

### Part II: Christine Korsgaard's Procedural Realism and the Space of Public Reasons

5. Procedural vs. Substantive Realism	171
6. The Moral Law	174
7. The Space of Public Reasons	179
8. Character and Moral Pluralism	184

## CHAPTER 6

## SELF-KNOWLEDGE IN THE LIGHT OF A DANCE

### Part I: First-Person Authority and the Deliberative Attitude

1. Presentation	198
2. The Deliberative and the Theoretical Attitudes	202
3. The Transparency Condition	205
4. Avowals and the Goal of Psychoanalytic Treatment	209
4.1 The Ambiguity of Avowals	210
4.2 The Goal of Psychoanalytic Treatment	211

### Part II: Acknowledgment and Self-Knowledge as an Achievement

5. The Notion of Acknowledgment	214
6. 'Being forced to'	218
7. Receptive Passivity and Double Permeability	220
8. Receptive Passivity and Psychic Health	225
8.1 The Experience of Dancing	225
8.2 Receptive Passivity and the Theoretical Attitude	228
8.3 Psychic Health	230
8.3.1 Expression	230
8.3.2 Inner Figures	231
9. Conclusion	234

## CHAPTER 7

### CONCLUSION: TOWARDS EXPRESSIVE AWARENESS

1. The Moral Question	236
2. The Divided Conception of the Self	239
3. The Frailty of Principles	244
4. Expressive Awareness and the Three Poles of Harm	247
4.1 The Victim	249
4.2 Third Agents	251
4.3 The Torturer	254

## PRESENTATION

**1. A Certain Kind of Human Suffering.** This essay is in search of some light as to how one should relate to a certain kind of human suffering. The word 'should' undoubtedly raises a normative issue, and also the use of 'one' to refer to the agents who may be subject to a normative demand. Choice of such words is hardly neutral and betrays a certain philosophical stance, which, I hope, later considerations will reasonably ground. We may examine this picture, to begin with:



What do you see in it? A naked girl screaming and walking in the middle of a road with some soldiers in the background? Do you see this girl in company with some of other children running away from somewhere, perhaps, the village being burned by some bombs, napalm bombs as we know? Do you see the innocence in their bodies and faces, the isolation and abandonment in the way they walk, which is even emphasized by the presence of a girl who caringly holds the hand of, may be, her small brother? Do you see, by contrast, the soldiers as armed and powerful, as members of an army, as not deserving neither the word 'innocence' nor the word 'isolation'?<sup>1</sup> Do you see the village through the smoke? It looks as if it had vanished: these children do not seem to have a way back home. Do you feel tempted to cry: '*This shouldn't happen!*'? But what does

---

<sup>1</sup> I will dwell on the soldier's experience in chapter 3. There, we will see the rather dramatic sense in which the word 'isolation' does, after all, apply to the life of a soldier, although not so much the word 'innocence'.

'this' refer to and what's the strength of 'shouldn't'? Let us begin with the word 'this', since, by examining its content, we may also learn about the nature of 'shouldn't'.

We may thus ask: does 'this' refer to this particular girl, to this particular party of children, to what happened to the people in this specific village on this particular occasion? It seems clear that whenever an agent may cry 'This shouldn't happen!', the reference of 'this' is hardly confined to the particular case, but tends to generalize in two directions. On the one hand, 'this' is meant to embrace not only the Vietnamese girl, but a number of other situations that may be relevantly similar and have taken (or will take) place at some other point of time or space. Yet, the expansion of the agent's concern beyond the particular occasion only makes sense if she assumes that her injunction does not derive from some idiosyncratic feature of her psychology, but voices a concern that *anyone should have*. Otherwise, the agent should be interested in providing shelter, food, and consolation only for this particular girl, and show no regard for the other children in the picture or in other villages or for any other people in a relevantly similar situation. By perceiving such an attitude as narrow, as missing an essential point in the outcry, we acknowledge that the agent's concern is not idiosyncratically hers; this points to a second direction in which 'this' generalizes, that is, it not only expands beyond the particular situation, but also beyond the particular agent who may issue the injunction. And it seems that we can only intelligibly generalize in the former direction if we are prepared to generalize in the latter one as well, whereas the obverse is uncontroversially true. So, 'This shouldn't happen!' seems to place a constraint upon us which does not depend either on the details of our idiosyncrasy or on the peculiarities of a given situation, but presupposes some commonalities among situations and agents which justify the generality of the normative constraints. Whether such a demand may ultimately be justified and what reasons for such a demand may look like, are central issues to be explored throughout this essay in dispute with a specific view about morality.

**2. The Moralistic View.** Our moral practices appear as the most robust attempt to articulate the way one should respond to the kind of human suffering that the picture of the Vietnamese girl brings to the fore, namely, the *harm* that we cause to each other. Such a harm surely goes far beyond the pain inflicted on the victim's body, even though it is already a matter of elucidation what else may be actually involved. It is, in any

case, clear that avoidance of such a harm and provision for an appropriate response lies at the core of the injunction: 'This shouldn't happen!'.

Such an injunction expresses a matter of importance and some may, from this fact, derive the conclusion that all kinds of moral harm are equally important; as a result, they may regard avoidance (and prevention) of any such harm as the most important goal that any decent person must pursue. A certain view about the role that morality should have in our lives will thereby emerge, namely: a view that regards moral reasons as being of supreme significance, so that, in case of conflict any other sort of reason should yield. We may call it '*the moralist view*', which could be presented as committed to:

(MV1) moral reasons should, in case of conflict, override any other sort of reason that one might have, and

(MV2) being moral is, to say the least, a necessary condition for leading a valuable life.

These two conditions put a significant pressure on the agent and, derivatively, on morality. It is not just that being moral may turn out to be hard on some specific situations, as it will certainly be; but the fact that moral reasons, no matter how significant they are, should always come first. This constraint pictures the life of a moral agent in such a way that the demands of morality will quite often conflict with her other projects and attachments, if not in the whole, perhaps in the details as to how they might be more meaningfully pursued and cultivated. In the light of such conflicts, the agent will surely feel tempted to either challenge the relevance of morality or, less costly to her own self-image, revise her view of the situation in such a way that she will no longer regard herself as being under the *obligation* to undertake a rather demanding course of action. We may thus see how, on the moralist view, the notion of obligation

should become central to morality,<sup>2</sup> as well as the risk of self-deception.<sup>3</sup> To put it another way, we may say that the moralist interprets 'shouldn't' in the injunction as deriving from the set of moral obligations that must guide an agent's life, so that certain actions should be prevented and, in any case, adequately punished and repaired.

The moralist is often convinced that the best we can do to counteract our temptation to disregard the demands of morality, is to ground our moral obligations, as well as the supreme importance of morality, on principles that no rational agent could challenge. In the age of natural sciences, the moralist is, nevertheless, confronted with rather specific difficulties to carry out her project. For natural sciences seem to tell us that the world as it is, independent of our response, is dispossessed of any axiological property. There is no place in such a world for the good or the bad, for cruelty or generosity, for courage or cowardice, or even for moral harm itself. Where could, then, the moralist ground our moral obligations and, ultimately, the very need of a moral response? If such grounds are not to be found in the world, as it is independent of us, we should rather elaborate such foundations on the basis of the way *we respond to it*. After all, the starting point of our reflection was already a response, namely: *the experience*

---

2 Williams (1985) presents morality as peculiar institution shaped in the light of a certain notion of obligation: "The important thing about morality is its spirit, its underlying aims, and the general picture of ethical life it implies. In order to see them, we shall need to look carefully at a particular concept, *moral obligation*. The mere fact that it uses a notion of obligation is not what makes morality special. There is an everyday notion of obligation, as one consideration among others, and it is ethically useful. Morality is distinguished by the special notion of obligation it uses, and the significance it gives to it. It is this special notion that I call 'moral obligation'... The philosopher who has given the purest, deepest, and most thorough representation of morality is Kant. But morality is not an invention of philosophers. It is the outlook, or, incoherently, part of the outlook, of almost all of us." (p. 174). In a similar vein, Christine Korsgaard regards the notion of obligation as constitutive of our humanity: "And Nietzsche was right when he warned the enemies of obligation not to think of it lightly because it was born in pain and ugliness. Obligation is what makes us human. Or anyway, so I will argue." (Korsgaard (1996), p. 5)

3 "The human being feels within himself a powerful counterweight to all the commands of duty, which reason represents to him as so deserving of the highest respect - the counterweight of his needs and inclinations, the entire satisfaction of which he sums up under the name happiness. Now reason issues its precepts unremittingly, without thereby promising anything to the inclinations, and so, as it were, with regard and contempt for those claims, which are so impetuous and besides so apparently equitable (and refuse to be neutralized by any command). But from this there arises a *natural dialectic*, that is, a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations, that is, to corrupt them at their basis and to destroy all their dignity - something that even common practical reason cannot, in the end, call good." (Kant (GMM), pp. 17-9). Cf. Williams (2002, ch. 5) for a particular case of self-deception regarding the alleged obligation not to lie under any circumstances.

*of being affected in a certain manner by a particular case of harm.*<sup>4</sup> Kantian approaches seek to ground our moral obligations on such basis, that is, they grant that the world in itself is dispossessed of evaluative properties and, consequently, to assume that morality should be grounded on those responses of us that satisfy some normative criteria that every rational agent must accept. The notion of *agreement* among rational agents is thereby emphasized, but also the notion of *procedure*. For Kantian approaches aim at meeting the metaphysical demands of a disenchanted conception of the world by relying on some version of the distinction between form and content, namely: by appeal to a *procedure* which every rational agent may apply and such that it determines all moral obligations without relying on any substantive moral claims. By this means, being moral becomes inextricably associated with a certain conception of what a rational moral agent may consist of which focuses on (a) the ability to reach an agreement with every other rational agent in virtue of their common capacity to follow a specific procedure and (b) the ability to be faithful to whatever principles and obligations a proper application of such a procedure may deliver. We may see (a) and (b) as central elements in any Kantian approach to morality, no matter how differently the details may be worked out within each particular project. Hence, I will hereafter use phrases such as '*Kantian approaches*' or '*the Kantian approach*' to refer to a set of theories or practices (since a Kantian approach may be tacitly present in our understanding of certain practices) that *endorse* (or take for granted) *the moralist view and seek to ground morality in the light of a conception of moral agency and rationality inspired in (a) and (b)*.

I feel, nevertheless, quite dissatisfied with Kantian (and, in general, moralist) approaches to the harm that we may cause to each other.<sup>5</sup> A first source of my dissatisfaction is in itself a moral consideration, namely: Kantian approaches may fail to

---

4 A relevant question, which will in due course be addressed, is whether the fact to which that experience is a response can be identified independently of our response, and also the challenge that an appropriate answer to that question may place to morality.

5 Other moralist approaches, like some consequentialist views, may crucially differentiate themselves from the Kantian approach, and still be affected by my line of reasoning insofar as they may share some fundamental assumptions with the latter, like a conception of the self as divided into reason and passions; or the related conviction that hypothetical, distant, deliberation is preferable to a more engaged kind of reflection, since the latter may easily be contaminated with one's passions and inclinations. Yet, I will not, in this book, address any issue that might specifically concern consequentialism or any moralist view other than the Kantian approach. For, even if such views may share with the Kantian approach some fundamental assumptions, the particular way in which they are articulated would require an equally specific response which goes well beyond the scope of this essay.

encourage a morally adequate response to the harm that we cause to each other; and this should certainly count as a *reductio* inasmuch as the need of an appropriate response was what supposedly motivated them in the first place. This failure is, nevertheless, associated with my second source of dissatisfaction, namely: the moralist significantly (and unduly) reduces the conditions under which the life an agent may make sense. The moralist view derives from the perception of an almost permanent conflict between the demands of morality and the demands of happiness (as they put it) or meaningfulness (as I would rather prefer to say), together with the attempt to ground that, in the light of such a conflict, rational agents should always opt for morality.<sup>6</sup> I do not mean to deny that there may be plenty of such conflicts and also that some such conflicts may be really tragic to the people involved. It does not follow, however, that our lives should be conceived of mainly in the light of such a conflict, as if we were at all times forced to choose between what makes sense to us and what morality demands from us. I will, moreover, argue that the moralist insistence on obligation and on the supreme importance of morality (as it is expressed in (MV1), leads us astray from our initial response to the picture of the Vietnamese girl: 'This shouldn't happen!'; for, in some relevant circumstances, the moralist view favors, rather than deters, the production of harm.

**3. A Cultural Stereotype.** One might claim the Kantian approach constitutes the *received view* about morality. Yet, the role of such an approach is deeper than that of a view that one might defend when asked. This model constitutes a *cultural stereotype* in the light of which we *actually* shape and examine our lives. I hope the feasibility of this claim and the relevance of the distinction on which it rests, will become clear as the discussion proceeds. We will thus see how the way we try to make sense of several aspects of our experience of harm is conditioned by such a model. This conditioning is so deep that we may identify our moral experience with a Kantian characterization of it, so that we may have a hard time to even make room for the intelligibility of an alternative description of some fundamental experiences concerning harm. I will try to motivate the need for such alternative description by pointing to reasons that may appeal even to those -which, if I am right, are many- who are convinced that there is no conceptual

---

<sup>6</sup> "To assure one's one happiness is a duty (at least indirectly): for, want of satisfaction with one's condition, under pressure from many anxieties and amid unsatisfied needs, could easily become a great *temptation to transgression of duty*." (Kant (GMM), p. 4:399)

room for it. This is one of the reasons why I begin by taking morality back to the experience of harm and seek to show why the Kantian approach may, in some relevant circumstances, favor the production of harm. This will certainly count as a serious challenge to such an approach (and, therefore, to a fundamental stereotype of our culture), whose point of departure was precisely to ensure that harm was avoided or, at least, appropriately punished and repaired. In the light of this, some may feel encouraged to explore an alternative model, and an initial step in that direction is to allow for a gap between our experience of harm and a deeply entrenched characterization of it. Once this gap is open, there will at least be the hope that an alternative approach might help us to deal more appropriately with harm. At a later stage in the discussion, the perplexity may turn out to be not so much how there could be a non-Kantian description of the experience of moral harm, but rather how we could have missed it in the first place. And we may, as a result, be inclined to inquiry why and when we were led astray. A number of fears will, in due course, be mentioned and characterized to shed some light on this particular perplexity.

Some might reply that, even if I succeeded in my challenge, it would only prove that, as it has been actually implemented in our culture, the Kantian approach encourages the production of harm; but this will fall short of showing that such consequences are *essential* to this approach. What is, however, the point and content of this reply? It may just stress that the Kantian approach does not lead to inappropriate consequences when practiced by rational agents. We cannot, though, find much consolation in this thought because the Kantian approach was meant to provide a response to the harm that human beings cause to each other. And we are now asked to accept that the Kantian approach only favors that end if handled by rational agents, and not by human beings as they are to be found in our culture. Of course, the Kantian approach could be vindicated for another culture, but this by itself would surely weaken its stake. For the Kantian is supposed to address all rational agents while relying on the assumption that human beings are, in general, rational enough to be moved in the right direction by the Kantian line of reasoning. My claim is that this assumption cannot be granted, just the other way round: we have serious reasons to believe that, in some relevant circumstances, the Kantian approach favors the production of harm.

Hence, insofar we are more concerned with harm than with preserving the Kantian approach (which everybody is, as we shall see), then we have a really serious

reason to look for an alternative model that may help us to prevent and deter harm. In any case, I hope that, as we go through the details, we may understand why the Kantian model is not only morally disreputable, but conceptually untenable, and see how the alternative model may be able to meet some vexed metaphysical and epistemic concerns. A question that should be asked is why the alternative model should care at all about meeting these last concerns. We are looking for a model that should help us to shape our individual and social lives in a way that appropriately responds to the experience of harm. Why should it be so important that this model meets some rather abstract perplexities? It is crucial for the Kantian approach to deal with such worries because it addresses rational agents and assumes that such agents will only be motivated to behave morally by being shown that being moral is a necessary condition for being rational. Consequently, if the Kantian approach were in conflict with some central elements of our world-view, it could longer not be presented as rational. Hence, there is much pressure for the Kantian approach to form a part of a comprehensive and robust picture, but it is unclear why the alternative model should be forced to satisfy that demand. An initial motivation derives from the rather circumstantial fact that we are challenging the Kantian approach; and the latter rests on a rather architectonic conception of rationality, which involves the need to articulate an overall convincing picture. In such dialectical context, some may accept a challenge to the Kantian approach to morality without, at the same time, doubting the conception of rationality that may come with it. They may, thereby, expect that any defense of an alternative model should show how this model satisfies the demands that the Kantian approach may fail to meet. Yet, beyond this context of discussion, it is unclear why any alternative model should worry about such demands. Even if it turns out that it should worry, the nature and source of their concern may be rather alien to those of the Kantian approach. More specifically, I will favor a model that focuses not only on the harm that we cause to each other, but also on the harm that an agent may cause to herself; for there is the hope that the conditions under which the latter may be repaired, may give us a hint as to how we may more efficiently respond to the former. Within this alternative model, it is a certain kind of attention, and not so much a system of obligation, that lies at the core of a more positive response; this is why I will refer to this alternative approach as *'the attentional model'*. A feature of this model is that it requires that agents should not abstract away from their emotions, projects and

engagements, but, on the contrary, put them into use in her attempt to discern the morally relevant aspects of any given situation. This will come as a challenge to the Kantian picture of what should count as a rational agent and, relatedly, diminish the pressure to provide groundings for morality that every rational agent ought to acknowledge.

There are, in any case, several ways in which we may access the Kantian approach as a cultural stereotype. Some explicitly Kantian theories of morality might, to this purpose, be examined, but one should also track Kantian assumptions in philosophers who do not regard themselves as Kantian at all; and some literary works may indeed be explored to discern in what fundamental respects they may have been conditioned by such a stereotype, and the same applies a practices and institutions of rather disparate kinds. In this essay, I will mainly focus on some explicitly Kantian views as well as on the study of some experiences of harm in order to examine to what extent the Kantian stereotype may hinder (or, instead, favor) a proper understanding of such experiences and their significance in our lives.

I must say that, in general, Kantian theories are not meant at all to make a certain cultural stereotype explicit, but rather to ground our moral obligations; although such a grounding may, in turn, serve to the purpose of promoting and strengthening that stereotype itself. The entire process may, nevertheless, turn out to be question begging, since that grounding may ultimately rely on some experiences whose perception has already been molded by the stereotype at stake. This kind of argumentative flaw will be detected both in the content of the assumptions that are explicitly regarded as uncontroversial, and in the way the morally relevant features of a given situation are identified. To this purpose, I will try to go beyond the skinny examples and thought experiments often provided by academic philosophers, and study some central experiences of harm, like torture, genocide and war, as they have been reported by some individuals that were directly involved in some such experiences. This methodological proposal immediately raises the issue as to the authority of such voices: whether they are not too close to the experience to be reliable, what to do when they conflict with each other, or whether one should accept everything they say on a certain matter. These concerns are certainly relevant, but we should not assume that only an impartial view could tell us what is really at issue in such experiences. Although my worry will not so much be about the notion of impartiality itself, but about a certain

understanding of it, namely: the one that depends on the idea that, in order to be really impartial, agents should abstract away from their respective character and deliberate only on the basis of some rational procedure.<sup>7</sup> At some point, I will challenge the view of an agent's character as composed by a set of elements that may only pollute her reasoning and suggest that, if the notion of impartiality is to play a role in moral deliberation, we should rather explore how such a notion may be of relevance within a less encapsulated conception of the resources to which an agent may legitimately appeal in her deliberation about how to lead her life.

**4. Thought Experiments and Character.** The Kantian conception of impartiality and rationality surely favors a methodological approach that should focus at most on thought experiments and rather skinny examples; the study of more full-fledged cases is to be avoided because, otherwise, one may easily be engaged emotionally and, thereby, acquire a biased picture of the moral situation at issue.<sup>8</sup> I am certainly convinced that thought experiments and skinny examples should play a role in ethical and meta-ethical deliberation, but I will, in chapter 1, argue that thought experiments can hardly be the main element in one's philosophical diet, since, in my view, the virtues of such a tool can only be delivered if adequately combined with the study of more full-fledged experiences. To this purpose, I will challenge the relevance of a particular thought experiment (namely, John Rawls' original position) in the light of what one may learn from a rather specific experience, namely, Primo Levi's shame after his liberation from Auschwitz.

Rawls presents the original position as a representational device, as a thought experiment, designed to determine the principles of justice. He stresses, however, that

---

7 This is, indeed, a central point in Rawls' original position (cf. Rawls (1999, p. 11; 2001, sec. 2; 1993, p. 305), which may, in turn, be construed as a variant of Kant's insistence on the supreme principle of morality being derived by merely a priori means (see Kant (GMM), 4: 389).

8 Kant (GMM) dismisses the relevance of examples, no matter whether full-fledged or skinny, to specify the supreme principle of morality: "Nor could one give worse advice to morality than by wanting to derive it from examples. For, every example of it represented to me must itself first be appraised in accordance with principles of morality, as to whether it is also worthy to serve as an original example, that is, as a model; it can by no means authoritatively provide the concept of morality.... Solely from the *idea* of moral perfection that reason frames a priori and connects inseparably with the concept of a free will. Imitation has no place at all in matters of morality, and examples serve only for encouragement, that is, they put beyond doubt the practicability of what the law commands and make intuitive what the practical rule expresses more generally, but they can never justify setting aside their true original, which lies in reason, and guiding oneself by examples." (p. 21)

the original position could only play such a role if it satisfies the demands of reflective equilibrium. More specifically, he claims that the principles that the original position may deliver, as well as the particular judgments that may follow from them, should match “our considered judgments once they have been pruned and adjusted.”<sup>9</sup> It can be easily be argued that such a match is only possible if we assume that, when agents deliberate appropriately, two kinds of judgments *must* come together; if not generally, at least when justice is at issue:

*Hypothetical C-Judgments:* An agent's *present* considered judgment about a particular case concerning justice which she is *not actually* confronted with.

*Actual C-Judgments:* An agent's considered judgment when she *actually* confronts such a case.

where 'c-judgment' stands for 'considered judgment'. The first kind of judgment is *hypothetical* because the agent is not assessing a situation she is actually confronted with, but a situation that she (or someone else) might eventually confront; by contrast, the second kind of judgment appears as *actual* insofar as the agent is presently confronting the situation under evaluation. The original position is surely concerned with hypothetical c-judgments, even though of a rather specific kind, as we shall soon see, and it is quite clear that the original position could only pass the reflective equilibrium test on the assumption that

*The Matching Assumption:* if agents deliberate appropriately, there should not be any gap between their hypothetical c-judgments and their actual c-judgments.

For, otherwise, there is no reason why our judgments in the original position should track both our hypothetical c-judgments and our actual c-judgments, as reflective equilibrium seems to demand. I will argue however that, in some significant cases, there are mismatches between an agent's hypothetical and an agent's actual c-judgments which do not derive from any flaw in deliberation. Moreover, I will conclude that, at least in some such cases, it is the agent's actual c-judgment that should prevail. This

---

<sup>9</sup> Rawls (1999), p. 18.

will come as a challenge to the Matching Assumption and, as a result, the role that Rawls ascribes to the original position in his theory of justice will be doubted. To develop my line of reasoning, I will mainly rely on Primo Levi's description of his experience of shame as a survivor of Auschwitz, which he conscientiously describes in his last book, *The Drowned and the Saved*. We may reasonably regard this text as expressing Levi's considered judgment on his (and other survivors') experience of shame; and, since his analysis in that text is commonly presented as a central contribution to the discernment of some crucial aspects of a rather significant experience of harm, we may also view his considered judgment as particularly authoritative. It may be argued that the main point of his reflection on shame is to stress the mismatch of his actual and his hypothetical c-judgments with regard to issues that are central to policies of justice and reparation, and also that his actual ones should prevail. It would thus follow that the Matching Assumption should be questioned. One might certainly doubt the relevance of Levi's c-judgment, though, in my view, this should be done for some *specific* reason, since, otherwise, we run the risk emptying the demand that the original position should pass the reflective equilibrium test.

This line of reasoning may also serve as a vindication of a methodological approach where careful examination of full-fledged experiences and actual experiments play a significant role. This is the inspiration that will guide my reflection throughout this essay, but there are some other issues that the study of Levi's shame will raise and are to be resumed in later chapters. First of all, Levi's reflection on shame fosters the need to elaborate an alternative account of the self and its deliberative capacities that should be able to make sense of the fact that, in some relevant cases, actual c-judgments prevail. The distinction between a merely knowing that such and such is a fact and becoming sensitive to such fact will thereby become central, even if it is manifestly at odds with the conception of rationality and the self presupposed in Rawls's original position. And, related to that, I will raise a second issue concerning the asymmetries between the first-person and the third-person perspective with regard to moral issues. The Kantian notion of impartiality seems to rule out any such asymmetries, whereas they are arguably constitutive of our moral practices. To put it in a nutshell, I will conclude that there are cases where it may be rational that an agent accuses herself of having done something shameful, even though no third party could legitimately make such an accusation or, in other words, I should conclude that there is more to morality

than third-person accusation.

After these methodological remarks, it seems quite natural that reflection on the most abstract issues in this essay should be closely connected to the examination of some rather specific experiences of harm. My first goal will be to articulate an analysis of what is essentially involved in that experience; such an endeavor will occupy chapters 2 and 3. After that, I will make a case against the way in which Kantian approaches assume that one should respond to the harm thus characterized. This will be the subject matter of chapters 4 and 5; whereas the two final chapters (i.e, chapters 6 and 7) will have a more positive approach and it is there where I will sketch an alternative approach to morality and self-knowledge, which will be mainly based on the exercise of a certain kind of attention. In the light of this, chapter 2 will focus on a paradigmatic (and most devastating) experience of harm: torture.

**5. Torture and the Loss of Confidence in the World.** In *The Mind at its limits*, Jean Améry reflects on the subjective condition of the victim; and it is my purpose to take his experience as the starting-point of our inquiry into the nature of torture. There is surely the worry that Améry's experience might be too idiosyncratic, so that we could hardly learn for him about the nature of torture itself.<sup>10</sup> This raises again the issue about the way in which the relevance of a particular experience is to be assessed. The first thing to notice is that this worry also affects (and, presumably, more severely) examples and counterexamples that might be presented in the light of some thought experiments or skinny examples, except, of course, if we were in the business of providing necessary and sufficient conditions for a certain X. Once such a project is dropped, as it seems it should concerning any matters of importance, then examples and counterexamples are to be assessed in their *relevance* and, in this respect, they are not methodologically better off than the appeal to any full-fledged experience. In either case, it seems that a central question in such an evaluation is how any such examples or experiences may allow us to *make sense* of one or another aspect of the experience of harm. 'Making sense' is surely relative to an audience<sup>11</sup> and involves some circularity, since the precise

---

<sup>10</sup> This worry being increased by the possibility that the emotional intensity of his experience as a victim of torture might have easily distort his elucidation of it. Yet, as mentioned in section 3, emotional detachment can also operate as a severe bias; I will argue that a certain kind of emotional involvement may be required for a proper understanding of certain facts.

<sup>11</sup> Williams (2002) makes this point as follows: "There may be a suspicion that the idea of 'making sense' is being used here in a promiscuous way, in particular -and this is a specially nasty offence among some philosophers- because it seemingly runs together fact and value. It

boundaries of the 'we' whose understanding of harm is supposed to be significantly favored by Améry's experience, should be already be fixed by appeal to some moral intuitions and this, in turn, is surely connected to an agent's capacity to see some experiences of harm as central or paradigmatic. I hope, though, that the study of Améry's experience will appear as pertinent and justified to quite an extended and indeterminate audience, once I suggest how his reflections on torture may illuminate some other experiences of harm, as well as the ways in which one should relate to them.

Torture has apparently two poles: the torturer and the victim. It is easy to show, however, why a third pole is necessarily involved, namely: what I will refer to as 'third agents'; the appeal to such third party is present both in the victim's experience and in that of the torturer. The latter presents the question that she raises in the interrogation as urgent and necessary to protect us from an attack, while the answer is almost inevitably interpreted as a betrayal. This way the initial sympathy of third agents for the victim's plight is reverted into distrust, and the torturer's action appears as a necessary maneuver of self-defense. But the victim also makes a claim to third agents and on similar basis. To experience the world as hospitable involves, according to Améry, the expectation that

(E1) no one will hurt me,

but also that,

(E2) if I were after all hurt, someone else would come and protect me.

The loss of confidence in the world, which Améry regards as central to his experience as a victim of torture, takes place when this second expectation, where third agents are certainly involved, is experienced as ungrounded. There will be some discussion as to how exactly the content of such expectations is to be specified; this will turn out to be a

---

does have different implications in different connections, but there is no muddle involved. The basic idea is that it makes sense to a certain person (or group, etc.) *that P*. We can take 'P' as the story, and then the idea is that it makes sense (to that person) that things should work out as the story says they worked out." (p. 235). A few pages later, Williams claims: "What makes sense of the past to us may not make sense of it to others. This applies to people in the past: we know, historically, that their interpretation of their past differed from our interpretations of both their past and ours. It applies to people in the future." (p. 258). This remark can certainly be taken as an illustration of his previous general point.

privileged occasion to see how the Kantian stereotype may stand in the way of a proper understanding of the victim's condition and, in the end, of the experience of harm. The most extended debate will concern, however, the issue as to whether the loss of confidence that some victims experience should be construed as a merely irrational response to a trauma or instead placed in the space of reasons. I will conclude that we should opt for the latter horn of the dilemma and provide some evidence to back up the claim that not only the victim, but all of us have reason to lose the confidence in the world. And, nevertheless, only those victims who do not expect their harm to be looked after, are bound to lose their confidence in the world. This will invite the distinction between merely *knowing that* such and such is fact and *becoming sensitive to* it. Third agents may easily know certain facts that might justify the loss of confidence in the world, but *a mere declarative awareness* of such facts is insufficient to produce such an effect; a different kind of contact with such facts is needed, like the one that some victims may experience. The latter may acquire what I shall describe as an *expressive awareness* of the harm inflicted upon them and thereby become sensitive to such facts to the point of their lives is severely disfigured by it. My point is that this expressive awareness, contrary to what the Kantian approach would suggest, should not be interpreted as blinding or distorting, but, on the contrary, as enhancing our capacity to grasp the significance of some facts that third agents are bound to displace to the realm of the imaginary (i.e., to relate to them as if they were not really facts, but creatures of their imagination), in an attempt to keep their lives going on as usual. I will thus conclude that it is third agents who are prey of an illusion, and not those victims who lose their confidence in the world, since the latter are expressively aware of the facts that the former in a relevant sense deny.

**6. The Real and the Imaginary.** War is also a paradigmatic experience of harm. In my approach to that matter, I examine in chapter 3 the soldier's experience in the light the testimonies collected by Svetlana Alexievich in her book, *Zinky Boys. Soviet Vocies From a Forgotten War*.<sup>12</sup> Here, Soviet soldiers reflect upon their experience in the Afghanistan war, at the end of last century. The relevance of these particulars reports and the way they have been structured in my line of reasoning should be assessed, once again, by their ability to make sense of some other experiences of harm. Some such connections

---

<sup>12</sup> See Alexievich (1992).

are drawn in this essay, although, needless to say, many more are required to confirm my view and also to enrich it with further nuances and connections.

As the soldier arrives in the battlefield, he (I will, in chapter 3, use 'he' instead 'she' as the neutral pronoun whenever I refer to soldiers, given that the soldiers I will be talking about are all men) is shocked by the fact that he can hear the bullet hitting his mate's body and, nevertheless, this experience, both his being shocked and the bullet, appears to him as unreal. The sound of a bullet hitting his mate's body is so *strange* that his mind interprets it as an event within a dream; but why should such a sound be strange at all? Didn't he know about it? Haven't we all heard that noise in movies, read about it in novels? I want to explore such a strangeness in the light of the notion of a human world which was introduced in the previous section. We may regard the human world as hospitable, as homely, as opposed to the battlefield, where the expectations that constitute our confidence in the world are challenged. And this may help us to understand the nature of the soldier's strangeness. When the soldier goes for the first time to the battlefield, his sense of reality is still shaped by the hospitable world he assumed to dwell in before his departure, namely: a world where bullets do not hit human bodies. The bullet hitting a mate's body seems so strange because such a fact *is excluded* from the homely world he has just left behind and still permeates his experience:

“When a bullet hits a person you hear it. It's an unmistakable sound you never forget, like a kind of wet slap. Your mate next to you falls face down in the sand, sand that tastes bitter as ash. You turn him over on his back. The cigarette you just gave him is stuck between his teeth, and it's still alight. The first time it happens you react like in a dream. You run, you drag him, and you shoot, and afterwards you can't remember a thing about it and can't tell anyone anyway. It's like a nightmare you watch happening behind a sheet of glass. Your wake up scared, and don't why.”<sup>13</sup>

At the outset, the soldier experiences the bullet and his mate's dead body as a nightmare, as alien to his conception of what may be a fact, but he soon realizes that such circumstances are facts that belong to a different world. We might thus say that he

---

13 Alexievich (1992), p. 16.

perceives his life as divided into two worlds: home and the battlefield. Some may reply that there is *only one world* and, therefore, that home and the battlefield should rather be construed as two aspects (or, perhaps, regions) of *a single world*. And, therefore, an appropriate description of such unity should show how any two aspects or regions, including home and the battlefield, do relate to each other. The soldier's experience is in need of explanation precisely because it seems to resist that obvious truth: what happens in the battlefield is so strange that he cannot experience it as real, as an aspect of the world that he inhabited before his departure. How can we make sense of this experience? It is true that those who stay at home, away from bullets, *know that* in the battlefield soldiers are injured and killed, and also that bullets make noise as they hit our bodies; but there must be another sense in which they do not know, in which they *are not aware* of what is *actually* taking place, in which they do not understand that the confrontations the news talk about are not fictions, stories invented to entertain, but facts that involve actual injuries and deaths. Those who feel away from the battlefield know that there people kill and die; nevertheless, what happens there comes to their awareness as if such deaths did not really occur and the bullet that hurts belonged to a theatrical representation or a fiction. So, it seems that the distinction between knowing that and being sensitive to may also play a role in our understanding the soldier's experience and, indeed, the parallel distinction between declarative and expressive awareness.

It does not take much time before the soldier realizes that the battlefield is not a dream and, thereby, that the bullet really hurts and dead bodies are hard to be dragged. Yet, this ability to perceive the battlefield as real, comes with a transformation that the soldier will only later become fully aware of, namely: his old self has vanished and a new self has replaced it. How are to make sense of this experience? The moral nature of the divide between the homely world and the battlefield will help us to understand it. In fact, I will argue that the confidence in the world forms a part of his identity and this is shown by the fact that, whoever loses it, regards their life as severely damaged and often as a kind of death. One could then say that the expectations that are constitutive of a human world are so deeply rooted in our identity that, in order to perceive the denial of such expectations as real, one must become a different, and also damaged person.

As soon as the soldier comes back to what he still regards as home, he realizes

that this transformation has taken place, and irreversibly so. His life seems to have changed for ever, there seem to be no actual homecoming because there is a divide between him and those who stayed away from it: there is no way in which they could understand his experience in the battlefield, there is no way in which they would share a common, homely, world. It seems, then, that the world of the soldier who returns is neither the battlefield, since war is not a place to stay, nor his old home, since he has lost his old confidence in the world, but a world poisoned by an endless attempt to come back home.

**7. The Divided Conception of the Self.** Once the experience of harm has been characterized in chapters 2 and 3, there is still the question as to how one should relate to it. This essay is inspired by the hope that by looking carefully into the details of the experience of harm, we may get some significant hints as to how such a question should be answered and the way a certain kind of awareness of that answer may contribute to shaping our lives. An aspect of one's relation to harm on which I agree with the Kantian approach, is that the experience of harm *calls for a response* and, consequently, that the absence of it should be regarded as a denial. Hence, the issue in the remaining chapters will not so much be whether harm requires a response, but instead how we may determine what sort of response it imposes upon us. To this end, I will characterize the Kantian approach in some more detail and argue why I think such an approach is both morally counterproductive and philosophically untenable. In the light of such a challenge, an alternative model will emerge. If at the core of the Kantian approach lies a distrust for one's natural inclinations that leads to *a divided conception of the self*, the point of the attentional model will be to show how such a division might be overcome and the benefits, moral and otherwise, that may be derived from an emphasis on a certain kind of integration..

The Kantian approach places obligations at the center of morality, as we have seen; and this involves, in their view, that they should also be at core of any valuable life. Yet, if an emphasis on obligations is at all required, it is because there is a tendency to disregard them and yield to the impulse of our natural inclinations. It is of much importance, then, that a moral agent should continuously keep her natural inclinations under surveillance; as a result, the agent should conceive of herself as divided between, on the one hand, her rational parts with which she identifies and

constitute what we might call 'her true self' and, on the other hand, her natural inclinations which the true self perceives as alien, that is, as obstacles or accidental favorers in the pursuit of a rational (and moral) life. But what should guide the life of the agent if her natural inclinations are to be distrusted? It seems that only a system of principles that her true self should endorse might do the job. Such principles should not be fixed on the basis of any idiosyncratic features of her, which necessarily belong to the alienated parts of the self, but on those features that all rational agents necessarily share, so that any obligations that could thereby be grounded should concern every other rational agent as well. There is, however, a further demand that such principles must meet, namely: they must be consistent with the disenchanted conception of the world. It follows that both the content of such principles and the foundations that the Kantian provide for them cannot, at any stage, presuppose that there are moral features in the world as it is in itself. In chapter 5, I argue that this requirement cannot be coherently satisfied; whereas in chapter 4, I dispute the role that principles should play in moral deliberation. Let us begin with chapter 5 first.

**7.1 The Limits of Moral Projectivism.** If there are no moral features in the world, then our belief that there are, which seems to be presupposed in our moral judgments, must be regarded as *a projection* upon the world of the ways we are inclined to respond to it. We may thereby derive a projectivist view about morality, which Kantian approaches tend to endorse. A presupposition of that view is that our moral beliefs and judgment do possess *a certain content* which is inadvertently projected upon the world, and also that such a content can be individuated without ascribing moral features to the world, since, otherwise, the projectivist view could not be coherently articulated. In chapter 5, I will, nevertheless, challenge what I regard as the most promising attempt to meet both demands, namely: a dispositionalist account of response-dependent properties. More specifically, I will question David Lewis' dispositionalist account of color terms and argue how my line of objection might apply to any dispositionalist account of moral features. I must stress that, even if my line of argument were correct, one would not thereby be entitled to conclude that moral projectivism is false, since the proposed line of argument has only a transcendental scope and, consequently, the strongest claim that one may derive from it reduces to this: the content of moral judgments can *only* be fixed *if* we assume that moral features are properties of the world. And, needless to say,

the truth of a conditional does not at all make its antecedent true.

Much in my challenge rests on the idea that thick moral terms, like 'humiliating' or 'generous', have a certain *narrative discipline* which is absent in some paradigmatic response-dependent properties like 'disgusting' or 'amusing'. I argue that we may significantly disagree on our moral judgments, but such a disagreement will, in any case, be subject to a certain discipline: it has to be justified by appeal to some other moral features of the action or situation at stake. To put it another way, we may say that moral features form a reticulum and a certain disagreement can only be intelligibly identified as moral if the views in dispute appeal to certain features in that reticulum for a defense. My disagreement with a racist could only be moral if the latter meant to justify her claims by appeal to properties that are morally significant, like differences in intelligence or social abilities, and, consequently, if she were only prepared to change her view by appeal to features of that nature. It would not make sense that someone were racist in the morning, but not so in the evening, unless this change could be explained in that specific way. This discipline parallels that of colors, which also form a reticulum, and is alien to properties like 'disgusting' or 'amusing' about whose individuation no similar constraint is imposed. My point is that we can only make sense of this narrative discipline by assuming that the features at stake are properties of the world in itself. Dispositionalist accounts surely constitute a serious attempt to avoid that conclusion; I will, though, try to show why they are bound to fail.

The narrative discipline of moral features has also some implications as to whether the normativity of moral judgments can be established within the narrow boundaries of moral projectivism. Even if we might, for the sake of argument, assume that the content of such judgments could be fixed within such boundaries; there is still the question as to whether the normative value of moral judgments can be determined in that restricted manner. At this stage, I will discuss a certain Kantian approach, namely; that of Christine Korsgaard and, more specifically, some of her attempts to specify a procedure such that (a) satisfies the metaphysical demands of moral projectivism and (b) tracks our moral intuitions as to what moral judgments one should endorse. This is why she presents her view as a procedural kind of realism: it is *realist* insofar as she takes it that there is an answer to the normative question, but it is *procedural* because it does not rely on the assumption that there are moral properties in the world that the procedure should supposedly track. The normative value of a moral

judgment is ultimately determined by the outcome of the procedure, and not by something else that the procedure might more or less skillfully track. In chapter 5, I will comment on two aspects of her procedure, say, her distinction between private and public reasons, and her notion of a unifying agency. I will argue however that, insofar as we may accept that there are moral dilemmas and conflicts among non-moral values that cannot be resolved in a principled manner, as Korsgaard herself seems to assume, then the closer a procedure is to meet condition (b), the more it departs from condition (a). And this upshot is grounded in the idea that, in order to satisfy condition (b), one must at some point assess the relative significance of any conflicting reasons and it is hard to see how this could be done without appealing to thick moral concepts in a more or less explicit manner. The problem is that, if I am right, such concepts (and the corresponding properties) cannot be individuated within the boundaries of moral projectivism whereby condition (a) would remain unfulfilled.

And, yet, given what natural sciences tell us about how the world is in itself, how could we coherently claim that moral features are properties of it? The answer to this question very much depends on a certain image of what a conception of the world looks like. Following up on Barry Stroud,<sup>14</sup> we may say that projectivism assumes as a *bipartite image* of how our world-view is formed, namely, as a result of the combination of what there is in the world in itself (say, W) and our biological and psychological make up (say, H). From this perspective, it seems that my line of reasoning leads to the rather implausible conclusion that one should at least assume that moral features are properties of W. But this is not the conclusion I would like to endorse; on the contrary, I would rather interpret my case against moral projectivism as a *reductio* of that bipartite image or, more specifically, as an invitation to revise the assumption that such an image is imposed upon us by the discoveries of natural sciences. To put it another way, I should stress that, by claiming that moral features are properties of the world, I am not thereby committing myself to the claim that they should be properties of W. That there is margin of maneuver for such a position is a question that I have explored somewhere else.<sup>15</sup>

**7.2 The Role of Principles.** We may now go back to chapter 4 where I cast doubt on another central assumption in the Kantian approach, namely: that the set of

---

14 See Stroud (1999, ch. 1-2).

15 See Corbí and Prades (2000) and Corbí (2003).

obligations that must govern the life of a moral agent are assumed to derive from some moral principles that every rational agent should endorse. The emphasis on principles comes often with the further conviction that there is a relevant sense in which there is only *one appropriate response* to each morally relevant situation, and that principles should not leave undetermined which such a response should be. These two convictions are so deeply entrenched within our culture that one tends to feel at a loss whenever someone claims that principles do not play a significant role in moral deliberation and, moreover, that, in many significant cases, there is no single response that should count as appropriate, but different responses all of which are inappropriate in one or another respect and such that the perception of this dissatisfaction appears as constitutive of one's sensitivity to harm. And, yet, I will argue for such a view, which we may call '*particularist*'. It is not may stance that we should give up any kind of moral deliberation or that there is no real normative issue involved, no legitimate question concerning the appropriateness of our response, but that such issues should instead be addressed from a perspective quite alien to the Kantian approach, namely: a perspective which distrusts the stress on principles and emphasizes the role of perception and character in moral deliberation.

To this purpose, I will vindicate, in chapter 4, the irreducible moral significance of the experience of *coming across* a particular moral situation P; such an experience involves not only P, but an agent A who is placed in a certain position with regard to that P. It is easy to see that, if an agent's actual position is to be at all morally significant, it must include where he is physically, but also what his projects, engagements and commitments are; a doctor is not in the same moral position with regard to a sick person as a layman. I will thus argue that an agent's *character* (which, following on from Bernard Williams,<sup>16</sup> includes a rather complex variety of elements, some of which possess a clear normative import) is a factor that contributes to fix what situations an agent may come morally across. A crucial point is how an agent relates to her own character when she is engaged in first-person deliberation. I will argue that, in contrast with what the Kantian approach tends to assume, she cannot conceive of the features in her character as mere facts to be contemplated within the set of principles that should guide her deliberation. This will impose, as we shall see, a severe limit to the role that principles may play in first-person moral deliberation.

---

<sup>16</sup> See Williams 1981, p. 5, Williams 1973, p. 115-116.

**7.3 Guilt as an Immoral Emotion.** The true self was supposed to govern on its natural inclinations in the light of a system of principles which every rational agent should endorse and such that their content and the procedure by which they were grounded must be consistent with the disenchanting conception of the world. We come from seeing, however, some rather general reasons why such an ambitious project founders. There is no procedure that could reasonably track our moral intuitions and also honor the demands of moral projectivism, but, even if that problem might be solved, the principles that should as a result be endorsed could hardly guide a moral agent in the complexities of her moral life. That much for the theoretical worries that may assault the Kantian approach, let us now dwell on the reasons why we may think that such an approach is also *morally counterproductive*.

The Kantian approach tends to see guilt as a rather virtuous moral emotion. It is presented as moral because, in contrast with shame, it is claimed to be autonomous, that is, mature agents feel guilty only at the infringement of a principle that she actually endorses; and it is regarded as virtuous because it allegedly acts as quite a severe deterrent in the production of harm and encourages the agent to fulfill her moral obligations. I argue in chapter 4, however, that these two convictions rests on a significantly distorted view about the structure of guilt. Kantian approaches certainly emphasize that one should distinguish guilt as it is experienced in childhood, just out of fear to those who have power upon oneself and, therefore, guided by the principles and instructions that those people might command, from a mature guilt where the agent experiences such an emotion only inasmuch as she may become aware of having infringed a principle that she endorses. Only the latter kind of guilt is presented as properly autonomous and thereby as recognizably moral. It is hard to see, however, how this analysis of guilt can account for some rather uncontroversial features of such an emotion, namely: (a) that is directed against oneself, (b) that the entire value of oneself is called into question, and (c) that a healthy adult agent may feel easily guilty even if she does not think that she is. I will argue, following up on Richard Wollheim,<sup>17</sup> that such features cannot be explained unless we assume that guilt involves an accusation. It is not in general connected to the fact that an agent might know that she has infringed a certain principle that she endorses; she must instead feel, at some level of awareness,

---

<sup>17</sup> See Wollheim (1999, ch 3, Gabriele Taylor and Bernard Williams defend a similar view: cf. Taylor (1985, ch. 3-4) y Williams (1993, ch 4 and endnote 1)

accused by an agency of having infringed a certain principle, no matter whether she endorses it or not. The grounding for such a claim comes both from its capacity to account for features (a)-(c), and also from some phenomenological experiences, which, even though they may not be quite obvious for some people, are extensively present in classic novels and movies, and are also accessible to everyone through dreams or some kind training.

Such accusing agency is the result, as we shall see, of a series of introjections and projections. An external critic A may be introjected by the agent at a certain point for reasons that I will immediately discuss, and give rise to an inner figure where some attributes of A are preserved and others distorted and enhanced. The agent will then, almost inexorably, project such a figure upon another external critic B that she may encounter, and this process should go on in an endless sequence of projections and introjections. The agent will, as a result, conceive of herself as divided into two parts: a severe judge and the condemned self; and such that the latter will interpret any attack launched by an actual external critic as being ultimately launched by her inner severe judge. That much I agree with Wollheim's approach, although I find the details of his account quite unsatisfactory because, if we are to conceive of the self as being really divided, we still need an explanation of why the inner figure is endowed by the condemned self with authority so that anger is not perceived as the kind of response that she deserves. Here it is where the expectations of a human world will come back to the stage. I will argue that, contrary to Wollheim's view, the accusing inner figure is not set up in an attempt to keep the external critic under control, but as a manifestation of her extreme power over the agent. On this view, it is the perception of the external critic's power that leads the agent to endow her with authority; and here the mechanisms of self-deception parallel those that may lead third agents to side the torturer and disregard any evidence that may suggest that the latter's action was monstrous and unjustified. The child who depends in every respect upon her parents would rather believe that their actions are human and justified, since, otherwise, she might be paralyzed with fear and helplessness. Hence, whenever the child feels accused or attacked, she would rather think that she has done something wrong and, consequently, assume that, if he had behaved properly, he would not have been attacked. This will provide her with a sense of protection and control over her own life; and here it is how the external critic's attack is reverted against oneself, that is, as a

result of endowing the powerful with authority. The internalization of any such external critic should be proportional to the agent's fright at being caught in the midst of a course of action that the external critic might disapprove of. Yet, if guilt is set up in an attempt to prevent the attack of the powerful by adapting to her wishes, once they have duly adorned with the beauties of reason, then one can hardly expect that, in case of conflict between the victim's claim and the dictates of the powerful, guilt should favor a moral response as opposed to mere submission to what the agent will almost inevitably perceive as the legitimate demands of the authorities. Only exceptionally it may occur (as in Claude Eatherly's case, as we shall see) that an agent's sense of guilt should lead her to side the victim and confront the pressure that inexorably third agents and the executioners exercise. This capacity to distance oneself from such a pressure should count as a major and exceptional achievement and not as a likely result of the experience of guilt.

The divided conception of the self which Kantian approaches present as a conceptual tool to keep one's natural inclinations under control and let one's life be strictly guided by the principles of morality, emerges, in the light of this analysis, as a theoretical rationalization of the way in which guilt structures the self. For the condemned self has all reason to keep its inclinations under control, since, otherwise, it may be unable to promptly adapt to the exacting demands of its severe inner judge; and the emphasis on principles should, instead, be regarded as a projection on the side of the condemned self of her need to believe that the severe judge is not arbitrary, but fair. So, we have some reason to think that a life guided by such a conception, even if meant to defend morality, will in some relevant circumstances favor the production of harm. We should not expect, as a result, that third agents whose lives have been shaped in this manner will meet the victim's legitimate demands, except when the required response would nicely fit with the authorities' needs and demands. It seems then that, if the moral question 'How should one relate to harm if the human world is to be restored?' has any hope of obtaining an appropriate response on the side of third agents, it should be by moving away from such a model and cultivating a kind of attitude that might lead to a more integrated experience of the self.

**8. The Moral Question and Self-Knowledge as an Achievement.** Richard Moran proposes a rather novel and illuminating account of how first-person authority may be

consistent with self-knowledge as an achievement. His approach emphasizes two aspects that have been neglected within the Cartesian model, namely: *the role of practical commitment in self-knowledge* as well as *the permeability of psychological dispositions to deliberation*. Both aspects stress how self-knowledge is constitutively connected with an agent's capacity to shape his life in a reflective manner and, consequently, with her ability to articulate a certain response to harm. There is then some hope that a proper discussion of Moran's approach to self-knowledge may shed some light on how the moral question should be addressed.

The starting point in Moran's approach is the experience of an agent as being divided between her deliberative abilities and her psychological dispositions. This contrast is clearly perceived in the plight of the akratic gambler, who despite his best intentions to the contrary, can't help gambling; he somehow observes his passion from a distance, as he might contemplate someone else's behavior or the nuances in the bark of a tree. There is, then, a sense of alienation because his psychological dispositions are significantly impervious to his intentions and deliberations. The issue is how his psychological dispositions could become more permeable to his reflective attitudes and, more specifically, whether there is a certain kind of self-knowledge that may play a relevant role in such a process. This seems to be a crucial assumption in the development of psychoanalytic therapy, and Moran emphasizes that his approach may serve to identify the specific kind of awareness that may possess such healing effects. He appeals to the distinction between the deliberative and the theoretical attitudes to this purpose, as well to the associated notions of permeability, transparency, and avowal. In chapter 6, I will argue however that, even if these notions point in the right direction, they are in need of substantial modification if they are to apprehend the relevant kind of awareness. And the ultimate reason for such an inadequacy will be that Moran's proposal, like Kantian approaches, still presuppose the distinction between a true (and deliberative) self and its psychological dispositions. It is true that he is not of the view that an agent should just keep her natural inclinations at bay, but rather that she should learn to *tame* them, that is, that they should be *permeable* to what her deliberations dictate. But still there is a part of oneself that should obey and such that it may be alienated, and another one which governs and constitutes the true self. I will argue though that, only if we get rid of this divided conception of the self, will we be able to understand why a certain kind of awareness may have a significant therapeutic

value. To this purpose, I will take advantage of Bernard Williams' notion of acknowledgment,<sup>18</sup> as well as Simone Weil's distinction between two notions of necessity,<sup>19</sup> to introduce the notion of 'receptive passivity' *a kind of attitude an agent may adopt towards herself* and such that it will produce some repairing effect by favoring a more integrated experience of herself. The exploration of such an attitude will derivatively suggest a more integrated conception of how an agent's deliberative capacities and her psychological dispositions relate. The notion of receptive passivity was precisely introduced to make sense of such an attitude, that is, an attitude that is meant to bridge the gap between an agent's deliberative attitude and her psychological inclinations. The experience of the graceful dancer served us a model to ascertain what such an attitude may consist of.

Submission to the order in the music is essential to the dancer's experience. Although submission in this case, as opposed to what happens in the akratic gambler's case, does not come with degradation, but may significantly contribute to her development and flourishing. To this end, the dancer should not try to move her body in the light of some specific rules, but rather pay attention to the music, as well as to the emotions and bodily experiences that may derive from it, and then let herself go. What she might learn by this means is not to be detailed in a report, as a mere declarative kind of awareness she might acquire, but lead her to move and experience her body in a graceful and expressive way. The music, her emotions, her bodily experiences and movements form thereby a *gestalt*, a unit of intelligibility, which may be qualified as more or less graceful, or intriguing, or creative, and also as expressing sorrow, or cheerfulness, or curiosity, and so on. It is essential to this point that we should not interpret the dancer's attitude in the light the image of as the severe judge who wants to keep all her actions under control; after paying attention to the different aspects that her activity may involve, there is a stage at which she must *trust herself and let herself go* in the hope that she may thereby perceive and express the order in the music more deeply and gracefully. And, indeed, a similar suggestion applies to the way in which attention is focused on the piece of music, since there are ways which are rigid, detached or fearful, and others that are more open and exploratory as to what she may encounter. It is the latter kind of attention that I associate with receptive passivity and whose cultivation may increase the dancer's sensitivity, that is, her capacity to become

---

18 See Williams (1981, ch. 9-10; 1985, ch. 10 and postscript; 1993, ch. 4; and 2002, ch. 8).

19 See Weil (1963), 38-44.

expressively aware of the order in the music, in her bodily experiences and movements.

The overall cultivation of this attitude may have a therapeutic effect by bringing about a more integrated perspective upon oneself. Receptive passivity is in itself an experience of integration, but there is a less obvious, although most significant, contribution of the exercise of such a capacity, namely, a gradual alteration of the landscape of inner figures in the light of which the agent leads her life. If the divided conception of the self, encouraged a view of one's true self as a severe judge who should keep one's natural inclinations under control; the mere exercise of receptive passivity upon one's experiences and the world expresses and shapes a friendlier attitude toward oneself. One thereby learns to look at oneself in a more trustful manner. One's emotions and natural inclinations no longer appear as a threat that the effort of will should keep at bay, but as clues as to what kind of ingredients a meaningful life should have and how they may be reasonably concocted. This is, indeed, an endless process of discernment where attention to the nuances and commonalities of every situation play a crucial role, although such an activity is not experienced as a burden, that is, as a task that one must perform before the real life takes place, since it is life itself that develops in that attentive and discerning, but not detached, way. In chapter 7, I will though return to a less joyful aspect of human life, the production of harm, and see whether the cultivation of receptive passivity may be of significance to articulate a response to the moral question. Given the asymmetries that are constitutive of the experience harm, the issue should briefly addressed from three different perspectives: the victim, the torturer, and third agents.